DEVELOPMENT OF AN INTELLIGENT SYSTEM TO MONITOR VISUAL ATTRIBUTES OF FOOD IN REAL-TIME DURING FLUIDIZED BED DRYING

BY

ANTHONY CHUKWUEMEKA IHEONYE

Department of Bioresource Engineering

Faculty of Agricultural and Environmental Sciences

McGill University

Sainte-Anne-de-Bellevue, Québec, Canada

October 2022

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of

Doctor of Philosophy (Ph.D.)

in

Bioresource Engineering

This Thesis is dedicated to my Parents,

for their unfailing love, support, and encouragement.

ABSTRACT

Most food dryers are controlled by monitoring real-time changes in moisture content of the food. However, the food property that attracts consumers or dissuades them from purchasing dried products is not moisture content; instead, it is the visual appearance. To significantly reduce food loss and waste, we need dryers that monitor moisture content and visual appearance. Current fluidized dryers are unable to monitor visual attributes in real-time. Hence, this research attempts to answer the question, ***"is it possible to monitor real-time changes in color, texture, and size of food during fluidized bed drying?"*** To answer this question, an image acquisition system, a classical computer vision model, and two deep learning models were built. These models were developed using images of green peas captured during fluidized bed drying experiments conducted at 50, 55, and 60^{o}C. The operation of the classical model and the first deep learning model (Unet-Xception) involved two steps – image segmentation and visual attribute prediction using the segmented output. While the second deep learning model (Xception) had only one operational step - visual attribute prediction using unsegmented images.

Segmentation results for the first two models were evaluated using Mean Intersection-over-Union (MIoU) and Categorical accuracy. The classical model was overwhelmed by challenges that negatively affected its segmentation quality. Some of these challenges include overlapping peas within the image, unequal distance of the peas from the camera, and non-uniform changes in color and texture across the surface of individual peas. These challenges were overcome by Unet-Xception, resulting in an improved MIoU of 0.9464. In addition to the higher MIoU, Unet-Xception also identified objects of interest (OOI) that were missing within the ground truth masks. This discovery implied that both deep learning models would perform better in practice, than during model evaluation. This expectation turned out to be true.

Color was monitored by predicting CIE L^* a^* b^* values. Surface texture was measured by predicting contrast, correlation, entropy, and homogeneity. While equivalent diameter, ferret diameter, filled area, perimeter, and roundness were used to monitor size. Unet-Xception and Xception significantly outperformed the classical model. Furthermore, Unet-Xception produced the smoothest real-time trend.

The coefficient of determination (R^2) achieved by each model is reported in the form [*classical model, Unet-Xception, Xception*]. *a** and *b** indices were very good predictors of color, with model R^2 values of [0.7719, 0.8851, 0.8675] and [0.8694, 0.8981, 0.9064], respectively. Homogeneity was the best parameter for monitoring texture, with R^2 values of [-0.8660, 0.8017, 0.7845]. Finally, the following values were obtained for size: [-1.0278, 0.7926, 0.8748] for equivalent diameter, [-1.6417, 0.8482, 0.8802] for ferret diameter, [-0.3663, 0.8391, 0.8901] for filled area, and [-1.5425, 0.825, 0.8421] for perimeter.

This adaptable and novel solution proves that it is possible to monitor real-time changes in visual attributes of food during fluidized bed drying. Incorporating this solution into current food dryers could lead to consistent product quality and significant reduction in food waste and loss.

RÉSUMÉ

La plupart des séchoirs alimentaires sont contrôlés en surveillant les changements en temps réel de la teneur en humidité des aliments. Cependant, la propriété alimentaire qui attire les consommateurs ou les dissuade d'acheter des produits séchés n'est pas la teneur en humidité, mais plutôt l'aspect visuel. Pour réduire considérablement les pertes et le gaspillage alimentaires, nous avons besoin de séchoirs capables de surveiller à la fois la teneur en humidité et l'apparence visuelle. Les sécheurs fluidisés actuels sont incapables de surveiller les attributs visuels en temps réel. Par conséquent, cette recherche tente de répondre à la question « est-il possible de surveiller en temps réel les changements de couleur, de texture et de taille des aliments pendant le séchage en lit fluidisé ? » Pour répondre à cette question, un système d'acquisition d'images, un modèle de vision par ordinateur classique et deux modèles d'apprentissage en profondeur ont été construits et évalués. Ces modèles ont été développés à partir d'images de pois verts capturées lors d'expériences de séchage en lit fluidisé, menées à 50, 55 et 60^{o}C. Le fonctionnement du modèle classique et du premier modèle d'apprentissage en profondeur (Unet-Xception) impliquait deux étapes : la segmentation de l'image et la prédiction des attributs visuels à l'aide de la sortie segmentée. Alors que le deuxième modèle d'apprentissage en profondeur (Xception) ne comportait qu'une seule étape opérationnelle : la prédiction des attributs visuels à l'aide d'images non segmentées.

Les résultats de segmentation pour les deux premiers modèles ont été évalués à l'aide de l'intersection moyenne sur l'union (MIoU) et de la précision catégorielle. Le modèle classique a été incapable de résoudre des défis qui ont affecté négativement sa qualité de segmentation. Certains de ces défis incluent des pois qui se chevauchent dans l'image, une distance inégale entre les pois et la caméra et des changements non uniformes de couleur et de texture sur la surface des pois individuels. Ces défis ont été surmontés par Unet-Xception, ce qui a permis d'améliorer le MioU de 0,9464. En plus du MIoU plus élevé, Unet-Xception a également identifié des objets d'intérêt (OOI) qui manquaient dans les masques de vérité terrain. Cette découverte impliquait que les deux modèles d'apprentissage en profondeur fonctionneraient mieux dans la pratique que lors de l'évaluation du modèle. Cette attente s'est avérée vraie.

La couleur a été surveillée en prédisant les valeurs CIE L^* a^* b^*. La texture de surface a été mesurée en prédisant le contraste, la corrélation, l'entropie et l'homogénéité. Alors que le diamètre équivalent, le diamètre du furet, la zone remplie, le périmètre et la rondeur ont été utilisés pour

surveiller la taille. Les résultats obtenus ont indiqué qu'Unet-Xception et Xception ont largement surpassé le modèle classique. De plus, Unet-Xception a produit la tendance en temps réel la plus fluide.

Le coefficient de détermination (R^2) atteint par chaque modèle est reporté dans l'ordre suivant [modèle classique, Unet-Xception, Xception]. Les indices a* et b* étaient de très bons prédicteurs de couleur, avec des valeurs R^2 pour les modèles de [0,7719, 0,8851, 0,8675] et [0,8694, 0,8981, 0,9064], respectivement. L'homogénéité était le meilleur paramètre pour surveiller la texture, avec des valeurs R^2 de [-0,8660, 0,8017, 0,7845]. Enfin, les valeurs suivantes ont été obtenues pour la taille : [-1.0278, 0.7926, 0.8748] pour le diamètre équivalent, [-1.6417, 0.8482, 0.8802] pour le diamètre du furet, [-0.3663, 0.8391, 0.8901] pour la surface remplie, et [-1.5425, 0,825, 0,8421] pour le périmètre.

Cette solution adaptable et innovante prouve qu'il est possible de suivre en temps réel l'évolution des attributs visuels des aliments, lors du séchage en lit fluidisé. L'intégration de cette solution dans les séchoirs alimentaires actuels pourrait conduire à une qualité de produit constante et à une réduction significative du gaspillage et des pertes alimentaires.

ACKNOWLEDGEMENTS

My heartfelt thanks go to the Blessed Virgin Mary who has made the completion of this work possible through Her Motherly intercession on my behalf, Her endless stream of graces lavished on me, and the opportunity She granted me to learn from amazing intellectuals and friends, some of whom are listed below. One of these intellectuals is my supervisor Professor Vijaya Raghavan.

I would forever be indebted to Professor Raghavan for his steadfast and all-round support, encouragement, patience, and academic guidance. Thank you so much Professor for all our discussions, for always providing constructive feedback, and for numerous opportunities that has helped me become a better researcher and person. You enhanced my knowledge and experience in food drying technology, and when it was time to integrate computer vision and deeplearning to our research, you made it possible for us to collaborate with one of the best researchers in the field – Professor Frank P. Ferrie.

I got to know Professor Ferrie while I was starting out in Deep Learning (DL). He graciously invited me to be a member of his Artificial Perception Lab (APL), provided direction and excellent resources that deepened my knowledge of deep learning. In APL, I had the opportunity to work with and learn from kind and excellent DL and Computer vision researchers including Dr. Olivier Cormier, Michael Smith, Andrew Phan, and Zhonghao Zhao. I would forever remain grateful to Professor Frank and the members of APL. They made my DL journey fulfilling and memorable.

Professor Valérie Orsat improved my understanding of physico-chemical properties of biological products. Thank you so much Professor Orsat for your engaging lectures, patience, unwavering support, encouragement, and priceless feedback. Your lectures and advice helped me develop the computer vision system to monitor various visual properties of food. I would never forget your kindness.

I am so thankful to Mr. Yvan Gariepy and Dr. Darwin Lyew for their unflinching support. Every graduate student in Dr. Vijaya Raghavan's Post-Harvest Technology Laboratory has beautiful stories on how Yvan and Darwin helped them succeed in their research work. Thank you so much for your technical support. You ensured that all the materials I needed for my research was provided. Thanks for the countless brain-storming sessions. You were always present during the storm. I also deeply acknowledge the priceless contributions from the faculty and staff of the

Department of Bioresource Engineering and the Macdonald Campus office of Graduate and Post-doctoral Studies (GPS).

I am profoundly grateful to Dr. C. E. Chinweze. He stoke the passion for agricultural engineering in my heart, during my final year of undergraduate studies. I attribute my drive to developing intelligent food drying systems to him. His final year undergraduate course on food drying and storage filled me with inspiration and renewed purpose. Thank you so much Dr. Chinweze for your continuous support, advice, and motivation during my research work.

Completing this thesis would not have been possible without the help of my Lab mates including Dr. Sellam Perinban, Suhasini Srinivasan, Angela Arevalo, Shuyao Wang, Xin Dong and Jolvis Pou. I appreciate all our productive discussions and your constructive feedback. My research work cut across food drying technology, computer vision technology, and deep learning. While progressing through each research stage, I routinely met challenges. During such times, I was fortunate to have priceless, thorough and most times perceptive-shifting discussions with many kind-hearted and excellent researchers, including Dr. Yvonne Goh, Dr. Karolin Dietrich, Dr. Christopher Nzediegwu, Dr. Nafiseh Yavari, Dr. Samuel Ihuoma, Dr. Okenna Obi-Njoku, Dr. Ogan Mba, Dr. Christopher Kucha, Dr. Kosoluchukwu Ekwunife, Mfon Essien, Amanda Boatswain, Joba Purkaystha, Aisha Abudu, Babatunde Onadipe, Nastaran Alizadeh, and Shafieh Salehinia. I am so grateful for the time we shared, for your support and helpful feedback.

Beyond McGill University, I also had very good friends who supported me with their time, kind words, encouragement, motivation, and generosity. They made the journey worthwhile and memorable. Thank you so much to Maureen Pouliot, Hélène Gosselin, Daiva Blynas, Kris Gravenor, Howard Fox, Benson Chukwu, Oumama Alaoui, Deepikasai Bhemarasetty, Ali Asgar Rampurwala. I deeply appreciate the friendship of Mr. & Mrs Gaston LeBlanc and Family. Every time I spent with them was like heaven on earth. The peace, joy and love they radiated made the day beautiful.

Finally, I will forever be indebted to my Parents Professor & Mrs A. E. Iheonye, to Engr. & Mrs. Lucky Eke and to Mr. & Mrs Cyracius Amuchie. Thank you so much for your daily prayers, financial support, and words of advice, inspiration, and encouragement, I am so grateful. To my lovely sisters Chinwendu Eke, Obioma Onuoha, Chinonye Chijioke, Onyinyechi Iwuala,

Assumpta Iheonye, and Eucharia Iheonye, thank you so much for your unwavering support, prayers, and encouragement, I love you all so much.

THESIS FORMAT

This thesis is submitted in the format of papers suitable for journal publication. This thesis format has been approved by the Faculty of Graduate and Postdoctoral Studies, McGill University, and follows the conditions outlined in the Guidelines: Concerning Thesis Preparation, which are as follows:

“As an alternative to the traditional thesis format, the dissertation can consist of a collection of papers of which the student is an author or co-author. These papers must have a cohesive, unitary character making them a report of a single program of research. The structure for the manuscript-based thesis must conform to the following:

1. Candidates have the option of including, as part of the thesis, the text of one or more papers submitted, or to be submitted, for publication, or the clearly duplicated text (not the reprints) of one or more published papers. These texts must conform to the "Guidelines for Thesis Preparation" with respect to font size, line spacing and margin sizes and must be bound together as an integral part of the thesis. (Reprints of published papers can be included in the appendices at the end of the thesis).

2. The thesis must be more than a collection of manuscripts. All components must be integrated into a cohesive unit with a logical progression from one chapter to the next. In order to ensure that the thesis has continuity, connecting texts that provide logical bridges between the different papers are mandatory.

3. The thesis must conform to all other requirements of the "Guidelines for Thesis Preparation" in addition to the manuscripts. The thesis must include the following:

a) A table of contents
b) An abstract in English and French
c) An introduction which clearly states the rational and objectives of the research
d) A comprehensive review of the literature (in addition to that covered in the introduction to each paper)
e) A conclusion and summary

4. As manuscripts for publication are frequently very concise documents, where appropriate, additional material must be provided (e.g., in appendices) in sufficient detail to allow a clear and precise judgment to be made of the importance and originality of the research reported in the thesis.

5. In general, when co-authored papers are included in a thesis the candidate must have made a substantial contribution to all papers included in the thesis. In addition, the candidate is required to make an explicit statement in the thesis as to who contributed to such work and to what extent. This statement should appear in a single section entitled "Contributions of Authors" as a preface to the thesis. Since the task of the examiners is made more difficult in these cases, it is in the candidate's interest to clearly specify the responsibilities of all the authors of the co-authored papers”.

CONTRIBUTIONS OF AUTHORS

These are the list of manuscripts (published and prepared for submission) and conference proceedings whose content are used for writing this thesis.

1. **Iheonye, A. C**; Gariepy, Y.; Raghavan, V. Computer vision for real-time monitoring of shrinkage for peas dried in a fluidized bed dryer. Drying Technology 2019, 1-17. DOI: 10.1080/07373937.2019.1649277. (Published) [Chapter 3]
2. **Iheonye, A**.; Raghavan, V.; Ferrie, F.; Gariepy, Y.; Orsat, V. Monitoring visual properties of food in realtime during food drying – a review. (Submitted for publication) [Chapter 2]
3. **Iheonye, A**.; Ferrie, F.; Raghavan, V.; Gariepy, Y.; Orsat, V. A Two-pass deep learning method for monitoring visual attributes of food during fluidized bed drying. (Submitted for publication) [Chapter 4]
4. **Iheonye, A**.; Raghavan, V.; Ferrie, F.; Orsat, V.; Gariepy, Y. An intelligent single-pass system for monitoring visual attributes of food during fluidized bed drying (To be submitted) [Chapter 5]
5. **Iheonye, A**.; Raghavan, V.; Ferrie, F.; Orsat, V.; Gariepy, Y. Developing a two-pass and a single-pass AI-driven system for monitoring real-time changes in visual attributes during fluidized bed drying. Presented at the CSBE/SCGAB 2022 Annual Conference, Charlottetown, Prince Edward Island, Canada, 2022. [Chapter 4 & 5]
6. **Iheonye, A**.; Ferrie, F.; Orsat, V.; Raghavan, V. Deep learning-based method for real-time monitoring of visual attributes during fluidized bed drying. Presented at the IDS 2022 – 22nd International Drying Symposium, Worcester, Massachusetts, USA, 2022. [Chapter 3 & 4]
7. **Iheonye, A**.; Gariepy, Y.; Raghavan, V. Real-Time Monitoring of Fluidized Bed Drying Using Computer Vision. Presented at the 2019 ASABE Annual International Meeting, Boston, Massachusetts, USA, 2019. [Chapter 3 & 4]
8. **Iheonye, A**.; Gariepy, Y.; Raghavan, V. Computer vision for real-time monitoring of shrinkage during fluidized bed drying. Presented at the CSBE/SCGAB 2019 Annual Conference, Vancouver, BC, 2019. [Chapter 3 & 4]

The research work reported here was conducted by Anthony Chukwuemeka Iheonye and supervised by Professor Vijaya Raghavan of the Department of Bioresource Engineering, at McGill University, Quebec, Canada. The contributions to this research work could be grouped into five broad parts:

1. Development of an image acquisition system
2. Development of a classical computer vision pipeline
3. The use of a classical CV system to monitor real-time and offline visual attributes of green peas during Fluidized bed drying.
4. Development of deep learning solutions for monitoring visual attributes of green in real-time during fluidized bed drying.
5. Reviewing the manuscripts

The first three parts were conducted at Professor Vijaya Raghavan's Post-Harvest Technology Laboratory located at the Macdonald campus of McGill University. While the fourth part of the research was mostly conducted at Professor Frank Ferrie's Artificial Perceptron Laboratory, Centre for Intelligent Machines, McGill University.

Professor Vijaya Raghavan and Mr. Yvan Gariepy provided scientific advice for developing the online and offline image acquisition systems. Professor Valérie Orsat, Professor Vijaya Raghavan, and Mr. Yvan Gariepy provided scientific advice on selecting the visual properties that were monitored and on developing the classical computer vision pipeline for monitoring visual attributes of peas in real-time during fluidized bed drying.

Professor Frank Ferrie and Professor Vijaya Raghavan provided scientific guidance on developing methodologies to acquire ground truth data from real-time images of green peas. Their scientific advice was also very significant when developing, evaluating, and analyzing the performance of the two-pass and single-pass deep learning solutions. Both solutions were mostly developed using high-speed computers located at Professor Frank Ferrie's Artificial Perceptron Laboratory, while the rest were developed on ComputeCanada HPC cluster using the access provided by Professor Vijaya Raghavan. Finally, both Professor Frank Ferrie and Professor Vijaya Raghavan were directly involved in reviewing and editing the manuscripts.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

Think about the last time you visited the store to buy fresh foods. Can you remember the last time you bought apples? How did you decide which apple to pick? Like most consumers, we inspect the visual attributes of the fruit – the color, size, texture, and shape. Occasionally, we even hold the apple within our palm, pressing it a bit with our thumb to ascertain how hard it is. At the end, if we are attracted by the fruit's appearance, we purchase it, else we walk away. Most times our decision to buy a particular food product is not only based on one physical feature, but a combination of two or more features. The same criteria apply to dried foods. Consumers are attracted to dried food products that are as close as possible to their fresh counterpart, both in physical appearance, and nutritional content. During drying, the moisture within the food is reduced to a safe level. By so doing, the shelf-life and stability of the food increases. The drying process decelerates biochemical, microbial, and chemical reactions that cause food spoilage. With the loss of water, the weight of the product is reduced, which in turn improves product handling and reduces shipping cost.

It is estimated that by the year 2050, the world population would grow to ten billion people. Food drying would play a pivotal role in ensuring food security. However, before we get to 2050, food drying is needed today more than ever before, in order to meet the demands of one billion people who would go to bed tonight hungry, and over 66 million primary school children in developing countries who would walk long distances to school tomorrow without breakfast.[1,2] Drying is needed because although there are so many in dare need of food, 1/3 of food produced annually is either lost or wasted. [1-5]

In developing countries, food is mainly lost. A huge percentage is lost between the time of harvest to the time the food gets to the processing facility for drying [1,3,4] For the ones that make their way into the facility in good condition, a significant portion is still lost during the drying process, due to lack of both adequate drying equipment and specialization.

Over 80% of food produced in developing countries is done by subsistence farming. The farmers produce high quality fresh harvest due to their extensive years of experience, and improved skills acquired by partnering with local and international agricultural organizations, such as FAO and

IITA. However, three months down the line, the fresh produce are scarce in the market due to spoilage. By then, most of the unattractive dried substitutes would have been infested by mold. On rare occasions, you would find a couple of farmers who can afford the cost of appropriate drying equipment, but without thorough understanding on how the drying process interacts with the food material, most farmers end up with substandard products.

Food drying is a complicated and multi-factored thermal process. The process is unsteady, highly nonlinear, and consists of successively interconnected phases.[5] If we could develop a system that adequately takes care of this complexity, while giving the farmer the breathing room to concentrate solely on food production, then we would have contributed immensely towards solving the problem of food loss and insecurity in developing countries. Furthermore, this improved system could drastically reduce the problem of food waste in developed countries such as Canada, the United States of America, and many countries in Europe.

The end goal of drying is to reduce the moisture content to a safe level, with minimal change to the appearance and nutritional quality of the food. However, this is achieved in developed countries. Yes, most drying systems in developed countries succeed at bringing the moisture content of the food to a safe level, however such systems are not optimized to consistently produce dried foods that meet the physical and nutritional demands of the market. The food processor knows that food stores would only accept products that are attractive to consumers, hence processors only supply their best products, and the rest (not withstanding that their moisture content are at safe levels) are either converted to animal feed or buried in landfills as waste. Even while in the store, some of the dried foods are wasted. Customers choose only the best, while rejecting the ones that are unattractive. Also, even after purchase, while the products are still in good condition, some are routinely dumped into the garbage, due to unappealing appearance, or when the consumer finds a more attractive one on their next visit to the store.

To solve the problem of food loss and waste, our drying systems should be such that they can consistently produce dried foods that have long shelf-life, are physically attractive, and high in nutrient content. However, this has not been the case, because it has been difficult (and to some extent impossible) to fully understand the complex underlying mechanisms that take place during the drying process. We know that one of those processes is heat and mass transfer - heat moves into the product, and because of that, the moisture in the food leaves as vapor. However, we are

yet to fully understand other processes, such as simultaneous coupled momentum, several phases of transformations; various chemical and bio-chemical reactions taking place within the product; component migration; and sometimes abrupt surface hardening. All these non-linear and interconnected processes play a role in the visual appearance of the final product. To control the quality of the product, we need drying systems that take all these processes into consideration. Incidentally, this goal is yet to be achieved for fluidized bed dryers.

Current drying systems were designed and optimized to monitor and control process parameters that directly affect the heat and mass transfer process. These parameters are linked to the drying kinetics of the food (the relationship that shows the reduction in moisture content or moisture ratio over time). For instance, current convective dryers are controlled by monitoring the temperature, velocity, humidity, and pressure of the hot air, so that the moisture within the food is reduced to a safe level. In so doing, we end up with a product that has a long shelf life. But as we discussed earlier, focusing on moisture content alone is not enough. We need drying systems that could also monitor and control real-time changes in visual attributes of the food. Once this is achieved, the quality of the food would always meet the consumers expectation, and food loss and waste may soon become a thing of the past.

To achieve this goal, we need models that can identify accurate relationships between inputs and outputs of the drying process.[6] For instance, during convective drying, we would need a relationship that connects both process parameters and drying kinetics to the changes in color, texture, size, and shape of food product. With such a model we could further optimize the drying process, which in turn would lead to effective and timely product quality control. However, due to the complex and nonlinear processes that occur during drying, it can be cumbersome, computationally expensive, and sometimes practically impossible to model such relationships using mathematical, statistical, numerical, and analytical techniques.[5] Fortunately, all these setbacks are disappearing because of advances in technology and AI.

With the introduction of high-speed computing infrastructures, improved computer vision (CV) technology and the emergence of deep learning, it is now possible to develop such relationships. In recent times, numerous researchers have taken advantage of these advanced techniques to improve the drying process for convective dryers,[7-16] microwave-vacuum dryers,[8,17] and osmotic-convective dryers,[8,18,19] to mention a few. For instance, by integrating computer vision

technology, some researchers were able to monitor real-time changes in the moisture content and some visual attributes of foods.[9,20] When using CV technology, a camera is placed over the stationary food, capturing images of the food as it dries. Then, with appropriate CV algorithms, information on the color, texture, size, or shape of the food are obtained.

Unfortunately, there are still two efficient drying systems that are yet to *fully* benefit from these advanced techniques. They are the fluidized bed and spouted bed drying systems. These dryers are used both in the food and pharmaceutical industries due to their high rate of heat and mass transfer, uniform temperature distribution, high thermal efficiency, effective material mixing and handling; and their suitability for processing heat sensitive solids.[21-23] On a positive note, there have been two novel applications within the pharmaceutical industry, where CV technology was used to monitor the size of pharmaceutical products during fluidized bed drying. Kadunc et al. [6] developed a CV system to monitor the thickness of pharmaceutical tablets during spray coating in a fluidized bed dryer. Another company, Innopharma Technology®, developed a CV solution to monitor the size of agglomerates during the granulation process. In both cases, the systems were effective in monitoring the size of the pharmaceutical product in real-time

However, to the best of our knowledge, there is no published literature in which CV technology has been used to monitor visual attributes (color, texture, size, and shape) of food products, in realtime, during fluidized bed drying. Hence, the overall objective of this research work is to ***"develop an intelligent system to monitor visual attributes of food in real-time, during fluidized bed drying"***. To achieve this goal, there are a couple of challenges that need to be resolved in order to use CV technology and deep learning techniques to monitor visual attributes of food during fluidized bed drying. Let us discuss these challenges.

1.1. Unresolved Challenges during fluidized bed drying

To appreciate these challenges better, before discussing them, let us briefly describe how CV technology is used to monitor visual attributes of food, within other drying systems that are not fluidized bed dryers. Afterwards, we can highlight the conditions unique to these drying systems that make it possible to integrate CV technology. Figure 1.1 is a sketch of a convective dryer. During drying operations, the food materials are placed on a perforated tray to permit airflow

across the food surface. To monitor the size, color, and texture of the food, using CV technology, our first step would be to capture an RGB image containing the food materials.

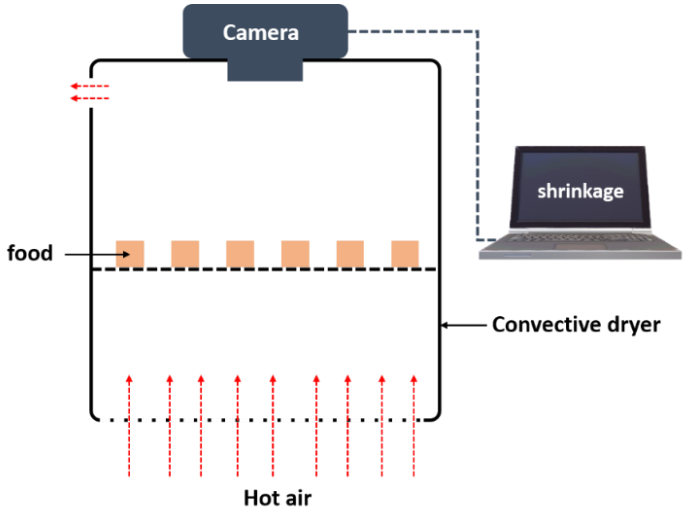


Figure 1.1 Convective drying system monitored using CV technology

After image acquisition, the image is fed to CV algorithm. The first task by the CV algorithm is performed on the image. The algorithm segments the image by isolating the food (the object of interest) from the flat tray (the background). Isolating the objects of interest would enable us to compute the size of individual food material. To isolate the food, the RGB image is converted to a grayscale image, which is a matrix of pixel intensity values ranging from 0 to 255. Where 0 is black, 255 is white, and the numbers in-between are different shades of gray. For the isolation to work, there must be significant contrast between the color of the food product and that of the tray holding them, so that when the RGB image is converted to grayscale, the food product and the tray would display different shades of gray. To extract the object of interest (food) from the background, we apply a threshold value (intensity value) that separates both of them. The result would be a binary image made up of pixels values of 0s (black) and 1s (white). The white region on the image represents areas on the RGB image that were occupied by individual food products, while the black region represents the background (the drying tray). To compute the size of the food, we sum the number of pixels within each region of interest. By repeating these steps, the average size of the food can be monitored over time. To compute the accurate size of individual products it is crucial that they lay separate from each other. The positions occupied by the white region within the binary image, correspond to the positions occupied by the food product on the RGB image. Hence, to compute the color and texture of the food, we determine the average color

and texture within those regions of interest. There are key features unique to convective dryers that make it possible to monitor visual attributes of food using CV technology.

First, the food product is static throughout the drying process. This makes it possible to capture and analyze images of the same products over time. Next, the products are at the same distance from the camera (Figure 1.1). Due to the uniform distance, it is easier to determine their true size. Thirdly, the products are spaced out, there are no overlaps, hence, it is straight-forward to distinguish the food from the background. Most of these favorable features are absent during fluidized bed drying.

During fluidized bed drying, the food products are in continuous motion, tumbling up and down throughout the drying process (Figure 1.2a). So, when you capture an RGB image of the food, such as the peas shown Figure 1.2b, you would notice a lot of overlapping peas, which makes it difficult to isolate each of them from the background. If we cannot isolate them, we cannot compute their visual parameters using CV algorithms

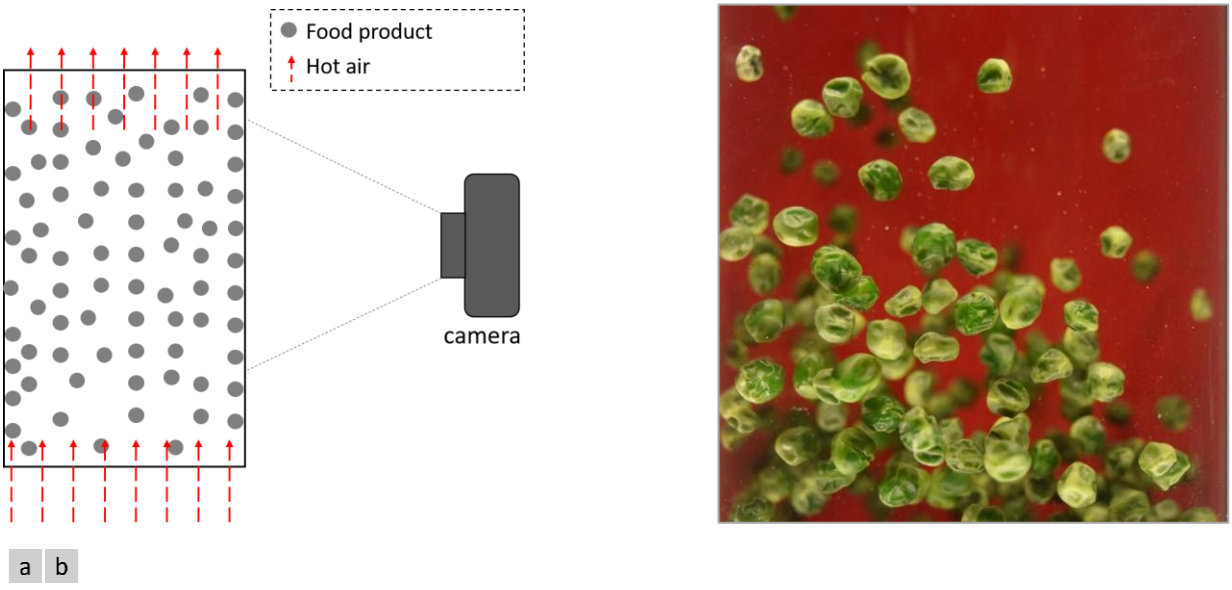


Figure 1.2 (a). Uneven distance for food products from the camera's lens during fluidized bed and spouted bed drying. (b). Image of peas during fluidized bed drying

The second stumbling block that makes it really challenging to segment images captured during fluidized drying, is that the products within the image are not at the same distance from the camera. This variation in distance makes it difficult to monitor shrinkage. For some food products, such as green peas that have high initial moisture content (over 68% wb), as they lose water during

fluidized bed drying, their shape continually changes. This continuous change in shape, for randomly moving food material, makes it impossible to analyze real-time images, using the CV methodology discussed earlier. Suppose all the peas were uniform in shape (e.g., spherical), even if they were touching themselves, we can still isolate them from the background by fitting an ellipse around the boundaries of each pea. In so doing, we can connect the outline of each pea, which in turn would help us compute their size. However, when we have overlapping products with irregular shapes, it poses a huge challenge to image segmentation.

One other challenge that makes it difficult to use CV technology during fluidized bed drying is that the surface texture of some food materials continually changes as drying progresses. This non-uniform transition in texture is displayed in Figure 1.3.

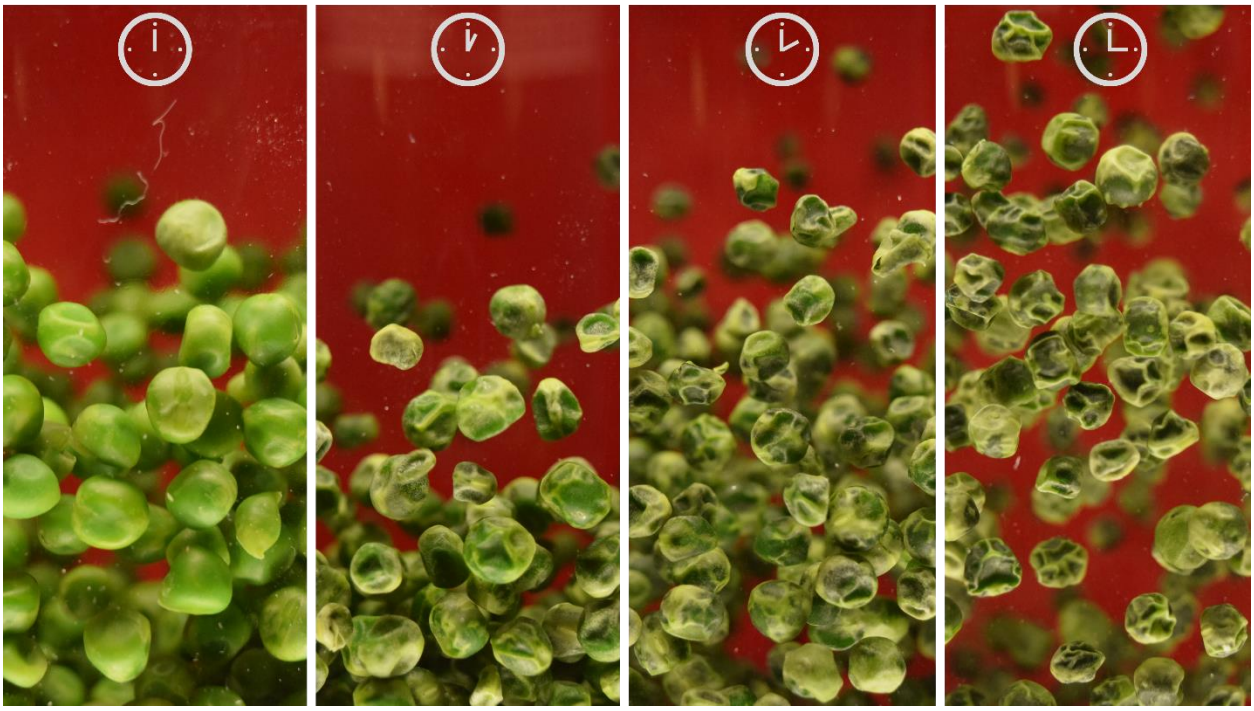


Figure 1.3 Image depicting the real-time change in visual attributes of green peas, during fluidized bed drying.

Previously, we mentioned that a threshold is usually set to differentiate the objects of interest from the background. However, changes in texture may pose a hinderance towards achieving this goal. When peas are dried in a fluidized bed, at the start of drying all the peas appear similar in color. At this stage, applying a threshold would show us areas of the image that are occupied by peas and

areas occupied by the background. However, as drying progresses, the surface of the peas gets rougher, the surface starts to form wrinkles, and the color of the wrinkle gets darker with time. At a point, the threshold would not be able to differentiate the pixel intensity of the background from that of the darkened wrinkles on the pea, thus making is difficult and even sometimes impossible to isolate the peas.

Just like texture, the color of some food changes during drying. The change is also non-uniform across the surface of individual food materials. This variation in color makes it even more challenging to isolate the object of interest using only thresholds. For instance, while drying peas, at some point, the surface of the individual peas develops different shades of green, ranging from very light green to grayish green, as shown in Figure 1.3. During threshold-based segmentation, after converting the image to grayscale and applying a threshold, some regions on the surface of the peas would have the same pixel intensity as the background, hence making it impossible to properly segment the pea. Without proper segmentation, computing visual attributes becomes problematic.

Current CV technique for monitoring visual attributes is based on a process we would refer to as the ***"Two-pass"*** framework. This framework comprises of two steps – image segmentation followed by visual attribute prediction. Once an image of the food is fed into the visual attribute monitor, the first step the monitor undergoes is to segment the image, so as to determine the regions or pixel locations that are occupied by food materials. The result of segmentation is a binary image made up of 0s and 1s. As we mentioned previously the food material are represented by 1s and the background by 0s. After segmentation, the monitor returns visual attributes of the food, by computing the visual properties of pixel locations on the RGB image whose corresponding position on the binary image has an intensity value of 1.

In summary, to use the two-pass approach to monitor real-time changes in visual attributes of food, during fluidized bed drying, there are four challenges we need to overcome.

1. Non-uniform distance of the food material from the camera lens
2. Overlapping objects of interest
3. Variation in color and texture on the surface of individual objects of interest, as drying progresses.
4. Variation in shape and size as drying progresses, and

5. Clustering of objects of interest, especially at the initial stages of drying.

The first two challenges listed above are not dependent on the type of food being dried. Irrespective of the food, during fluidized bed drying, once an image is captured, we would observe overlaps and non-uniform distances amongst the objects of interest. What this means is that the two-pass prediction framework used to process convective drying images, would fail when applied to images captured during fluidized bed drying.

Furthermore, if we observe an image captured during convective drying, we would discover that each of the food materials are fully visible on the RGB image. Also, since all the food materials are at the same distance from the camera (Figure 1.1), all the food products within the image are used to compute visual attributes. However, the case is different for fluidized bed drying, because certain regions of overlapped food material would not be visible within the image. The partially visible products would give substandard readings for product size, shape, color, and texture. The variation in distance would also affect size prediction. Both hurdles could be overcome by computing real-time changes in visual properties, using only food material that meet two criteria. First, all the regions belonging to the food material are fully visible on the RGB image. Second, the material is at a known distance from the camera. These two criteria would become the foundation upon which we built the hypothesis for this research.

Our hypothesis states that ***"Hence, visual attributes of food dried in a cylindrical fluidized bed dryer can be monitored in real-time, by computing the average visual attributes of individual food material that are touching the wall of the drying chamber facing the camera."***

Figure 1.4a sheds more light on this hypothesis. The distance between the camera and the cylindrical drying chamber is measurable, so, the true size of all food material touching the cylindrical wall can be measured accurately. Secondly, during fluidized bed drying, the products mix while they jump up and down within the drying chamber. Due to the mixing action, at each point during the drying process, there would always be a new layer of food product touching the cylindrical wall. Hence, to ascertain the visual property of the food at time t, all we need to do is capture multiple consecutive images, at time t, segment each of them to isolate only the food material on the surface of the wall, then compute the average visual attribute, using the segmented output.

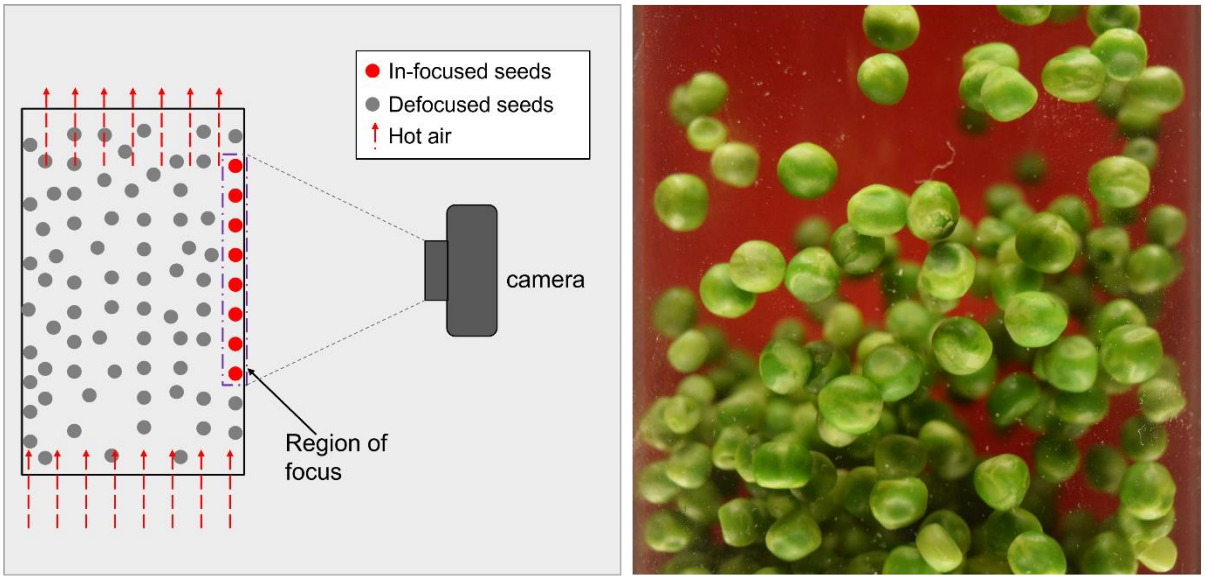


Figure 1.4 *(a) sketch of a fluidized bed showing peas (infocus peas) touching the wall of the drying chamber, and peas behind the infocus peas called the defocused peas. (b) real-time image of peas*

Another reason we propose to compute visual attributes using only food materials touching the drying chamber, is that even if these individual food materials are overlapped by neighboring materials, the ones overlapping them would be behind the object of interest. Hence, all the regions of our object of interest (that is, the food material touching the drying chamber) would be visible on the captured image, thus, increasing the probability of accurate segmentation.

Now that we have settled on what our objects of interest are, a follow-up question could be *"how would the CV system detect the food materials that touch the cylindrical wall?"* That is where our first specific research objective comes in. The objective is ***"to develop an image acquisition system for capturing real-time images of peas during fluidized bed drying"***. Chapter 3 describes how we achieve this objective and results we obtained. We chose to use green peas as the food material for our study, because it exhibits the last three challenges that are listed previously. The challenges are:

1. Variation in color and texture on the surface of individual objects of interest, as drying progresses.
2. Variation in shape and size as drying progresses.

3. Clustering of objects of interest, especially at the initial stages of drying.

One or a combination of these challenges are exhibited by most food materials – grain, seeds, nuts, and vegetables, during fluidized bed drying. Hence, if our solution works for green peas, it could be adapted to monitor visual attributes of other food materials. This image acquisition system should be such that it differentiates objects of interest from other objects within the image. The system should also prevent motion related blur that is unique to fast moving objects.

Once the real time image is captured, the next step is segmentation followed by visual attribute computation. Hence, our second specific objective was ***"to develop a classical computer vision pipeline to monitor real-time changes in size, color and texture of green peas during fluidized bed drying".*** Chapters 3 and 4 provide details on how this objective was achieved and the results that were obtained. We chose to develop the pipeline using classical CV techniques, so that once our second specific objective was achieved, we could ascertain how computer vision methodology (which is the current state of the art), performs when applied to images captured during fluidized bed drying.

However, we anticipated that the three challenges listed above may act as stumbling blocks for the classical CV approach. Most likely, these challenges would make it almost impossible to properly segment real-time images. To solve this problem, we could apply deep learning methods. Hence, our third specific objective was ***"to develop a convolutional neural network (CNN) for real-time segmentation of green peas during fluidized bed drying"***. Chapter 4 presents a deep dive into how this objective was achieved.

In line with the two-pass framework, immediately after image segmentation, our next step would be to compute visual attributes using the segmented output. Hence, the CNN used for image segmentation would not stand alone. It must be coupled to another model that predicts visual attributes from the segmented output. Therefore, once we had developed a CNN that was effective at image segmentation, our next specific objective was ***"to develop another CNN for predicting real-time changes in size, color and texture of food materials from segmented images".*** The methodology for achieving the objective and the results obtained are also reported in Chapter 4.

If everything went as planned, at this point in the research, we would have developed an AI-driven solution that predicts visual attributes of peas using the two-pass framework. Once this goal is achieved, the natural question would be *"can visual attributes of food be predicted without*

segmenting the captured image? Can we develop a model that predicts visual attributes of food without performing image segmentation?" To answer this question, our final specific objective was ***"to develop a Single-pass solution that predicts visual attributes of food, during fluidized bed drying, using unsegmented images"***. Chapter 5 describes our journey towards achieving this goal.

By accomplishing these five specific objectives, we will have improved the state of the art by achieving three novel milestones (Figure 1.5) within the food industry. At this point in the research, none has been achieved, hence they are assigned a value of 0% in their progress bar.

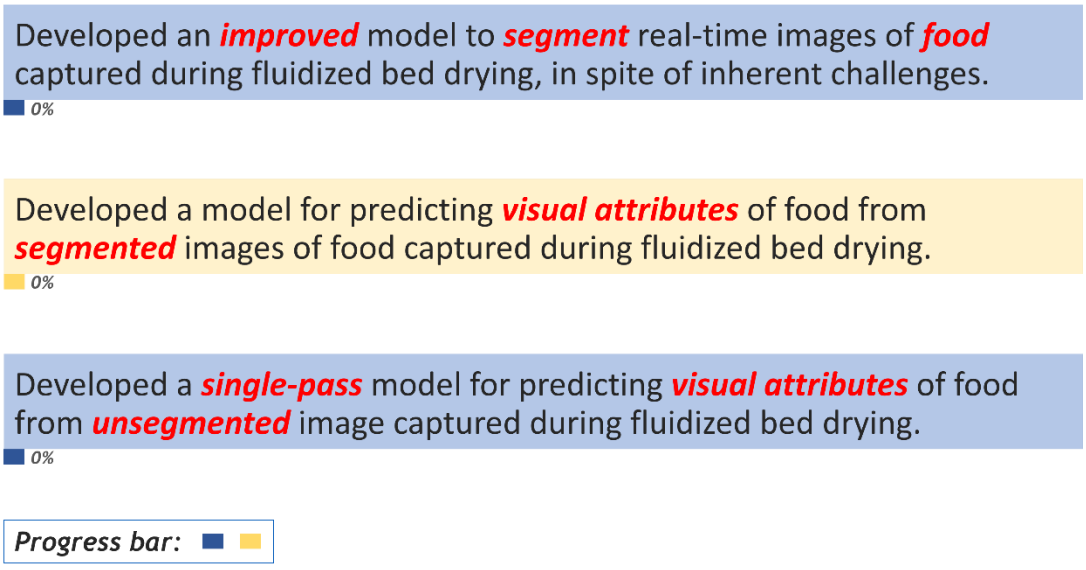


Figure 1.5 Milestones to improve the state-of-the-art

As we progress through this thesis, at the end of each chapter, we would revisit to these three milestones, to ascertain how much progress we have made towards improving the state of the art for monitoring visual attributes of food during fluidized bed drying.

Once these three milestones are achieved, we will have produced novel two-pass and single-pass solutions for monitoring visual attributes of food in real-time during fluidized bed drying. When coupled to a control system and integrated with the current drying fluidized bed dryer, we would end up with an intelligent drying system that could consistently produce dry food that is visually appealing and has a long shelf life.

Additionally, both novel visual attribute monitoring system could be a priceless resource for research when optimizing the fluidized bed drying process. With a two-pass or single-pass system,

the optimization process could go beyond drying parameter and moisture content of the food. Instead, researchers could now be able to investigate how different drying conditions would affect moisture content, color, texture, size and even shape of the food product.

CONNECTING TEXT

Before diving into objective one, that is, developing an image acquisition system for capturing real-time images of green peas during fluidized bed drying, it would be beneficial to review previous research works where computer vision technology was applied to monitor visual attributes of food during various drying processes. From this review, we would get to know the food products they were worked on, the visual attributes that were monitored, and the design considerations that were made when developing each CV system. We would also get to learn why each system succeeded and conditions in which they would fail. The wealth of knowledge derived from such a review would be a guiding light for us as we journey towards achieving our research objectives. Hence, in Chapter 2 we present this review.

CHAPTER II

LITERATURE REVIEW

MONITORING VISUAL PROPERTIES OF FOOD IN REALTIME DURING FOOD DRYING – A REVIEW

2.1. ABSTRACT

Annually, one third of the food produced globally is lost or wasted.[1-5,24,25] One effective technique to solve this problem is by developing dryers that consistently produce dry foods that are visually appealing and have a long shelf life. The beating heart of such dryers is a computer vision (CV) system that monitors visual attributes of the food, in real-time, during the drying process. These attributes are predicted from images of the food, using a Two-pass framework that starts with image segmentation. Unfortunately, there are currently no real-time CV systems for monitoring visual attributes of food during fluidized bed drying. This setback was linked to unresolved challenges encountered while segmenting real-time images of the food. To provide solutions to this lingering problem, this paper reviewed various research publications where CV systems were used to monitor visual attributes of food, during various drying processes, ranging from convective drying to microwave vacuum drying. Various CV systems were investigated and categorized based on the layout, the state and patten of motion, and the behavior of food materials within the image, from which visual attributes were predicted. The input images for the different CV systems were categorized as being either a static-type or a chaotic-type image. Suggestions on how the classical CV approach could be adapted to monitor the color, texture, size, and shape of food during fluidized bed drying were provided. Additionally, two improved deep learning-based methods were discussed. When such AI-driven solutions are merged with a control system, then integrated with fluidized bed dryers and other drying systems, this union could open the gateway to intelligent drying, where dryers consistently produce high quality dry foods. By extension, the consistency in product quality could reduce global food losses and waste significantly.

2.2. INTRODUCTION

In July 2022, global population stood was about 7.8 billion. By the end of 2050, it is projected that this number would grow to almost 10 billion people.[5,26-30] This significant jump in world population would mean more mouths to feed. The FAO estimates that food production should grow by 50% to ensure food security come 2050. [29,30]

One way to ensure that everyone would have enough food to eat could be by cultivating more. To achieve this goal, more fertile land would have to be converted to farmland. In some regions, increase in farmland could put more pressure on limited water resources. More financial investment would be needed for equipment, labour, rent, insurance, and repairs, to mention a few. In summary, it is going to cost us much more if our only solution is to cultivate more. Interestingly, there is a cheaper and more effective way to meet our future food need.

As we have mentioned previously, every year, over one-third of the food produced globally goes to waste or is lost [1-5,24,25]. A huge portion of food loss and waste usually occurs between food harvest and purchase. Occasionally, the food is thrown into the garbage by consumers, even after it has been purchased, especially, when they discover more attractive products on their next visit to the food store.[26] By eliminating or significantly reducing food loss and waste, we would be on course towards providing more food, as we journey towards the year 2050. Interestingly, food drying could play a pivotal role in achieving food sufficiency.

Drying has been a popular and effective method for prolonging the shelf life and maintaining the quality of food, such as vegetables, grains, meat, nuts, and seeds. During the drying process, as moisture within the food is reduced to a safe level, microbiological activities are reduced, and enzymes are inactivated. Thus, the dry product is more stable, has a longer shelf life, requires lesser storage volume and cost less to transport.[31,32] However, if the drying process is not well controlled, the organoleptic properties and nutritional content of the food could be adversely affected. Thus, resulting in heterogenous shrinkage, casehardening, shape distortion, development of surface cracks, loss of flavour, and changes in color due to enzymatic and non-enzymatic browning. [33,34]

Due to these unwanted effects, when consumers visit the food market to buy dry foods, they carefully inspect the visual attributes of the products. They examine the color, visual texture, size, and shape of the dried food. If the product is attractive, the consumer picks it up from the food

shelf, then pays for it. On the other hand, if the food is not appealing, customers walk away from it, thereby leading to increased waste, even after the food has made it to the market. Let us briefly describe the state of the food before it is are fed into the dryer.

Batches of fresh food are usually supplied to commercial food processors (drying companies) at different initial moisture contents. Notwithstanding the discrepancies in moisture content, the goal of the processor is to dry the food to a specific final moisture content. To achieve this objective, the processor uses an important tool known as the drying curve. This curve is unique to the food being dried. For instance, millet would have a different drying curve than carrot. The curve shows how the moisture ratio of the food reduces over time, when dried at known drying conditions. Two examples of such conditions for convective dryers are air temperature and air velocity. As far as the equation of the curve and the initial moisture content of the fresh food are known, the food processor could easily compute how long it would take to dry the fresh food to a desired final moisture content. Likewise, once we know the initial moisture content, with the drying curve equation, the moisture content of the product can be computed at any given stage during the drying process. With the drying curve, the food processor would consistently end up with the desired moisture content.

Unfortunately, the same cannot be said for consistency in achieving the desired visual appearance. Inconsistency in appearance is a major reason why a product that has been dried to a safe moisture content could still be rejected by consumers (all because of poor visual appearance). However, if there were dryers that could monitor and control both moisture content and visual appearance, the problem of dry food waste could soon be a thing of the past. Such improved dryers would consistently produce dry foods that have long shelf life and are visually appealing. Thereby immensely reducing global food loss and waste.

In recent years, various researchers have developed Computer Vision (CV) systems for monitoring visual attributes of food during drying. This technology monitors the color, visual texture, size, and shape of foods, from images of the food captured at difference stages during the drying process. This visual attribute monitoring system is made up of two broad parts – an image acquisition system and an image processing system. The first part comprises of a digital camera for capturing images of the food; a lighting system to ensure uniform illumination across the surface of the food; and an image acquisition software for operating the digital camera. On the

other hand, the second part is made up of computer hardware and software for processing the captured image, to compute or predict visual attributes of the food captured in the image.

Beyond food drying, CV systems have also been used for quality assurance of agricultural and food products.[12] The advantages of using CV system to ascertain the quality of agricultural products were explained in detail by Gunasekaran,[35] Zheng et al.,[36] and Du and Sun [37,38]. With CV technology, the quality of food products can be investigated non-destructively and without touching the product, thus, maintaining product quality and improving accuracy and efficiency of the process[35]. Furthermore, excellent reviews have been written on how CV systems have been used to inspect, sort, and grade various grains, vegetables, and fruits [39-42].

Integrating CV technology into current food dryers leads to improved, predictable, and consistent product quality. Traditional methods of measuring quality attributes were not originally designed for food drying.[12] Take for instance the process of measuring product volume. The measurement would have to be conducted using a Stereopycnometer or by applying the volume displacement method. In both cases, volume estimation would have to wait till after drying is completed.

Another example of an incompatible quality estimation technique was the use of a *L*a*b** colorimeter [43,44] or spectrophotometer [45] to measure product color. When using either equipment, a sample of food must be retrieved from the dryer, each time a measurement is to be taken. Hence, in addition to being destructive, for products with varying surface pixel intensity, results generated by this method are not fully representative of the whole food. On the other hand, computer vision technology is non-destructive, contact-less, and from the captured image, the visual attributes of the whole food could be computed in one shot.

In his comprehensive review[12] titled *"Computer vision for real-time control in drying"*, Martynenko described the evolution of CV technology within the food drying industry. This technology was first introduced as an offline system for evaluating food quality during drying. The offline CV system is a stand-alone set-up. It is not coupled to the dryer. To evaluate the color, texture, shape, or size of the food, or to investigate the effect of various drying conditions linked to visual attributes, a sample of the food had to be extracted from the dryer and placed in the offline system for image acquisition and processing. Martynenko discussed various forms of drying where the offline CV system was applied. Some of them include convective drying[8,46-48], microwave-vacuum drying[49], and osmotic drying[50,51]. This CV system was also used within the

pharmaceutical industry to compute the size of granules during fluidized bed granulation. Some of the foods whose visual attributes were studied using the offline CV system include apple [8], carrot [49], hawthorn fruit[46], kiwifruit[50], pumpkin[51], and soybean[47].

Since visual attributes of the food had to be assessed outside the dryer, the attributes obtained by the offline system only described the state of the food at the point of sample extraction. Once the sample had been retrieved from the dryer, the operator had no idea how the drying conditions were affecting the visual attributes of food that was still in the dryer. Occasionally, by the time an updated product state or quality information was obtained, the damage had already occurred. To solve this problem, one may suggest that the operator retrieve food samples more frequently. Unfortunately, in the case of convective drying, frequent extraction would mean opening and closing the dryer so many times during the drying process. This method would introduce errors and inefficiencies to the drying process. Hence, with the offline system, it is impossible to monitor real-time evolution of visual attributes during drying. The drive to make up for this limitation gave birth to "online" CV systems, also referred to as the second evolutionary phase in CV technology.[12]

With the online CV system, researchers were able to obtain granular and real-time information on visual attributes of food while drying was taking place. Unlike the offline system that was a separate unit from the dryer, the online CV system on the other hand was an integral part of the dryer. The camera and lighting system were configured in such a manner that made it possible to capture illuminated images of the whole food in real-time, while the food was in the dryer. Currently, during convective drying, researchers can measure moisture content of the food in very close intervals, as close as 30 seconds. With the advent of online CV systems, it has become possible to capture images of the food that corresponded with each 30 second time step, such that after a 4-hour drying run, the researcher had 480 images, each corresponding to one of the 30-second time step. So, with these images, for each time step, there would be a corresponding value for color, texture, size, shape, and moisture content of the food.

With such granular detail, it becomes possible to develop mathematical relationships between drying parameters and one of the visual attributes, or between moisture content of the food and one of the visual properties. The introduction of machine learning further improved the landscape. For instance, using feedforward neural networks and the massive training data captured during the

drying process, various researchers were able to train networks, whose inputs were system parameters, and the predicted outputs were moisture content, various visual attributes, and other parameters specific to the food. Later in this review, we will delve more into some of these research findings.

Due to its capacity to capture visual attributes in real-time, online CV technology provides researchers with better and more detailed understanding on how different drying conditions affected visual attributes of numerous food products. From real-time plots of color, texture, size, or shape, it is easier to visualize how product quality changed or degraded over time. Researchers were able to identify critical points beyond which visual appearance of the product ceased to be attractive to consumers.

Another key point to note about online CV technology is the wide door of opportunity it opens for drying system optimization. At each time step, during the drying process, real-time data on drying system parameters, moisture content, and various visual attributes of the food are available. With this abundant, recurrent, and valuable stream of data, it becomes possible to optimize the drying process to produce dry food that has long shelf life and is visually appealing, while providing the same energy efficiency.

In a detailed review [12], Martynenko outlined multiple food drying research where online CV systems were applied to monitor real-time changes in visual attributes, as well as investigated the effect of various drying conditions on one or more visual attributes. Some of the drying methods that were in the review include convective drying,[14-17,21,52-54] superheated steam drying [17,21,54,55], and infra-red assisted convective drying [21,55]. Coupled with an online CV system, these dryers were used to investigate quality degradation of foods such as apple,[14,20] bell pepper,[52] blueberry [56], ginseng,[13,53] kiwifruit,[55] potato slices,[15,16] and shrimp[17,21,54]. With online CV technology, we can observe in real-time and within very close time intervals, how the color, texture, size and shape of the food changes as drying progresses. This recurrent stream of useful information helped propel computer vision technology into its third stage of evolution – the use of online imaging to identify critical control points for control and optimization of drying.[12] With such technology Martynenko [13] developed a control system for drying ginseng roots, using inputs on visual attributes produced by an online CV system.

Up to this point in our review, we have discussed some of the significant and comprehensive literature reviews and research works where computer vision technology was applied to some post-harvest operations. We started out by identifying different works that reviewed the benefit of CV technology. Next, we described how this technology has been applied in other areas of agriculture and food production. We concluded our discussion on the evolution of computer vision technology within the food drying industry. It is important to note that there is one common feature shared by all these CV systems – they all predict visual attributes from image(s) of the product. Therefore, the first step carried out by the various CV pipelines is image acquisition. Although they all share the same first step, their distinguishing feature lies in the second step known as image processing. It is during this second step that visual attributes are computed. Interestingly, the workflow for processing the image and thus computing visual attributes is highly dependent on the ***layout***, the ***state*** and ***patten of motion***, and the ***behaviour*** of food materials within the image.

These three factors (layout, state and patten of motion, and behaviour) determine how successful the CV system would be at predicting visual attributes of the food. To explain how critical these factors are, let us first describe the general process flow for predicting visual attributes using the classical computer vision approach. To compute the color, texture, size, and shape of a product, we first capture an image of the food. Since we plan to monitor color, the picture most likely would be captured in RGB format. After image acquisition, we move to image processing. The initial step during image processing would be to locate where the food materials are positioned in the image. This process of locating the objects of interest is known as image segmentation. The segmentation step could be based on thresholds, gradients, regions or on classifications.[57,58]

Segmentation produces a binary mask containing 0s and 1s, where 0 represents the background and 1 represents the food (object of interest). Using the binary mask and RGB image, the color, texture, size, and shape of the food material (within the captured image) are computed. In this review, we would refer to this two-part process of image segmentation followed by visual attribute prediction as the **"*Two-pass framework*".** Figure 2.1 describes this framework.

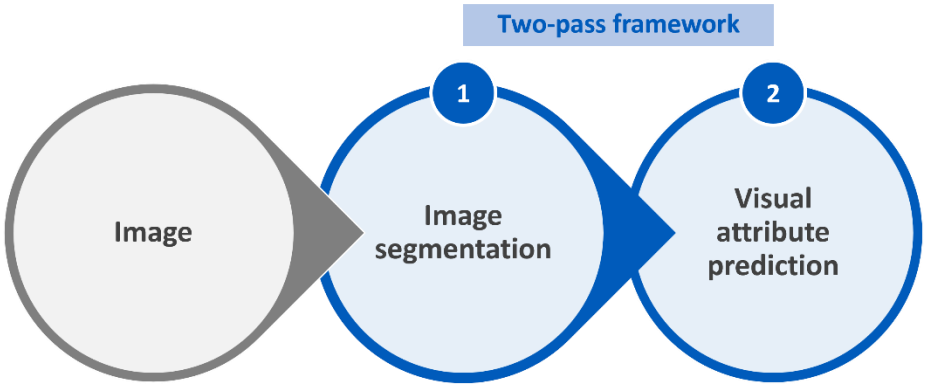


Figure 2.1 The Two-pass framework

To detect the objects of interest more easily, there must be significant contrast between the image background and the food materials. Such that when one of the RGB channels is examined, the pixel intensities within the regions occupied by the food would be significantly different from that of the background. For instance, if we were drying green peas, it would be better to use a red background than a yellow, black, or dark green background. With a red background, regardless of the drying stage, and non-uniform browning on the surface of individual peas, there would still be sufficient contrast between the green peas and the background.

Furthermore, since we are also computing size and shape, the system should be able to distinguish one pea from another. Why? Because the size of a pea is computed by counting the number of pixel positions within the region occupied by the pea. Hence, if the system is ineffective in distinguishing one pea from the other, when the peas are clustered or overlap themselves, the chances of counting multiple peas as one increases. To prevent such errors during convective drying, food materials are evenly spaced out on the drying tray, none touching the other. Also, during convective drying, since the food is placed on a flat mesh, the distance of the food from camera is known, hence, we can compute the actual size of the product.

With this being said, if we are to group computer vision systems based on the layout, the state and pattern of motion, and the behaviour of food materials within the captured image, each system could be grouped as either a ***"static-input"*** or ***"chaotic-input"*** CV system.

2.3. STATIC AND CHAOTIC INPUT CV SYSTEMS

Up to this point, we have been describing the static-input CV system, where the individual food products in the dryer are evenly spaced out on the drying tray; are at the same distance from the camera; and are motionless throughout the drying process. Most published work in which CV technology was applied to food drying, falls under this group. For all of them, their images were processed using the Two-pass framework. As we discussed previously, this framework starts with image segmentation, and ends with visual attribute prediction using the segmented output.

To the best of our knowledge, there is no published research where a CV system by-passed the segmentation step. Avoiding the segmentation step could save a lot of computational resource and time. However, to achieve this feat, instead of passing the image through a classical CV pipeline, it should be passed through a convolutional neural network that has a 3-D input layer and an output layer with nodes corresponding to the number of predicted visual attributes. In this review, we would refer to this new method as the ***"Single-pass"*** framework. We will discuss more about it towards the end of this review. Until then, let us continue our discussion on the classification of CV systems, based on the layout, the state and pattern of motion, and the behaviour of food materials within the captured image. We have talked about the Static-input CV system. Let us delve into the second group of computer vision systems referred to as ***"chaotic-input"*** CV systems.

2.3.1. Chaotic-input CV systems

Just like the static group, chaotic-input CV systems are also characterized by the layout, the state and pattern of motion, and the behaviour of food materials within the image. Dryers whose food materials are monitored by such CV systems exhibit the following traits. Throughout the drying process, the food materials are in random motion, jumping up and down within the drying chamber. While in motion, individual products collide with each other, forming clusters and overlaps. Considering all these factors, when an image of the food is captured, we notice that individual food materials are at different distances from the camera, thereby making it more difficult to monitor the size of the product. There are a lot of overlapping and clustered products, thereby posing a challenge to accurate segmentation. Also, unlike static-input CV system where the same surface of the food faces the cameras throughout the drying process, for the chaotic scenario, due to random motion and collision, different regions of the food product would appear on consecutive

images. To make matters worse, when we dry food materials that do not shrink uniformly, uneven shrinkage adds another layer of challenge when monitoring product size.

So, when we combine non-uniform distance, uneven shrinkage, clustering, and overlaps, of the food products, segmenting the captured image becomes a 3-D problem, as we would need to accurately compute the number of pixels within the object of interest, as well as how far each of the objects are from the camera. Hence, when computer vision pipelines for static-type images are used to process chaotic-type images, the pipeline fails miserably.

To the best of our knowledge, there are no publications that review computer vision systems based on the layout, the state and pattern of motion, and the behaviour of food materials within the captured image. Hence, that would be the objective of this review. We will discuss various research works where computer vision systems were used to compute, monitor, and control visual properties of food during drying. We start out with static-input CV systems and end with the chaotic version. Our discussion for each group would be divided into offline and online systems. Towards the end of this chapter, we would identify some chaotic-based drying processes that do not yet have published solutions for monitoring visual attributes of food dried in them. Finally, we will provide suggestions on how to develop such solutions. To kick off the discussion, let us review various publications on static-input offline CV systems.

2.3.2. Static-Input Offline CV Systems

Fernández et al. [8] investigated the effect of drying conditions on the color, shrinkage, and visual texture of apple slices, during convective drying. The CV system monitored color by computing the *L*, a* and b** indices of apple slices. For size, it monitored average radius, Feret diameter filled area, perimeter, and roundness. While for visual texture, the energy, contrast, entropy, homogeneity, and variance of the objects of interest were computed, from groups of images captured at different stages during the drying process. Their results show that hot air drying for 12 hours had significant effect on shrinkage, color, and texture of apple slices. All shrinkage-related parameters increased over the course of drying. From the plot of each visual parameter against the drying time, it was evident that there was significant increase in shrinkage parameters (except roundness) within the first 6 hours. *a** and *b** color indices increased as drying progressed, while *L** index remained almost constant. Entropy and contrast increased as drying progressed, meaning that the variation in pixel intensity on the surface of the apple slice increased over time. Unlike

entropy and contrast, homogeneity reduced over time. Being the opposite of contrast, consistent decrease in homogeneity confirmed the finding that there was increasing variation in pixel intensity on the surface of apple slices. Hence, the surface appeared rougher as drying progressed. The consistent reduction in homogeneity could also mean that the apple slices were not losing water uniformly across their surface. By the end of the research, Fernández et al. developed an effective offline CV to assess the quality of apple slices from static-type images captured at different stages during the drying process.

Nahimana and Zhang [49] investigated the effect of microwave vacuum drying on radial shrinkage, color, shape, rehydration capacity and carotenoid content of carrot slices. Color was monitored by computing the Hunter *L*a*b** indices. The shape descriptors were circularity, solidity, as well as major and minor axes. 90 g of carrot slices where dried during each experimental run. The drying runs were conducted at three power levels 1.3, 1.9 and 2.5 W/g. Other drying parameters included an operating pressure of 40 mbar (absolute pressure), turntable rotation speed of 5 rpm, and cabinet temperature of 33˚C. Similar to the work by Fernández et al., [8] visual attributes were also computed from static-type images. At regular intervals, carrot samples were extracted from the microwave and placed within an offline image acquisition chamber, where the images were captured. The color of the extracted samples was obtained using a colorimeter. Results of the research showed that for each microwave power level, the carrot core shrunk more than the cortex. There was significant change in shape, and the change mostly occurred during the first half of drying. For all microwave power levels, there was significant color difference between the cortex and core. The cortex was redder than the core, agreeing with the findings in literature [59]. Finally, it was concluded that microwave vacuum drying significantly reduced the Carotenoid content of carrot slices. Let us move from carrots to research conducted using hawthorn fruit.

Aral et al.[46] investigated the effect of air temperature and air velocity had on color, shrinkage, rehydration capacity, and microstructure of Hawthorn plant, during convective drying. 100 g of Hawthorn fruit was dried form 70 % (w.b.) to 10 % (w.b.). The study was conducted at three different level of air temperature (50, 60, and 70˚C), as well as three levels of air velocity 0.5, 0.9, and 1.3 m/s. The products were weighed hourly to monitor the moisture content. The color of hawthorn plant was captured twice – before drying started (when the product was fresh), and at the end of the drying process. The color values that were computed include lightness index (L*), greenish index (a*), yellowish index (b*), Total color difference (ΔE), and browning index (BI).

Shrinkage was measured by computing volume shrinkage. Since hawthorn fruit has a spheroid shape, the volume was computed using the geometric diameter of the fruit. Just like color, shrinkage was estimated at the beginning and end of drying. The drying time decreased with the increasing air temperature and air velocity. Amongst all the thin layer drying models that were tested, the Midilli et al.'s model best explained the drying behaviour of hawthorn fruit. As drying progressed, L^*, b^*, and ΔE increased, while a* decreased. The overall change in L^*, a^*, b^* and ΔE were highest at lower temperatures and velocities. The highest change occurred at 50˚C and 0.5 m/s, while the lowest change in color values occurred at 70˚C and 1.3 m/s.

Huang et al. [47] developed a computer vision solution to compute the color and moisture content of soybean. Static-type images were obtained using hyper-spectral reflectance imaging technique. During the experiment, ground truth data on color and moisture content were generated from mean reflectance and entropy data. Both data streams were extracted from hyper-spectral reflectance images. By the end of the study, the authors had developed a regression model for calculating moisture content and color of soybean. The model derived using mean reflectance performed better, achieving R^2 values of 0.862 and 0.971 for color and moisture content, respectively.

Zenoozian et al. [60] combined neural networks and image analysis to investigate how different drying conditions (temperature, air velocity, drying time, and sample thickness), and different osmotic treatments (use of sucrose solution and sorbitol glucose) affected the drying time and quality of osmotically dried pumpkins. They monitored the color, size (shrinkage) and deformation of pumpkin slices, at the beginning and end of the drying process. Product color was monitored by measuring total color change (ΔE), shrinkage was measured in terms of volume shrinkage, while deformation was monitored by computing the Heywood shape factor. At the start of drying, three thicknesses were used 20, 15 and 10 mm. In all cases, the length and width of the slices were 20 mm and 20 mm. Final volume was measured using the displacement method. The experimental results showed that deformation increased with drying rate. Also, high deformation occurrence at a lower temperature was noted. The data generated by the CV system were then used to train a feed forward neural network. The model inputs were air velocity, temperature, thickness, moisture ratio and drying time, while the outputs were total color change (ΔE), volume shrinkage and Haywood shape factor. When trained on data generated from experiments conducted using one of the three osmotic solutions (sorbitol, glucose, and sucrose), the model achieved the following coefficients of determination that ranged from: (0.9729 – 1) for shrinkage, (0.9531 – 1) for

Heywood factor, and (0.9813 – 1) for ΔE. The results imply that the AI solution could be an effective tool for ascertaining the quality of pumpkin before and after osmotic drying. Furthermore, when the model was trained on a combination of training examples obtained from drying experiments conducted using sorbitol, glucose, and sucrose, the coefficient of determination obtained for shrinkage, Heywood factor, and ΔE were 0.9026, 0.8783 and 0.9828, respectively.

Anther work that combined computer vision system and machine learning was conducted by Fathi et al.[50] The objectives of their study were two-fold. The first objective was to investigate if mass transfer kinetics (water loss and solid gain) and color changes of dried kiwifruit were dependent on the concentration, temperature, and duration of osmotic drying. Their next goal was to study how effective an ANN would be in predicting water loss, solid gain, and color changes during osmotic drying, given the operating parameters. During the study, kiwifruit slices were dried osmotically in sucrose at four concentrations (30, 40, 50, and 60˚ Brix) and at three temperatures (20, 40, and 60˚C). Drying was conducted at three immersion durations 0.5, 1, 1.5, and 2 hours. Images of the slices were captured before and after osmotic drying, to compute the color at both stages. At the end, the result obtained by the CV system was used to train a Feedforward neural network, made up of a single hidden layer containing 16 neurons. The inputs to the model were osmotic concentration, osmotic temperature, and drying duration. While the model outputs were water loss, solid gain, and total color change (ΔE). The model successfully predicted the three outputs, achieving coefficients of determination of 0.88, 0.92 and 0.994 for total color change, solid gain, and water loss, respectively. The result from the CV system and machine learning model both revealed that color change increased with the drying duration.

It is not only in food drying that static-type images are used as input for offline CV-systems, but also the images also been used when ascertaining the quality of extruded food. Gao and Tan[61] developed an image processing technique for analyzing the surface texture of expanded food. During their study, the system was built by correlating surface image features with texture related geometric properties. The geometric properties were obtained from SEM (scanning electron microscope) generated cross-sectional images of the food. Features acquired from images of the food surface, were highly correlated to the measurements of internal food structure computed from images acquired by SEM. For instance, the edge length, a texture-related geometric property was computed from the surface image, while the cell density was measured from the SEM image. There was a 0.97 correlation between the edge length and cell density. This finding implies that

the internal structure or textural properties of expanded products could be assessed from the surface of the product, without cutting it open to view the internal structure. Hence, reinforcing the benefits of CV system for food quality assessment.

Offline CV systems that work with static image input have also been used in the pharmaceutical industry. Such systems were employed by Možina et al[62] to monitor visual attributes of pharmaceutical pellets during the fluidized bed agglomeration process. Normally, images captured during fluidized bed drying should be categorized as chaotic-type images. Why? Because during this drying operation, the products are in random up and down motion, colliding with each other, and forming clusters and overlaps as they move within the dryer. However, to generate static type images during fluidized bed drying, Možina et al, extracted granule samples from the drying chamber, then captured and processed them offline within a stand-alone CV system. During their study, they investigated the use of digital visual imaging as a Process Analytical Technology (PAT) tool for monitoring pellet thickness and agglomeration during the coating process. Their goal was to develop a CV system to monitor spherical diameter, coating thickness and instance of unwanted agglomeration during fluid-bed pellet coating. The experiment was conducted using high precision bearing balls. To measure the parameters during coating, without disrupting the coating process, samples of the coated pellets were retrieved using a manual sampler located by the side of the product container. From the manual sampler, the pellets were discharged onto a horizontal vibrator that slid the pellets underneath the camera lens, pointing downwards towards the vibrator. Once the pellets were within the camera's field of view, images were captured. Afterwards, the objects of interest were segmented.

Using the segmentation results, circularity was computed. Objects of interest whose circularity was greater than 0.8 were used to compute the equivalent spherical diameter (ESD) and coating thickness estimation (CTE). The system was successful in detecting the onset of unwanted agglomeration. It achieved this feat by producing the distribution of ESD, from all the classes of objects whose circularity ranged from 0 – 1. To identify cases of unwanted agglomeration from an acquired image, the pellet size distribution was bimodal or sometimes multimodal. Finally, the system accurately estimated granule spherical diameter and coating thickness.

Kucheryavski et al.[63] also developed an offline CV system to monitor visual attributes of pharmaceutical pellets during the fluidized bed coating process. The quantitative description of the

pellets was obtained using wavelet decomposition and angle measurement technique (AMT). Both methods were used to develop relationships between image features and process parameters. The CV solution achieved the following results. It could segment and count individual pellets within the static-type image. With the computed number and measured weight of the pellets, the offline solution successfully calculated the average thickness of coating on the pellets. Later in this review, we will delve deeper into why computer vision technology has become an essential PAT tool to monitor agglomeration and granule coating.

Notwithstanding the success of offline CV systems, they only provide information about the food at very few stages during the process timeline. The data is collected either at the start or end of the drying or extrusion process, or at very few points during both processes. It would be better to have a solution that helps the end user observe and monitor real-time evolution of quality parameters. With this added benefit, food drying conditions could be adjusted in a timely manner to avert product degradation. As we mentioned previously, such real-time systems are referred to as online CV systems. This system is divided into two groups. The first group are online CV systems whose input are static-type images, while the second group receive chaotic-type images. Let us review various research work conducted using static-input online CV systems.

2.3.3. Static-Input Online CV Systems

Hosseinpour et al. published three studies[17,21,54] where online computer vision technology was used to monitor and investigate different visual attributes of shrimps during convective and super-heated steam drying (SSD). For the three studies, drying experiments were conducted at convective drying temperatures of 50, 70, & 90 °C, SSD temperatures of 110 and 120°C, and air velocities of 1, 1.5, and 2 m/s. For each drying run, images of the shrimps were captured at 1 min interval, till the mass of the product remained constant.

The objective of the first study[21] was divided into three parts. The first was to develop a CV system for real-time monitoring of morphological attributes of shrimps. The parameters monitored include the top and lateral view areas, average area, Feret diameter, perimeter, and roundness. They also monitored elongation, shape of the edge, and diameters parallel and perpendicular to the drying medium flow. The second objective was to investigate the effect of different air temperatures, air velocities, drying method (convective drying and super-heated steam drying) had on shrinkage and shape. Their last goal was to study the effect of drying medium direction on the

diameter of shrimp that was parallel and perpendicular to drying medium flow. Their results showed that there was consistent decrease in moisture content, drying rate, Feret diameter, perimeter, dimensionless top-view and lateral-view areas, average area, and diameters perpendicular and parallel to the drying medium flow.

Individual plots of moisture content against top-view area shrinkage, diameters parallel and perpendicular to drying medium flow, normalized Feret diameter, average area shrinkage, and dimensionless perimeter, showed a linear relationship, for moisture ratio > 0.1. Equations for those plots had R^2 values between 0.9756 and 0.9999. The temperature and the velocity of the drying medium significantly affected area shrinkage and Feret diameter. The shrinkage of diameter perpendicular to drying medium flow was significantly different from the shrinkage of parallel diameter. This discrepancy implies that the shrimp shrinkage during drying was non-isotropic. For most drying conditions, the diameter parallel to the drying medium flow was greater than the perpendicular diameter shrinkage. Higher temperatures produced dried shrimps with more irregular edge deformation. Finally, at elevated temperatures, SSD produced more porous shrimps than convective drying. This study demonstrates that static-input online CV systems can be used to effectively monitor real-time changes in shape and size of shrimps. It also confirms that this tool can provide deeper insight on how drying parameters affect both visual attributes. This brings us to the end of the first study by Hosseinpour.

The second study by Hosseinpour et al.[17] concentrated on shrimp color. The first objective of this work was to develop a computer vision system for monitoring real-time changes in color attributes of shrimps during convective drying and super-heated steam drying. The color attributes monitored include lightness (L*), redness (a*), yellowness (b*), total color change (ΔE), Hue angle, browning index (BI), and chroma (CH). The next objective was to model color parameters based on drying system variables, and moisture ratio based on color parameters and drying time. From the graph of color parameters against moisture ratio, the following conclusions were evident. Lightness index (*L**) decreased with moisture content, while the other color indices increased. Increase in drying temperature decreased lightness index (L*) and hue angle, while it increased other color parameters. Shrimps that were dried in the superheated steam dryer retained more of their color, compared to those dried in the convective dryer. Increasing temperature of the drying medium decreased L* and H°, while other color attributes increased.

The fractional conversion model was adequate in predicting color data. The velocity of the drying medium significantly affected L*, ΔE and BI. At the end of their research, a quadratic regression model was developed to accurately model the relationship between moisture ratio and the color parameters and drying time. The accurate prediction of color parameters meant that this CV system could be deployed to monitor and control the shrimp color in real-time. After their work on monitoring product color, Hosseinpour et al. moved on to their third study.

This research was focused on visual texture. The goal of the study[54] was to develop an image processing technique for obtaining accurate textural features of shrimps during convective drying. Inline with their objective the textural features had to be invariant to translation, scale, and rotation of the input image. The next objective was to correlate the image textural property with moisture content and geometric attributes such as perimeter, top-view area, feret diameter, and lateral-view area. The computer vision system was based on a combination of pseudo Fourier-Merlin transform, Radon transform and Fourier spectrum based-fractal dimension. Two parameters were indicative of the visual textural state of the shrimps during drying. The first parameter was the normalized Fourier spectrum-based fractional dimension (FD_h), followed by the normalized invariant moment $(S_{2,0})$. Both visual textural parameters were significantly affected by temperature of the drying medium. On the flip side, velocity of the drying medium had no effect on the textural parameters. During this study, there was a linear relationship between the normalized Fourier spectrum-based fractional dimension and on both the moisture content and geometric attributes of the shrimps. On the other hand, the normalized invariant moment, $S_{2,0}$, was correlated to the moisture content and geometric parameters using a cubic model. The researchers concluded that the fractional conversion and zero-order model were the best models for real-time and online monitoring of visual texture of shrimps during convective drying and super-heated steam drying. At this point, we conclude our discussion on shrimps, let us discuss an application of online CV system during potato drying.

Yadollahinia and Jahangiri [15] investigated the effect different air temperatures, air velocities, and air direction had on shrinkage of potato slices, during convective drying. The shrinkage parameters that were monitored include area, perimeter, major and minor diameters, diameter parallel and perpendicular to airflow, roundness, and elongation of the slice. Drying experiments were conducted at air temperatures of 60, 70 and 80^{o}C, and air velocities of 0.5 and 1.0 m/s. The results

showed that shrinkage had an almost linear relationship with the moisture content. Airflow direction had significant effect on shrinkage of the parallel and perpendicular diameters. Based on the range of drying conditions that that was evaluated, air velocity had no significant effect on drying time and shrinkage. Area and perimeter shrinkage were significantly affected by air temperature, but not by air velocity. These findings were evident on the graph of real-time change in visual attributes.

From the graph of area shrinkage versus moisture ratio, it was noticeable that potato slices dried at 60˚C and 70˚C had significant reduction in area, due to upward bending of the slice as moisture ratio went below 0.1. On the other hand, due to case hardening, great reduction in area was not observed for potato slices dried at 80˚C. Also, total shrinkage decreased as air temperature and velocity increased. On the other hand, rehydration rate increased with increasing air temperature and air velocity. Unlike potato slices dried at 60˚C and 70˚C, the elongation and roundness of slices dried at 80˚C were not affected by air temperature. Finally, due to isotropic shrinkage and absence of upward bending phenomena, potato slices dried at 80˚C appeared more marketable and were easier to pack. All these quality-determinant results were detected courtesy of the online CV system that was capable of monitoring potato slice degradation, from static-type images. After concluding this research, Yadollahinia et al. conducted another study[16] on potato shrinkage. However, this time around, they included a new shrinkage parameter – Feret diameter. Except for roundness, all other size-related features - area, perimeter, major and minor diameters, Feret diameter, and parallel and perpendicular diameters continuously reduced as drying progressed. At this point, we have come to the end of our discussion on potato slices, let us discuss another instance of research on apple slices.

The study by Wang et al[14] was aimed at ascertaining how accurate an imaging technique is at estimating real-time changes in bulk volume of apple slices during convective drying. The study was conducted at 40 and 80˚C. The initial hypothesis was that bulk volume of cylindrical apple slices could be estimated using their diameter and the thickness. Measurements were made from real-time images of apple slices. Throughout the drying process, image-generated diameter agreed with the values obtained by a calliper. Conversely, mid-way into the drying process, image-generated thickness did not agree with corresponding calliper measurements. The disparity was attributed to non-uniform shrinkage and bending phenomena. A Linear regression model that relates CV generated diameter to bulk volume was developed. The imaging technique reliably

monitored bulk volume using real-time changes in the diameter of apple slices captured during the drying process. This successful outcome was observed for drying runs conducted between 40 - 80°C and for moisture content ranging from 9.1 to 0.25 g/g.

Chen and Martynenko[10] utilized computer vision techniques to evaluate real-time effect of temperature on shrinkage and color of two blueberry varieties, during convective drying. For their investigation, they set up two cameras. One was placed directly over the blueberries to capture the top view from which the average equivalent diameter was computed. The second camera was placed by the side to capture the side view of blueberries, from which the average height of the berries was extracted. The average equivalent diameter and average height from each pair of images were then used to compute the volume of the blueberries at each stage during the drying process. Volume shrinkage was calculated as the ratio of instantaneous volume to initial volume. The *L*a*b** color indices were computed from the images captured by the side camera. These values were used to compute Total color change, ΔE. From the real-time plot of ΔE, for different air temperatures, it was evident that the drying air temperature had a significant effect on ΔE. Within the first 10 minutes of drying of wild blueberries, ΔE for 70, 80 and 90°C had risen beyond 20, while wild blueberries dried at 40 and 50°C were still below 5. From the graphs also, it was noticed that for both types of blueberries, there was a major change in color within 10 - 30 mins of drying. Notably, the change was significant for blueberries dried at 60, 70, and 80°C. Hence, real-time monitoring of color can be used as a tool to detect early signs of quality degradation during drying. Blueberries dried at 40°C retained most for their original color and texture. However, the drying time was the longest (88 hrs for highbush blueberries and 62.1 hrs for wild blueberries). The lengthy drying time may be a setback for industrial applications. Interestingly, from the plot of real-time color change, it was evident that a shorter drying duration and minimal color change could be achieved with an air temperature of 50°C.

Another study that used static-input online CV systems was conducted by Romano et al.[52] They examined the possibility of using a CCD camera and laser diodes emitting wave lengths of 532 and 635 nm, to monitor the moisture content of bell pepper during convective drying. Their second objective was to investigate if real-time images from a CCD camera could be used to monitor the color of bell pepper during drying. The combination of CCD camera and laser diodes turned out to be a fast and reliable method for monitoring moisture content in real-time, during convective

drying of bell peppers. Also, analysis of images captured with the CCD camera, showed that L^* and a^* indices of bell pepper were accurately monitored.

Some of these researchers have also conducted studies where they complemented online CV systems with machine learning solutions, to monitor visual attributes, as well as investigate how such attributes were affected by drying conditions. Nadian et al. [20] conducted one such study. Their first objective was to investigate the effect of air temperature, air velocity, and slice thickness on the color of apple during convective drying. To achieve this goal, they developed an online CV system that captures static-type images, from which appropriate visual attributes were extracted. Their next objective was to develop a neural network that was trained on real-time color kinetics data and drying system operating parameters. The function of the network was to predict real-time color values and moisture ratio of apple slices during convective drying. During the study, drying experiments were conducted at air temperatures of 50, 60 and 70^{o}C, and air velocities of 1, 1.5 and 2 m/s. The color parameters that were acquired from real-time images include lightness (L^*), Redness (a*), Yellowness (b*), and Total color change (ΔE). The CV system effectively tracked real-time changes in color. Air temperature and velocity had significant effects on color parameters.

After investigating the effects of drying conditions on product color, the training examples captured by the CV system were used to train a feed forward neural network. The network predicted L^*/L^*_0, a^*/a^*_0, b^*/b^*_0, and ΔE. The inputs were air temperature, air velocity, slice thickness and time. The model achieved the following R^2 values for each ratio: $L^*/L^*_0 = 0.9304$, $a^*/a^*_0 = 0.9242$, $b^*/b^*_0 = 0.9246$, and $\Delta E = 0.9497$. Such high R^2 values demonstrate the effectiveness of complementing CV systems with deep-learning solutions, when monitoring visual attributes of food during drying. The results of the study show that this solution is dependable, cost effective, and non-destructive. With this study on visual attributes of apple slices, we come to the end of our review of static-input online CV systems. Let us switch our focus to online CV systems that work with chaotic-type images.

2.3.4. Chaotic-Input Online CV Systems

As we discussed previously, the features that differentiate chaotic-type CV systems from their static counterparts, is the ***layout***, the ***state*** and ***patten of motion***, and the ***behaviour*** of food material within the image. For chaotic-input CV systems, the images are captured when the food materials

are in random up and down motion. Hence, when the image is observed, we would notice overlaps, non-uniform distance of the individual products from the camera, and product clustering. A drying process that provides such chaotic condition is fluidized bed drying. When images are captured in real-time, all the features unique to chaotic-type images are present. We can observe them in Figure 2.2 that shows images of green peas captured during fluidized bed drying.

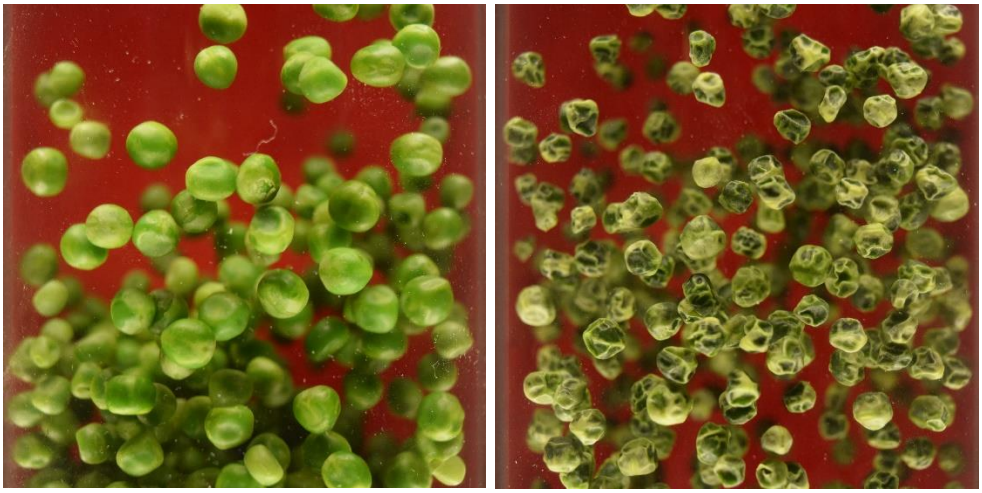


Figure 2.2 Chaotic-type images of green peas captured during fluidized bed drying

Many industries make use fluidized bed dryers. However, in this review, our discussion will be limited to the pharmaceutical and food industries. In both sectors, it has become very important to monitor visual attributes of products, in real-time, while the fluidized bed dryer is in operation. In recent years, researchers within the pharmaceutical sector have developed various computer vision solutions to monitor visual attributes during the granulation and coating processes. Both processes are part of the pharmaceutical tablet or drug manufacturing process. Granulation and coating are conducted within a fluidized bed dryer.

The formulation for producing pharmaceutical tablets usually comes in powdery form. Before compacting the powder into tablets, we first determine the flowability of the powdery formulation. Flowability of the formulation is obtained by measuring its angle of incidence. If the angle is too low, it means the formulation has high flowability, hence, the powdery formulation is very good for direct compression into tablets. On the other hand, if the angle of incidence is too high, it means the flowability is poor. In most cases, the powder would cause arching, bridging or ratholing within the tablet press. To prevent these undesirable outcomes, before compressing the formulation, we would first pass the powder through two processes, after which it would be introduced into the

tablet press. The two processes are known as granulation and coating.[64] Granulation is done for a couple of reasons. First, to improve flowability of the formulation and secondly, to make the formulation more compressible. The third reason is to reduce segregation and eliminate excessive numbers of fine particles, and finally, to spread the API (Active Pharmaceutical Ingredient) evenly within the formulation [62-64].

There are two types of granulations – dry and wet granulation. During wet granulation multiple individual powdery particles are bound together by a wet binding agent to form larger size, free-flowing and dust-free granules. Wet granulation is carried out using either a fluidized bed granulator, a planetary mixer granulator, or a rapid high shear granulator.[64] Currently, various computer vision systems have been developed to monitor visual attributes of granules during fluidized bed granulation. The key visual parameter monitored by CV systems is the size of the granules. Once the agglomerated granules attain the expected size, the binder spray nozzle is turned off, and granules are left to dry while suspended in the stream of hot air flowing through the fluidized bed. Once drying is completed, the granules could be coated immediately while they are still inside the fluidized bed granulator. The granules are coated for key reasons,[65,66] some of them include:

a. To increase the size of granules
b. Modify surface color
c. Delay release of Active Pharmaceutical Ingredient (API)
d. Mask the taste of the granule
e. Protect the granules from sunlight and moisture, and
f. Prevent degradation of the API

To achieve these objectives, it is important to apply the correct amount of coating. So, how do we determine when we have sprayed enough coating on the surface of the granules? They are achieved through the measure of coating thickness. Coating thickness could be measured manually by scanning electron microscopy (SEM). However, the method is too expensive and is not suited for real-time monitoring. The measurement could also be conducted using the disintegration and dissolution method.[22] However, the process is offline, destructive and time consuming. Furthermore, when dissolving the coated pellets, there may be loss of pellet core material. This loss would affect the accuracy of the result.[22]

Fortunately, within the last ten years, significant progress has been made towards developing real-time CV solutions for monitoring coating thickness. Watano et al. [67] developed a CV system for online and in-process monitoring of granule size distribution and shape during fluidized bed granulation. The system was used to monitor the granulation of lactose and cornstarch, mixed at 7:3 by weight. The mixture was agglomerated by spraying an aqueous mixture of hydroxypropyl-cellulose solution. Real-time images of the granules were captured by an image acquisition probe and processed using an image processing system. Granule size was monitored by computing Ferret diameter, while circularity and aspect ratio were used to monitor shape.

During the study, the graph of CV-generated mass median diameter versus moisture content agreed with data obtained through sieve analysis. The graph correctly showed that granule size increased as granulation progressed. Similarly, the individual graphs of circularity against operational moisture content and aspect ratio against operational moisture content, both agreed with the results generated by the image processing system. Likewise, the offline result obtained through manual analysis of granule microphotographs (produced after drying), also agreed with the data generated by the CV system. Circularity and aspect ratio increased with moisture content, meaning that the granules became more spherical as granulation progressed. With this CV solution, the fluid bed operator would have timely access to real-time data needed to achieve consistent end-product quality.

During coating, challenges such as unwanted agglomeration could occur. Unfortunately, this problem may not be noticed till the end of the coating process, by then it is too late to make any changes. To solve this problem, Možina et al.[62] investigated the use of real-time image processing as a Process Analytical Technology (PAT) for spotting out such anomalies during the fluid-bed pellet coating process.

Building on the success of Watano[67], Oman et al.[22] investigated the possibility of using visual inspection technique as a PAT tool for real-time monitoring of coating thickness during the pharmaceutical pellet coating process. Let us discuss how their CV pipeline works in real-time. First, at each interval, the system captured and segmented images of coated granules. From the segmented output, the Equivalent Spherical Diameter (ESD) and Coating Thickness Estimation (CTE) were computed. Both values were then used to monitor coating thickness. The results generated from the study were in good agreement with a popular study on automatic visual

inspection of pharmaceutical pellets during the coating process.[68] Thus, this new method by Oman et al.[22] was suitable for in-process and online monitoring of coating thickness.

In 2014, Kadunc et al.[6] conducted a study that reinforced how effective online CV systems are in monitoring visual attributes of coated granules, using chaotic-type images captured during the coating process. During their study, they monitored coating thickness from ESD and CTE. Interestingly, when their online CV-based results were compared to those generated using the spectrophotometric analytical method, there was statistically acceptable agreement between both results. The online method achieved a R^2 value of 0.998, when predicting coating thickness estimation (CTE).

Considering that CV technology has been successfully used to monitor pharmaceutical pellets during fluidized bed granulation and coating, there is the strong possibility that the same technology could be adapted to monitor real-time changes in visual attributes of **food** during fluidized bed drying. Unfortunately, to the best of our knowledge, this aspiration is yet to be achieved. We are yet to come across any research publication that monitored real-time changes in visual attributes of food during fluidized bed drying. Due to this setback, one may ask "why hasn't such an advancement been made?"

This setback could be attributed to two unique challenges associated with chaotic-type images of food captured during fluidized bed drying. The challenges in question are non-existent or not prominent for chaotic-type images of pharmaceutical products captured during granulation and coating. Whenever theses challenges appear within an image, current segmentation algorithms for agglomeration and coating are mostly ineffective at identifying the pixel positions occupied by individual food materials. Unfortunately, going by the Two-pass framework, if we can't segment the image, computing visual attribute becomes unattainable. So, what are these challenges that have made effective segmentation of chaotic-type images of food a mirage? Our answers would be described using Figure 2.3 of our preliminary investigation.

This figure contains chaotic-type images of green peas captured at different stages during fluidized bed drying. As we observe the figure from left to right, we would notice that as drying progressed, the texture and color distribution on the surface of individual peas increasingly becomes non-uniform. This non-uniformity in color and texture is the major reason why the classical CV

technique used for granulation and coating fails to segment real-time images of food captured during fluidized bed drying.

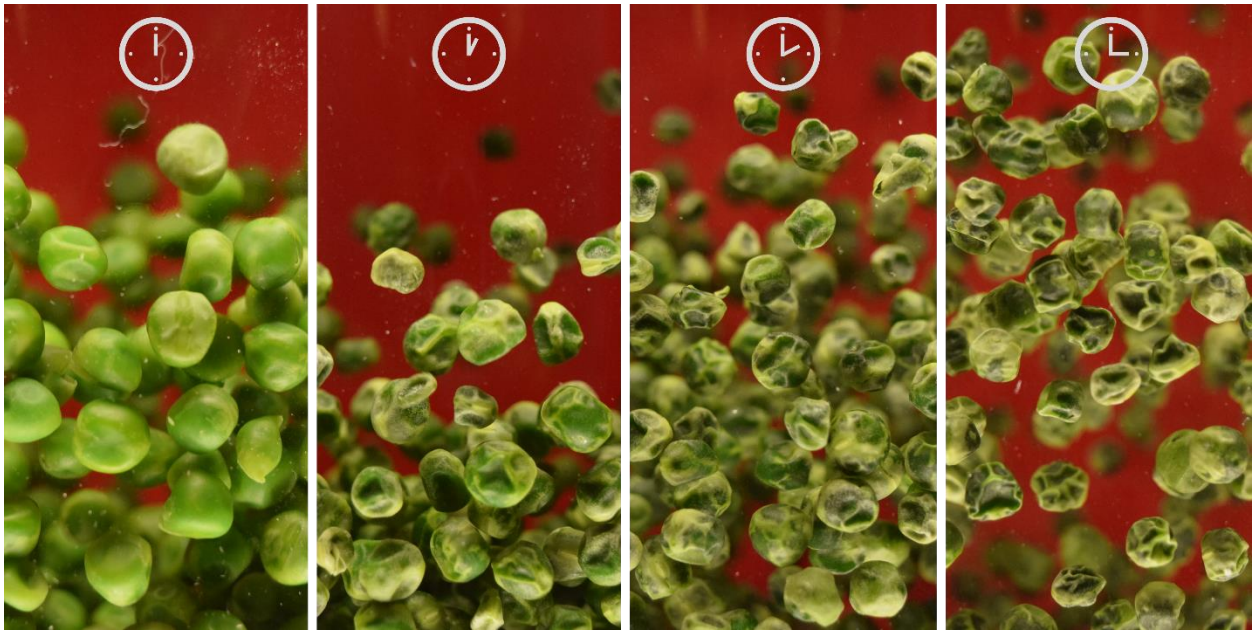


Figure 2.3 Images of green peas captured at difference stages during fluidized bed drying.

To understand how this problem impacts segmentation, let us take a step backward and compare similarities between the two classes of chaotic-type images we have discussed so far in this review. Once we consider what chaotic-type images of food and pharmaceutical products have in common and include the problem of uneven texture and color into the mix, the effect of non-uniformity becomes conspicuous. So, what do they have in common? Both images contain objects of interest that overlap themselves, that are clustered, and are at different distances from the camera. According to the two-pass framework, to compute visual attributes, our first step is always *image segmentation*. After segmentation, the output mask is used to compute visual attributes of the food within the RGB image. However, one question stands out.

"Considering that the peas in Figure 2.3 are overlapped, and sections of some of the overlapped peas are hidden by one or more peas in front of them, how do we compute actual product size, at different stages during the drying process?"

While we ponder that question, here is another difficulty we need to solve as we journey towards developing a solution for monitoring visual attributes of food, in real-time, during fluidized bed drying. The second question is:

"Since the peas are at different distances from the camera, how do we compute the actual size of the peas?"

This question and the one before have the same answer or solution. This solution was applied by various researchers when developing CV systems for monitoring pellet thickness, during fluidized bed coating [6,22,62,68]. During fluidization, the food materials mix up as they randomly jump up and down within the drying chamber. Due to this mixing action, there would be a continuous and replenishing layer of peas on the wall of the drying chamber that faces the camera (Figure 2.4). In this review, we would refer to this layer of peas as ***"infocus peas"***, while the other peas behind the infocus ones would be referred to as the ***"defocused peas".*** Additionally, due to the mixing action during fluidization, each time an image is captured, a new set infocus peas would appear in the image. Hence, as we know the distance between infocus peas and the camera, and as we also know that a different set of infocus peas appear each time a new image is captured, we could compute the average visual attributes of the peas, for a particular time step, by doing the following. At the specified time step, capture multiple (eg. fifteen) images of the food material. The average visual attribute of the peas at that instance would be the mean visual attribute of all the infocus peas within all fifteen images.

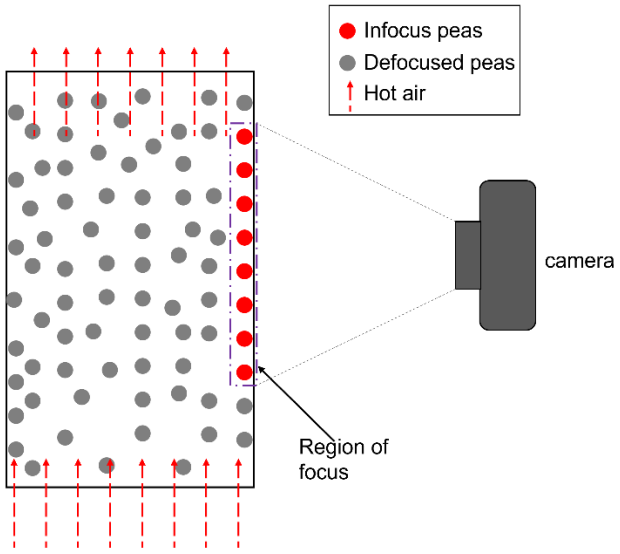


Figure 2.4 Infocus and defocus seeds during fluidized bed drying.

Since the infocus peas are not blocked by any other peas, the infocus peas are the objects of interest that would be used to compute color, texture, shape, and size of the peas during fluidized bed drying. Hence, the goal would be to develop an image segmentation pipeline that receives a real-time RGB image containing infocus peas, defocus peas, and image background, then produces a mask containing only infocus peas. This method was applied by other researchers for monitoring granulation and coating.[6,22]

Unlike pharmaceutical pellets, the case is different for food materials. As we mentioned earlier, segmentation of chaotic-type images of food is significantly affected by the non-uniformity in color and texture of individual food materials within the image. The color and texture on the surface of individual pharmaceutical pellets are mostly uniform. Hence, during image segmentation, it is more straightforward to obtain complete outlines of infocus pellets. However, for food materials, the non-uniformity in texture and color, combined with overlapping and clustering, poses a hinderance to detecting the complete outline of infocus peas. Occasionally, due to sharp changes in texture and color, the gradient intensity along the surface of infocus peas would have similar intensity values as the edges of some defocused peas. When such a threshold is applied, sections of infocus peas and defocused peas would appear on the segmentation mask.

Also, due to clustering of infocus peas whose surface color and texture is non-uniform, the gradient intensity at the edge where two clustered infocus peas meet, would have a lower value than the value obtained at their unclustered edges. This low gradient intensity most of the times would tally with values associated with defocused peas, and sometimes the value would be the same as some regions in the background. Hence, the segmented result would contain incomplete sections of infocus and defocused peas.

Normally, image segmentation is based on thresholds, gradients, regions or on classifications [57,58]. To develop a CV system for chaotic-type images of food, multiple segmentation methods may have to be integrated. However, even after combining multiple techniques, the classical computer vision approach may still run into problems. It may not be very effective in segmenting overlapping and clustered food materials. To improve segmentation results we would need to integrate deep learning techniques, by introducing semantic segmentation models, such as U-Net,[69] DCAN, [70,71] and DeepLabV3 [72] to mention a few.

In line with the Two-pass framework, a deep learning model for predicting visual attributes would comprise of two sub-models. The first sub-model would be for segmentation. The input layer of this model receives an RGB image containing infocus peas, defocus peas and image background. At the output layer, it would produce an RGB image containing only the infocus peas. This result is then fed into the second sub-model that predicts one or more visual attributes. This second sub-model could one of the state-of-the-art convolutional neural networks, such as ResNet-101[73], Inception-v4[74], VGG-16[75], GoogleNet[76], or Xception[77], to mention a few. Figure 2.5 describes the layout of the Two-pass deep learning solution.

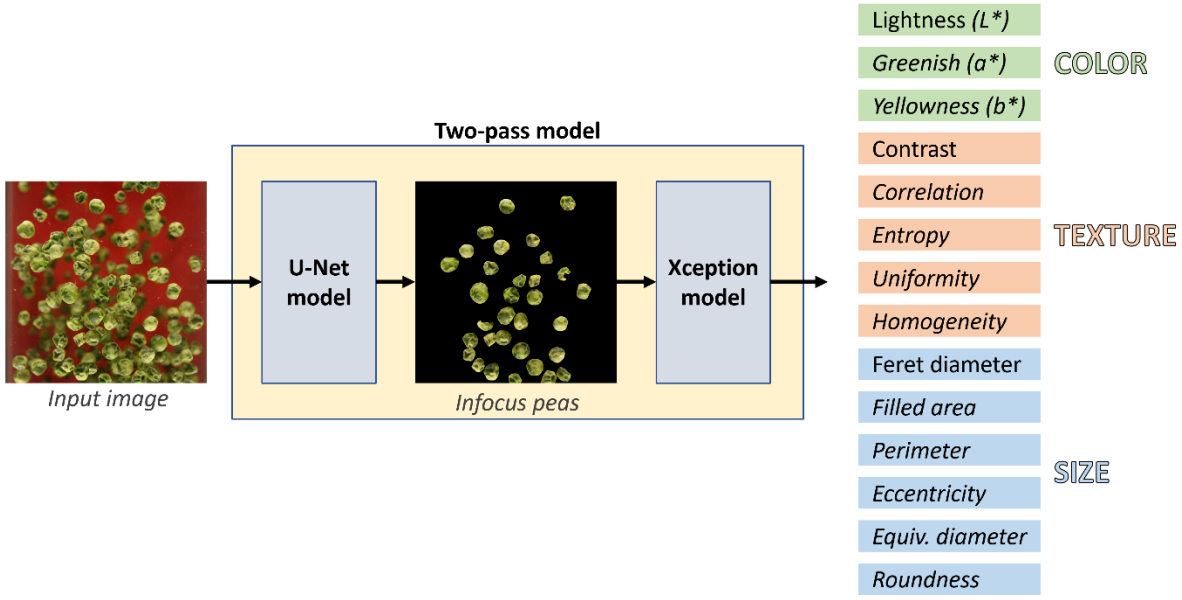


Figure 2.5 The Two-pass deep learning model for predicting visual attribute of food during fluidized bed drying

Each of the sub-models would be trained separately, then merged. The segmentation model would be trained using complete real-time images, while the second sub-model would be trained using RGB images of infocus peas. Due to the ability of CNNs to learn spatial and cross channel relationships between image inputs and expected outputs, deep learning models could significantly improve segmentation results and visual attribute prediction. Figure 2.6 describes the *"train and merge"* training protocol.

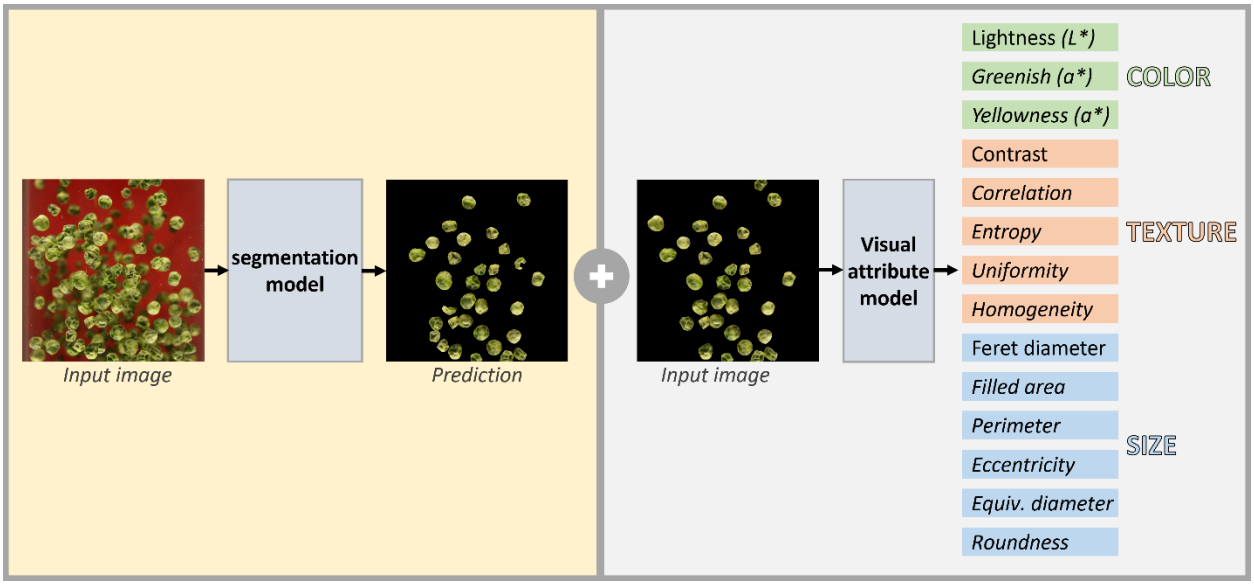


Figure 2.6 Training protocol for the Two-pass deep learning model

To learn vital relationships needed to detect infocus peas (objects of interest) within the original image, the deep learning model would require numerous examples of real-time images and their corresponding ground truth masks. Usually, these ground truth masks are manually annotated using different annotation tools, such as LabelMe®.[78] The manual process makes it difficult to come up with sufficient training examples. To overcome this challenge, we apply image augmentation. During this process, various random transformations are applied to the original images and their corresponding ground truth masks, to produce realistic variants that increase the number of training examples.

Some of the image transformations applied during augmentation include left-right flip, up-down flip, random rotation, and random crop. Figure 2.7 contains some realistic variants obtained through image augmentation. If the texture and color of the food material is part of the visual attributes that would be monitored, transformations such as adjusting the brightness, the saturation, or the contrast of the RGB image may not be applied. Because it is unlikely that images with such modified brightness, saturation and contrast could be obtained during actual drying runs.

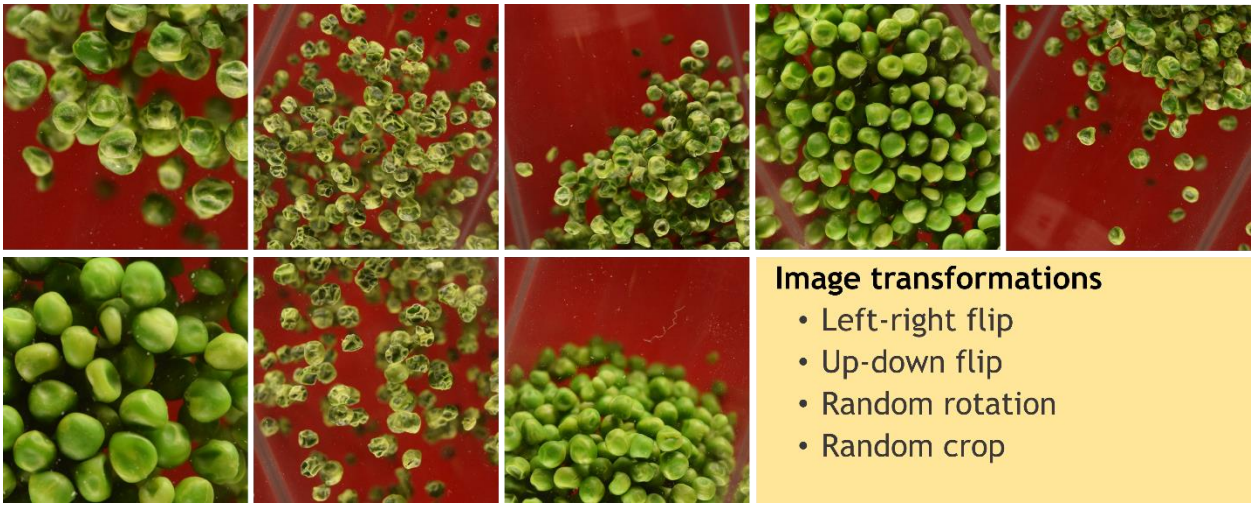


Figure 2.7 Realistic variants generated through image augmentation

2.4. SINGLE-PASS FRAMEWORK

So far, we have been discussing how visual attributes of food and pharmaceutical products could be monitored in real-time using the Two-pass framework. However, one may ask:

"*Is it possible to develop a **Single-pass** solution that by-passes the image segmentation step? Is it possible to have a model that receives real-time image of the food, and computes visual attributes values, without segmentation?"*

To the best of our knowledge, there is no publish work on such developments, both for static-type and chaotic-type images. When developed using the deep learning approach, a single-pass solution could significantly reduce computational load, as the solution would require only one model, unlike the Two-pass solution that requires a model for segmentation and another for predicting visual attributes from the segmented output.

Since CNNs are very effective in learning spatial and cross-channel relationships between images and expected regression outputs, experimenting with state-of-the-art models mentioned previously can be investigated. While training such model, the input would be the original and augmented RGB images, while the ground truth would be various continuous variables that represent visual attributes. Color could be represented by Lightness (L*), greenish index (a*), yellowish index (b*) and total color change (ΔE). Visual texture could be measured in terms of contrast, correlation, energy, entropy, homogeneity, and uniformity. While Feret diameter, filled area, perimeter, and

equivalent diameter could be used to monitor product size. Figure 2.8 describes the Single-pass model.

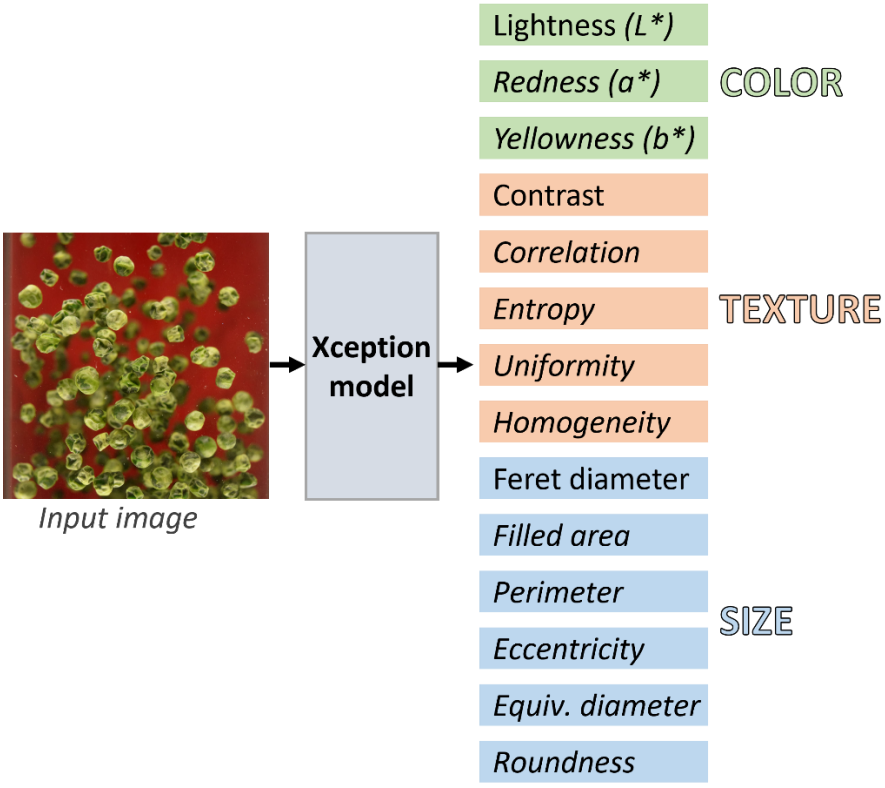


Figure 2.8 The Single-pass solution for predicting visual attributes of food during fluidized bed drying

This section on the Single-pass solution brings us to the end of discussion on how real-time changes in visual attributes of food could be monitored during fluidized bed drying. However, before we conclude this chapter, let us look at the big picture again. Why do the Two-pass and Single pass solutions matter? What role would they play in the broad scheme of things?

We discussed why pharmaceutical based CV systems would fail when segmenting images of food captured during fluidized bed drying. If segmentation results are poor, visual attribute prediction would also be poor. Hence, it would be impossible to achieve consistency in dry product quality, leading to increased food loss and waste. The Two-pass deep learning solution could solve this problem, by improving image segmentation, and by extension visual attribute prediction. The single-pass solution goes a step further; it would cut down the computational load, cost, and time, which in turn, could significantly reduce the cost of such intelligent system.

When the Two-pass or Single-pass system is coupled with a control system and integrated with current drying systems (fluidized bed dryer and other dryers), we would end up with a dryer that can monitor and control the color, texture, size, shape, and moisture content of the food. Hence, we could end up with a food dryer that consistently produces dried foods that are visually attractive and have long shelf life.

2.5. CONCLUSION

Food drying could be an effective method to significantly reduce global food waste and loss. To achieve this goal, we would need dryers that ensure no food is lost during drying or wasted after drying. This means the dryer should have the capability to consistently produce dried foods that are visually appealing and have long shelf life. Researchers have achieved this milestone for several drying processes, using computer vision (CV) technology that can monitor real-time changes in visual attributes of food. The CV system computes visual attributes from images of the food captured during the drying process. However, such an online solution has not been developed for fluidized bed drying of food products.

This setback was attributed to challenges in segmenting real-time images of food captured while the products were randomly jumping up and down during fluidized bed drying. To come up with an effective solution, this chapter reviewed various computer vision systems based on the ***layout***, *the* ***state*** *and* ***patten of motion***, *and the* ***behaviour*** *of food materials* within the image used for computing visual attributes. Based on these three criteria, each CV system was characterized as either a Static-input or Chaotic-input CV system. This review shows that mostly static-input CV solutions have been used to monitor visual attributes of food during drying. The only application of chaotic-input CV system was in the pharmaceutical industry, for monitoring the granulation and coating process.

Building on the success of pharmaceutical CV solutions, as well as the challenges this same system would encounter when applied to food materials, this review offers improved approaches for monitoring visual attributes of food during fluidized bed drying. Two of these solutions include a Two-pass deep learning model and a single-pass deep learning model. Both have the potential to significantly improve visual attribute prediction, during fluidized bed drying of food. When combined with a control system (powered by reinforcement learning), the merger would ensure

that dryers consistently produce dry foods that are visually appealing and have a long shelf life. Hence, dry food waste and loss could someday become a distant memory.

CONNECTING TEXT

In Chapter II, we reviewed various research works where computer vision technology was used to monitor visual attributes of food during drying. We pointed out that based on published literature, there is no chaotic-input CV system to monitor visual attributes of food in realtime, during fluidized bed drying. The next chapter works on solving this problem.

To achieve such improvement, the work in the upcoming chapter would be geared towards developing a chaotic-input CV system to monitor visual attributes of green peas, in real-time, during fluidized bed drying. This CV system would be made up of three broad parts:

a) An image acquisition system for capturing chaotic-type images of green peas
b) A CV pipeline to segment the chaotic-type images, and
c) A CV pipeline for computing visual attributes using the segmented output.

By developing such effective CV system for monitoring visual attributes of peas, using chaotic-type images, we would make significant progress towards achieving the first and second milestones needed to advance the state of the Art (Figure 2.9).

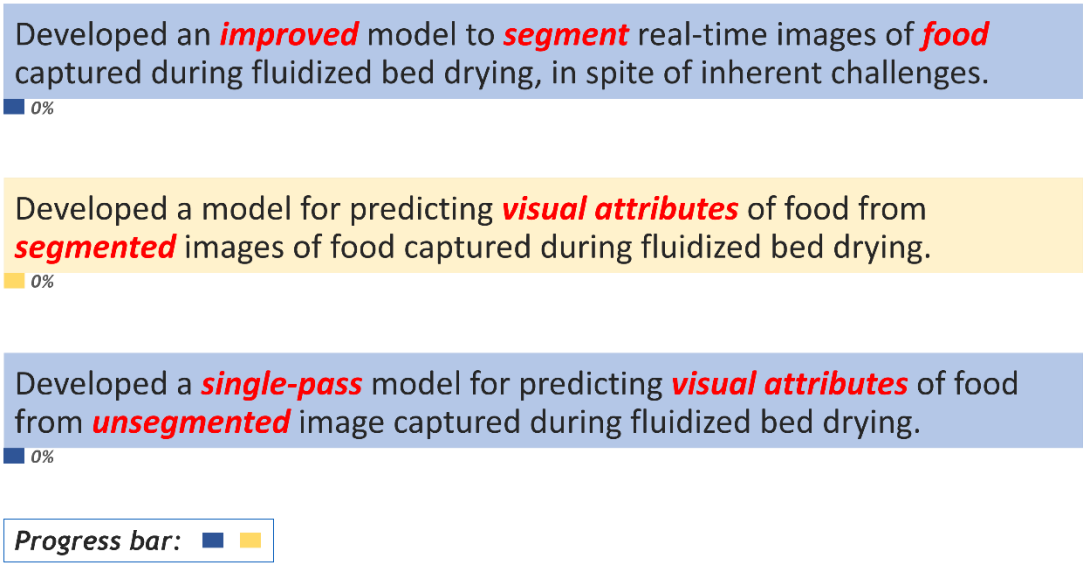


Figure 2.9 Milestones to improve the state-of-the-art

CHAPTER III

COMPUTER VISION FOR REAL-TIME MONITORING OF SHRINKAGE FOR PEAS DRIED IN A FLUIDIZED BED DRYER

3.1. ABSTRACT

This research work addresses the use of computer vision (CV) to monitor in real-time, the shrinkage of peas drying in a fluidized bed dryer. The experiment was conducted at a constant air velocity of 10 m/s, and at temperature levels of 50, 55 and 60^{o}C, respectively. Shrinkage was monitored separately with a real-time and an offline CV system. The monitoring was carried out by measuring the area, perimeter, and the diameter of peas at specific intervals during the drying process. A real-time CV algorithm has been developed to compute the average area, perimeter, and diameter at each interval, by detecting and analyzing only seeds that were within the region of focus of the camera. The result from the offline method showed a steady increase in shrinkage as drying progressed. Similarly, the results from the real-time shrinkage curve, from 0 to 15 minutes were similar to those on the offline curve. The real-time classical CV algorithm encountered figure-ground separation problems, especially when the peas were overlapped, clustered together, or variation in color and texture across the surface of individual peas. At the end of the study, it was noted that areas of improvement which could enhance the results of the real-time method are needed.

3.2. INTRODUCTION

Think about the last time you visited the food store to buy fruits or even vegetables. Let us say you were in the store to buy apples; how did you decide which apple to pick? If you are like most consumers, your eyes are first drawn to apples that are attractive in color. After moving your eyes back and forth through the pile of apples, finally, your eyes are fixed on a couple of apples that are attractive in color.

After inspecting the color, we move on to assess the texture. The texture describes how smooth or wrinkled the outer surface of the apple is. Most customers do not wish to buy wrinkled apples, it sends the message that the apple has lost its quality attributes while in storage. Once we are done with color and texture, we swiftly move on to examine the size and shape. Occasionally, we grab

the apple within our palms, then press it a bit with our thumb, to have an idea how crunchy it could be inside. So, within few minutes, while standing in the food store, "color, texture, size, shape and hardness" are the physical characteristics that help consumers decide which fresh food to select. [7-10]

A similar scenario plays out when consumers visit the food store to purchase dried foods. Be it nuts, seeds, fruits, or even dried vegetables, you would realize, that within a few seconds, the potential consumer finds him/herself comparing the physical characteristics of the dried product, to the physical characteristics of the same product when it was fresh. Most consumers are attracted to dried food products that are similar in color, texture, size, shape, and aroma to its fresh counterpart. Once the consumer spots out any pack of the dried product that exhibit such similarity, this resemblance is quickly associated to high nutrient content, longer shelf life and sometimes high rehydration capacity when used for cooking.

Take for instance that the customer came to the food store to buy dried onion flakes. However, instead of finding transparent bags filled with thinly cut, purple strips of dried onions, what was found on the shelve were bags filled with deep brown, wrinkled, shriveled, and shrunk onion strips. What do you think would be the consumer's conclusion? He/she would swiftly conclude that the onion has lost a lot of its nutritional compounds during the drying process. And since it looks so shriveled, shrunk, and wrinkled, the dried product may be difficult to rehydrate when used for cooking.

Food stores are aware that their customers are mostly attracted to dried foods whose physical appearance meets a certain threshold. Consequently, to maintain consumer patronage, and avoid losses, these stores set a minimum quality benchmark when accepting supplies of dried products from food processors. Due to this minimum standard, it has become indispensable for food processors to integrate into their drying process, techniques that could help them monitor the changes in the physical appearance (color, size, shape, and texture) of the food during the drying process. Integrating such visual monitoring systems has the tendency to increase demand for dried products and mitigate the risk that could arise from losses due to products that do not meet the agreed benchmark. Also, with such visual feedback systems in place, the food processor would be able to adjust the drying conditions (such as air temperature, air velocity, and air pressure) in a timely fashion, so that the dried product would turn out as expected.[11-13]

As per the current trends, it is now possible to monitor changes in physical characteristics of food during drying, by using Computer Vision (CV) systems. This technology is made up of a lighting system that adequately illuminates the food material; and a digital camera that captures and transmits images of the food to a computer. Within the computer, the images of the food are extracted from the background using appropriate CV algorithms.[10,13-19,21,22] After extraction, these algorithms analyze the image of the food, then interpret and display their physical properties for that point during the drying process.

CV technology has been used in various drying systems, such as convective,[8,13-17,20,21,46,47,51-54] microwave,[49] fluidized-bed,[6,22,26,62,63,79] and osmotic-convective drying systems[50,51]. The application of CV systems in the drying industry has evolved over the years.

First, this food imaging and analysis technique started out as an offline method for evaluating the quality of dried foods, at specific times during the drying process. It concentrated mainly on building relationships between the visual appearance of the food (color, size, texture, and shape) and the food quality attributes such as moisture content, density, and porosity, at different stages during the drying process.[6] In the area of convective drying, Madiouli et al.[80] studied the shrinkage kinetics of banana slices. During this study, they developed a relationship between the shrinkage and porosity of the banana slices. Ramos et al.[81] also investigated the radial shrinkage in grapes. While carrying out their analysis, they developed a correlation between shrinkage and moisture diffusivity for grapes. In addition to monitoring shrinkage, Zenoozian et al.[51] investigated how osmotic plus convective drying affected the color and textural changes in pumpkin. One more application of offline imaging in convective drying was the study on how drying temperature and shape influences shrinkage in potato and cauliflower stems. During this research, Mulet et al.[19] compared three potato shapes – cubes, parallelepipeds, and cylinders. They observed that the shape of the potato affected its shrinkage. For the cube, which had uniform sides, the shrinkage was uniform on all sides, however for the parallelepipeds and cylinder that had sides with different lengths, the shrinkage at the longer side was less than that at the shorter ends. This discrepancy was linked to the influence of core drying during shrinkage. For the cauliflower stems, the difference in shrinkage between directions was greater than those observed for potato. This variation between cauliflower and potato was attributed to the effect of fiber alignment along the axis of the cauliflower stems. Besides convective drying, offline imaging has been used to monitor the changes in physical characteristics of food in a microwave dryer.

Huang et al.[47] used hyperspectral imaging technology to capture images of soybean at different stages during the microwave drying process. With these images, they were able to study the relationship between color change and moisture content of the soybean seeds. At the end of the study, Huang and his colleagues concluded on the existence of a correlation among reflectance, entropy, and moisture content. Still on microwave drying, a study was carried out to investigate the effect of microwave-vacuum drying on the shrinkage, shape, and color of carrot. Nahimana and Zhang[49] monitored the changes in the physical properties of the carrot, as the drying conditions changed. Just as it has been used to monitor microwave drying, the offline imaging technique has also been applied to the fluidized bed drying process. This technique has been widely used to monitor the drying of powdery/granular substances. Kucheryavski et al. [63] investigated the change in shape and size of pellets during the coating and drying process in a fluidized bed. During their study, they examined the effect the coating substance and drying time had on the thickness of the coated pellets. With the offline imaging technique, the researchers were able to monitor the increase in pellets' thickness as the coating substance was sprayed on the surface of the pellets.

Although the offline imaging method came with a lot of advantages, it also came with major challenges, because of its configuration. The offline imaging system was configured to be a separate unit from the drying equipment. When it was time to analyze the properties of the food product, the food processor had to open the dryer, collect some samples, then move the samples to a separate enclosure that housed the lightening system and camera. The image capturing was carried out in this enclosure. Afterward, analysis followed. This sequence of operations made the offline imaging technique slow. But in addition to being a slow process, the frequent opening and closing of the dryer induced fluctuation in the drying conditions, such as the temperature, velocity, pressure, and relative humidity of the hot air within the dryer. This variation in drying conditions introduces errors and uncertainty to the quality of the dried food.[12] Another significant setback with the offline imaging system is that the information we get from the analysis does not tell how the physical features of the food are changing at every point during the drying cycle, instead it only gives us information on the state of the food at the point when the sample was removed from the dryer. Take for instance that we want to monitor the color of carrots dried in a convective dryer, by collecting samples from the dryer every 15 minutes for examination. If this process were carried out 5 times throughout the drying duration (i.e., 0, 15, 30, 45, and 60 minutes), you would be able

to compare the color of the carrot at 5 stages. However, what was the color of the carrot at the 50th minute? This would be hard to assess. Supposing the color of the carrot at the 50th minute was unacceptable, we would only have discovered it after 60 minutes, by then, the damage would have been done. To solve this problem, it would be nice to have an imaging system that provides real-time feedback on the changes in the physical features of the food as drying progresses. The drive to create this solution was what gave birth to the use of real-time/online imaging in food drying.

There are three features that makes the real-time imaging technique unique. First, the camera and illumination system are integrated with the dryer. The camera lens points directly on the food that is being dried to capture images of the food even while they are in the dryer, thus eliminating the errors and uncertainty introduced by frequent opening and closing of the dryer. Secondly, since we did not need to open the dryer anymore, it became feasible and attractive to immediately analyze the images as they are captured. It is then possible to capture and analyze a series of images as the drying progresses. For example, by capturing an image of food every second, we can extract information that goes beyond just relating the physical features of the food to its quality attributes (as it is done in offline imaging). With real-time imaging, we are now able to develop smoother trends that show the minute-to-minute changes in both the physical features and quality attributes of the food. With such trends, the real-time imaging technique can become a vital tool that relates image features to quality degradation.[10,12,16,53,82] Real-time imaging has been extensively applied to convective drying.

Chen and Martynenko[10] used this imaging technique to investigate the effect of temperature on shrinkage and color of two blueberry varieties. After conducting a statistical analysis, they concluded that the drying temperature had a significant effect on color. Interestingly, this same conclusion was also detected on the graph of color change versus time. From the graph, it was observed that when the blueberries were dried at 55 ^{o}C, within the first 30 minutes, the blueberries had undergone a huge change in total color difference (ΔE) of more than 20. Whereas, when the blueberries were dried at 50^{o}C, within the same time interval, the total color difference was not even up to 5. Thus, real-time monitoring of color using computer vision is an effective tool for monitoring the quality degradation of blueberries even while drying is taking place. The same technique was applied by Yadollahina and Jahangiri[15] to monitor the shrinkage of potato slices, dried at three different temperatures - 60, 70 and 80^{o}C. During their study, they developed a relationship that connected moisture content of the potato slices to the shrinkage associated with

its area and perimeter. They concluded that the airflow direction within the dryer has a significant effect on the shrinkage of the parallel and perpendicular diameters of the potato that were dried at 60 and 70 ^{o}C, but not for those dried at 80^{o}C. They arrived at this conclusion after analyzing consecutive images captured using real-time imaging. Similar research was conducted by Hosseinpour et al.[17] on shrimps. These researchers examined the effect of air temperature and air velocity on the color change kinetics. This was done by observing how the lightness, redness, and yellowness indices, together with the total color difference, chroma hue angle and browning index changed during convective and superheated steam drying. These researchers concluded that increase in drying temperature decreased the lightness index and the hue angle but increased the other color indices. Also, the change in moisture ratio of the shrimps had a stronger relationship to the lightness index (L), than it had to the reddish (a*) and yellowish (b*) indices. These relationships were easily spotted out on the graph of L, a* and b* versus time. In a similar manner, Nadian et al.[55] improved the quality of dried kiwifruit after monitoring its shrinkage and color change. With the information they gathered during the analysis, they optimized the drying process by reducing drying time and color change. Apart from convective drying, real-time computer vision technology has also been applied to fluidized bed drying, especially in the pharmaceutical industry, where it was used to monitor the morphological features and flow characteristics of powder and tablets.

This approach was used to determine the particle velocity and voidage profile in a spouted fluid bed coating device.[83] It was also applied by Watano and Miyanami[67] to monitor the granule distribution and shape of pharmaceutical tablets. Kadunc et al. employed the same technique to monitor the increase in the size of pharmaceutical tablets during spray-coating in a fluidized bed. The morphological results obtained from this imaging method agreed with the values that were obtained by using the conventional offline testing method.

To the best of our knowledge, the use of the real-time CV technology[6] for monitoring the fluidized bed drying process has only been applied in the pharmaceutical and powder industry. We have been unable to find any publication on its application in food drying.

3.2.1. Hypothesis

While carrying out this research, our hypothesis was built on three considerations:

a. To determine the average morphological values (area, diameter, and perimeter) of peas moving randomly in a fluidized bed. The first step would be to capture an image of the peas while they are in motion. After that, produce the outline of each pea, then count the number of pixels within the outline of each pea. The average number of pixels would give us the average area of the peas within the image. Likewise, the average perimeter would be the average number of pixels along the outline of the pea; and the average diameter of the peas would be the same as the average diameter of a circle that is equal in area to the average area of the peas.[58]

b. During the drying process, as the peas lose water, they reduce in size. By capturing and analyzing consecutive images of the peas, the reduction in size over the drying period would give us the relationship between shrinkage and time.

c. Take for instance that we have two peas of the same size in front of a camera lens. If one of the peas is closer to the lens than the other, and an image containing both peas is captured, you would notice that the closer pea would look larger (and thus have larger morphological values) than the pea behind it. To monitor shrinkage, we need to know how far each pea (used in computing shrinkage) is from the lens. Once we know their individual distance from the lens, we can obtain their actual size. Hence, in this work, the depth of field of the camera would be set to contain only those peas that are touching the interior wall of the fluidized bed chamber. Therefore, we assume that during fluidization, as the peas move up and down, a continuous layer of peas would always touch the inner wall of the transparent cylindrical drying chamber.

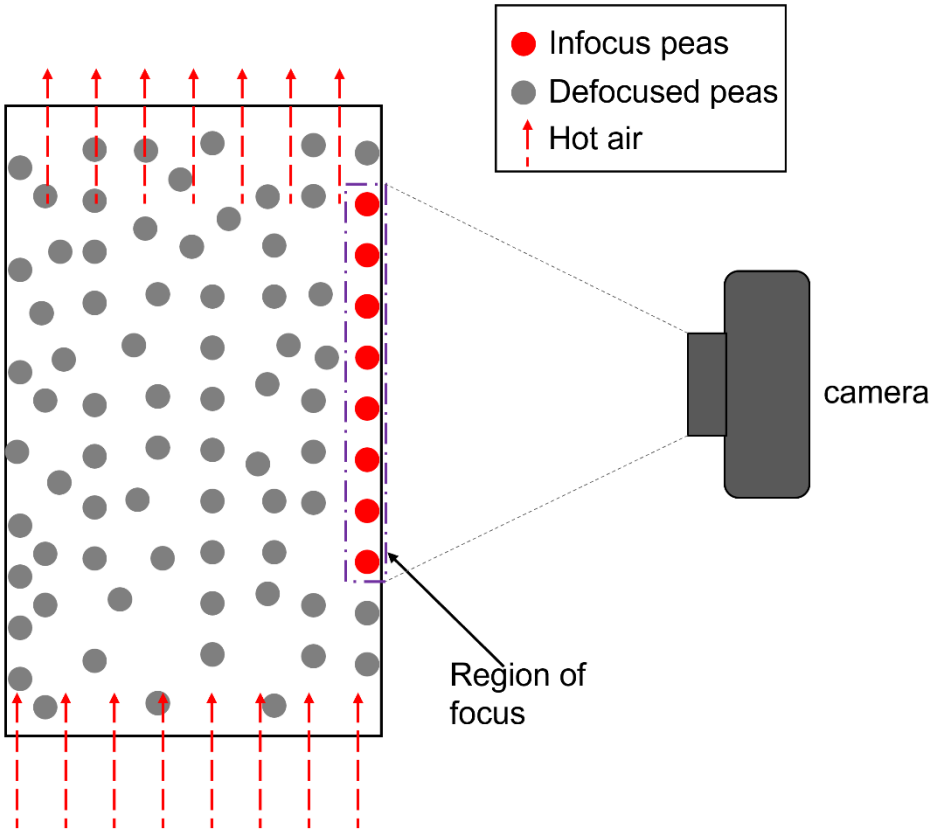


Figure 3.1 Infocus peas on the surface of the cylindrical drying chamber.

3.2.2. Objectives

The objective of this work is to develop a computer vision system for monitoring real-time shrinkage of peas during fluidized bed drying, by examining the average change in area, diameter, and perimeter of the peas during the drying process. The next objective is to evaluate the performance of the real-time CV system, by comparing its results to those obtained by an offline CV system, in order to come up with techniques and new areas of research that could improve real-time monitoring of shrinkage during fluidized bed drying.

3.3. MATERIALS AND METHODS

3.3.1. Sample Preparation

The peas (Pisum sativum spp.) that were used for the experiment were purchased from Marché Richelieu, a nearby supermarket in Ste-Anne-de-Bellevue, Montreal, Canada. The peas came in

packs of 700 g. After purchasing the peas, each pack that was not being used was stored in the refrigerator at 4°C. Before drying the peas, they had to undergo some pre-drying steps.

First, they were conditioned at room temperature for about four hours. After conditioning, the peas were sorted by removing all the broken peas. The reason for eliminating these peas was to prevent them from introducing error during the image analysis process. If a pea that is split in half rotates insides the fluidized bed, when the images of the peas are captured, the broken peas would display varying shapes, depending on its orientation. At some point it would be displayed on the image as a circle, some other time it could be a semicircle or even an ellipse. As the shape changes, the area also changes, which in turn would affect the average area of the peas in the dryer. That is why broken peas had to be removed from the pack of peas that were used for the experiment. After eliminating the broken peas, the next step was to determine the size distribution of the peas.

The size distribution was estimated by measuring the average diameter of 100 randomly selected peas. The diameter of these peas ranged from 8-12 mm. Subsequently, the initial moisture content of the peas was determined by drying five sets of peas, 20 g each, for 48 hours in an oven set to 65°C. At the end, the average moisture content (MC) was computed as 76% (wb), using Equation 3.1 shown below:

$$\text{MC (wb)} = \frac{M_w}{M_w + M_s} \times 100, \tag{3.1}$$

where M_w is the mass of moisture in the sample measured in grams (g), M_s is the mass in grams (g) of dry solid in the sample. The moisture ratio of the peas was computed using Equation 3.2:

$$\text{MR} = \frac{M_t - M_e}{M_0 - M_e} \times 100, \tag{3.2}$$

where M_e and M_0 are the equilibrium moisture content and initial moisture content respectively, on wet basis.

3.3.2. Drying Equipment

The dryer that was used for the experiment was a fluidized bed type, designed and built in our shop, and it has all the key features. The fluidized bed dryer is equipped with a heater, a blower, air flow gate system and a cylindrical drying chamber. The heater was connected to the inlet of the blower using an insulated duct. Once the dryer was switched on, the blower (Lesson Electric,

Wisconsin, USA) sucked in ambient air across the hot fins of the heater (Ceramic heater, Cranbury, USA). As the air passed through, its temperature was increased to the desired level (50, 55 or 60^{o}C). The temperature was maintained using a PID (proportional-integral-derivative) controller (Omega CN132-12V, Canada). Once the air left the heater, it flowed into the blower inlet, moved to its outlet, and from there it got into the drying chamber, through an insulated pipe that connected the blower outlet to the drying chamber. This chamber was made from a cylindrical glass tube, 30 cm in length, with an internal diameter of 10 cm and a 2.5 mm thick wall. Along the pipe that connected the blower and the drying chamber, a manually operated gate system was set up which helps in varying the velocity of the hot air entering the drying chamber. Both the bottom entrance and the top end of the cylindrical drying chamber were covered with a circular sheet of polypropylene mesh. The sheet was used to facilitate streamline flow of air through the drying chamber, but at the same time, prevent the peas from falling out or escaping through the top during fluidization. Furthermore, the fluidized bed dryer used for this experiment had a re-circulatable feature. This feature reduced the energy needed to heat up the drying air. To achieve this reduction, as the hot air streamed out of the drying chamber, instead of escaping it to the environment, part or all the hot air was re-channeled back to the heater. In the current study, over 90% of the air were recirculated.

At various locations within the dryer, there were sensors to monitor the air temperature, humidity, and velocity. The information collected by these sensors were fed to the computer for further analysis and evaluation using a data acquisition system (Agilent 3470A, Keysight Technologies, California, USA) and a data collection software (Benchvue Data Acquisition, Keysight Technologies, California, USA).

3.3.3. Drying Experiment

The experiment was designed using a full factorial design, with one factor, at three levels, replicated three times. The study investigated the effect of three temperature levels, 50, 55 and 60^{o}C on the drying and shrinkage of peas during the fluidized bed drying process. For each of the experimental runs, 350 g (±0.01g) of peas were dried from an initial moisture content of 76% (wb) to a final moisture content of 24% (wb). At an interval of fifteen minutes, while the peas were fluidizing within the cylindrical glass column, ten images of the randomly moving peas were captured at a resolution of 2,992 x 2,000 pixels. These images were used to determine the average

morphological values (area, Ferret diameter and perimeter) of the peas for each interval during the drying process. These values were computed using a conversion factor of 0.00312 mm^2/pixel, 0.05583 mm/pixel and 0.05617 mm/pixel for area, Ferret diameter and perimeter, respectively. By comparing the average values at each interval, the authors were able to determine the shrinkage of the peas throughout the drying process.

After capturing the 10 images, the mass of the peas together with the cylinder was measured using a digital weighing balance (Denver Instrument PI-2002, New York, USA) with ±0.01 g sensitivity. The mass obtained from the balance was used to determine the moisture content of the peas, by substituting the measured values into Equation 3.1. Next, a tablespoon full of randomly selected peas were taken from the cylinder for offline image acquisition and analysis. The offline results were used as a control for evaluating the performance of the real-time method. The offline image was captured at a resolution of 1,920 x 2,560 pixels, and the area, Ferret diameter and perimeter of the peas were computed using a scaling factor of 0.01255 mm^2/pixel, 0.11204 mm/pixel and 0.11226 mm/pixel, respectively. After acquiring the offline image, the peas that were captured were returned to the cylinder for further drying. This process of real-time image acquisition and analysis, moisture content determination, and offline image capturing and analysis, was repeated every 15 minutes, until the peas got to a final moisture content of 24% (wb). For all the drying runs, the velocity of the air was maintained at 10 m/s.

3.3.4. Computer-Vision System

The computer vision system was made up of hardware and software components. The hardware component was comprised of a Digital SLR Camera (Nikon D5300, Nikon Inc., Japan) for real-time image acquisition. The camera was mounted on a 15 mm LWS Rail system (F&V, Illinois, USA). With this rail system, the lens was centered along the line of symmetry of the cylindrical drying chamber. The distance between the camera lens and the drying chamber was fixed at 170 mm throughout the experiment. The second hardware item was an Infinix Note 4 Pro mobile phone (Infinix Mobile, Shanghai, China), used for offline image acquisition. Then a lighting system to provide adequate illumination for the real-time image acquisition. This system comprised of two R300 SE® LED ring lamps (18.7W, F&V USA), having a color temperature of 5,600 K and illuminance of 1187 Lux. The ring lamps were placed horizontally at the top and bottom of the cylindrical drying chamber with the bulbs pointing into the cylinder. The last

hardware component was an offline illumination chamber (Figure 3.2), used for offline image acquisition.

Figure 3.2 Offline illumination chamber

The illumination chamber had one R300 SE® LED. It was positioned horizontally, above the peas to provide uniform illumination during the image acquisition step.

During this study, two software packages were used, digiCamControl and MATLAB R2019a. digiCamControl is a free open-source windows app which was used to control the Digital camera. Through this app, we were able to set the shutter speed, exposure, aperture and capturing frequency of the camera. MATLAB (R2019) on the other hand, was used to develop the image processing algorithms.

3.3.5. Computer Vision Process

To monitor the shrinkage of peas, the process was divided into two phases – image acquisition and image analysis. During the first phase, the system captures images of the peas, while they are in random motion within the fluidized bed. Section 3.3.6 describes the conditions that were put in place to capture images that provided sufficient information for carrying out the next phase of the computer vision process known as *image analysis.* The end goal of the image analysis phase was to extract vital information (such as edges) from each of the infocus peas. Once the edges of these peas are obtained, then their average morphological values (area, perimeter, and diameter) at each stage during the drying process can be computed.

3.3.6. Image Acquisition

As the title suggests, the image acquisition phase, is the stage when real-time images are captured. During this phase, three conditions must be met to ensure that the images would be suitable for analysis. First, there must be adequate and uniform lighting within the drying chamber. Adequate lighting would ensure that the features (color, edge, and texture) of the peas and the color of the background are conspicuous on the captured image. During the experiment, adequate lighting was achieved by using two ring lights (described in section 3.3.4).

Since the peas needed for analysis were those that were touching the glass cylinder, our next step was to create a unique feature that distinguished the peas on the surface of the cylinder from those behind them. To achieve this distinction, the depth of field of the camera was set in such a way that the peas which were on the wall of the cylinder looked sharp (infocus), while the others behind them were blurred (de-focused). Adjusting the depth of field also helped to solve the problem of optical distortion that occurred along the curved ends of the cylinder. The depth of field blurred out all the peas that were located at both ends of the cylinder. Thus, excluding them from the pool of peas that would be used for computing real-time shrinkage. After distinguishing the infocus

peas from the defocused ones, the blurred out the peas that were affected by optical distortion, an additional step was taken to make the outline of the infocus peas easily detectable, so as to enhance the computation of their average morphological values. The outline of the infocus peas was made more conspicuous by pasting a red sheet of paper on the outer wall of the cylinder as shown in Figure 3.3 below:

Figure 3.3 Component of the real-time image acquisition system for fluidized bed drying

During the preliminary experiments, the authors discovered that a red background produced better contrast than a black or blue background, due to the green color of the peas. The red background also produces better segmentation results for the following reasons: The first reason was the green color of the peas; second, the segmentation was carried out on the green channel of the captured RGB image. Since the peas are green, it would be straightforward to identify the surface and edges of the peas by identifying areas on the image that is green. Lastly, as the peas dried and shrank, the color of the wrinkles on their surface slowly moved from green to grayish brown. Hence, if a black or blue background was used, when the green channel is extracted, the wrinkled surface of the pea (which is now grayish brown) may not be detected as part of the pea.

3.3.7. Image Analysis

At this stage, the image acquisition stage has just taken place. The camera has captured ten images of the peas. The CV algorithm now needs to analyze the images and determine what the average area, average diameter and average perimeter of the peas are at that stage during the drying process. Before computing these values, the algorithm distinguishes the three components of the image (infocus peas, out-of-focus peas, and red background), then it produces an image that contains only the in-focused peas. From this new image, the area, perimeter, and equivalent diameter of each pea is computed. Once this is done, it is then possible to determine what the shrinkage of the peas are at that stage during the drying process, by comparing the average values at that point to the values obtained during the preceding interval. As discussed earlier, the CV algorithm distinguishes the focused peas from the non-focused one through a process known as segmentation – separating the object(s) of interest from the background of an image. In this research work, the objects of interest are the peas that are in focus, while the background is made up of the out-of-focus peas and the red paper that was pasted on the glass cylinder.

There are at least four segmentation methods, any of which can be used in computer vision systems. The segmentation method could be based on thresholds, gradients, regions or on classifications [57,58]. What determines which method is applied would be the unique feature within the image that differentiates the object of interest from its background. For instance, consider the threshold method, which has been widely applied to numerous drying experiments.[15,17,23,24,49,80,81,84,85] This method would work well if there were significant contrast between the color of the object of interest, and that of its background. Secondly, if there is more than one object of interest, like in our case where there are numerous peas, the threshold method would work if the peas were lying on the same surface, not touching each other (no overlaps), and the surface of the peas share a similar color intensity that contrasts with that of the background. Thus, when the green channel is observed, the red background would be displayed as deep gray (since it contains little or no green in it), while the green color of the peas would tend towards white. At this point, there is significant contrast between the color of the peas and that of its background, so once you apply a threshold (a pixel intensity that separates the object of interest from the background), the color of the peas would be displayed as white, while the background would appear as black. This method was effective in segmenting the peas that were captured offline (Figure 3.4), since the image met most of the criteria stipulated above. We only encountered

challenges as the surface color of the peas changed, such that different sections of the peas displayed different color tones. However, as far as the peas were not touching each other, the algorithm was robust enough to handle the variations.

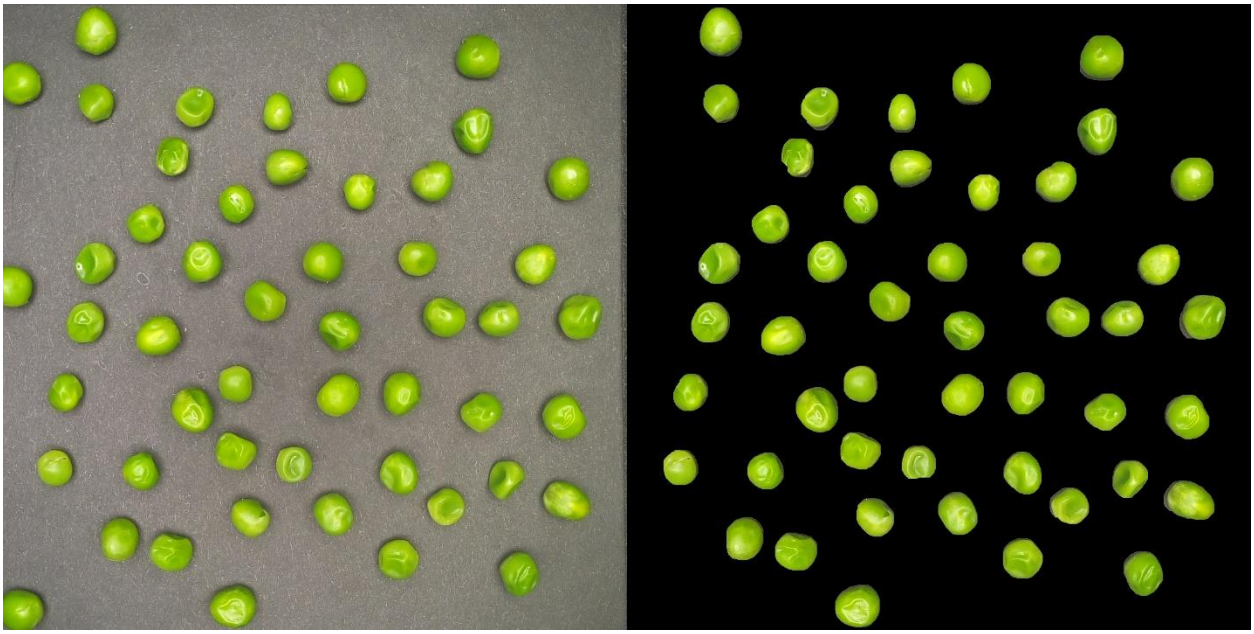


Figure 3.4 Acquired offline image and the segmented output.

But in the case of real-time image segmentation, the threshold-based method would not be effective, due to color variations on the surface of the peas as drying progresses. When the green channel of the RGB image is observed, the color of the blurred defocused peas would be similar to some areas on the infocus peas. Therefore, when a threshold is applied, the resulting image would display both regions of the infocus and defocused peas that meet the threshold. This problem was solved by the gradient-based method.

During gradient-based segmentation, a filter is moved across every pixel position on the image. As it travels within the image, it computes the gradient at each position by comparing the intensity value at each pixel position with that of its neighbors. As the gradients are being determined, the algorithm searches for pixel positions on the image where there is significant spike in gradient. Most times the spike occurs at the boundary that separates the peas from the red background. This boundary is the edge of the pea.

Observe pea E in Figure 3.5B; you will notice its center region has a similar pixel intensity as some areas on the background (such as point F). Hence the threshold method would not segment this image properly.

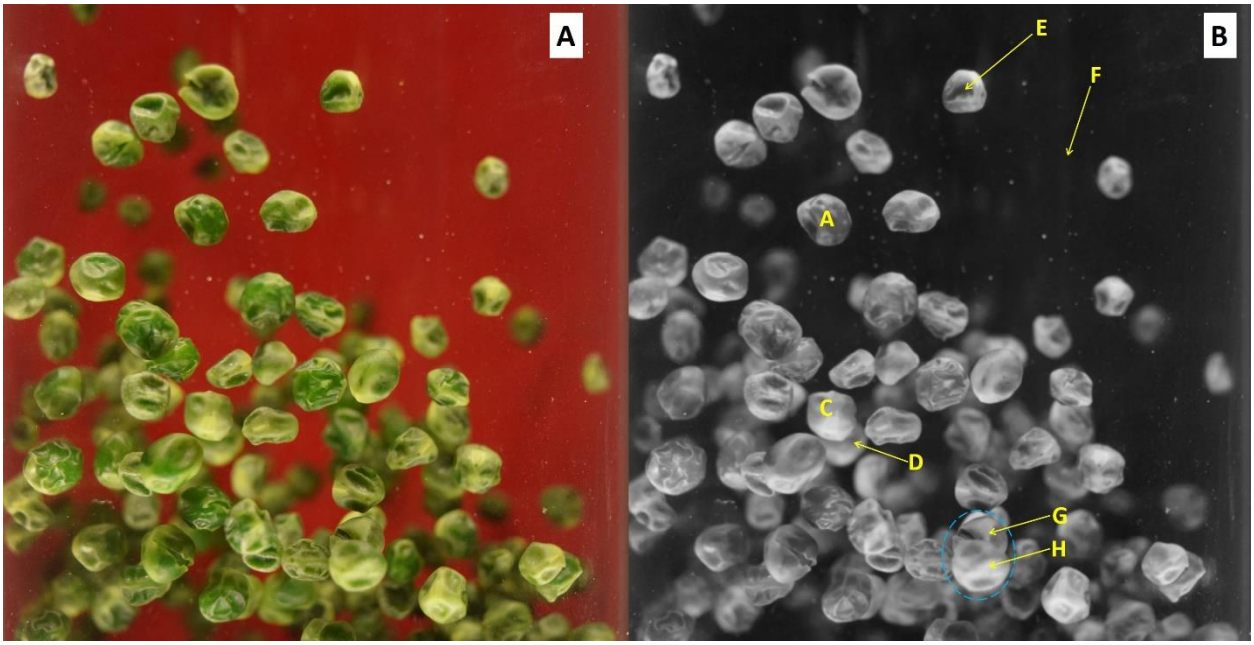


Figure 3.5 (A) RGB image of peas in the fluidized bed. (B) Green channel of the RGB Image

However, if the image was segmented based on edges, we would be able to obtain the outline of the pea (labelled E) using the high gradient along the edges of the pea. Additionally, the gradient-based method can be used to distinguish in-focus peas from those that are out of focus. The gradient intensity at the edges of the infocus peas would be higher than their defocused counterparts. Thus, while designing the CV algorithm, a benchmark (gradient intensity) was used to identify the edges of the peas that were in focus. Furthermore, the same technique was used to separate overlapping peas, especially when the pea behind the in-focus pea was out of focus. For instance, observe peas C and D in Figure 3.5B, pea C is in focus, while D is defocused. With the gradient-based method, pea C was chosen as an object of interest, while D was left out. Now that we have seen how the gradient-based method can be used to identify the infocus pea, let us stitch together all the steps and sub-steps that were used to work on the capture image (Figure 3.6a), until a new image containing only the in-focus peas (Figure 3.7F) was obtained. These steps are outlined in the flow chart (Figure 3.6) and in Figure 3.7 below.

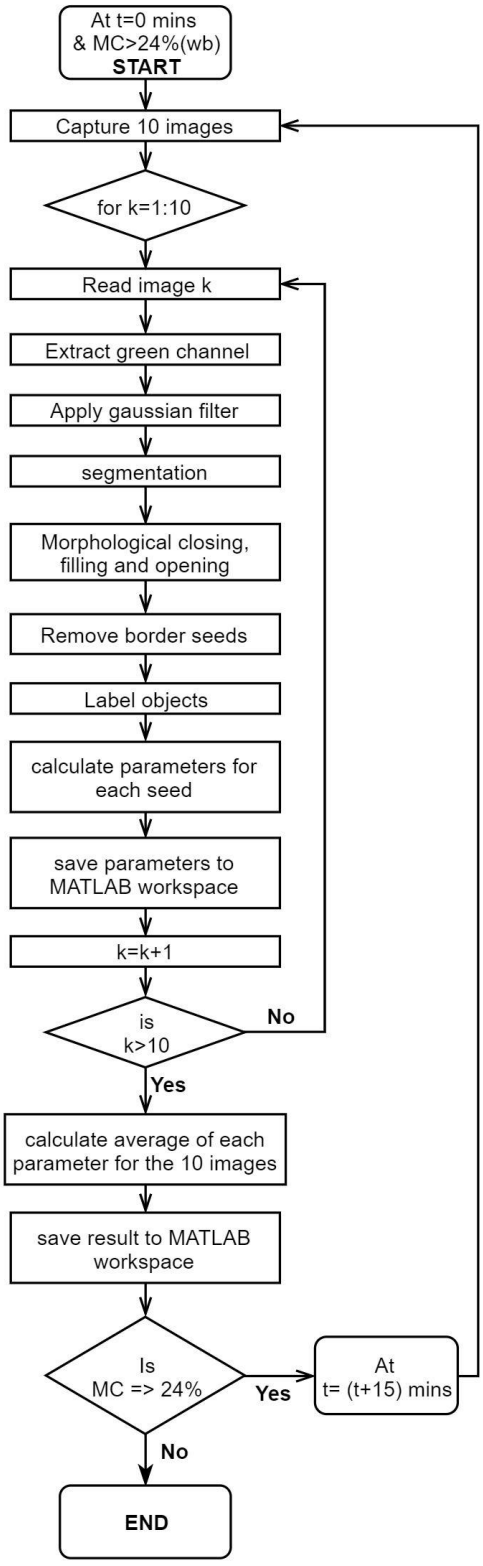


Figure 3.6 Process flow for real-time image acquisition and analysis

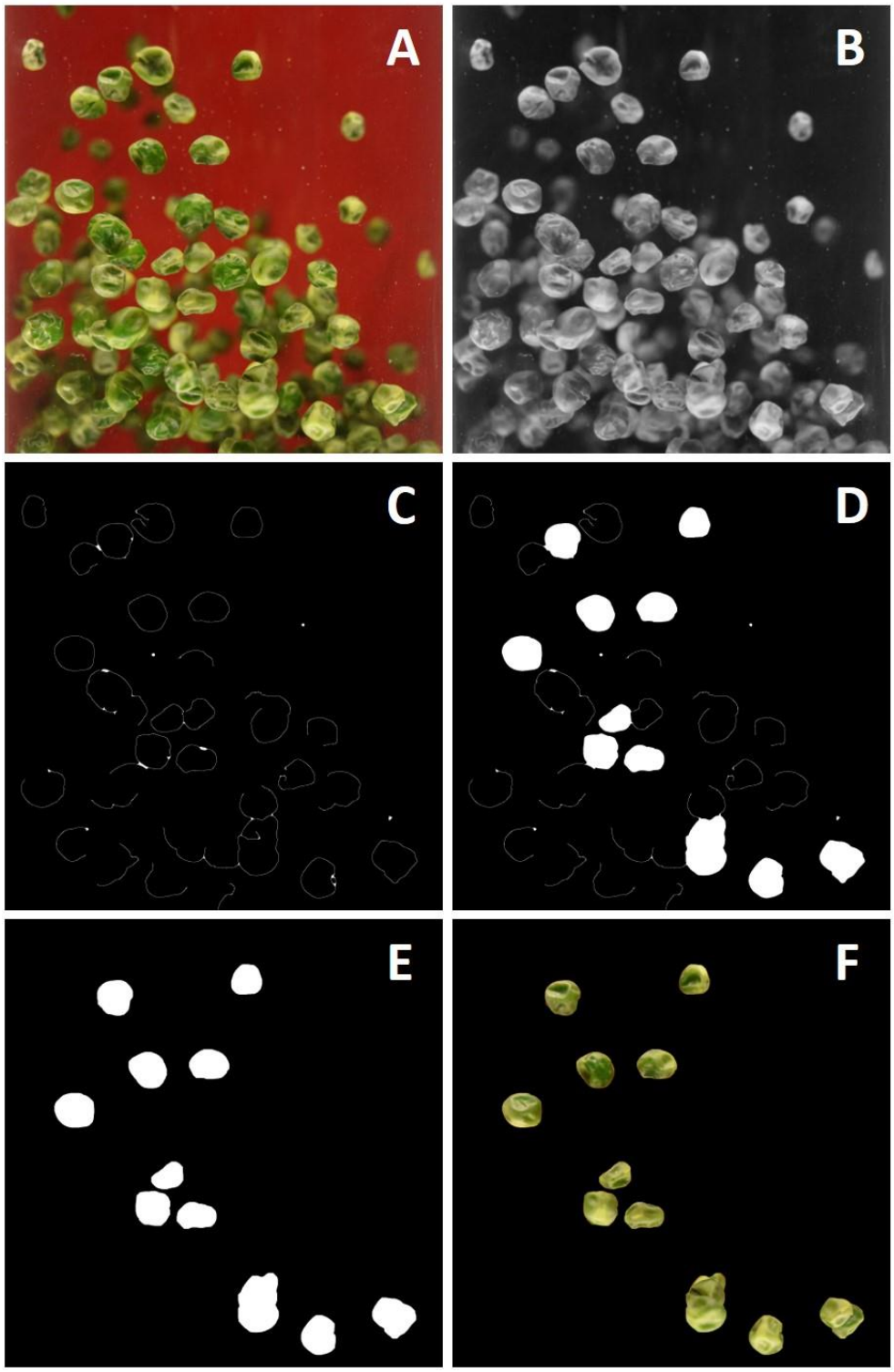


Figure 3.7 *The image analysis steps: (A) RGB Image, (B) green channel of the RGB image, (C) closing operation on edge diagram, (D) filling operation, (E) opening operation and removing border peas, and (F) infocus peas.*

The first step of the image analysis phase was to read the RGB image. After that, the green channel was extracted from the RGB image. Next, the noise within the image was removed using a gaussian filter that had a standard deviation of 1.7. Given that some of the images contained infocus peas that were overlapped behind by out-of-focus peas, it was counterproductive to denoise the image with an averaging filter. When an averaging filter was used, it successfully eliminated the noise from the image, but in addition to that, it smoothened the sharp edges of the infocus peas. Thus, making it difficult to identify the peas' true edge. Hence, the denoising was done using an averaging filter.

After denoising the image, the next stage was to segment the image using the gradient-based method. An important point to note is that this method works very well when the in-focus pea is surrounded by the red background. When the threshold value is applied, the outline of the pea is produced. However, the case is different when a side of the in-focus pea is overlapped behind by a defocused pea. At the region where the overlap occurs, the edge of the infocus pea would have a gradient intensity that is lower than the values obtained at its other edges that are not overlapped but are only surrounded by the red background. The gradient value of the overlapped edge would be similar to the gradient value of the edges of some defocused peas. Consequently, when a threshold is applied the overlapped edge of the infocus pea would not be displayed, because its gradient intensity would be lower than the assigned threshold. Hence the algorithm was designed to accommodate the low gradient intensity that occurs at the edge of overlapped infocus peas. An example of this is shown in Figure 3.8. The algorithm was able to identify peas G, E, I, and F, as in-focus peas, notwithstanding that they were overlapped behind by defocused peas.

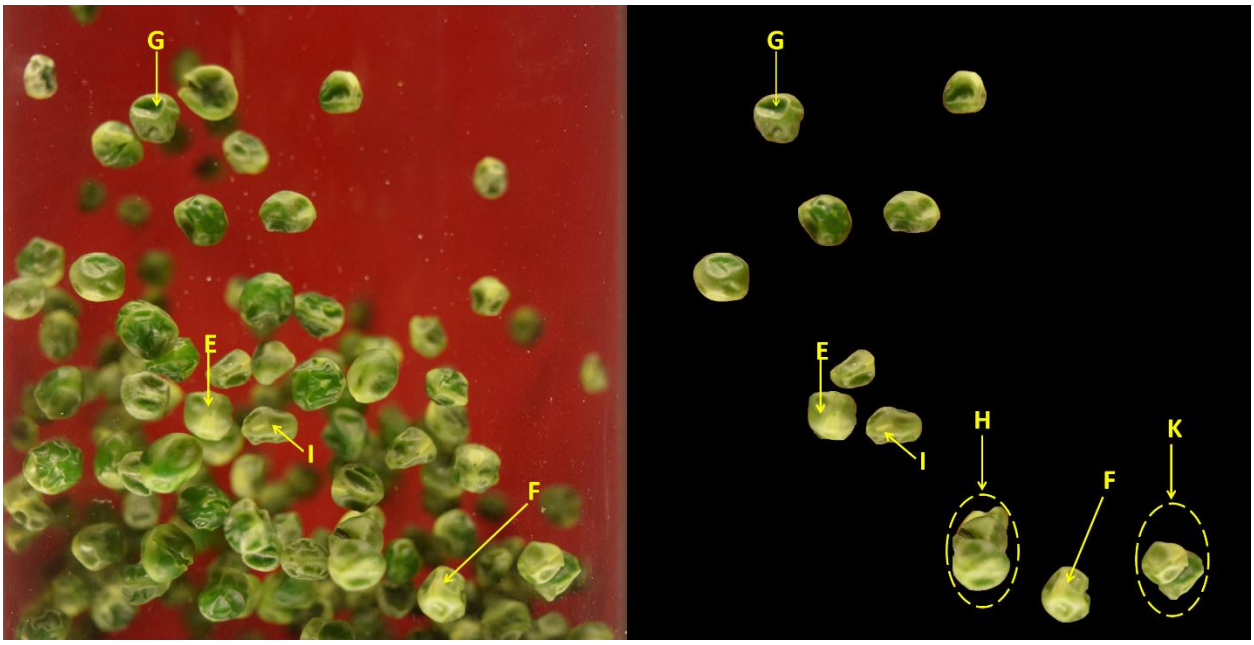


Figure 3.8 (A) RGB image showing captured pea. (B) Image of infocus peas.

The method outlined above works well only if the infocus pea is overlapped by a defocused pea. But we would have a problem when the in-focus pea in front is overlapped behind by another in-focus pea. In this case, when we apply the previous method, it would produce the outline of both peas, because their gradient values are above the upper threshold. An example of this scenario is also shown in Figure 3.8, for peas H and K. The overlapped peas are all in focus.

To solve this problem, towards the end of the image analysis, the algorithm would estimate the circularity or roundness (R) of all the segmented peas, using Equation 3.3. [8,15,53,58]

$$R = \frac{4\pi A}{P^2} \tag{3.3}$$

Where A is the area of the object of interest and that of a circle that has the same perimeter (P) as the object. The value of roundness is 1 for a circle (the maximum value) and $\pi/4$ for a square. During this study, objects of interest that had a roundness below 0.9 were designated as overlapping peas. Such peas were not used to compute the average values for area, equivalent diameter, and perimeter.

At this point, we have produced the outline of the in-focus peas. Because of the applied threshold, the output image was a binary image, made up of ones (“1s”) and zeroes (“0s”). The “1s” identify the object of interest (the regions in the image that met the set condition), while the “0s” designate

the background of the image. Since the binary image only shows the edges of the infocus pea, if we use this image the way it is, the algorithm will calculate the area for each of the in-focus pea, by counting only the number of pixels that make up the outline. However, the true area should comprise of both the pixels on the outline and those within it. Thus, before calculating the area, the region within each outline had to be filled up with pixel values of 1.

To ensure that only the region within the outline is identified as an object of interest, the CV algorithm was designed to fill only outlines that form a closed loop. That is, there should be no gap (missing pixel) along the thin line that makes up the outline. However, after the segmentation process, some of the outlines of the in-focus peas had missing pixels along them. To solve this problem a morphological closing operation was carried out before the outlines were filled.

The closing operation was executed using a flat disk-shaped structuring element (SE), having a radius of 7 pixels. Gonzalez et al.[58] described morphological closing of an image A by a structuring element B, as the dilation of A by B, followed by an erosion of the result by B. During this operation, as the SE travels along the peas' outline, it dilates the outline with extra pixels. Once the SE gets to a section of the outline that has a gap (missing pixels), as far as the size of the mixing pixels is less than the radius of the SE (7 in our case), the gap is filled. If not, the outline is deleted. After dilating the outline of the infocus peas, the SE journeys again along the outline, to reduce it to its initial thickness. Hence, at the end of the closing operation, we now have closed loop outlines. At this point, closed outlines are filled with pixels, so that the area of the peas could be computed.

Following the above, a morphological opening operation is carried out to disconnect closely aligned outlines that were connected during the morphological closing operation, because the space between the individual outlines were less than the size of the SE used for the closing operation. During the opening of an image A by B (denoted as $A \circ B$), the image is first eroded with the structuring element B. Afterwards, the result from the erosion is dilated by B.[58] This process is expressed in Equation 3.4. The opening operation was carried out with a flat disk-shape SE, having a radius of 25.

$$A \circ B = (A \ominus B) \oplus B \tag{3.4}$$

After completing the opening operation, the algorithm removes the objects of interest that are at the border of the image. Once this is done, the remaining objects on the image are labelled, and

the various morphological parameters are computed for each pea in the image. The parameters that are computed are the Area (*A*), the Ferret/Equivalent diameter (*FD*), Perimeter (*P*) and Roundness (*R*). Some of these parameters are defined below:

- Ferret diameter: Diameter of a circle that has the same area (*A*) as the object of interest [6,8,16,22]

$$FD = 2\sqrt{\frac{A}{\pi}} \tag{3.5}$$

- Roundness: Ratio of the area (*A*) of the object of interest and that of a circle that has the same perimeter (*P*) as the object.[8,15,58]

$$R = \frac{4\pi A}{P^2} \tag{3.6}$$

Shrinkage in area, Ferret diameter and perimeter were calculated as follows:[53,55]

$$Area\ shrinkage, \qquad A/A_0\ (\%) = \frac{A_0 - A}{A_0}, \tag{3.7}$$

$$Ferret\ diameter\ shrinkage,\ D/D_0\ (\%) = \frac{D_0 - D}{D_0}, \tag{3.8}$$

$$Perimeter\ shrinkage, P/P_0\ \ (\%) = \frac{P_0 - P}{P_0}, \tag{3.9}$$

where *A*, *D* and *P* are the current average values for the area, Ferret diameter and perimeter. While the subscript "$_0$" indicates their corresponding initial values.

3.4. RESULTS AND DISCUSSION

3.4.1. Drying Effects on Moisture Ratio

The relationship between moisture ratio and drying time is shown in Figure 3.9. As expected, the moisture ratio reduced exponentially with time. Peas dried at 50 and 55^{o}C took about 75 min to get to 24% moisture content (wb). While those dried at 60^{o}C had a shorter drying time of about 58 minutes.

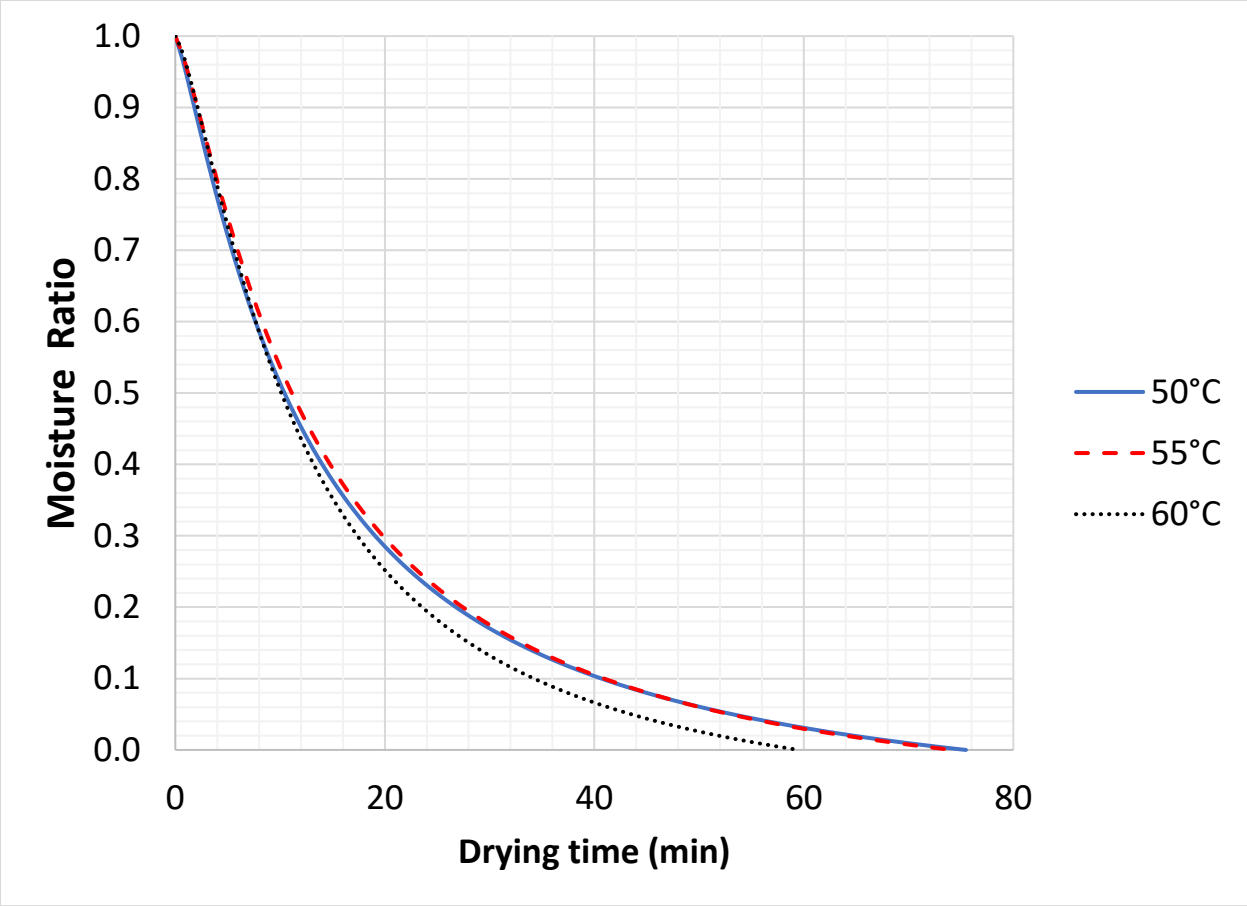


Figure 3.9 Moisture ratio vs time for 350 g of peas dried at 50, 55 and 60°C

3.4.2. Drying Effects on Area Shrinkage

The shrinkage kinetics of peas, dried at three different temperatures 50, 55 and 60°C are shown in Figure 3.10. The figure displays offline and real-time results for experiments carried out using 200 g and 350 g of peas, respectively. From both curves, the result from the offline method showed that the area of the peas reduced steadily as the drying progressed.

A similar trend can be observed along the first section for each of the real-time curves in Figure 3.10. Within the first 30 minutes, the area of the pea was reduced to 65% of its initial value. However, the level of shrinkage at this point for the real-time method was higher than the results obtained from the offline method. Both values should be similar because the peas that were captured during real-time imaging, were from the same batch as the peas that were used for offline imaging.

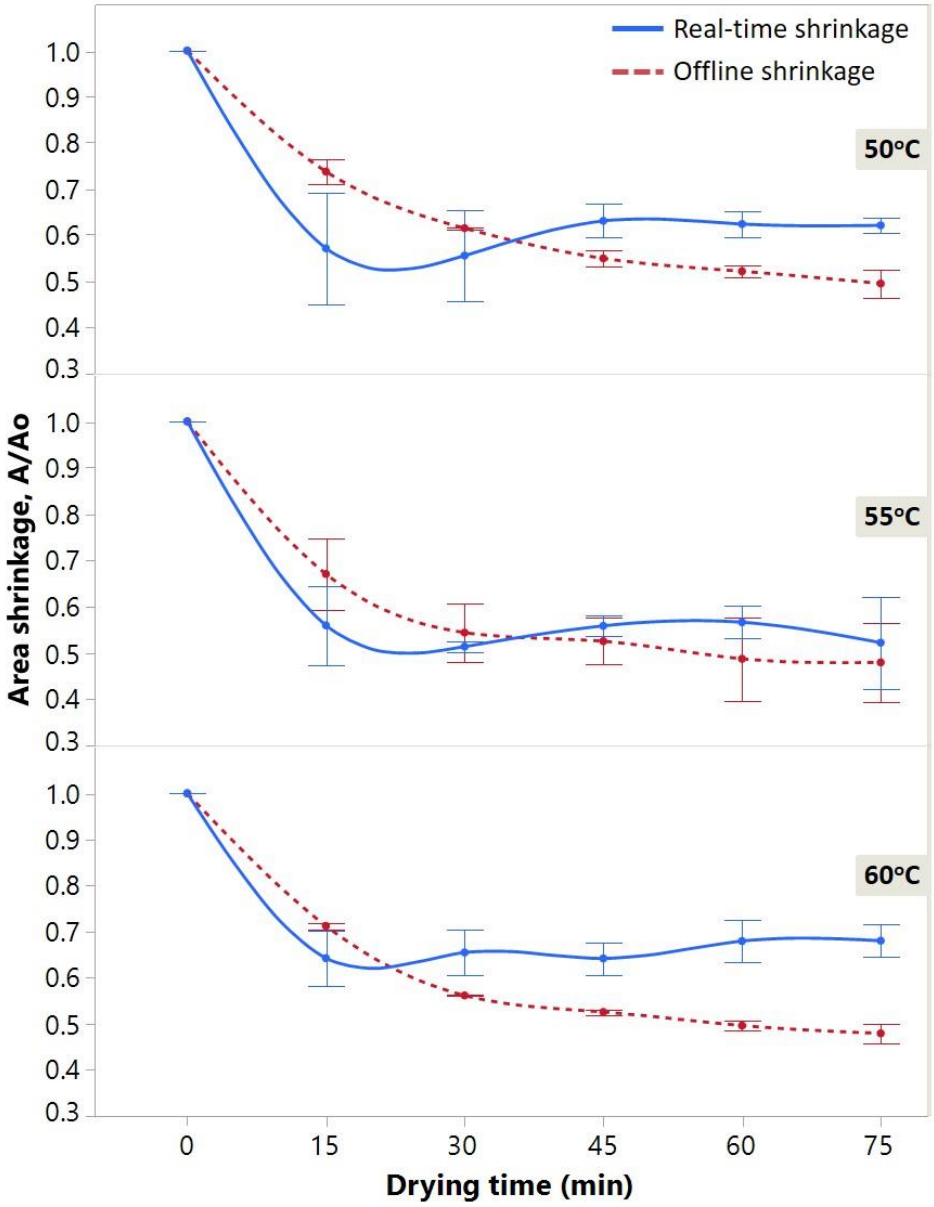


Figure 3.10 Real-time and offline area shrinkage of 200 g of peas dried at 50, 55, and 60°C.

Since offline image acquisition was done when the peas laid still on a flat surface with each pea standing separately on its own, not touching its neighbor, and all the peas were at the same distance from the camera, the shrinkage obtained from this process was used as the reference standard. So, why is the real-time result at 15 minutes higher than the offline result? The answer lies in how far the in-focus peas were from the camera, when the image was captured. Were they touching surface of the glass, as it was assumed in the hypothesis? If they were touching, the real-time shrinkage at 15 minutes would have been the same or very close to the offline result.

However, when the drying experiment was carried out with 200 g of peas, during the fluidization process, the flow pattern was such that by the time the peas had risen up to the section of the cylinder that was within the field of view of the camera, most of the in-focus peas were no more resting on the interior surface of the glass cylinder. More of the in-focus peas were farther away from the surface, so they appeared smaller than the very few in-focus peas that were on the surface. Therefore, the average area was smaller than the actual value. A description of this explanation is shown in Figure 3.11. Peas A and B are all in focus, however, peas B are farther away from the camera. When the average area is computed, the reduced size of peas B would decrease the average cross-sectional area, thereby increasing the area shrinkage.

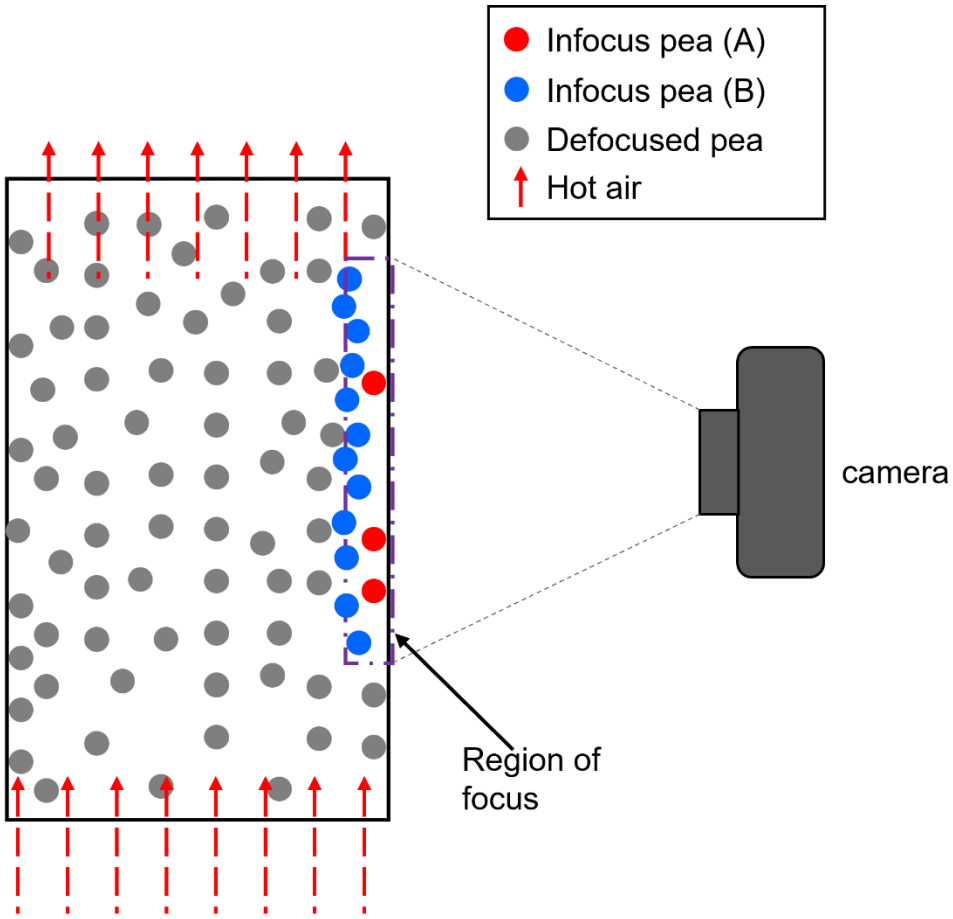


Figure 3.11 Image depicting how the non-uniform distance of the in-focus peas from the lens can affect real-time shrinkage results.

The same discrepancy between the real-time and offline shrinkage values, can also be observed in Figure 3.10, for peas that were dried at 55 and 60°C. The previous explanation for the peas dried

at 50°C also applies to these two cases. To solve this problem, the mass of the peas was increased from 200 g to 350 g. So that by the time the peas rise to the section of the cylinder, where they can be viewed through the lens of the camera, most of the in-focus peas would still be touching the wall of the cylinder. Since the diameter of the drying chamber is fixed, as the mass of the peas increases, so will its height. With the increased height, during fluidization, as the air starts to lift the peas, before they are scattered in random directions, they would first be moved up together as a cylindrical heap of peas, rubbing the wall of the glass, as they travel upwards. Increasing the mass of the peas to 350 g improved the real-time shrinkage results. This improvement can be seen in Figure 3.12.

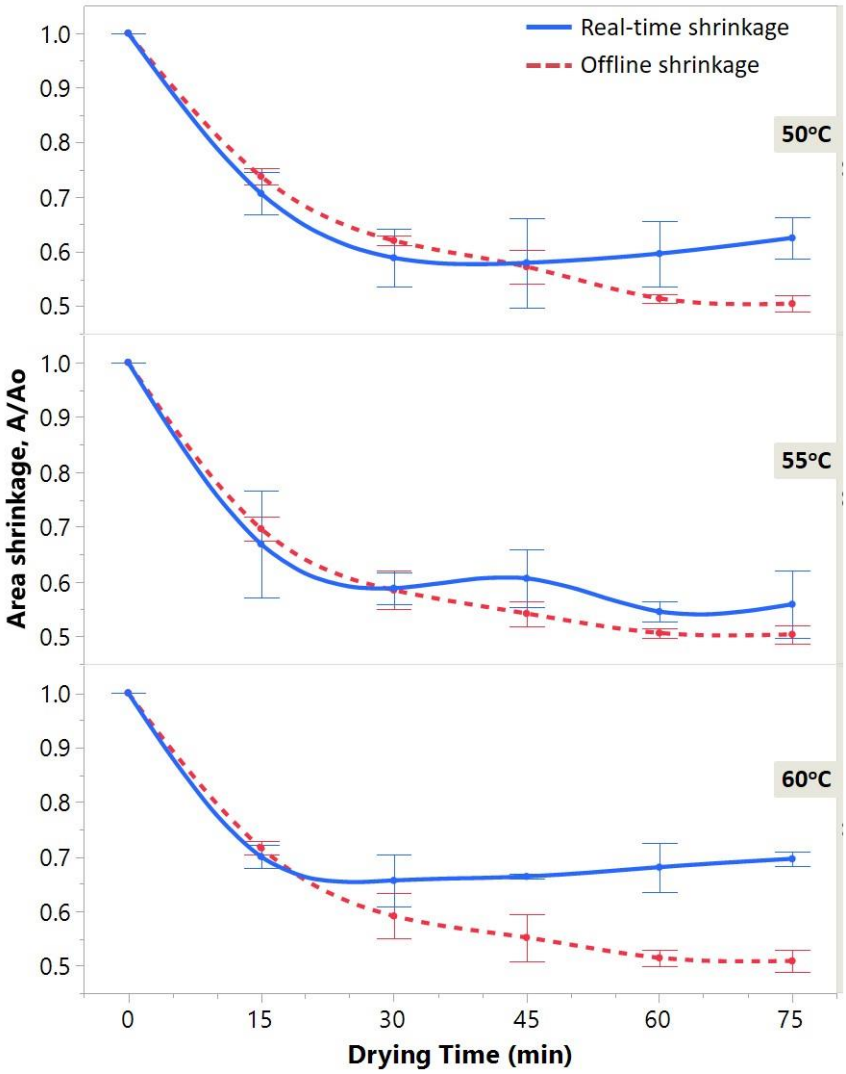


Figure 3.12 Real-time and offline area shrinkage kinetics of 350 g of peas dried at 50, 55, and 60°C.

In Figure 3.12, the real-time shrinkage at 15 minutes is most accurate for peas dried at 60^{o}C, followed by 55 and lastly 50^{o}C. The reason for this disparity could be attributed to the effect of air temperature on the texture of the seed coat, and on the shrinkage of both the seed coat and cotyledon. Fifteen minutes into drying, the seed coat for peas that were dried at 60^{o}C was harder in texture than those dried at 55 and 50^{o}C. As the pea case-hardened, the seed coat deformed/wrinkled by a lesser degree, compared to seeds that were dried at lower temperatures. So, when the texture and size of the three sets of peas (50, 55 and 60^{o}C) were compared, it was observed that due to the effect of casehardening, the peas dried at 60^{o}C were larger and occupied more volume than the other two sets. That is why the real-time shrinkage from 0 to 15 minutes was more accurate for peas dried at 60^{o}C, than it was for those dried at 55 and 50^{o}C. Next, let us discuss why the other section of the real-time curve lay above the offline curve.

Observe the shrinkage curve for peas dried at 60^{o}C (Figure 3.12). Notice that from 0 to 15 minutes, both the offline and real-time curves were almost the same. From the 30th minute, the real-time values became higher than the offline values. This conspicuous discrepancy could also be attributed to the fluidization pattern within the cylinder, which in turn affected the distribution of the peas within the glass cylinder.

3.4.3. Drying Effect on Diameter Shrinkage

The shrinkage of diameter for peas, dried at three different temperatures 50, 55 and 60^{o}C is shown in Figure 3.13. At the three temperatures, the peas shrank in diameter, the same way they did in area. Thus, all the discussion about area shrinkage is also true for diameter shrinkage.

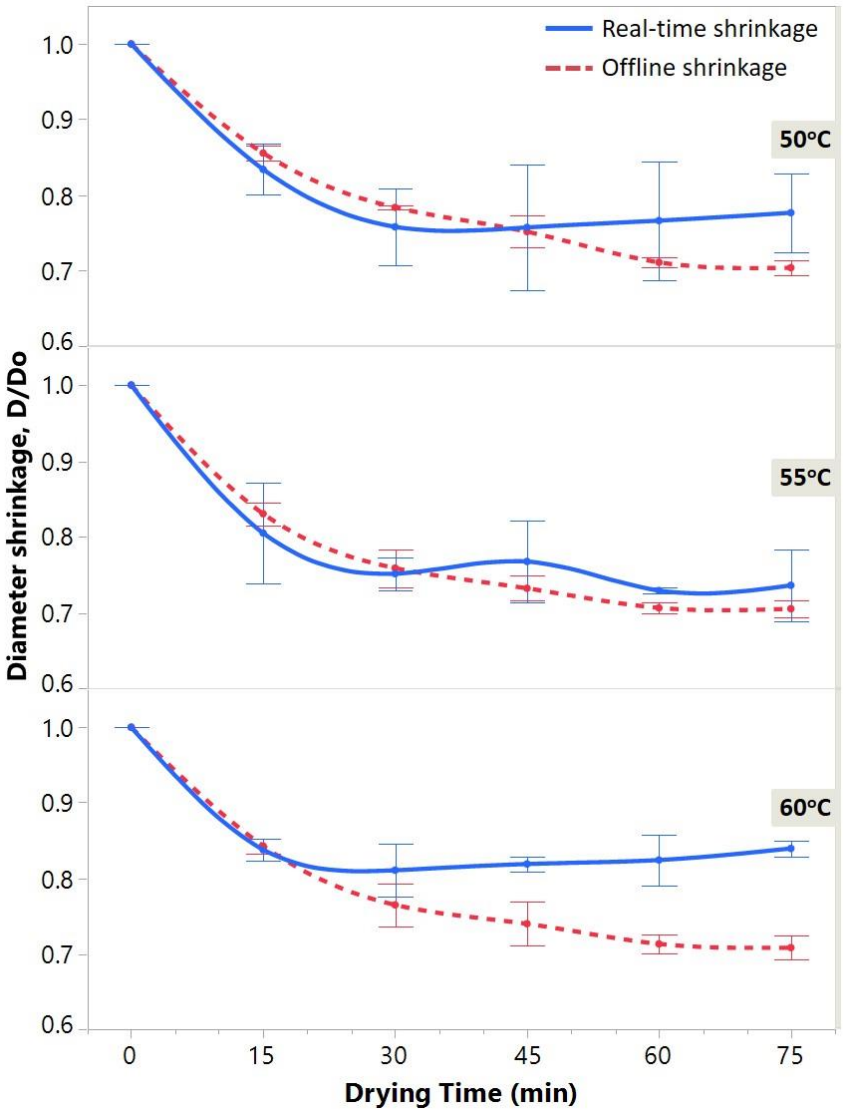


Figure 3.13 Real-time and offline diameter shrinkage kinetics of 350 g of peas dried at 50, 55, and 60°C.

3.4.4. Drying Effect on Perimeter Shrinkage

The shrinkage of perimeter for peas dried at 50,55 and 60°C are displayed in Figure 3.14. The results in the graph are similar to those obtained in Figure 3.12 (when the pea shrank in area). Thus, the discussions that were given for area shrinkage also apply to Figure 3.14.

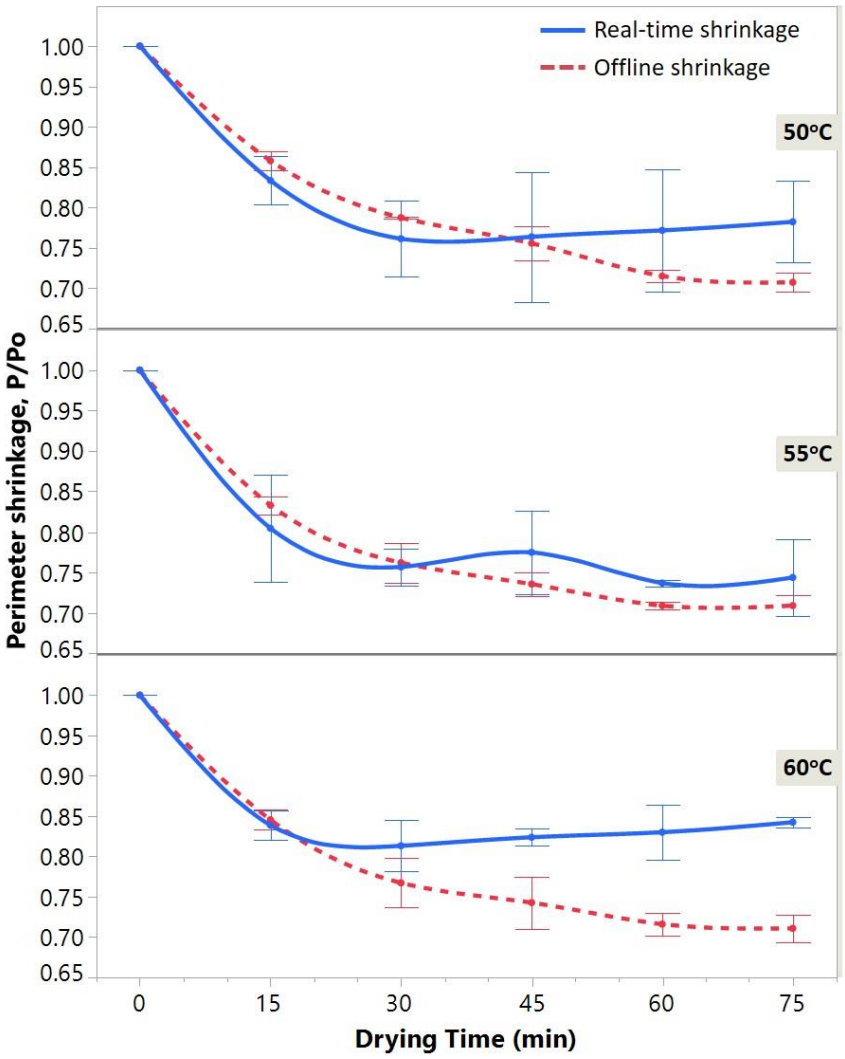


Figure 3.14 Real-time and offline perimeter shrinkage kinetics of 350 g of peas dried at 50, 55, and 60°C.

3.5. FUTURE RESEARCH

To solve the discrepancy between the offline and real-time results, the fluidized bed should be designed in such a way that all the in-focus peas would always touch the surface of the drying chamber, thus putting them all at a known distance from the camera lens. This improvement can be achieved by drying the pea in a spouted bed dryer that has a central draft tube. Due to the unique flow pattern within the spouted bed, the peas spouting out from the draft tube would hit the surface of the drying chamber, as they fall to the bottom of the draft tube. This design was applied by Kadunc et al.[6] to monitor the increase in pellet thickness, as the pellets were coated in a spout bed dryer. Their real-time results were similar to those obtained from offline image acquisition

and analysis. Another advantage of the spouted bed is that when the peas are ejected from the draft tube, they fall back to the base of the dryer at slower speed, than they would have in the fluidized bed (such as the one used for the current study). The reduced speed may result in fewer overlapped peas. This would enhance the quality of results obtained during the segmentation process.

In addition to improving the design of the dryer, another area for future research would be to feed the results from the experiment to a machine learning platform. The machine learning system could improve the segmentation process by identifying a more effective pattern to identify in-focus peas, even in overlapped situations. Additionally, if we feed the machine learning system with information that were acquired throughout the drying process, such as data that shows the trend in air temperature, air relative humidity, air velocity, color kinetics, texture kinetics, shrinkage and change in moisture, it would be possible to optimize the drying process, and open the door to customized drying – where a customer specifies how he/she wants the dried product to turn out (e.g. In terms of color, texture, shrinkage and moisture content), and the machine learning system would automatically adjust the drying system to produce dried products that exactly meets the customers' expectation.

3.6. CONCLUSION

In this work we investigated the shrinkage of peas during the fluidized bed drying process. The shrinkage was measured by computing the reduction in area, perimeter, and diameter of the peas. During the experiment, shrinkage was monitored using a real-time and an offline computer vision (CV) system.

A real-time CV system has been developed to estimate shrinkage by monitoring the reduction in size of peas that touched the inner wall of the drying chamber during fluidization. To ensure that the peas used for the real-time analysis met this criterion, the depth of field of the camera was adjusted to sharpen the edges of peas that were within 15 mm from the inner wall of the transparent drying chamber, and defocus (blur out) the other peas that were beyond this range. Consequently, the algorithm for the real-time CV system was designed to identify these sharp in-focus peas, and compute the shrinkage based on the information gotten from these peas.

On the other hand, the offline system monitored shrinkage by acquiring and analyzing offline images of peas collected from the fluidized bed dryer. The results of the offline CV system showed

a steady reduction in area, diameter, and perimeter throughout the drying process. Similarly, the results from the real-time CV system showed an overall increase in shrinkage throughout the drying process. However, at some points during the drying process, the real-time results deviated from the offline results. This deviation was attributed to the fluidization pattern within the fluidized bed dryer, which prevented most of the peas that were in focus from touching the inner surface of the dryer. Thus, creating non-uniform pea distance from the camera.

To solve this problem, it was recommended that the experiment be carried out in a spouted bed dryer, whose fluidization pattern would ensure that the peas are at known distances from the camera. Thereby reducing the discrepancy between the offline and real-time shrinkage results. In addition to improving the design of the dryer, it was also recommended that when the research is carried out in a spouted bed, the results should be fed to a machine learning platform for improved prediction and performance.

CONNECTING TEXT

Let us discuss the progress we have made towards improving the state-of-the-art for monitoring visual attributes of food, in real-time, during fluidized bed drying. The work described in this chapter borders on the first two milestones shown in Figure 3.15.

For the first milestone to be accomplished we needed to have developed two units – an image acquisition system and an image segmentation model. We have developed an effective image acquisition system that produces well illuminated images of peas. It was designed such that within each image, there was sufficient contrast amongst the infocus peas, the defocused peas, and the image background.

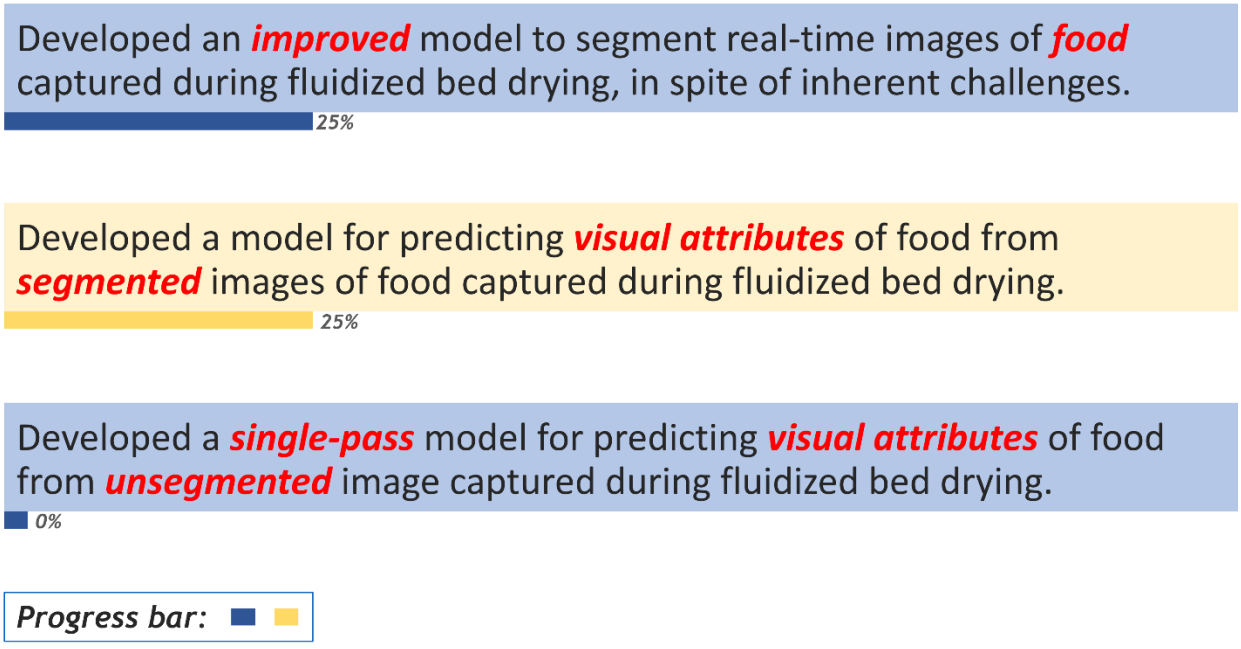


Figure 3.15 Milestones to improve the state-of-the-art

Also, the acquisition system was successful at preventing motion-related blur. After building the acquisition system, we developed an image segmentation pipeline, using the classical CV methodology. The segmentation pipeline detected some of the infocus peas, however, it omitted a lot of them. Hence, there is still a lot of work that needs to be done, in order to improve the segmented result, and by extension visual attribute prediction (the second milestone). Due to the results achieved so far, we assigned progress scores of 25% to the first and second milestones. Also, you would have noticed that the CV pipeline was only used to monitor shrinkage. Color and

texture were not monitored. Hence, in the upcoming chapter, we are going to build on the progress we have made so far by:

a. Using the classical CV pipeline to monitor color, size, and texture of green peas. Afterwards, the visual attribute predictions from the CV pipeline would be compared to the ground truth values, by computing coefficients of determination.
b. Next, we would improve the image segmentation results and visual attribute predictions by introducing deep learning techniques.

CHAPTER IV

A TWO-PASS DEEP LEARNING SYSTEM FOR MONITORING VISUAL ATTRIBUTES OF FOOD DURING FLUIDIZED BED DRYING

4.1. ABSTRACT

The visual appearance of dried foods affects consumer patronage. If the food is visually unappealing, consumers walk away from it, and the product ends up as waste. So, it has become crucial to develop food dryers that can monitor realtime changes in moisture content and visual attributes. Such advancement has not been effectively applied to fluidized bed drying of food. Hence, in this work, we developed an AI-driven solution for monitoring real-time changes in color, texture, and size of green peas, during fluidized bed drying. This solution was powered by *"Unet-Xception", a* deep learning model comprising of the U-Net and Xception architectures. U-Net segmented images of the peas, while Xception predicted visual attributes using the segmented output. In terms of segmentation quality, Unet-Xception achieved a Mean Intersection-Over-Union of 0.9464, surpassing 0.814 attained by a classical computer vision model. The AI-solution even detected additional objects of interest that were missing in the ground truth masks. For visual attribute prediction, Unet-Xception also outperformed the classical model. The coefficient of determination (R^2) obtained is reported in the form [classical model, Unet-Xception]. a^* and b^* indices were the best predictors of color, with R^2 values of [0.7719, 0.8851] and [0.8694, 0.8981], respectively. Homogeneity was the best parameter for monitoring texture, with R^2 values of [-0.8660, 0.8017]. Finally, the following values were obtained for size: [-1.0278, 0.7926] for equivalent diameter, [-1.6417, 0.8482] for ferret diameter, [-0.3663, 0.8391] for filled area, and [-1.5425, 0.825] for perimeter. As expected, the detection of additional objects of interest, slightly dampened the R^2. However, when Unet-Xception was used to monitor new fluidized bed drying experiments, it produced far smoother trends than the classical model. Hence, this adaptable and novel solution proves that it is possible to monitor real-time changes in visual attributes of food, during fluidized bed drying. Incorporating this solution into current food dryers could lead to consistent product quality and significant reduction in global food waste and loss.

4.2. INTRODUCTION

Tonight, over one billion people will go to bed hungry. By tomorrow morning, while we are on the dinning table eating breakfast, more than sixty-six million young children between the ages of six and twelve, living in developing countries, would have trekked long distances to school without having breakfast.[3,4,86] Yet, over one-third of the food produced annually is wasted or lost.[1-4,24] Assuming we put aside the effects of inflation, the problem of food insecurity in developing countries is not always that of food production, the problem has more to do with lack of adequate post-harvest processing and storage facilities.

In developing countries like Nigeria, during the harvest season, a lot of fresh food is brought to the market. However, three to four months down the line, due to lack of efficient processing and storage facilities, most of the food produced goes bad, thereby leading to reduced food supply and price hike. To solve this problem of food insecurity, we could develop efficient post-harvest processing systems that extend the shelf life of the food throughout the off-season period. Food drying could play a pivotal role in achieving this goal.

During drying, the moisture within the food is reduced to a safe level. By so doing, the shelf-life of the food is extended and the stability of the food in terms of quality improves. Drying decelerates biochemical, microbial, and chemical reactions that cause food spoilage. As the food loses water, the weight is also reduced, which in turn improves product handling and reduces shipping cost.

Think about the last time you visited the grocery store to buy dry foods such as fruits, nuts, spices, vegetables, or meat products. Imagine that you are walking along the aisle, heading towards the shelf that contains the dry product you wanted to buy. Try to visualize that brief period you spent inspecting the product, deciding whether to pick it from the shelf, or whether to look for a different brand. During that moment, what factors did you consider before making your choice? What did you look out for? What aspects of the dry product attracted you, or put you off? Imagine you were in the shop to buy cashew nuts, dried thyme, dry onion, or even dry fish. What would you look out for? It would not be surprising if your selection criteria were the color, the visual texture (how rough the surface is), the size, or the shape (how wrinkled it looks) of the product. Most customers would look out for these same features. If the product appears attractive, they pick it up, else they walk away from it.

Although one of the major aims of drying is to extend the shelf life of food, ironically, moisture content is not the food property that attracts consumers to the dried product. Just as we discussed, if the product is not visually attractive, no one buys it. Unfortunately, most food dryers in developing and developed countries are controlled only by monitoring real-time changes in moisture content of the product. But as we pointed out previously, monitoring moisture content alone is not enough. To ensure every fresh product that goes into the dryer comes out as visually appealing and is dried to a safe moisture content, we need dryers that can monitor real-time changes in both the moisture content and the visual properties of the food. Such drying systems have the potential to significantly reduce food loss and waste in developing countries, as well as food waste in developed countries, such as Canada, the United State, and Germany, to mention a few.

In a developed country like Canada, less than 3% of the population accounts for all the food produced and dried in the country. The reverse is the case in developing countries such as Nigeria, where over 80% of food produced is achieved by subsistence farming. In most cases, small-scale farmers also dry the food. The difference between Canada and Nigeria is the stack disparity in the quality of dried products that make it to the market. In Canada, most dry foods are produced by large scale food processors, who have specialized and standardized drying processes. Thus, most of their products are dried to a safe moisture content before they are supplied to retail outlets. However, the story is different in developing countries, where drying is conducted by so many small-scale farmers, most of whom are not specialized in the science behind food drying. Hence, while drying different food products, the farmer is overwhelmed by numerous unanswered questions, such as:

"What should my initial drying temperature be?"

"How long should I maintain this temperature?"

"When do I change the temperature, speed, or relative humidity of the air?"

"If the product comes in at a different initial moisture content, what should be the initial temperature, and speed of the air be?"

"What should the drying conditions be, to end up with a certain visual appearance and moisture content?"

Without answers to such questions, and understanding the reasoning behind each answer, drying would always remain a game of chance. The farmer arrives at the same final moisture content only if the fresh product is always supplied at the same initial moisture content, which is usually not the case. So, most times the farmer struggles between maintaining attractive attributes and drying the product to a safe moisture content.

The reverse is the case in developed countries where food drying is mostly done in large-scale by specialist with so many years of experience. While drying, they are guided by the product's drying curve. Once the dryer operator knows the initial moisture content of the fresh food and the equation of the drying curve, they can accurately calculate the time needed to dry the food to a specific final moisture content. They can compute what the moisture content would be ten minutes, twenty minutes or even hours into the drying process.

However, notwithstanding the benefits of the drying curve, we would run into problems when questions such as these arise: *"If the air temperature is reduced by 10^{o}C, what would be the effect on color, texture, or size of the product in the next 10, 20 or 40 minutes?"* Or assuming we use a microwave assisted convective dryer, the question could be *"if the microwave intensity was increased by 50 Watts/gram, what would be the color, texture, and size of the food in the next 30, 60 or 90 minutes?"*

During drying, as the food losses water, the operating parameters of the dryer such as temperature, speed, relative humidity of the air, and microwave intensity affect the rate at which the visual attributes of the food change. To ensure we consistently end up with food products that are properly dried and are visually attractive, we need dryers that can monitor and control moisture content and visual appearance, by manipulating the operating parameters in real-time, to achieve the desired quality. Such dryers are lacking both in developing and developed countries. Having such dryers, and even making them autonomous has the potential to reduce global food loss and waste significantly. If such drying systems were available, all the farmer needs to do is put in their best to food production and leave the rest to the dryer. When it is time to dry, they would feed in the fresh product into the dryer, select the expected quality from the display panel, and the intelligent monitoring and control system does the rest. It would independently control the drying conditions to ensure the desired quality is always achieved and guaranteed.

The first step to developing a dryer that controls moisture content and visual attributes, is to produce a sub-system that can capture real-time changes in color, texture, size, and shape of the food. With the advent of computer vision technology and the introduction of faster computing infrastructures, such advances are being made for convective dryers. During convective drying, the food products are spaced out on flat trays within the drying cabinet. To monitor visual attributes, as depicted in Figure 4.1 a camera is placed above the drying trays, capturing real-time RGB images of the food.

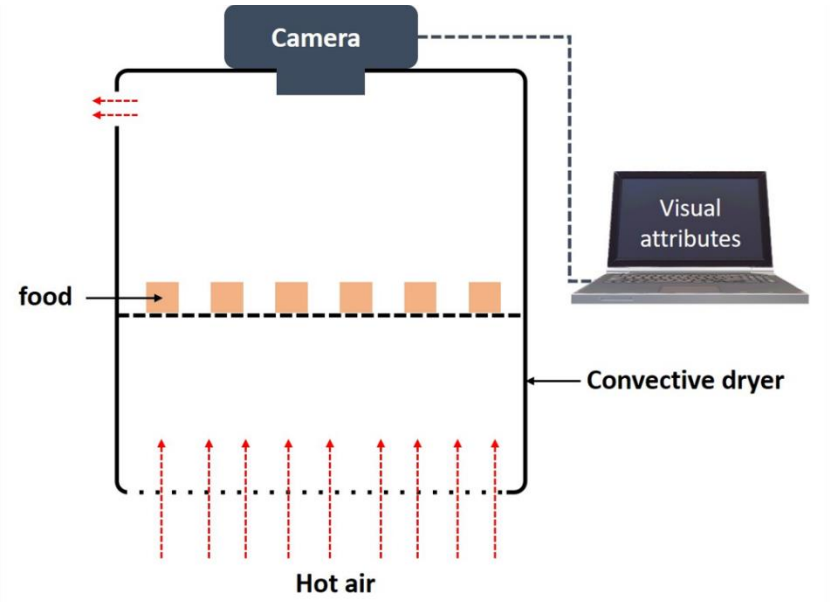


Figure 4.1 Monitoring visual attributes of food during convective drying

These images are segmented to produce a grayscale mask identifying where the food materials are located not the original image. The visual attributes are obtained by calculating the color, texture, size, and shape of the pixel positions within the RGB image that correspond to the regions of interest displayed in the grayscale mask. Some researchers moved a step further to develop control systems using real-time feedback from the computer vision system. [11-13,84,85]

Notwithstanding the advances made in monitoring real-time changes in visual attributes during convective drying, to the best of our knowledge, we are yet to see the same technological innovation applied to fluidized drying of food. One may ask why such a breakthrough is yet to happen. To answer this question, let us highlight two important factors that are critical to achieving such breakthroughs during fluidized bed drying. The first factor we would discuss is the individual

steps involved when monitoring real-time changes in visual attributes. The second critical factor is image-related challenges that have made it almost impossible to monitor visual attributes during fluidized bed drying. These challenges are unique features that differentiate images of food captured during convective drying, from those captured during fluidized drying. These features pose a lot of challenges during image processing. Having outlined the two determinant factors, let us discuss the first one – the steps for monitoring visual appearance.

The first step in monitoring visual attributes is to capture real-time images of the food. The image should be such that objects of interest (in this case the food material) would be distinct from the background. This sharp contrast helps in producing the segmentation mask. During convective drying, we achieve sufficient contrast in two ways. First the color of the drying tray must be different from the color of the food. Next, during drying, the individual food materials must be spaced out on the tray, making sure that no two products touch. Hence, once an image is captured the food would be distinct from the background (the tray). During fluidized bed drying, we can achieve the first condition by using a background that is distinct from the food. For instance, when drying green peas, we could use a red background.[26] However, we would run into problems when it comes to ensuring that the peas do not touch.

Unlike convective drying where the products lay static and are spaced out throughout the drying process, during fluidized bed drying, the food particles are randomly jumping up and down throughout the drying process. So, if you placed a camera in front of the cylindrical drying chamber, and captured images of the peas, you will notice the following characteristics, as shown in Figure 4.2.

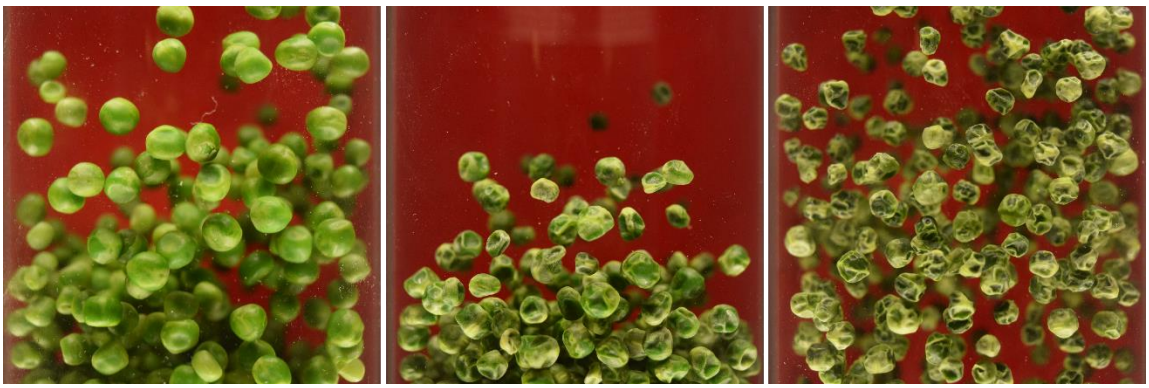


Figure 4.2 Images of green peas captured during fluidized bed drying [26]

First, the products are not at the same distance from the camera. Those that are closer look larger than those behind them. So how can we monitor size, if we cannot account for the distance of each pea from the camera lens? The next point to consider is that the peas are not spaced-out evenly, unlike food products captured during convective drying. Also, there are a lot of overlapping peas. Hence, the segmentation method applied to convective drying images would fail when used on images captured during fluidized bed drying.

Let us assume the peas captured during fluidized drying were evenly spaced-out, and we are to use a one of the classical computer vision techniques such gradient-based method for segmentation, the contrast between the green peas and the red background would make it feasible to compute gradients using one of the image channels, such as the green channel. Going by this non-realistic assumption, since the products are not overlapping, the gradient at the edges of the peas would be highest. Next, we apply a threshold and fill up the fully connected outlines to produce a binary image containing only the objects of interest. Unfortunately, in reality, due to the overlaps this series of steps would not be effective in segmenting images captured during fluidized bed drying.

Those are some of the stumbling blocks that may have made it almost impossible to develop a system for monitoring real-time changes in visual appearance of food during fluidized bed drying. So, the question remains “what is the way forward?”

An effective solution for fluidized bed dryers could be made up of three sub-systems, as shown in Figure 4.3. The first would be an image acquisition system for capturing real-time RGB images of the food during fluidized bed drying. The next would be a system for image segmenting. After that, there would be one more system to predict visual attributes using the segmentation mask and the colored image.

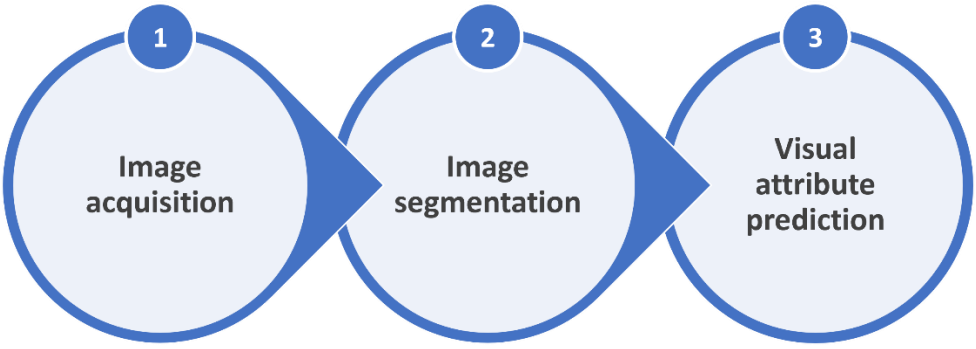


Figure 4.3 Operational steps for proposed intelligent monitoring system

Iheonye et al.[26] developed a prototype for achieving step one. Their image acquisition system was used to capture images of green peas during fluidized bed drying. Afterwards, they developed an image segmentation pipeline using classical computer vision techniques. Details for developing the acquisition system and the segmentation pipeline were thoroughly explained in their publication.[26] As we mentioned in the previous paragraph, two major differences between images captured during fluidized bed drying and convective drying, are the presence of overlaps and non-uniform distance of the objects of interest from the camera. These challenges become a source of difficulty when computing the actual size of the peas. To work around this challenge, Iheonye et. al. built their research on the hypothesis that during fluidization, since the green peas within the cylindrical drying chamber are in continuous random up and down motion, there would a replenishing layer of peas rubbing along the surface of the drying chamber facing the camera. The hypothesis is depicted in Figure 4.4b.

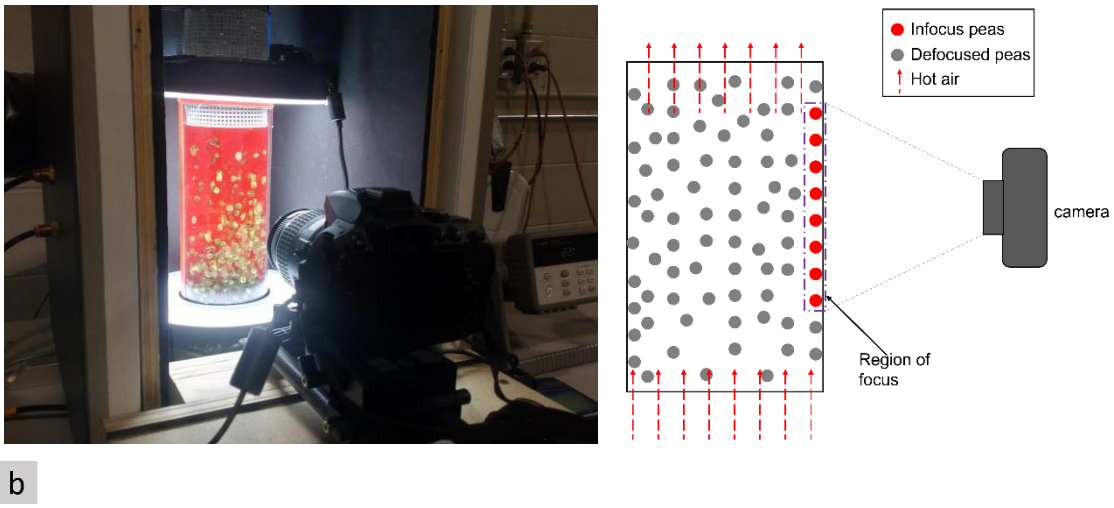


Figure 4.4 (a) Real-time CV system for fluidized bed drying, (b) image depicting infocus peas [26]

Due to the random motion, the food materials would be mixing within the drying chamber. Hence, there would be a different set of peas on the surface of the cylinder each time a new image is captured. Therefore, if multiple images are captured at different time steps during the drying process, the trend in visual attributes could be monitored by segmenting only those peas that are on the wall the cylinder. The average visual attribute for all the images captured at a specific time step would accurately represent the state of the peas at that point during the drying process. In their publication, the objects of interest (peas touching the wall of the drying chamber), were referred

to as "*infocus peas"*, while the peas behind the infocus ones were called the *"de-focused peas"* (Figure 4.4b). Their classical CV pipeline showed some promise in segmenting real-time images. Figure 4.5 and Figure 4.6 show some of the results. The figure contains three columns. The first column is the original image captured during the experiment; the second column are manually generated ground truth that was used to calculate how well the classical pipeline performed. While the last column contains segmentation results from the classical CV pipeline that was discussed in Chapter 3.

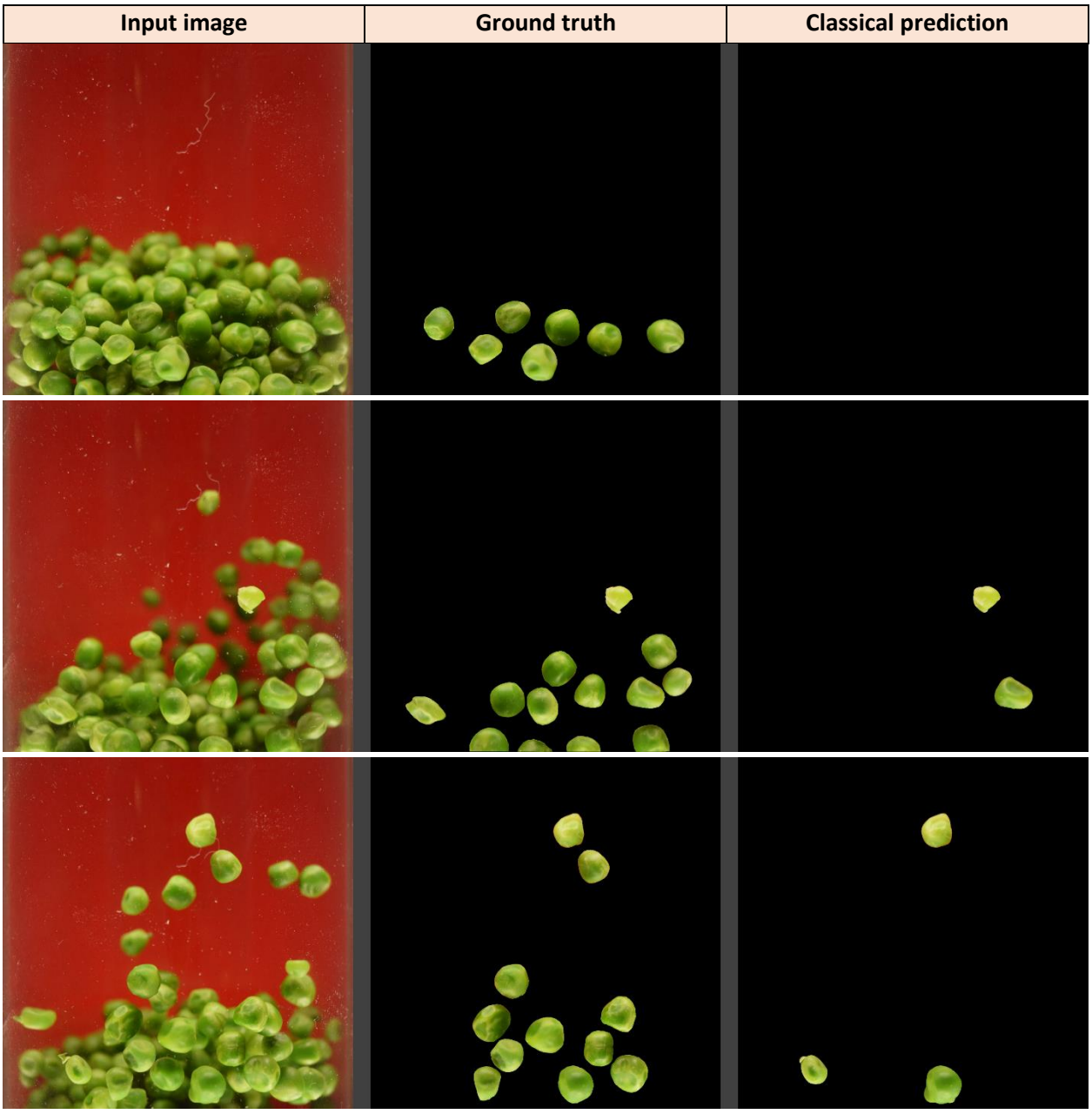


Figure 4.5 Original image, ground truth mask and segmented result using the classical CV pipeline.

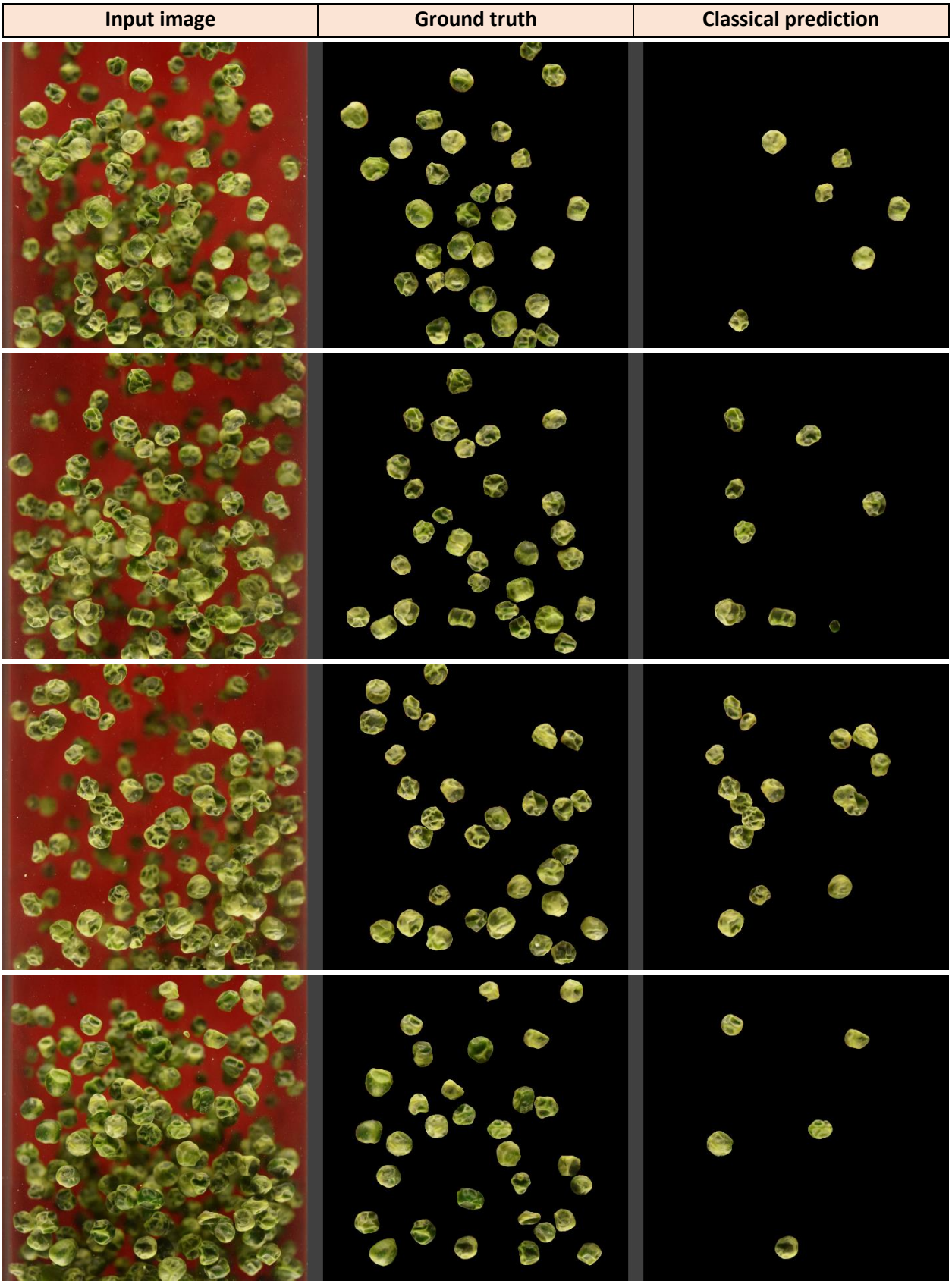


Figure 4.6 Original image, ground truth mask and segmented result using the classical CV pipeline.

After developing the segmentation pipeline, Iheonye et. al. used the segmented results to monitor product shrinkage. However, when we compare the ground truth in Figure 4.5 and Figure 4.6 to the segmented result from the CV pipeline, we would observe that the classical CV approach did not detect so many infocus peas. Hence, if the segmentation output is improved, visual attribute prediction would be enhanced. To improve segmentation and thus visual attribute prediction, we would need to go beyond the classical CV method. Therefore, this study was focused on applying deep-learning methods to improve image segmentation and visual attribute prediction.

Our first objective was to train a semantic segmentation model known as U-Net, using original images captured by Iheonye et. al, as well as images we obtained while conducting new fluidized bed drying runs. Our second objective was to investigate the performance of a popular convolutional neural network (CNN) known as Xception, for predicting real-time changes in color, texture, and size, using the segmented output from U-Net. Afterwards, we compared the performance of our deep learning solution to the classical CV technique, in terms of image segmentation and visual attribute prediction. Finally, we investigated how the classical and deep learning models performed when used to monitor real-time changes in visual attribute of peas, for new drying runs conducted at 50, 55 and 60^{o}C. By the end of this research paper, we hope to provide a tangible response to the research question ***"Is it possible to monitor real-time changes in visual attributes of food during fluidized bed drying, using images of the food captured during the drying process?"***

4.3. MATERIALS AND METHODS

In this section, we would discuss how we developed the system for monitoring real-time changes in visual attributes. This system monitors visual attributes using images of peas captured during the drying process. Our discussion would be broken down into four parts. Each part represents one of the sub-functions performed by the system when monitoring visual attributes. The sub-functions as depicted in Figure 4.7 include image acquisition, image preprocessing, image segmentation and visual attribute prediction. Let us start with image acquisition.

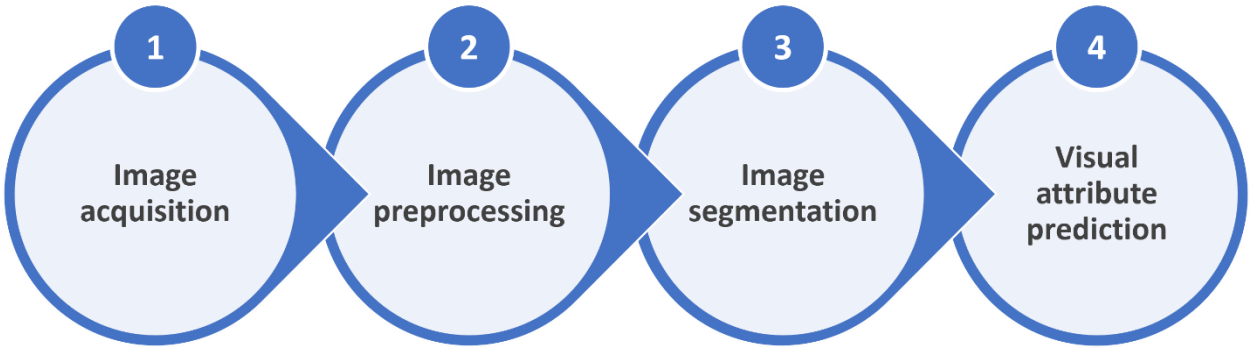


Figure 4.7 Detailed set of functions performed by the visual attribute monitoring system

4.3.1. Image Acquisition

In this study, real-time images of green peas were captured using the same Image Acquisition System (IAS) developed by Iheonye et al.[26] The IAS shown in Figure 4.8 captured images using a digital SLR Camera (Nikon D5300, Nikon Inc., Japan). The images were acquired during fluidized bed drying experiments conducted at 50, 55 and 60^{o}C. During each drying run, 10 images were captured every 15 minutes. The camera lens was centered along the line of symmetry of the cylindrical drying chamber, by mounting the camera on a 15 mm LWS Rail system (F&V, Illinois, USA).

Figure 4.8 Component of the real-time image acquisition system for fluidized bed drying [26]

To ensure that green peas within the image were distinct from the background, a red sheet of paper was pasted on the outer wall of the glass cylinder, as shown in Figure 4.8. During the preliminary experiments, green, blue, black, and red backgrounds were tested. The red background provided the best contrast, due to the green color of the peas.

Once adequate contrast between the peas and the background was achieved, the next task was to make the appearance of infocus peas (peas that touched the wall of drying chamber) distinct from the defocus peas (the peas behind the infocus peas). To achieve this goal, the aperture of the camera was increased to f/4.5. With this setting, the object of interest (infocus peas) appeared sharper while the details behind them appeared blurred. After making the infocus peas conspicuous, we had to the solve the problem of motion related blur identified around the edges of the peas. The blurry edges were due to swift random motion of the peas during fluidized bed drying. To eliminate this blur, the shutter speed of the camera was increased to 1/2500 second. However, with this high shutter speed the resulting RGB image appeared dark. To compensate for the reduction in illumination, two R300 SE ® LED ring lamps (18.7W, F&V USA) were positioned horizontally at the top and bottom of the cylindrical drying chamber, with the bulbs pointing into the drying chamber. The color temperature and illuminance of the lamps were 5,600 K and 1187 Lux, respectively. Examples of real-time images captured during the study are displayed in Figure 4.9.

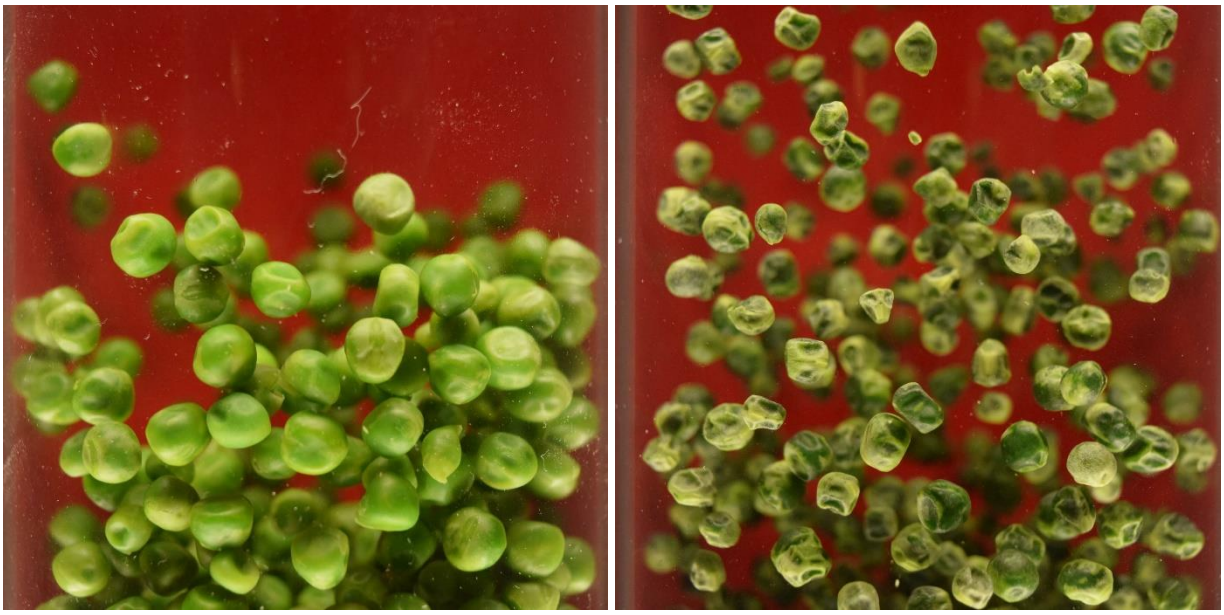


Figure 4.9 Images of peas capture during fluidized bed drying

More information on the image acquisition system can be found in the original publication by Iheonye et al.[26] For each of the drying experiments (50, 55, and 60°C), 350 g (±0.01g) of green peas were dried for 75 minutes. Each batch of green peas had an initial moisture content of 76% (WB). Ten images were captured at the start of drying, as well as at every 15 minutes afterwards. These images were later used to train a semantic segmentation model known as U-Net. However, before the images could be used to train the model, they had to be preprocessed.

4.3.2. Image Preprocessing

The image preprocessing step can be divided into three sub-steps – cropping, resizing and augmentation. Let us start with image cropping. During image acquisition, the size of the captured image was 2000 x 2992 pixels. The section of the image that was important for our study was the region within the cylindrical drying chamber, as it contains the green peas. Incidentally, since both the width and height of the cylinder that appeared in the image were 2000 pixels each, we cropped the image to 2000 x 2000 pixels. The cropping was achieved by removing the regions left and right of the cylindrical drying chamber. An example of the original image and the cropped version is shown in Figure 4.10.

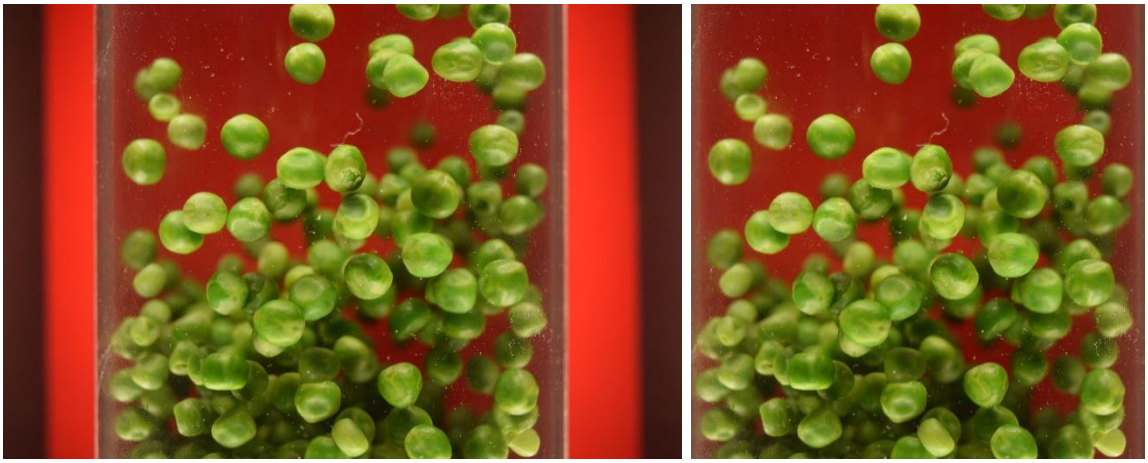


Figure 4.10 Original image and cropped version

Due to limitations in RAM size and the need to shuffle the training set at the start of each epoch, it was impractical to train the model using an input size of 2000 x 2000 pixels. However, since the width and length of the image were equal, it was feasible to resize the image to 1024 x 1024 pixels, without distorting the original shape of the green peas. After resizing, the second part of the image preprocessing step was completed. At that point, the images were ready to be fed into the model.

Training a semantic segmentation model such as U-Net, involves a two-stage loop, comprising of forward propagation and back propagation. During forward propagation, the input image is fed through a series of convolutional and deconvolutional layers. During this process, the model learns unique spatial and cross channel patterns regarding the image. At the end of the forward propagation step, the model produces a prediction. In our case, since we have only two classes – background (0) and infocus peas (1), the prediction was a 1024 x 1024 x 2 matrix. Once this matrix is produced, the next stage of the forward propagation phase would be to compute the cost between the predicted mask (matrix) and the ground truth mask. In order to carry out this cost computation, every input image must have a corresponding ground truth mask. Hence, before we get to the point of training the model, we need to produce all the ground truth masks.

These masks were produced manually using an image annotation software known as LabelMe®-v5.0.1.[78] Due to the labor-intensive nature of manual annotation, only images captured during the 60^{o}C drying experiment were annotated. By the end, we had sixty images and their corresponding ground truth masks. However, that number was insufficient to train, validate, and test the model. To automatically generate more images and masks, we moved on to accomplish the third sub-step of image processing known as image augmentation.

4.3.2.1. Image augmentation

Image augmentation is a process of generating more realistic variants of the original images and their corresponding masks. The new variants are produced by applying random transformations to the original training examples. The transformations applied in this study include random left-to-right flip, random rotation, random crop, and random up-down flip. Since we were to monitor real-time changes in color and texture, transformations such as adjusting the brightness, the saturation, or the contrast of the RGB image were not applied. Figure 4.11 contains some of the realistic variants obtained through image augmentation.

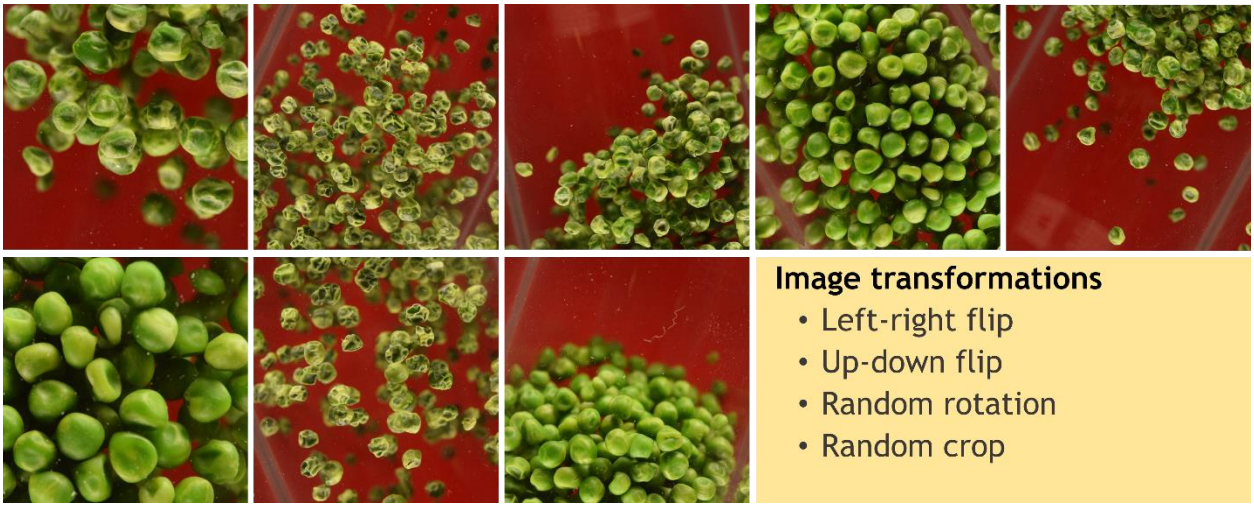


Figure 4.11 Realistic variants generated through image augmentation

To ensure that variants of a pair of original image and mask were either in the training set, the validation set, or the test set, before carrying out augmentation, the original images and their corresponding masks were split into training, validation, and test sets. However, at this point, one may ask "what is the splitting criteria? How is the training set, the validation set, or the test set to be determined?" One response to this question could be:

"We ensured that images captured at the six different time steps were well represented across the training, validation, and test sets". To achieve this goal, the original 60 images could be divided into six groups. These six divisions are depicted in Figure 4.12. Each column represents the ten images that were captured at 0, 15, 30, 45, 60, and 75 minutes time steps. To distribute the images across the training, validation, and test set, each of the six groups could be split into three. The first sub-division across the six groups would be assigned to the training set. The next subdivision would be assigned to the validation set. While the last subdivision would go to the test set. With this method, images captured are at different time steps and are well represented across the training, validation, and test sets.

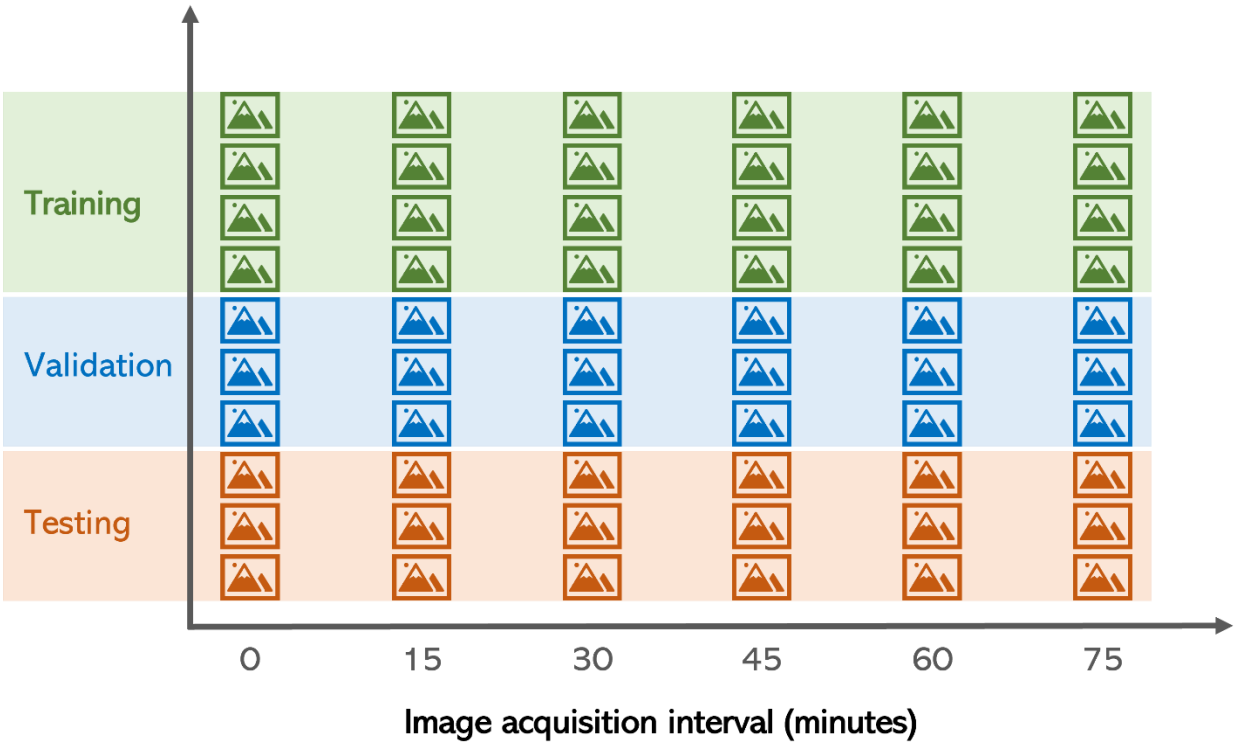


Figure 4.12 Dataset splitting based on acquisition interval

This method would be appropriate if the goal of this research was limited to image segmentation. However, in our case, we were also interested in predicting visual attributes using the segmented output. During our preliminary experiments, the dataset was split based on acquisition interval, the U-Net model segmented well. However, the visual attribute model performed poorly in predicting some of the attributes, especially those attributes whose distribution was skewed. Hence, to ensure the range of values for each visual attribute are equally represented across the training, validation and test sets, the images were split amongst the three sets, using a skewed visual attribute as the splitting parameter. This data splitting method is known as the stratified sampling method.[87]

During this study real-time changes in color was monitored by training the model to predict L*, a* and b* indices. Texture was monitored by predicting contrast, correlation, entropy, and homogeneity. While size was monitored using equivalent diameter, ferret diameter, filled area, perimeter, and roundness. The original sixty images and their corresponding masks were split across the training, validation, and test sets, using the stratified sampling method. [87] *a** index was the split parameter. The split ratio was 80% - 10% - 10%. After splitting the images within each set, each group was augmented using the method described previously. After augmentation, the

number of images in the training, validation and test sets were 2974, 905 and 732, respectively. The data pipeline for splitting and augmentation was developed using Python-v3.7. Some of the Python libraries that were used include Tensorflow-v2.7.0, Skimage-v0.18.0 and Panda-v1.40. Having augmented the dataset, we completed the image pre-processing step. Hence, we now discuss the third consecutive step performed by the visual attribute monitoring system. We refer to this step in Figure 4.7 as *"image segmentation"*.

4.3.3. Image Segmentation

Semantic segmentation was carried out using a deep learning model known as U-Net.[69]

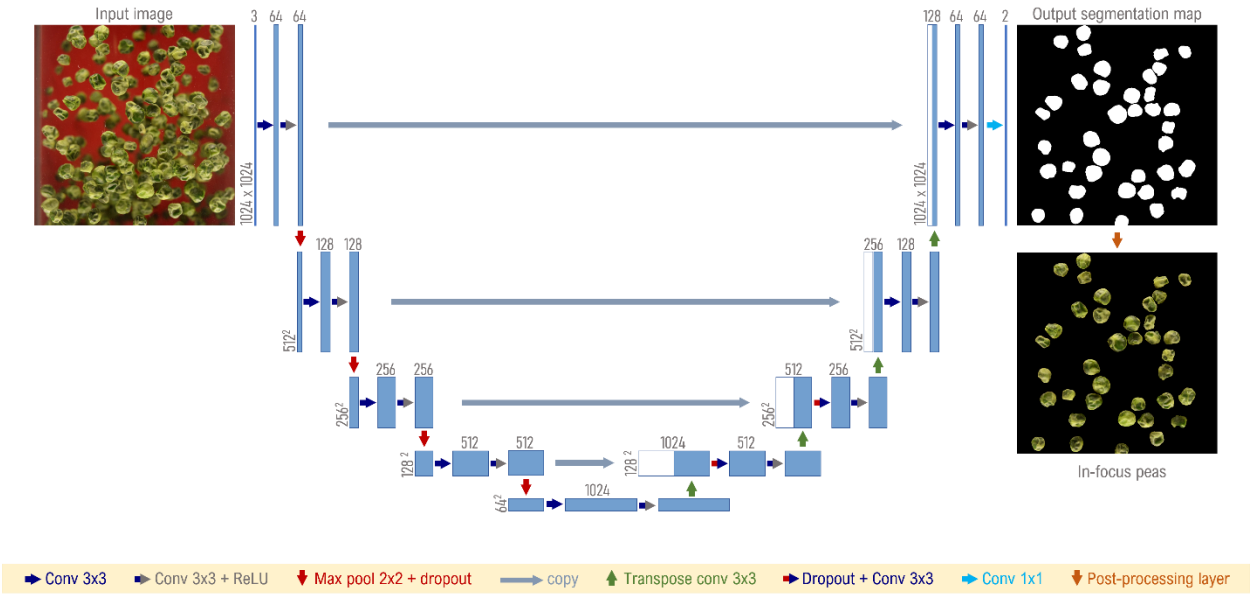


Figure 4.13 *The U-Net architecture. Each blue box corresponds to a multi-channel feature map. The number of channels is denoted on top of the box. The x-y-size is provided at the lower edge of the box. White boxes represent copied feature maps. The arrows denote the different operations.* [69]

This model consists of two sections: a contracting path (left) and an expansive path (right). The architecture of the contracting path is similar to that of a typical convolutional neural network. It is made up of repeated applications of 3x3 convolutions (with same padding), each followed by a ReLU activation unit and a 2x2 max pooling operation. These steps successively reduce the size of the convolutional layers, while increasing their feature maps.

For an input RGB image of size 1024 x 1024, the size of the last convolutional layer along the contracting path would be 28 x 28 with 1024 feature maps. With an increased number of feature maps, the neural network would learn unique spatial and cross-channel patterns regarding the image. With this information, it could successfully classify/detect the object of interest. However, due to the substantial reduction in the size of feature maps (from 1024 x 1024 to 64 x 64), the image losses its spatial resolution as it travels down the contracting path. Due to this loss, the network would perform poorly in classifying each pixel position as either a background or as an infocus pea. Notwithstanding this problem, the network would be able to detect that there are in-focus peas at various locations in the image, but it would be difficult to accurately identify the specific pixels that make up each in-focus pea.

To solve this problem, the authors of U-Net introduced an expansive path (the right region in Figure 4.13). As the output from the contracting path is fed into this region, the size of the feature map is upsampled using a transposed convolutional layer. At the same time, the number of feature maps is reduced by 2 (per step). Also, at each step along the expansive path, after the size and feature maps have been modified, the output is concatenated with the corresponding feature map from the contracting path. Afterwards, two 3 x 3 convolutions, each followed by ReLU activation units were applied. All these steps within the expansive path - upsampling, concatenation and convolutions help to recover significant information lost along the contracting path, while retaining the important cross-channel features of the image. Hence, at the end, the network can semantically segment the input image.

At the final layer, a 1 x 1 convolution was applied to map each of the 64-component vector to the desired number of classes. Afterwards, the SoftMax function (Equation 4.1) was used to compute the probability $\hat{p}_k$ that each pixel position within each feature map belonged to the positive class. The cost $J(\Theta)$ was computed using the cross-entropy cost function (Equation 4.2) [87]

$$\hat{p}_k = \sigma\big(s(\boldsymbol{x})\big)_k = \frac{\exp\big(s_k(\boldsymbol{x})\big)}{\sum_{j=1}^{k} \exp\left(s_j(\boldsymbol{x})\right)} \tag{4.1}$$

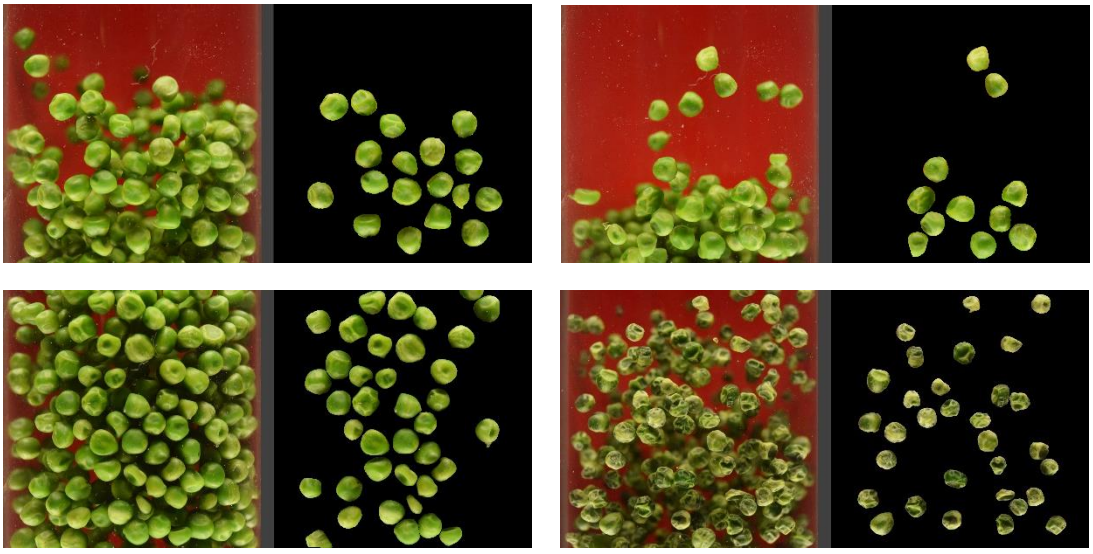
$$J(\Theta) = -\frac{1}{m}\sum_{i=1}^{m}\sum_{k=1}^{K} y_k^{(i)} \log\left(\hat{p}_k^{(i)}\right) \tag{4.2}$$

In these equations:

- K is the number of classes.
- $s(\boldsymbol{x})$ is a vector containing the scores of each class for the pixel position **x**
- $\sigma\big(s(\boldsymbol{x})\big)_k$ is the estimated probability that pixel position **x** belongs to class k, given the scores of both feature maps for that particular instance.
- $y_k^{(i)}$ is the target probability that the i^{th} pixel position belongs to the class K. It is either 1 or 0, depending on whether the pixel position belongs to the class or not.

In this study, we have two classes we were interested in, the background and the infocus peas. When an image is fed into the model, it generates a 3-D matrix containing probability or activation values ranging from 0 to 1. With the SoftMax activation function, the expectation is that a fully converged U-Net would return activation values between 0.5 and 1, for pixel positions that are occupied by infocus peas. While values less than 0.5 would be for positions occupied by the background. However, there were two major challenges that prevented optimal model convergence. What were these challenges and how were they solved?

Let us briefly observe Figure 4.14, it displays two images per row. Each picture contains a real-time image and a corresponding image of infocus peas. We would notice that the area or number of pixels positions occupied by the background (defocus peas and red paper) is larger than the number of pixels occupied by the infocus peas.



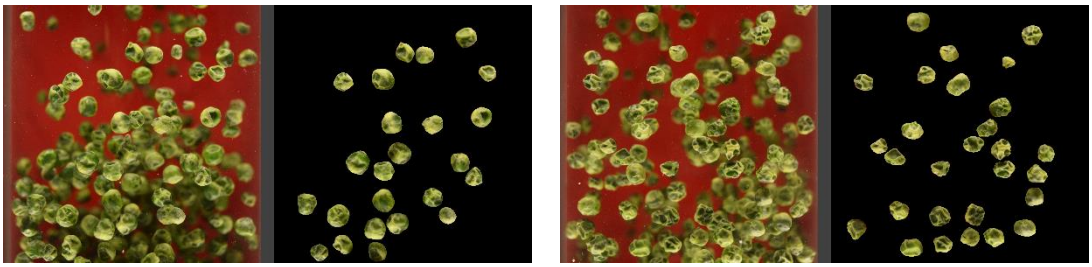


Figure 4.14 Image of infocus peas depicting class imbalance and varying background area across the training examples

This disproportion in area leads to class imbalance, thereby making it difficult for the model to properly classify the under-represented class, in this case the infocus peas. Although class imbalance is a challenge, it can be straightforward to deal with, if the area occupied by the background was the same for all ground truth masks. If that is the case, then all we need to do is to assign a weight to the under-represented class. For instance, if the number of pixels representing the background was three times the size of the infocus peas, we would apply a class weight of 0.6667 and 0.3333 to the infocus peas and the background, respectively. Unfortunately, in our case, this solution did not work, because the number of pixel positions representing the background varies significantly across the ground truth masks. This variance is further complicated with image argumentation, due to random cropping and random rotation. To solve this problem, we applied sample weights, instead of class weights. With sample weights, each training example would be assigned with a unique weight for the background class and the infocus pea class.

The second problem we experienced when training the network, was a challenge that was also was encountered by the inventors of U-Net. The U-Net model was originally built to semantically segment biological images, such as images of HeLa cells on glass. One of the initial challenges the authors encountered was that occasionally, the model wrongly classified the small boundary separating closely touching cells. Instead of classifying this boundary as a background, the model classified it as an object of interest (a HeLa cell). The same situation occurred in our study, when infocus peas were clustered, such as the one during the onset of drying. When this happened, the region between infocus peas were misclassified.

To solve this problem, the authors of U-Net introduced a weighed loss. During the training process, anytime a pixel position, along the boundary separating closely positioned cells, was misclassified as being an object of interest, a huge weight or penalty was applied to the activation value

associated with that pixel position. Figure 4.15 shows the weight map of a given training example. This map contains both the weights to offset class imbalance, as well as increased weights for boundaries of closely packed infocus peas. The increased values are depicted on the weight map by the yellowish color along the thin border separating the remarkably close peas. The closer the pea are to themselves, the greater the weight.

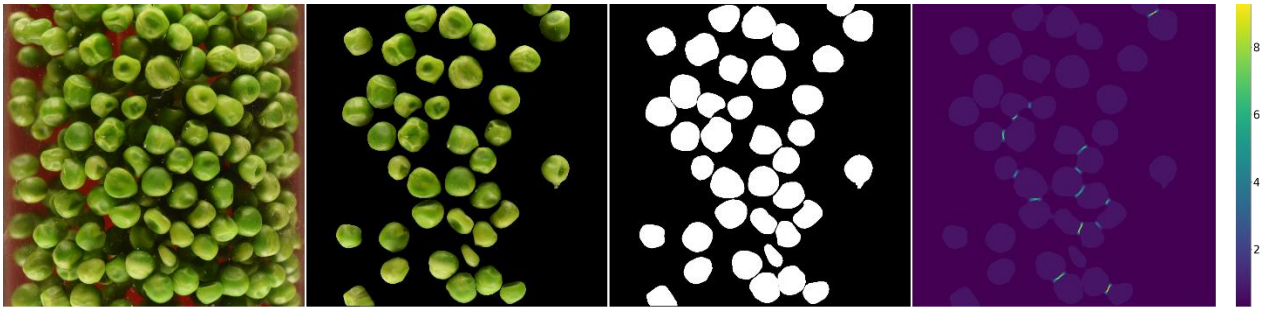


Figure 4.15 (a) original image, (b) image of infocus peas, (c) ground truth mask, and (d) image of weighted loss map

Applying such penalty to the activation value of a misclassified pixel increases the overall loss. Hence, after series of iterations and re-applied penalties, the network learns to identify the boundaries of closely packed infocus peas. The weighted loss was computed using Equation 4.3. [69]

$$\omega(\boldsymbol{x}) = \omega_c(\boldsymbol{x}) + \omega_0 \cdot \exp\left(-\frac{\left(d_1(\boldsymbol{x}) + d_2(\boldsymbol{x})\right)^2}{2\sigma^2}\right) \tag{4.3}$$

In this equation:

- ω_c is the weight map to offset class imbalance
- d_1 represents the distance to the border of the closest infocus pea
- d_2 is the distance to the border of the second closest infocus pea

Similar to the U-Net authors, while training U-Net, we set $\omega_0 = 10$ and $\sigma = 5$ pixels. At this point in the study, we had solved the problem of class imbalance and misclassification of the boundary of closely packed peas. Let us now discuss the hyperparameters and other settings that were put in place when training the model.

The U-Net model was trained using 1024 x 1024 images and masks, a batch size of 1 and a buffer size of 500. Back propagation was conducted using the Adam optimisation algorithm.[88] The momentum decay hyperparameters β_1 and β_2 were set to 0.9 and 0.999, respectively. For numerical stability, epsilon was set to 1×10^{-7}. The network was originally trained for 112 epochs, using an initial learning rate of 1e-4. With the help of a polynomial learning rate schedule callback, the learning rate decreased as the number of epochs increased. The dropout rate was set to 0.5 to prevent the model from overfitting the training set.

The cross-entropy function[87] (Equation 4.1) was used to compute the loss between the predicted and ground truth mask. Categorical accuracy and Mean Intersection-over-Union (MIoU) were the evaluation metrics. Equation 4.4 computes the IoU for a single class[87]:

$$iou = \frac{true\_postives}{(true\_postives + false\_positives + false\_negatives)} \tag{4.4}$$

Early stopping was applied, using a patience of 20. After the first round of training, the training set and validation set were combined to form new set of training examples. This combined set along with the original hyperparameters were used to train the model one more time. Afterwards, the model was evaluated using the test set. Once evaluation was completed, we were done with the image segmentation step. Let us conclude this *Materials and Methods* section by discussing the final sub-operation performed by the food monitoring system – visual attribute prediction.

4.3.4. Visual Attribute Prediction

Before we kick off our discussion on predicting visual attributes, let us briefly recap what the end goal of this research was. Our global objective was to develop an intelligent system that monitors real-time changes in visual attributes of food during fluidized bed drying. Looking at this system from the vantage point of an end-user, the system is made up two major components – an image acquisition system (hardware) for capturing real-time images of the food, and a deep-learning model for predicting visual attributes using images fed to it by the acquisition system.

Up to this point in the research, we had completed the first part. We now discuss the second part, that is developing the model for predicting visual attributes. The parent model, labelled as "Unet-Xception" in Figure 4.16, comprises of two sub-models: U-Net and Xception.

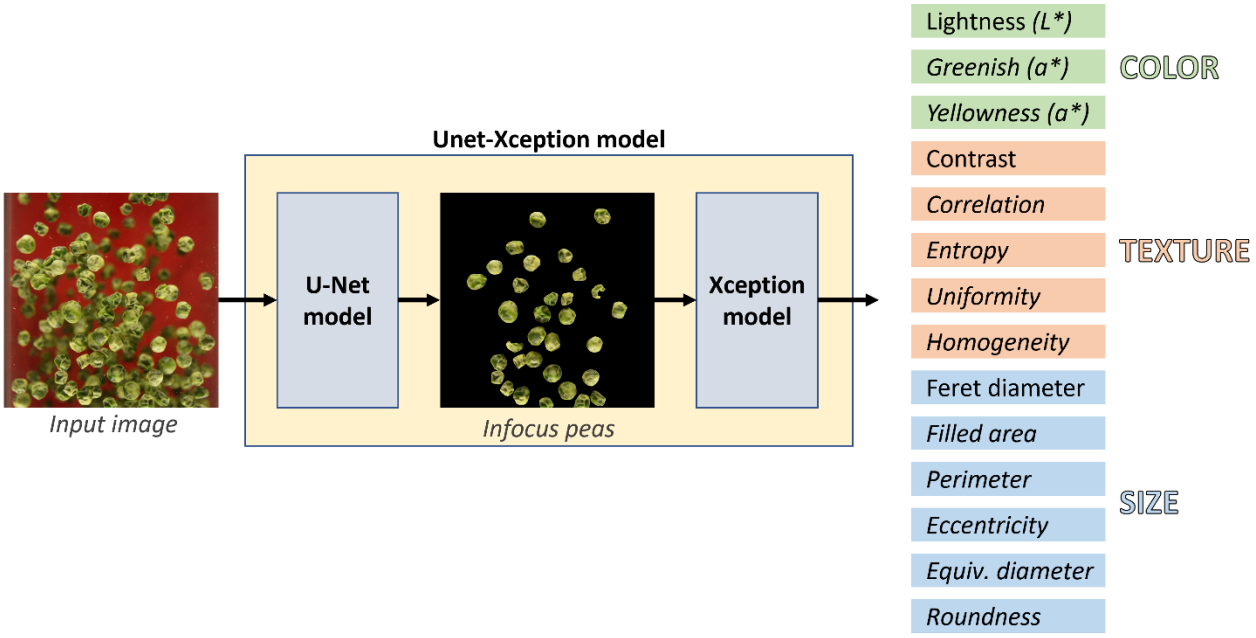


Figure 4.16 Image depicting the overall model made up of U-Net and Xception models.

Once the acquisition system captures an image, it feeds the image to U-Net for segmentation. The segmentation step produces an RGB image containing only the infocus peas. This output is then fed to Xception for visual attribute prediction. We have discussed how U-Net was trained using the complete RGB images that contained infocus peas, defocused peas, and the red background. We would shift the focus now to how Xception was trained. Combining the trained U-Net and a trained Xception would form the parent model "***Unet-Xception***". So, how did we train the Xception network?

Unlike U-Net that was trained with complete images, Xception was trained separately, using RGB images containing only infocus peas, because that is the form of input it would receive from U-Net when both models are coupled together.

4.3.5. Visual Attributes using Xception

Before delving deeper, let us briefly describe each of the visual attributes that was predicted and how their ground truth data was generated. Next, we go into more detail about the Xception network, and why we chose it over other popular CNN architectures. Finally, we discuss how Xception was trained. Within the model training section, we outline the hyperparameters, as well as the loss, optimization, and metric functions that were applied while training the model.

During this study, the Xception model was trained to monitor size, texture, and color of green peas, during fluidized bed drying. Each attribute was monitored using various visual quantities.

4.3.5.1. Texture

The surface texture of infocus peas was monitored by predicting the following statistical pixel intensity related measurements: contrast, correlation, energy, entropy, homogeneity, and uniformity.

Contrast

Measures the disparity in intensity between a pixel and its neighbouring pixel across the whole image. Contrast ranges from 0 to $(K-1)^2$. It is equal to 0, for an image containing the same intensity value at all pixel positions.[57,58] This visual property was computed using Equation 4.5.

$$contrast = \sum_{i=1}^{K}\sum_{j=1}^{K}(i-j)^2 \cdot P_{ij} \tag{4.5}$$

Correlation

This attribute measures how correlated a pixel is to its neighbour across the entire image.[57,58] Correlation ranges from -1 to 1. The value is 1 for a perfectly correlated image, and -1 for a perfectly uncorrelated image. This visual attribute was computed using Equation 4.6.

$$correlation = \sum_{i=1}^{K}\sum_{j=1}^{K}\frac{(i-m_r)(j-m_c)P_{ij}}{\sigma_r\sigma_c} \tag{4.6}$$

$$\sigma_r \neq 0;\ \sigma_c \neq 0$$

Entropy

Entropy measures the randomness of the pixel intensities within the region occupied by the infocus peas. The ground truth data was computed using Equation 4.7.[57,58]

$$Entropy = -\sum_{i=1}^{K}\sum_{j=1}^{K}P_{ij}\log_2 P_{ij} \tag{4.7}$$

Homogeneity

Homogeneity evaluates the spatial closeness of the distribution of elements in the **G** to the diagonal of **G**.[58] Homogeneity ranges from 0 to 1. The maximum value is obtained when **G** is a diagonal matrix.[57] Ground truth for homogeneity was computed using Equation 4.8. [57,58]

$$homogeneity = \sum_{i=1}^{K}\sum_{j=1}^{K}\frac{P_{ij}}{1+|i-j|} \tag{4.8}$$

Uniformity

It measures the uniformity of pixel intensity in the range [0, 1]. Maximum uniformity is achieved when all intensity values are equal, and it decreases from that point.[57,58] Uniformity was computed using Equation 4.9. [57,58]

$$uniformity = \sum_{i=1}^{K}\sum_{j=1}^{K}{P_{ij}}^{2} \tag{4.9}$$

where,

P_{ij} is the histogram of the intensity levels in a region,

K is the number of possible intensity levels,

G represents the co-occurrence matrix,

m_r and m_c are the mean computed along rows and columns of **G**,

σ_r and σ_c are the standard deviations calculated along the rows and columns of **G**

4.3.5.2. Size

The average size of the green peas at each time step during the drying process was monitored by predicting the following quantities: equivalent diameter, eccentricity, Feret diameter, filled area, perimeter, and roundness. Each of these properties are described below:

Equivalent diameter

The diameter of a circle with the same area as the region occupied by the infocus pea. This diameter was computed as:[58]

$$\sqrt{\frac{4 \times Area}{\pi}} \tag{4.10}$$

Eccentricity

The eccentricity of the ellipse that has the same second moment as the region occupied by the infocus pea. Eccentricity is the ratio of the distance between the foci of the ellipse and its major axis length.[58]

Ferret diameter

The longest distance between points around the convex hull contour of a region (infocus pea).[89]

Filled area

The number of pixels in a filled region (that is the region occupied by an infocus pea).[58]

Perimeter

The distance in pixels around an infocus pea.

Roundness

Ratio of area (A) of the object of interest and that of a circle that has the same perimeter (P) as the object.[8,15,53,58] Roundness was computed using Equation 4.11.

$$roundness = \sqrt{\frac{4\pi A}{P^2}} \tag{4.11}$$

4.3.5.3. Color

The color of the green peas was monitored using the CIE L^* a^* b^* model, also called CIELAB (CIE [1978], Robertson [1977]). This color space is a highly effective decoupler of intensity and color.[57] Intensity is represented by lightness L^*, while color is described by a^* and b^*. Lightness (L^*) ranges from 0 (black) to 100 (white). a^* is the redness-greenish index, it ranges from 127 to -128. The positive values indicate red, while negative values represent green. Finally, b^* is the yellowish-bluish index, ranging from 127 to -128. Where positive values represent yellow, and negative values represent blue. The conversion from RGB to L^* a^* b^* color space was conducted in a similar manner as descried by Leon et al.[57,90]

The pipeline for generating ground truths of visual attributes for each RGB image of infocus peas was developed using Tensorflow-v2.7.0, Skimage-v0.18.0 and Panda-v1.40. Let us now describe the Xception model.

4.3.6. Xception

While reading through this chapter, one may ask "Why use Xception? Why not ResNet, Inception-v4 or GoogleNet?" These other three networks are all CNN architectures that have proven to be effective in image classification. So, why did we choose Xception? During our preliminary trials, before deciding which network to use for the study, we considered three yardsticks, namely, the number of parameters within the model that need to be trained, the computational demand, and the overall performance of the network. Xception turned out to be the best for the study.

The unique feature of the Xception architecture that makes it remarkably effective and attractive is the nature of convolutional layers that make up the network. It uses a form of convolution known as ***Depthwise Separable Convolution***. This concept was brought back into limelight by Francois Chollet in 2016 when he developed the state-of-the-art image classifier popularly known as Xception.[77] With the help of Depthwise separable convolutional layers, Xception significantly outperformed the then state of the art image classifier known as Inception-v4.[74] Let us discuss how Xception archived that milestone.[87]

So far, regular convolutional layers use filters to extract spatial patterns and cross channel relationships. However, separable convolutional layers make the case that spatial and cross-channels patterns can be modelled separately, as depicted in Figure 4.17. [87]

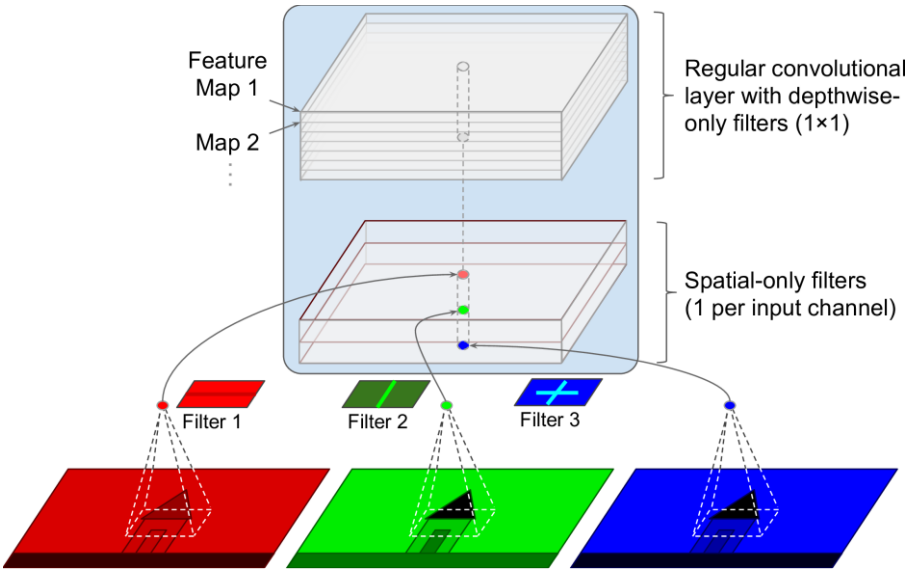


Figure 4.17 Depthwise separable convolutional layer [87]

This convolution process involves two steps, first, it applies a single spatial filter for each feature map in the input layer. Afterwards, 1 x 1 filters are applied to the output (pointwise convolution). The first step models spatial patterns, while the later models depthwise relationships. Notably, since this convolutional layer has only one filter per input channel, it is counterproductive to apply depthwise separable convolution to input layers with very few feature maps (such as the original RGB image that has only three channels). Applying it to very few channels would not produce significant cross channel features. This layer should be used only after the number of feature maps of the initial convolutional layer has been increased significantly.[87] Hence, the Xception network starts with two regular convolutional layers, and the rest are 34 depthwise separable convolution layers, in addition to a few max pooling layers. The architecture then ends with a global average pooling layers and a dense output layer. In practice, depthwise separable convolutional layers have proven to outperform regular convolutional layers.[87] Next, we would discuss how the model was trained.

4.3.7. Model Hyperparameters and Training

Xception was trained using the Adams optimization function.[88] The loss was computed using the Mean Squared Error (MSE) and Mean Absolute Error (MAE) functions. MAE was used to compute the loss for correlation and contrast, due to the presence of outliers in the dataset, while MSE was used for other visual parameters.

The Rectified Linear Unit (ReLU) activation function was used within the convolutional layers, while a linear activation function was applied at the output layer. At the end of each epoch, the model was evaluated using MAE and Coefficient of Determination (R^2). The model was trained for 153 epochs, using a single batch size, 2944 steps per epoch, image size of 1024 x 1024 pixels, and learning rate ranging from 1×10^{-4} to 1×10^{-7}. The initial weights of the model were those of an Xception model trained on the ImageNet dataset. Like U-Net, Xception was also trained using samples weights for each visual attribute.

During the first round of training, the model was trained to predict all twelve visual attributes. After the first round of training, we discovered the overall performance of the model was reduced, due its low R^2 value when predicting contrast, correlation, entropy, and roundness. Hence, during the second round, the model was trained without those four parameters. After this round of training was over, the original training set and the validation set were combined to form new training examples. The new set was then used to train the model one more time, using the same hyperparameters as before. However, this time the model was evaluated on the test set. In the result section, we show how the model performed on both the validation and test sets. The data input pipeline and the model architecture were developed and trained using the TensorFlow-v2.7.0 deep-learning library.

4.4. RESULTS AND DISCUSSION

In our introduction, we discussed the classical computer vision solution that was developed by Iheonye et al. for segmentation and visual attribute prediction. We highlighted some of the image-related problems their classical CV solution could not solve, hence the need to apply deep learning. Our discussion within this section, would focus on comparing the performance of the UNet-Xception model to the performance of the classical pipeline. This discussion is divided into two parts, image segmentation and visual attribute prediction. While reporting on image segmentation, we compare the performance of the U-Net model to that of the classical segmentation pipeline and discuss how U-Net responded in the face of challenges that were left unsolved by the classical pipeline.

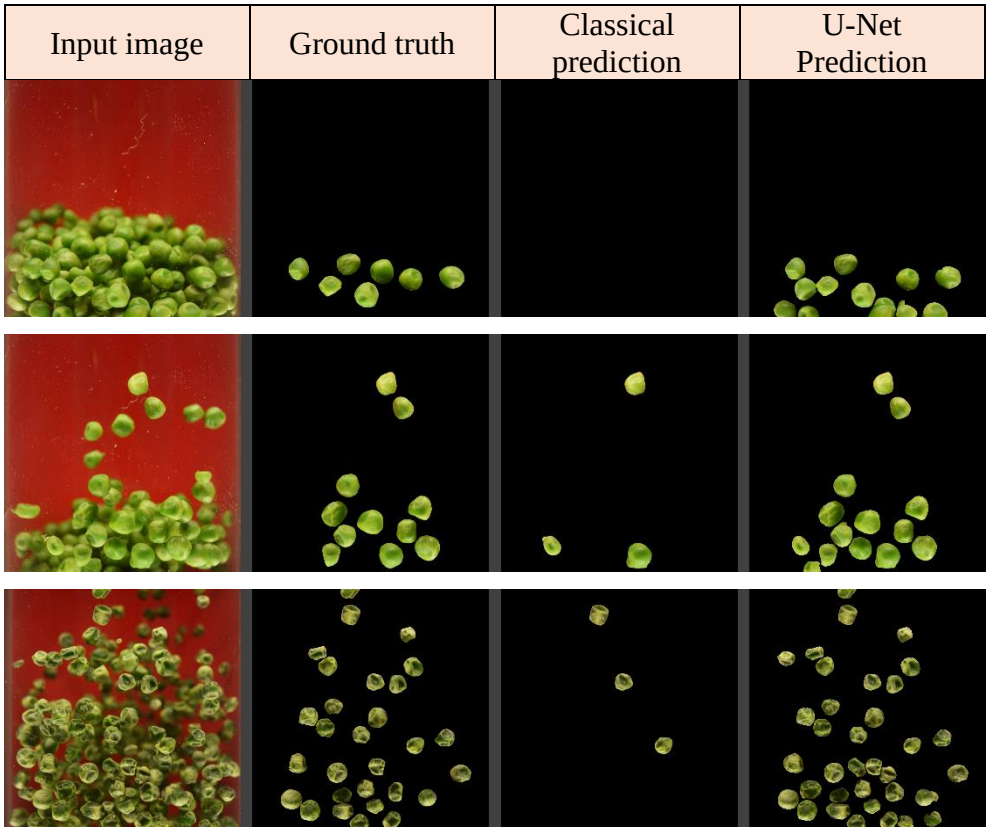
4.4.1. Image Segmentation

The Cross-entropy, Categorical accuracy, and Mean Intersection-over-Union (MIoU) for the U-Net model and the classical segmentation pipeline are displayed in Table 4.1 below.

Table 4.1. Loss and metrics for U-Net

Metric	**Classical**	**U-Net**	
		Validation	***Test***
MIoU	0.8190	0.9308	0.9464
Cross entropy	1.6037	0.1511	0.1249
Categorical accuracy	0.9005	0.9642	0.9725

U-Net significantly out-performed the classical model, producing a MIoU of 0.9464 compared to 0.81901 by the classical pipeline. The improvement in image segmentation can also be observed in Figure 4.18. This figure compares U-Net and classical pipeline predictions to the ground truth mask.



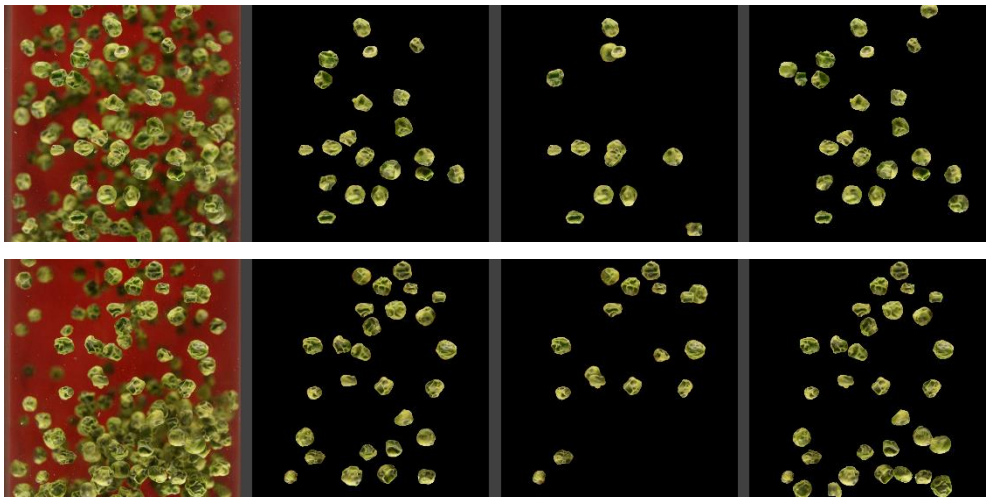


Figure 4.18 Segmented results: Input image, ground truth, classical results, and U-Net prediction

Looking through Figure 4.18, notice that U-Net segmented more infocus peas than the classical pipeline. Even when the green peas were clustered together and the color across their surface was uniform, such as in the first and second rows in Figure 4.18, U-Net was still able to segment the infocus peas. Another challenge that plagued the classical pipeline was the inability to properly segment regions within the original image that had overlapping infocus peas. U-Net performed well despite this challenge. Also, unlike the classical pipeline, U-Net was not negatively affected by variation in color and texture on the surface of individual peas.

There is another important discovery that was made. When we compare the Unet predictions in Figure 4.18 to their corresponding ground truth masks, we observe that U-Net discovered additional infocus peas that were missing from the manually annotated ground truth mask. This finding occurred for all U-net predictions in Figure 4.18. The additional infocus peas within the predicted mask provide more information that can help us understand the reason behind the unique trend that evolves within the learning curve (Figure 4.19).

Before explaining this unique pattern, let us highlight two important descriptors shown in the legend. Notice that the loss and metric terms include the words “predicted_mask” or “processed mask”. The predicted mask was the original output from U-Net before minor morphological operations were applied, while the processed mask was the result after applying morphological operations. Having explained both descriptors, let us return to our discussion on highlighting the unique patten that evolved within the learning curve (Figure 4.19).

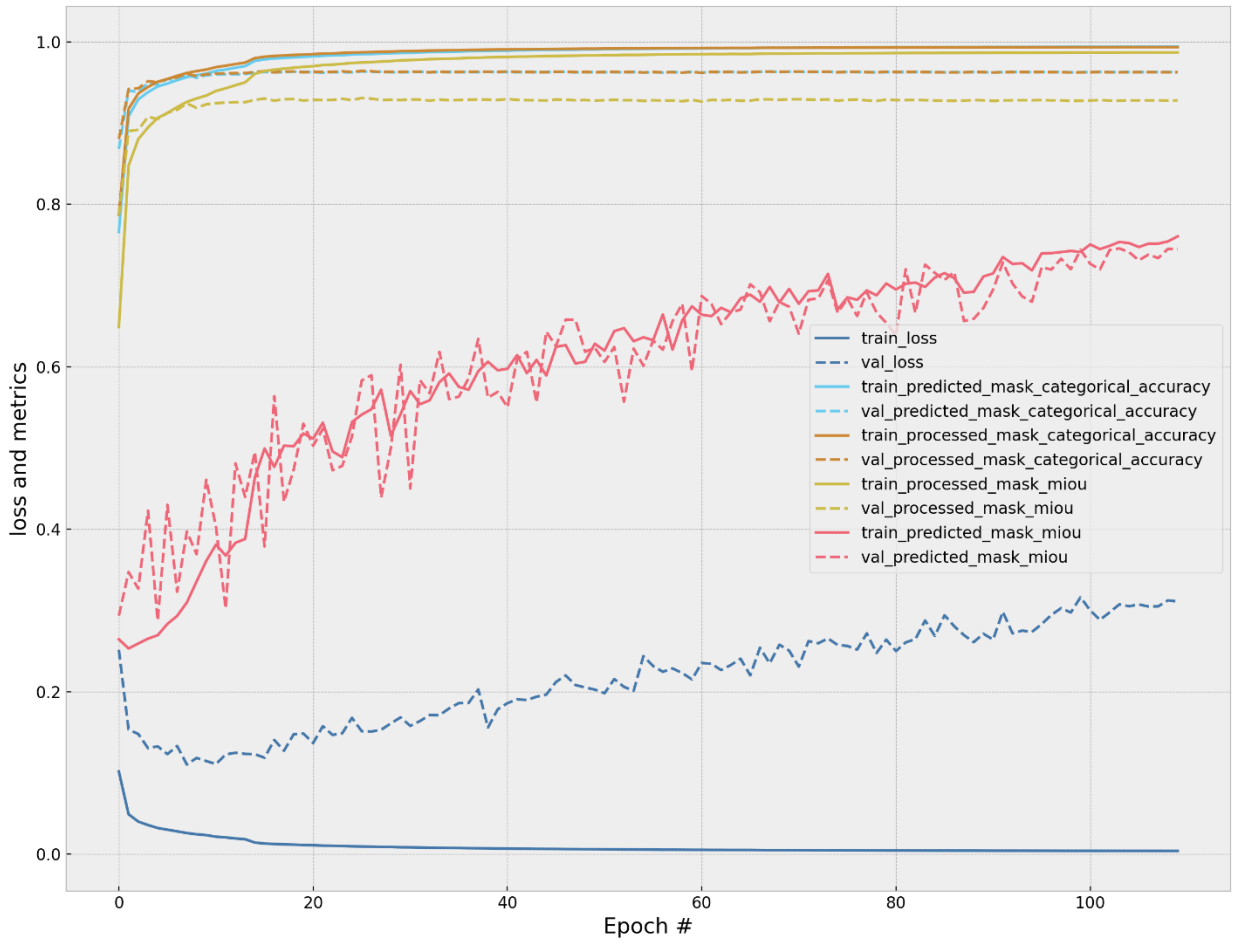


Figure 4.19 Learning curve generated while training the U-Net model

When we observe the validation loss curve in Figure 4.19, we will notice that from epoch 16, U-Net started trending towards overfitting the training set. If the model was evaluated by only comparing the training loss to the validation loss, this consistent increase in validation loss while the training loss was decreasing, meant that the model was performing well on the training set, but underperforming when presented with new data (validation set). Normally, this outcome is undesirable. However, in our case, it was a very good development. Let us explain why.

As the validation loss was trending upwards, the normal response should have been a corresponding reduction in validation MIoU. Interestingly, this was not the case. Instead, the validation MIoU kept trending upwards, along with the training MIoU. This simultaneous increase in validation loss and validation MIoU, meant that the segmentation result from the model contained most of the information present in ground truth mask, and even much more. A visual

explanation of this development is shown in Figure 4.20. This figure is made up four images, an input image, a ground truth mask, the predicted mask, and the processed mask. We notice that the predicted mask contains all the ground truth infocus peas, as well as additional infocus peas that were missing in the ground truth mask.

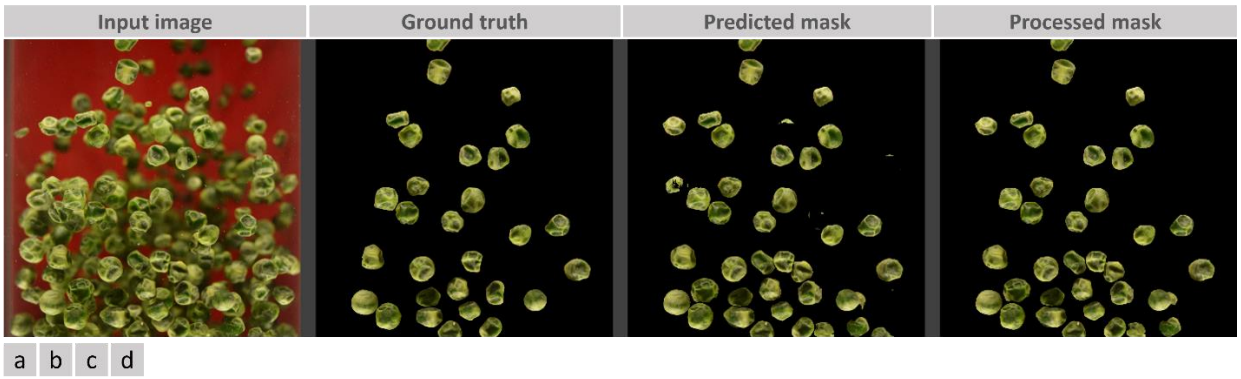


Figure 4.20 (a) input image, (b) ground truth, (c) U-Net predicted mask, and (d) U-Net processed mask

As the model was trained through a series of epochs, it learnt to include more ground truth infocus peas within the predicted mask. The inclusion of more infocus peas was responsible for the upward trend in validation MIoU. However, the presence of infocus peas that were omitted in the ground truth mask dampened the rate at which the validation and training MIoU increased. Examining the U-Net predicted mask (Figure 4.20c) shows that it also contains incomplete portions of some omitted infocus peas. These incomplete portions also contributed to the increased validation loss, and dampened MIoU growth.

Since visual attributes were meant to be predicted from the segmented results, using incomplete portions of infocus peas would lead to inaccurate visual attribute predictions. Hence, these incomplete peas were eliminated by passing the predicted mask through a custom TensorFlow morphological layer, thus producing the processed mask. Eliminating these unwanted regions significantly increased the MIoU for the processed mask. The fact that U-Net was able to detect omitted infocus peas, means that if the model is re-trained with updated ground truth masks, the processed MIoU may increase beyond 0.9464.

At this point in the research, we had developed a U-Net model that produced segmentation results containing most of the infocus peas. Once this was achieved, U-Net was coupled to a trained

Xception model for predicting visual attributes, using the processed mask from U-Net. Let us discuss how this combined model (UNet-Xception) performed in predicting visual attributes.

4.4.2. Visual Attributes Prediction

Our discussion on visual attributes is based on comparing the performance of the Unet-Xception model to that of the classical pipeline. As we mentioned in the materials and methods section, before merging U-Net and Xception, both models were trained separately. U-Net was trained to segment real-time images of peas, while Xception was trained to predict visual attributes, using RGB images of infocus peas. Afterwards, the merged model (Unet-Xception) was evaluated on the test set. Table 4.2 describes the performance of Xception and Unet-Xception in predicting twelve different visual attributes. The performance was measured by computing the Coefficient of Determination (R^2), using equations 4.12 to 4.15: [91]

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \tag{4.12}$$

$$SS_{res} = \sum_i (y_i - \hat{y}_i)^2 \tag{4.13}$$

$$SS_{tot} = \sum_i (y_i - \bar{y})^2 \tag{4.14}$$

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \tag{4.15}$$

where,

SS_{res} is the sum of squares of residuals,

SS_{tot} is the total sum of squares,

$\bar{y}$ is the mean of ground truth values,

$\hat{y}_i$ is a predicted values associated with a particular ground truth value, y_i.

*When the model's predicted values match the ground truth values, $SS_{res} = 0$ and $R^2 = 1$. If the predictions from the model are the same as a baseline model that always predicts $\bar{y}$, then $R^2 = 0$. However, if the model's predictions are worse than the predictions from baseline model, $SS_{res} > SS_{tot}$, resulting in a **negative** R^2.*

From Table 4.2, we would notice that in most cases, the Xception model slightly outperformed the combined model.

Table 4.2 R^2 values for classical, Xception, and Unet-Xception models in predicting twelve visual attributes

visual parameter	classical	Validation		Test	
		Xception	Unet-Xception	Xception	Unet-Xception
L*	-0.8660	0.7327	0.8141	0.7899	0.7731
a*	0.7719	0.9528	0.9321	0.8889	0.8669
b*	0.8694	0.9496	0.9486	0.8774	0.8521
Contrast	-3.5691	0.6421	0.2715	0.6446	**0.3128**
Correlation	-7.7618	0.5384	0.0769	0.3849	**0.1231**
Entropy	-1.4961	0.7450	0.0105	0.6262	**-0.0139**
Homogeneity	0.7987	0.9551	0.8341	0.8909	0.7234
Equivalent diameter	-1.0278	0.9254	0.7440	0.9226	0.7926
Ferret diameter	-1.6417	0.8927	0.7755	0.9152	0.8482
Filled area	-0.3663	0.8938	0.7434	0.9400	0.8391
Perimeter	-1.5425	0.8399	0.7499	0.8979	0.8250
Roundness	-8.0029	0.6532	-0.1473	0.6886	**0.3474**

This higher performance by the standalone Xception model was expected, especially when we consider that the predicted mask that was fed to the other Xception model that was part of Unet-Xception, contained more infocus peas than those in the ground truth masks.

To end up with an effective Unet-Xception model, we needed to have a U-Net model that could identify all or most of the infocus peas within the input image. So far, we had achieved this goal. In addition to obtaining a well performing U-Net, we also needed an Xception model that had been properly trained to predict visual attributes once the model received RGB images of infocus peas. At this point in the research, we are on course to achieving such high performing Xception. Why are we confident in our assertion?

Unlike the U-Net training scenario, where the ground truth masks do not always capture all the infocus peas within the original input image, the case is different when training Xception. The ground truth data of visual attributes were always accurate representations of the RGB image of infocus peas that was fed into Xception during training. How are we sure? Because the ground

truth visual attributes were generated from RGB images containing only infocus peas. Hence, by the end of the research, we had a U-Net model that was effective in segmentation, and an Xception model that predicted visual attributes well. Since the trained U-Net detected omitted infocus peas, and the Xception was trained using RGB images of infocus peas, this implies that the Unet-Xception model would predict more accurate results on visual attributes, than the ground truth values. That could be the reason behind the dampened R^2 values for Unet-Xception.

After training the models on twelve attributes, to further improve the performance of the Xception model, during the second round of training, contrast, correlation, entropy, and roundness were eliminated as inputs to the model. The reason for eliminating these parameters was due to the low coefficient of determination (R^2) of the model in predicting these attributes, as can be observed in bold red font, in Table 4.2. After Xception was trained with eight input parameters, the model was merged with U-Net to form the final version of the Unet-Xception model. The evaluation results after the second round of training are shown in Table 4.3.

Table 4.3 R^2 values for classical, Xception, and Unet-Xception models in predicting eight visual attributes

visual parameter	classical	Validation		Test	
		Xception	Unet-Xception	Xception	Unet-Xception
L*	-0.8660	0.7607	0.8227	0.8215	0.8017
a*	0.7719	0.9552	0.9338	0.9062	0.8851
b*	0.8694	0.9556	0.9503	0.8967	0.8981
Homogeneity	0.7987	0.9577	0.8405	0.9033	0.7947
Equivalent diameter	-1.0278	0.9319	0.7557	0.9298	0.8082
Ferret diameter	-1.6417	0.8954	0.7731	0.9241	0.8535
Filled area	-0.3663	0.8993	0.7472	0.9460	0.8627
Perimeter	-1.5425	0.8465	0.7493	0.9095	0.8152

At this point in the research, we have produced a Unet-Xception model comprised of highly performing U-Net and Xception models. Notwithstanding the reduced R^2 of the combined model, the high performance of the individual models meant that the Unet-Xception model would produce much smoother trends when used to monitor new drying experiments. The expectation that the combined model would monitor real-time changes in visual attributes much better than the classical model, was the second reason behind our claim that the reduced R^2 of model does not necessarily signify poor performance. Instead, it means that the model discovered more

information than was present in the ground truth data. This hypothesis was confirmed when the combined model was used to monitor new fluidized bed drying experiments, conducted at 50, 55 and 60^{o}C. For each of these drying runs, 15 images were captured every 5 minutes.

The performance of the Unet-Xception in predicting each visual attribute will be discussed using pairs of graphs. The first graph is a plot containing realtime changes in the visual attribute, predicted by Unet-Xception, for new drying runs conducted at 50, 55 and 60^{o}C. The second plot compares the Unet-Xception to the classical model, by displaying the real-time changes in the visual attribute for drying run conducted at 50^{o}C. While examining each of the plots that compare Unet-Xception to the classical mode, we notice, that for all visual attributes, the Unet-Xception model produced a smoother trend than the classical pipeline. This confirms that the deep learning model is more effective in extracting accurate information on visual attributes of the peas. Similar to the real-time trend produced by Unet-Xception for 50^{o}C, the model also produced smoother real-time trends for drying runs conducted at 55 and 60^{o}C. Let us take a look at the two classes of real-time plots that was produced for each of the visual attributes.

4.4.2.1. Color

Figure 4.21 is a graph of real-time changes in a*, predicted by Unet-Xception, for drying conducted at 50, 55 and 60^{o}C. For all the runs, a* increased as drying progressed, meaning that as the peas were drying their color changed from green to very dark green. This change in color could be due to degradation in chlorophyll on the surface of the peas. The dark coloration could have been caused by the Maillard browning reaction, due to the temperature of the drying air [17,92-94].

This upward trend in *a** was similar to the result obtained by Kaur et al.,[95] when drying green peas using a three-stage convective drying process. From Figure 4.21, higher air temperature produced increased a* values. For instance, 30 minutes into the drying process, for peas dried at 60^{o}C, the highest a* value was -10.93, unlike -15.52 and -14.44 that were captured for peas dried at 50, 55^{o}C. Also, throughout the drying process, the greenish index a* for peas dried at 50^{o}C maintained lower values than those dried at 55 and 60^{o}C. Therefore, to effect less change on the green appearance of peas, 50^{o}C would be preferable drying temperature. The higher temperature of 60^{o}C may have sped up the Maillard browning reaction, thus increasing the dark green colouration on the surface of the peas.

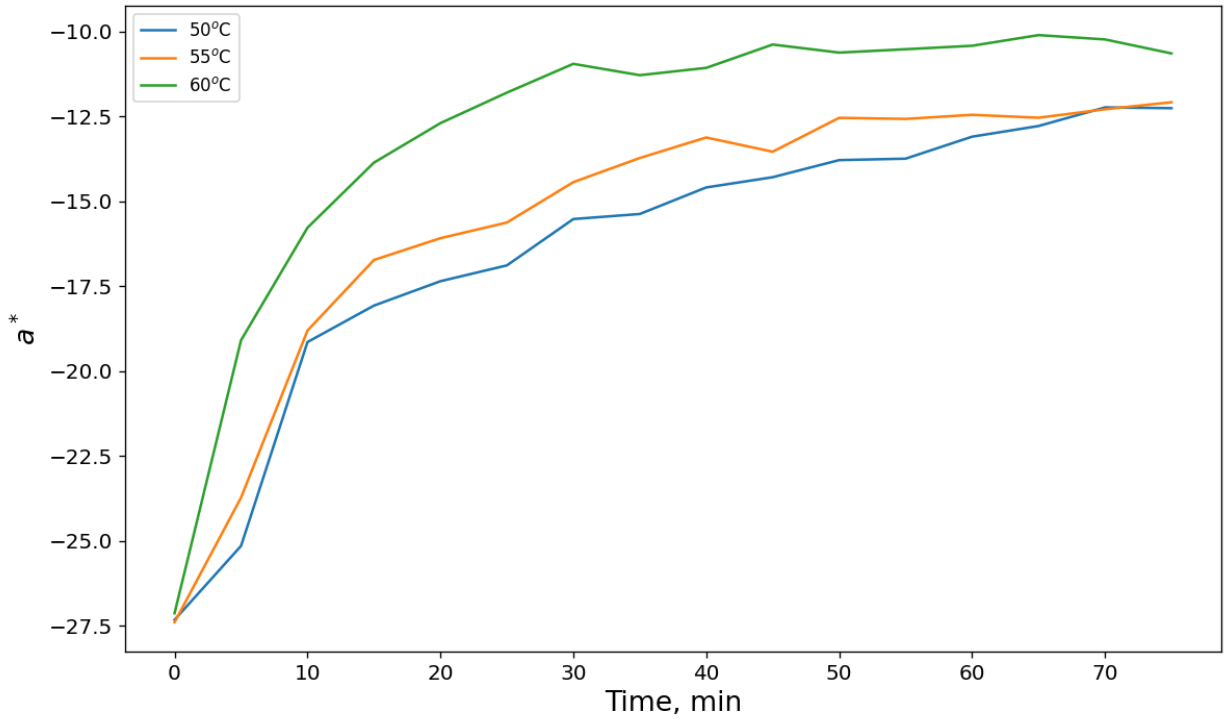


Figure 4.21 Real-time plot of a, predicted by Unet-Xception, for 50, 55 and 60°C drying runs.*

Figure 4.22 shows the plot for a* color index, produced by Unet-Xception and the classical model, for peas dried at 50°C. The figure shows that Unet-Xception produced a smoother plot of a* than the classical model. Both Figure 4.21 and Figure 4.22 show that with this AI-driven solution, it is possible to monitor real-time changes in color during fluidized drying of food.

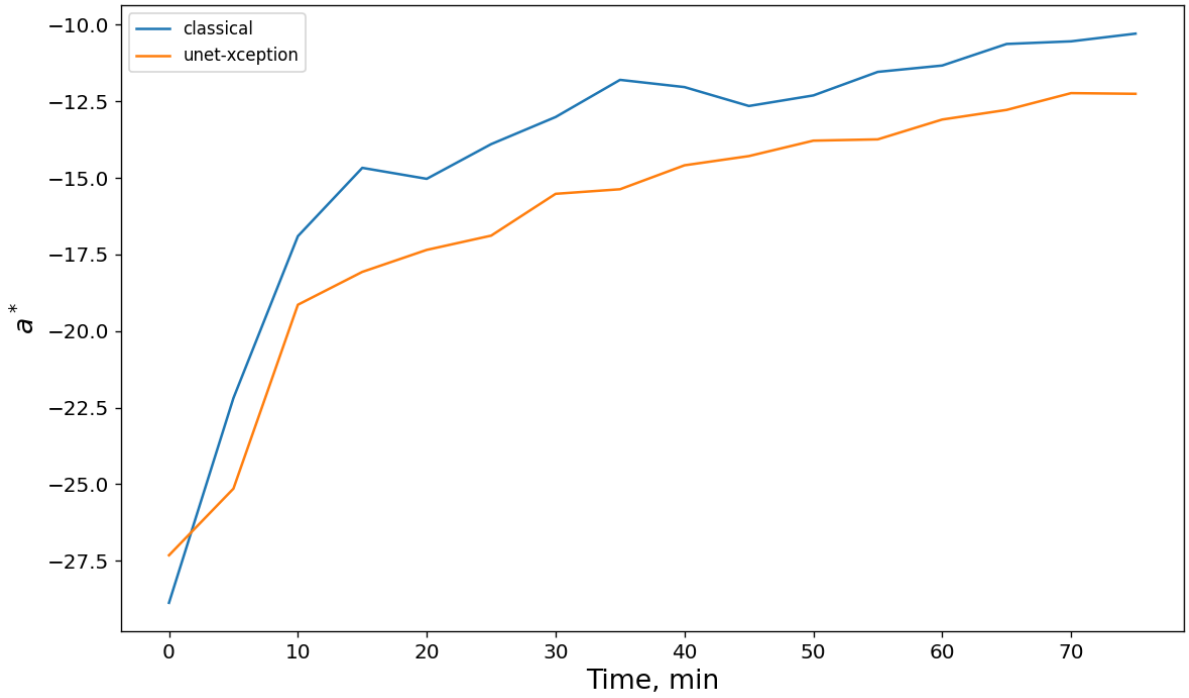


Figure 4.22 Real-time plot of a, produced by the classical pipeline and Unet-Xception model*

Figure 4.23 display plots of real-time changes in b*, produced by Unet-Xception, for peas dried at 50, 55 and 60^{o}C. For all the drying temperatures, the yellowish index b* of the peas decreased as the peas dried. The reduction in b* index could be related to Maillard browning reaction and the formation of brown pigment as drying progressed. The reduction in yellowish index may also be attributed to the concentration of β-carotenoids in the peas. The concentration may have increased due to moisture loss and subsequent shrinkage of the peas. [96,97]

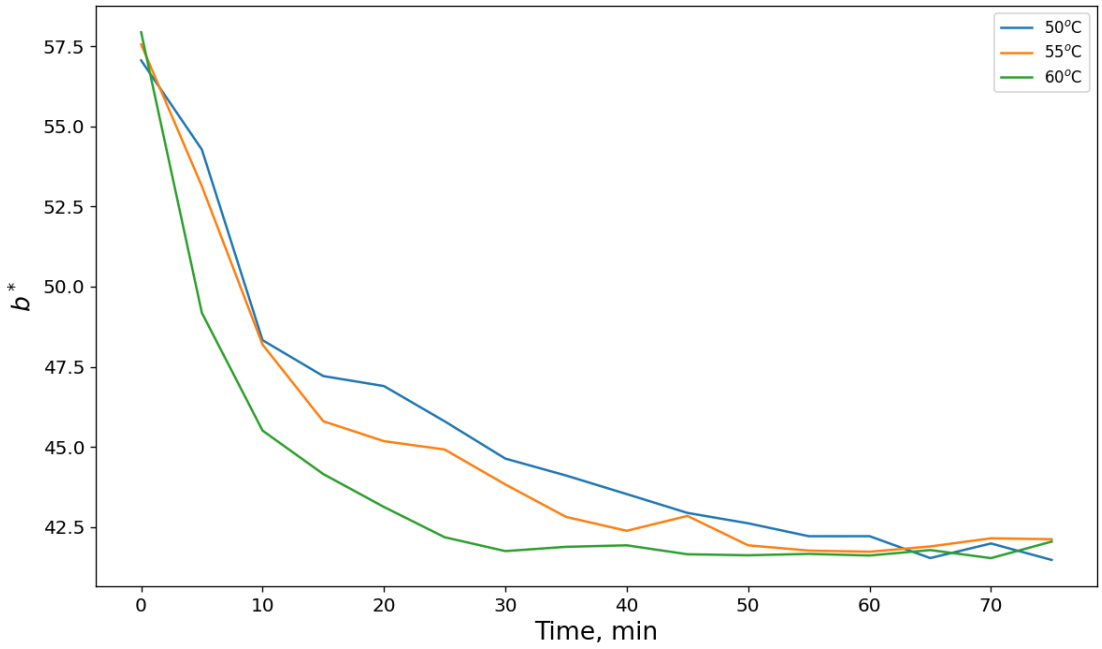


Figure 4.23 Real-time plot of b, predicted by Unet-Xception, for 50, 55 and 60°C drying runs.*

As shown in Figure 4.24, similar to the a* index trend, Unet-Xception produced smoother plot of b* than the classical model.

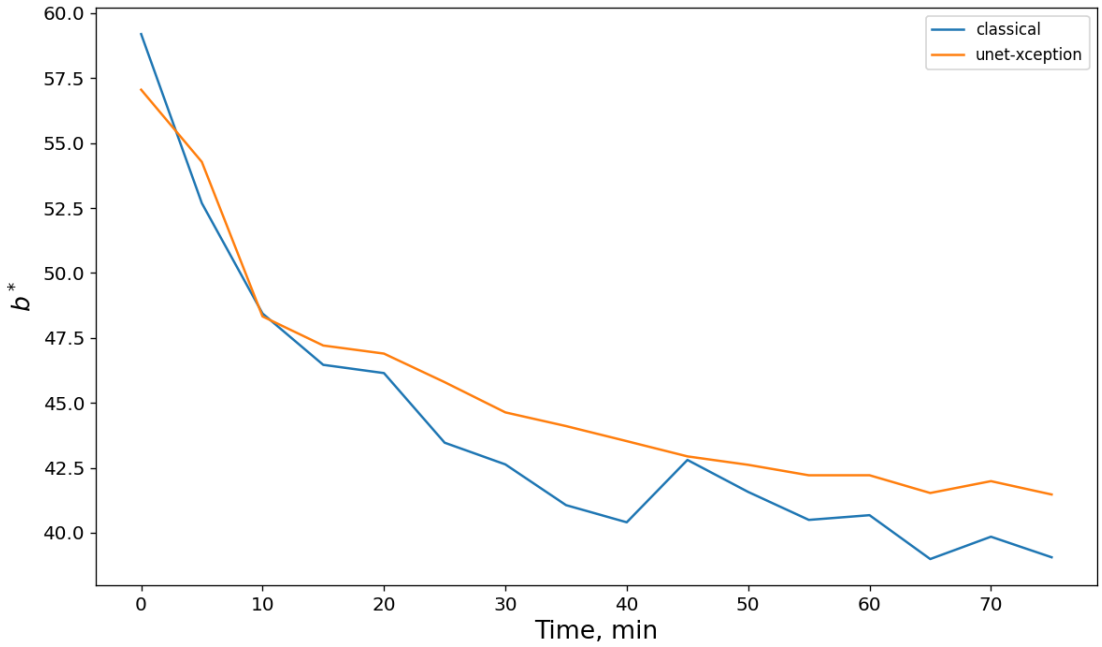


Figure 4.24 Real-time plot of b, produced by the classical pipeline and Unet-Xception model for 50°C drying run.*

The Lightness *L** represents the Luminance or the capacity of the food material to reflect light. [90,98] As show in Figure 4.25, for most part of the drying process the L* index trended downwards, along with moisture content. This reduction in L* meant the product turned darker as drying progressed. The same behaviour was observed by Zelinska et al. when drying green peas.[99]

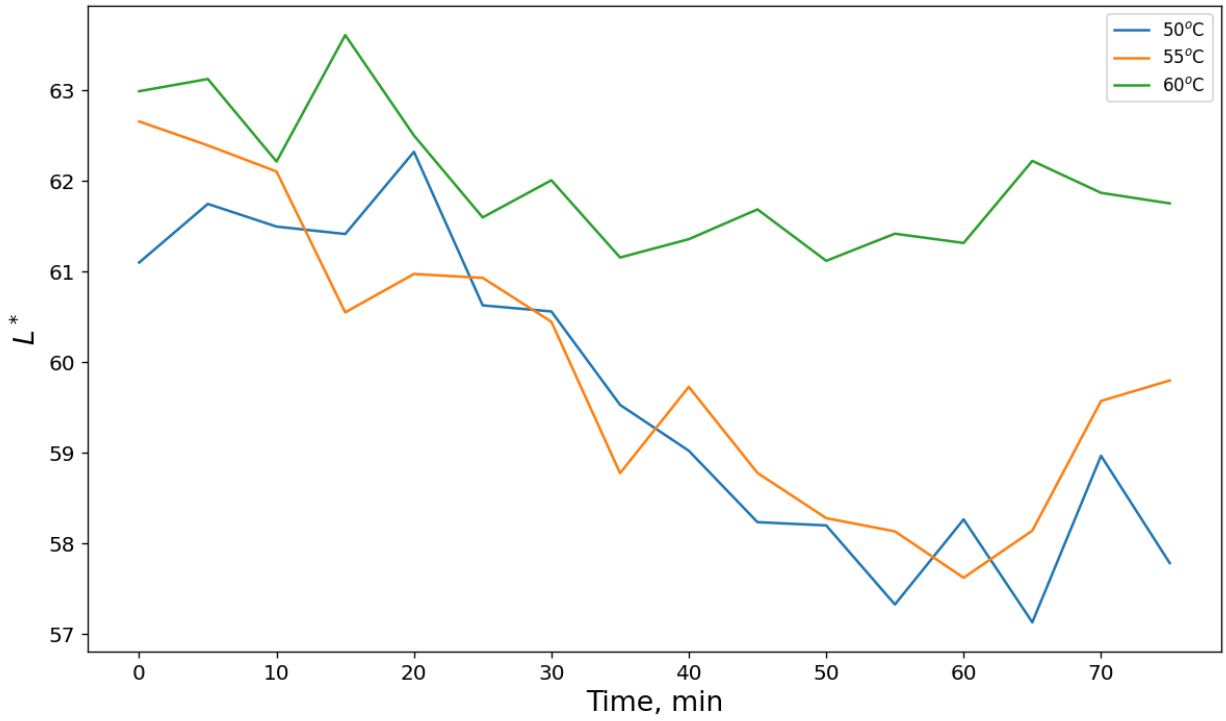


Figure 4.25 Real-time plot of L, predicted by Unet-Xception, for 50, 55 and 60°C drying runs.*

The reduction in *L** as drying progressed, could be attributed to degradation of chlorophyll into undesirable gray-brown compounds, such as pheophorbide or pheophytin.[99] The minor fluctuations in L* as it trends downwards, as well as the increase around the 65th minute, could be attributed to sample opacity resulting from structural shrinkage caused by moisture evaporation.[17,96]

*L** reduced more for green peas dried at 50 and 55°C, compared to those dried at 60°C; this outcome could be attributed to case hardening and significant moisture loss at the early stages of drying, for peas dried at 60°C. Due to case hardening, the peas dried at 60°C may have shrunk more uniformly compared to those dried at 50 and 55°C. As a result of uniform surface shrinkage, most regions on the surface of the peas were visible on the RGB image, hence both the lighter and

darker regions of the peas contributed to the overall L^* value that was obtained for peas dried at 60°C.

On the other hand, peas dried at 50 and 55°C shrunk more unevenly as they lost moisture, thus hiding some of the pea surfaces that were visible in real-time images captured at the earlier stages of drying. In all, Unet-Xception has shown to be more effective in monitoring real-time changes in L^* index. Furthermore, as shown in Figure 4.26, Unet-Xception produced smoother plots than the classical model.

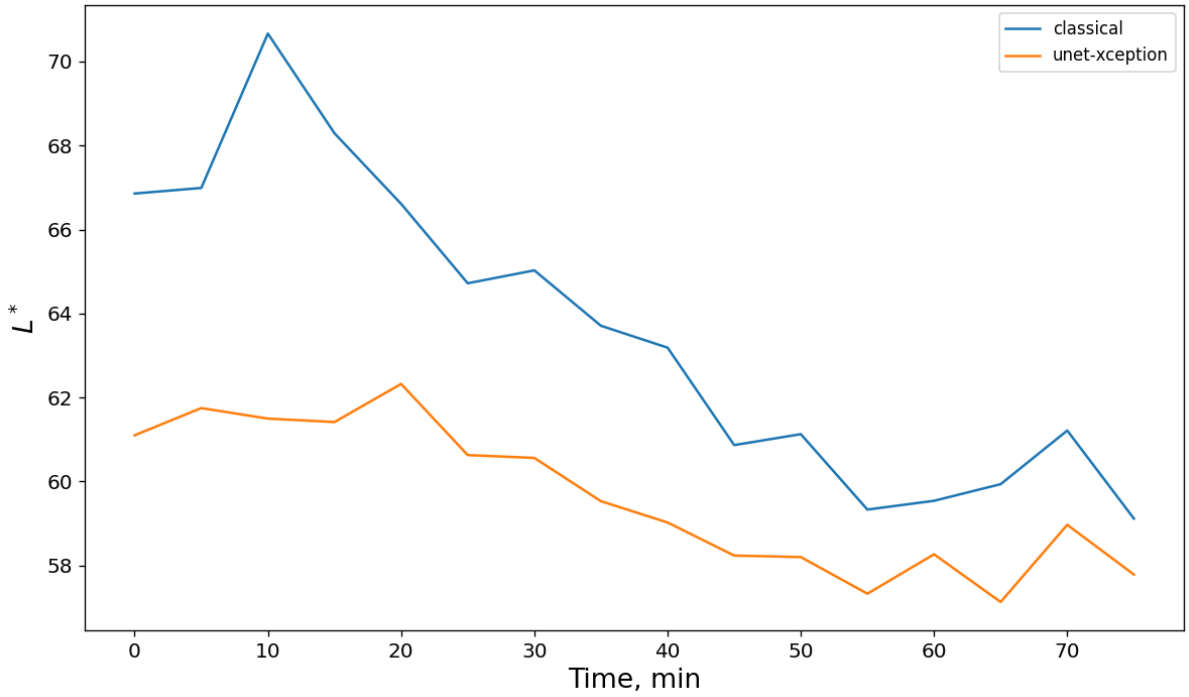


Figure 4.26 Real-time plot of L, produced by the classical pipeline and Unet-Xception model for 50°C drying run.*

4.4.2.2. Texture

Figure 4.27 displays a plot of real-time change in homogeneity, produced by Unet-Xception, for drying runs conducted at 50, 55, 60°C. As drying progress, the homogeneity value describes how rough the surface of the green peas has become. This roughness appears as disparity in pixel intensities across the surface of individual peas. The disparity in pixel intensity distribution increases as drying progresses. During drying, as the peas loose moisture, lighter and darker spots appear on the surface of the peas, the peas shrink, and sharp edges develop on the surface of the

peas. All these developments contribute to changes in homogeneity values. Among all the textural attributes that were predicted by the model, homogeneity produced the most meaningful trend.

As shown in Figure 4.27, for all air temperatures, textural homogeneity trended downwards throughout the drying process. A similar result was obtained by Fernandez et al. [11] when analyzing the effect of drying conditions on the shape, texture, and color of apples. In Figure 4.27, for peas dried at 60^{o}C, the major reduction in homogeneity occurred within the first 30 minutes, afterwards this visual property remained almost consistent till the end of drying.

For peas dried at 55^{o}C, the fastest drop in homogeneity occurred within the first 40 mins, while the homogeneity of the peas dried at 50^{o}C, trended downwards throughout the drying duration. For peas dried at 60^{o}C, the early significant drop in homogeneity, followed by almost consistent values, could be attributed to quick moisture evaporation, caused by high air temperature. This could also be responsible for the slower upward trend in greenish index, from the 30th minute onward, as previously shown in Figure 4.21, for peas dried at 60^{o}C.

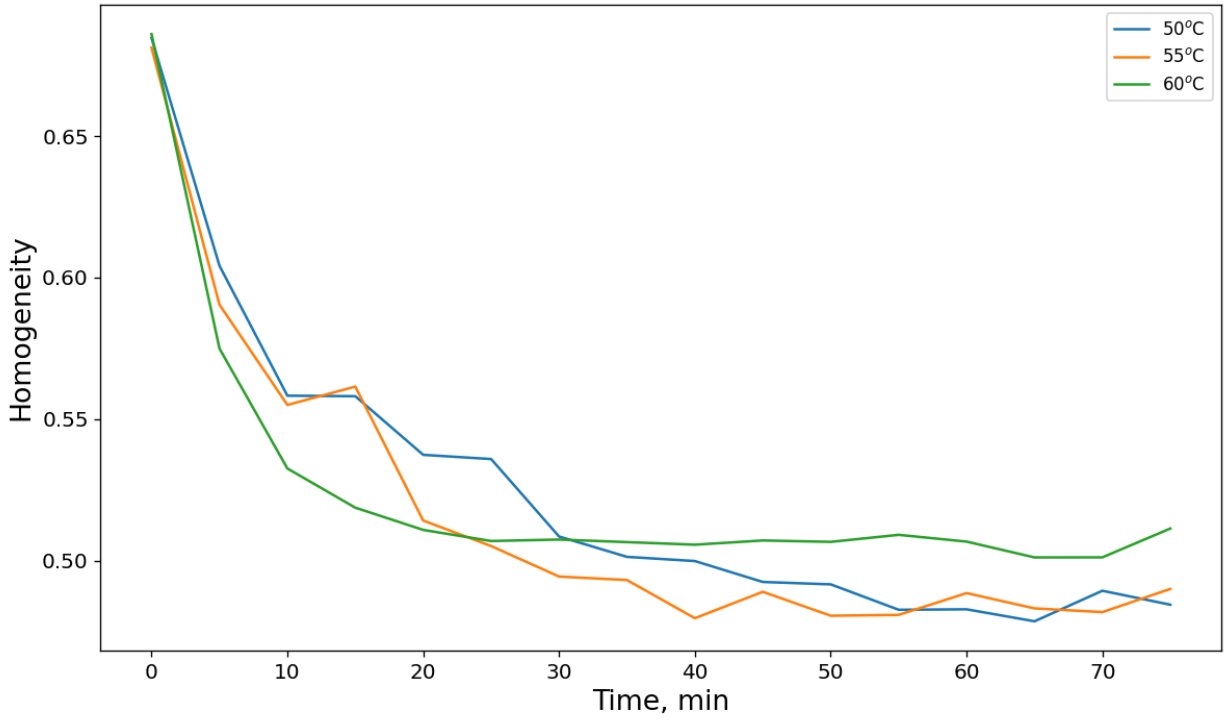


Figure 4.27 Real-time plot of homogeneity, predicted by Unet-Xception, for 50, 55 and 60^{o}C drying runs.

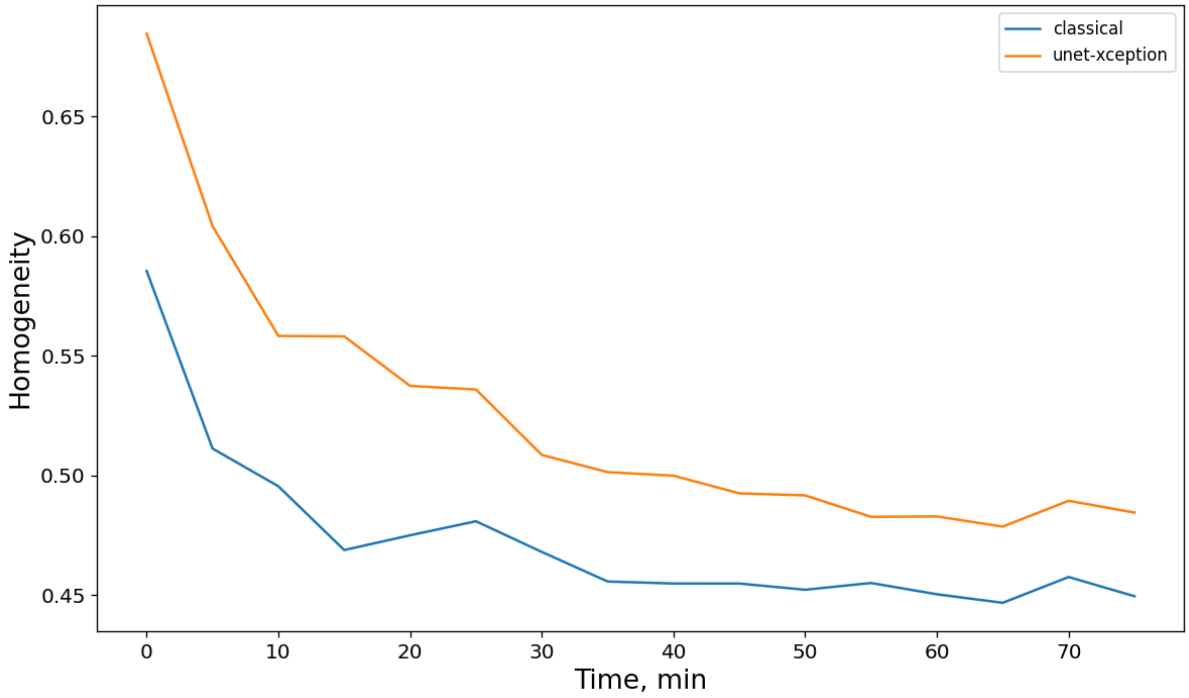


Figure 4.28 Real-time plot of homogeneity, produced by the classical pipeline and Unet-Xception model, for 50°C drying run.

4.4.2.3. Size

Figure 4.29 – Figure 4.36 contain plots of real-time changes in equivalent diameter, ferret diameter, filled area and perimeter. For each of the size-related attributes, when we observe the first plot that shows the real-time trend for drying experiments conducted at 50, 55 and 60°C, we would notice that for all the air temperatures, the size of the peas kept reducing as drying progressed.

On the other hand, while studying the second graph (Figure 4.30) that compares Unet-Xception to the classical model, we would notice that Unet-Xception also produced real-time trends that were smoother than those generated by the classical model.

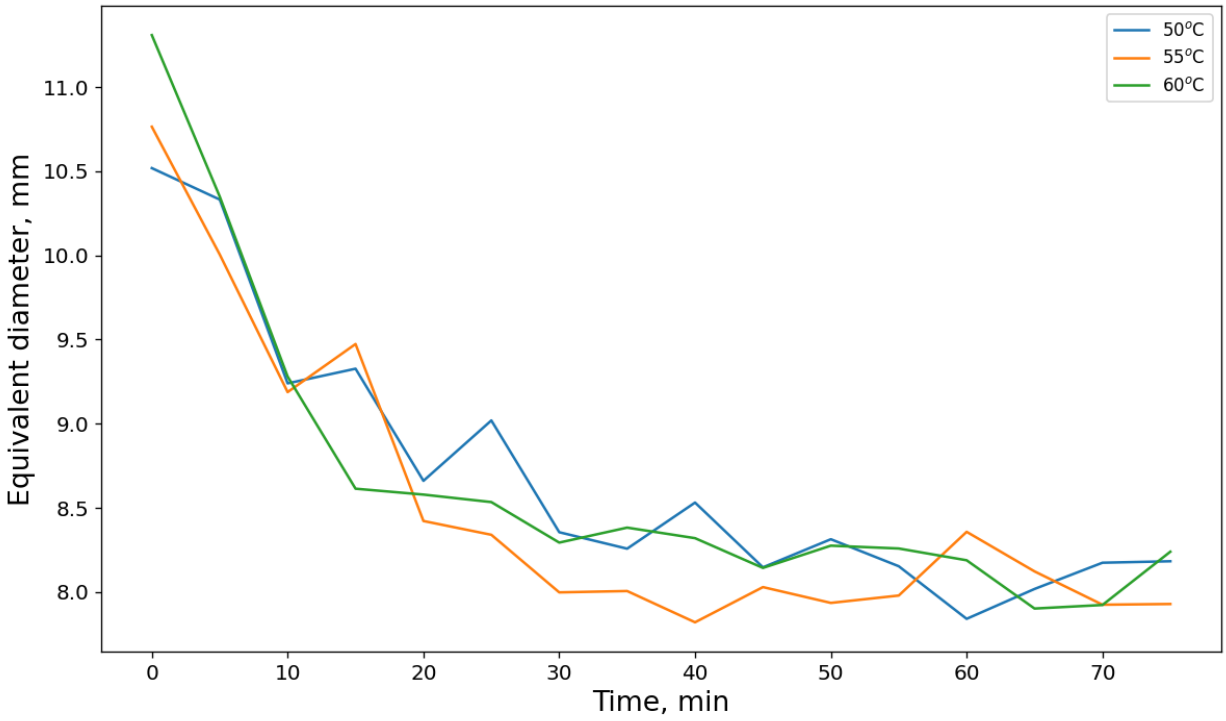


Figure 4.29 Real-time plot of homogeneity, predicted by Unet-Xception, for 50, 55 and 60°C drying runs.

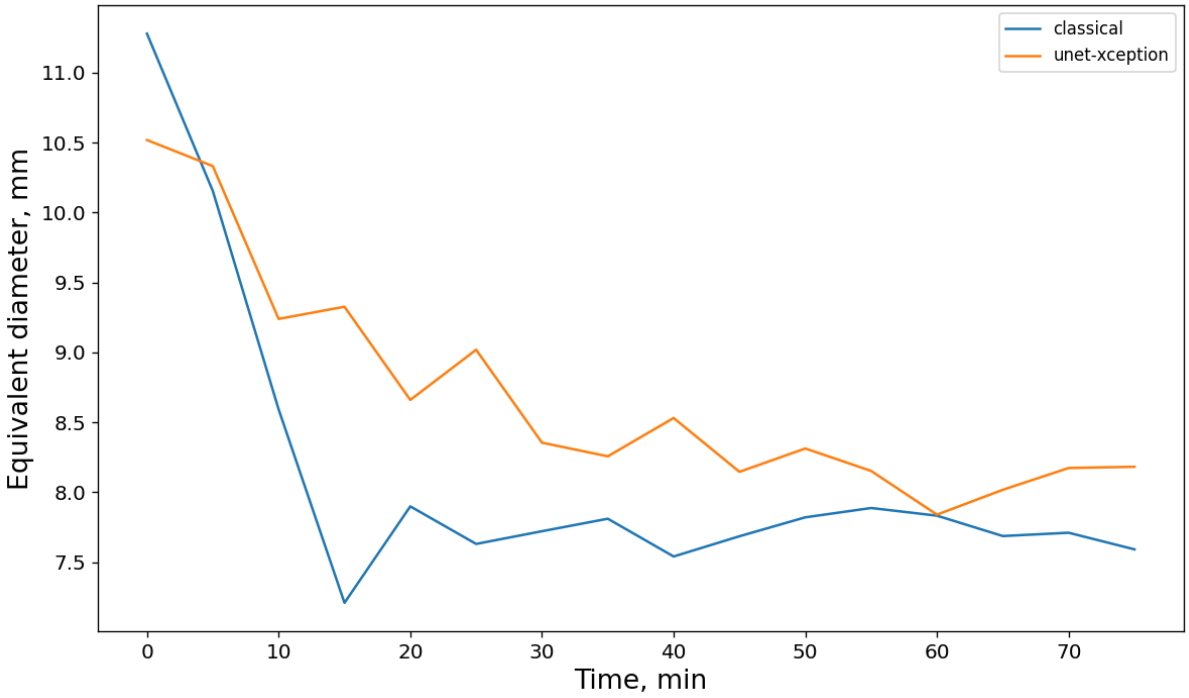


Figure 4.30 Real-time plot of equivalent diameter, produced by the classical pipeline and Unet-Xception model, for 50°C drying run.

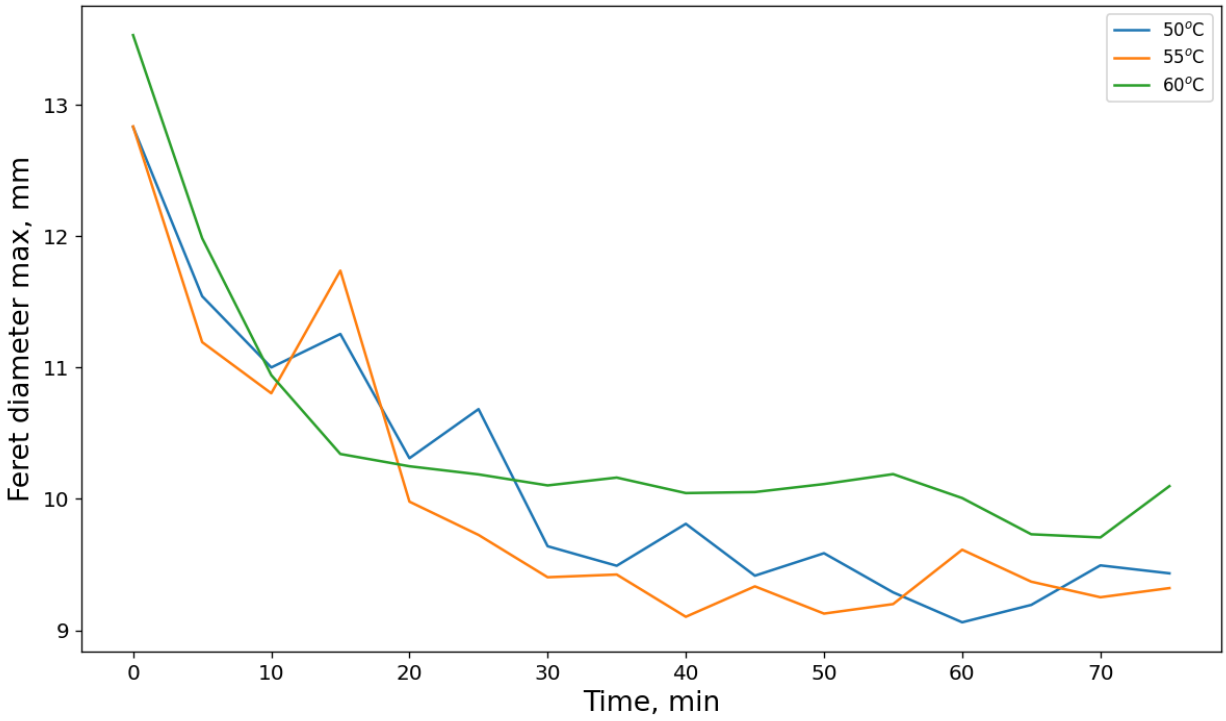


Figure 4.31 Real-time plot of predicted by Unet-Xception, for 50, 55 and 60°C drying runs.

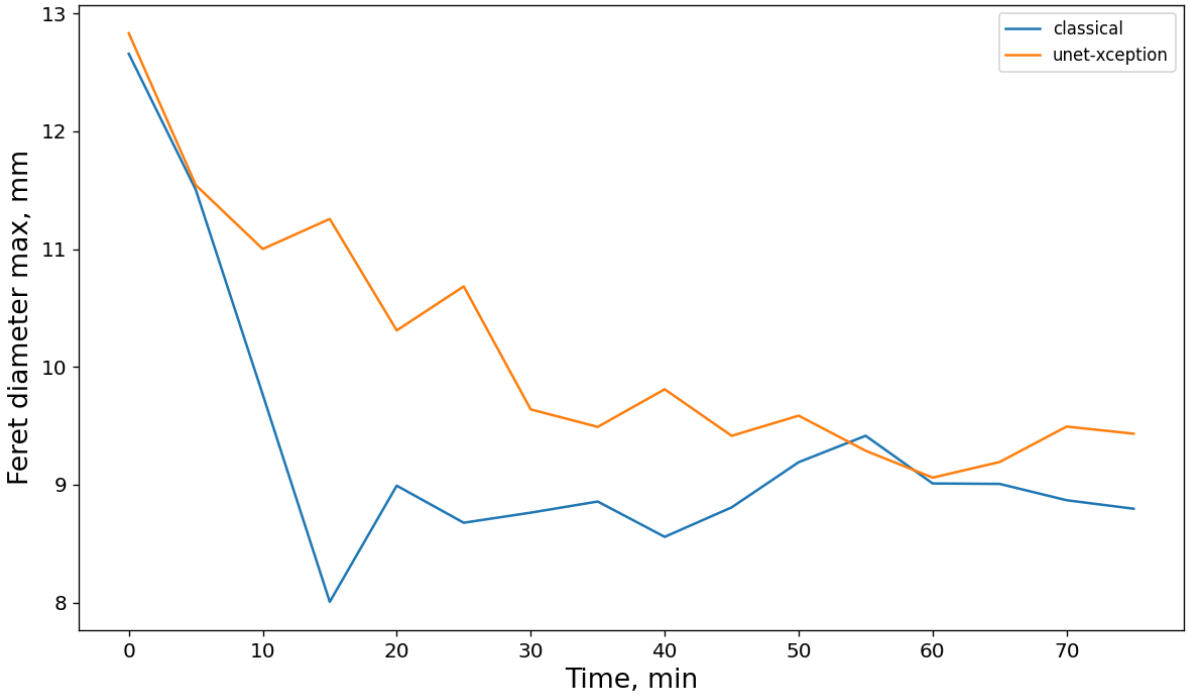


Figure 4.32 Real-time plot of ferret diameter, produced by the classical pipeline and Unet-Xception model, for 50°C drying run.

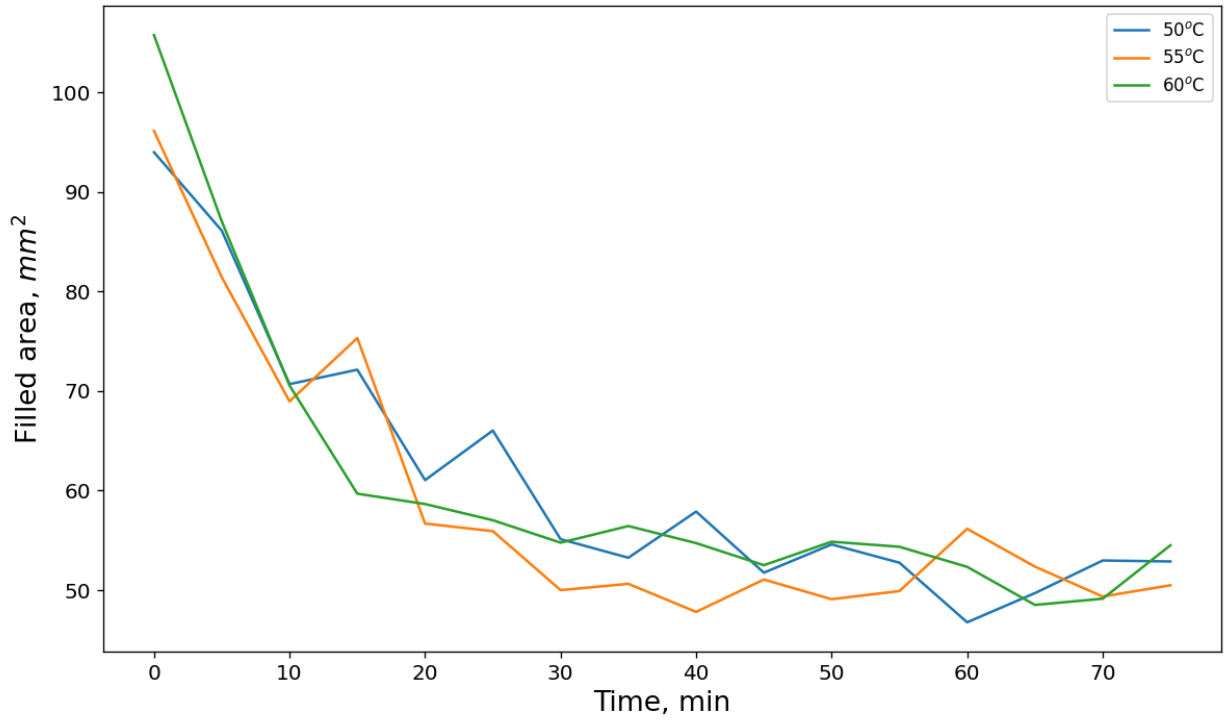


Figure 4.33 Real-time plot of filled area, predicted by Unet-Xception, for 50, 55 and 60°C drying runs.

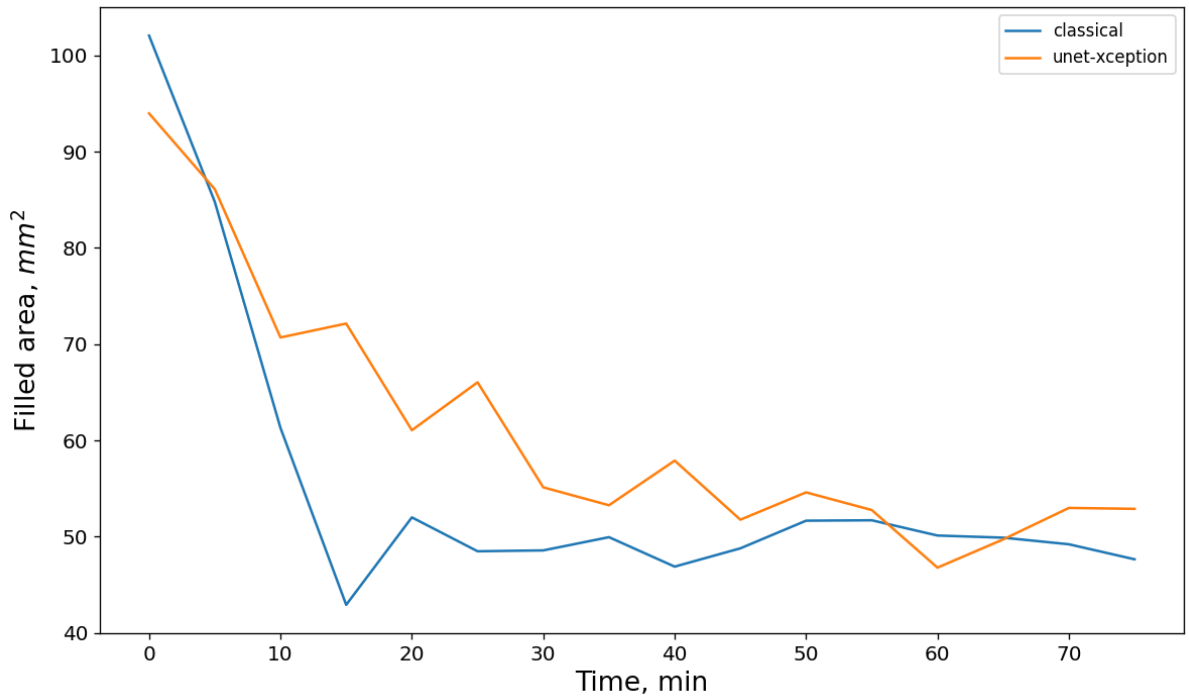


Figure 4.34 Real-time plot of filled area, produced by the classical pipeline and Unet-Xception model, for 50°C drying run.

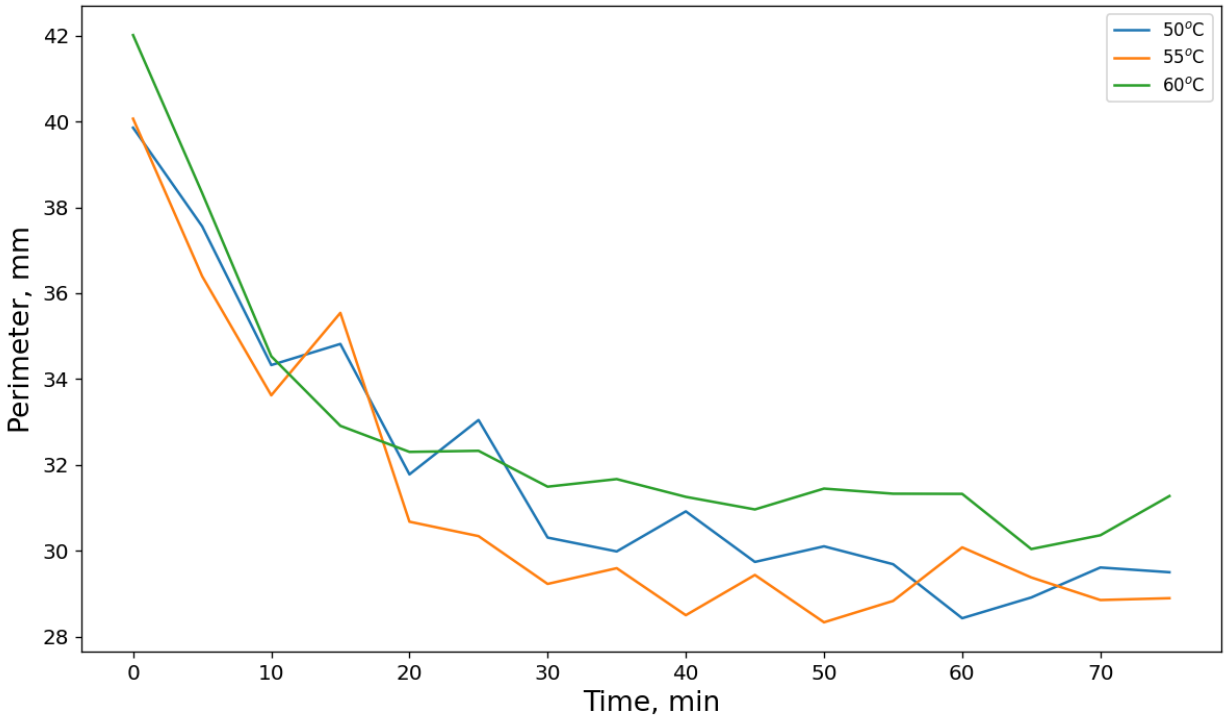


Figure 4.35 Real-time plot of perimeter, predicted by Unet-Xception, for 50, 55 and 60°C drying runs.

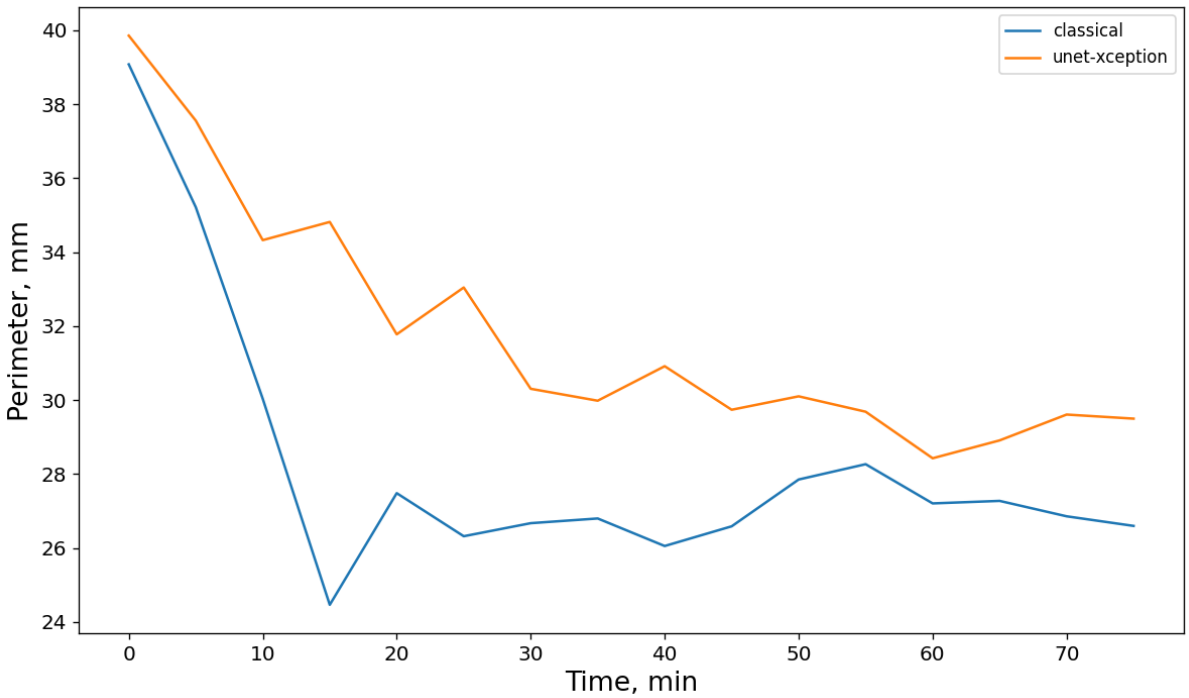


Figure 4.36 Real-time plot of perimeter, produced by the classical pipeline and Unet-Xception model, for 50°C drying run.

4.5. CONCLUSION

The main objective of this work was to develop an AI-powered solution for predicting visual attributes of peas during fluidized bed drying. Visual attributes were predicted from images of peas captured during the drying process. An integral part of the AI solution was Unet-Xception, a deep learning model, developed to solve the challenges encountered by a classical computer vision (CV) model. This CV model was previously used to predict visual attributes of peas, during fluidized bed drying.

The design of Unet-Xception and the classical CV model were based on the Two-pass framework, which involved image segmentation followed by visual attribute prediction. To predict visual attributes of food using this framework, first, the image containing the peas was segmented to produce an RGB image containing only infocus peas. Afterwards, visual attributes were predicted using the segmentation result.

While applying this two-stage process, the classical model encountered various challenges. Some of these challenges include the presence of overlapping infocus peas, variation in color and texture across the surface of individual peas, and clustering of uniformly coloured peas at the initial stages of drying. Due to these challenges, the classical model omitted too many infocus peas during segmentation, thus resulting in a low coefficient of determination when predicting most of the visual attributes.

The results of our research work show that Unet-Xception was able to learn important spatial and cross-channel relationships between the input image and infocus peas within it. Thus, the segmentation results from Unet-Xception included most of the infocus peas that were within the ground truth mask. Additionally, Unet-Xception also detected extra infocus peas that were missing in the ground truth mask. The detection of more infocus peas by Unet-Xception increased the MIoU to 0.9464, compared to an MIoU of 0.8190 obtained by the classical model. The presence of additional infocus peas within the predicted mask meant that if U-Net was retrained using updated ground truth masks, the MIoU may increase further.

Upon merging the trained U-Net to a trained Xception model, the combined model (Unet-Xception) achieved significantly higher R^2 values than those obtained when the classical model was used to predict visual attributes. Since coefficient of determination is a measure of how well the predicted values from the model mirror the growth truth values, we expected the R^2 values for

Unet-Xception to be slightly dampened, due to the presence of additional infocus peas within the predicted mask. However, notwithstanding the dampening effect, when Unet-Xception was used to monitor new drying experiments conducted at 50, 55 and 60^{o}C, it produced a much smoother real-time trend of visual attributes.

In our introduction, we asked the question ***"is it possible to monitor real-time changes in visual attributes of food during fluidized bed drying, using images of the food captured during the drying process?"*** The result of our study shows that, for the first time, it is possible to achieve this goal using Unet-Xception. *a* and b ** indices were the best indicators for monitoring color. Homogeneity was the best for surface texture, while size was measured using the equivalent diameter, Feret diameter, filled area and perimeter. The real-time feedback from this AI-driven solution could be fed to a control system, to modulate drying conditions, thus we would end up with a drying system that consistently produces dry food that is visually appealing and has a longer shelf life. Such an intelligent dryer could significantly reduce global food loss and waste.

4.5.1. Future work

After developing Unet-Xception using the Two-pass framework, a new research question sprung up – ***"is it possible to predict visual attributes of food, during fluidized drying, without segmenting the real-time image?"*** Meaning, can we develop a deep learning model that bypasses the segmentation step. Once the image is captured, it is fed to the model, and the values of visual attributes are produced at the other end of the model, without the image going through the segmentation step. Developing such a model could significantly reduce the number of model weights that need to be trained. Also, the computational resources during deployment could be less than what is required for the two-stage counterpart Unet-Xception. Therefore, our future work would focus on developing a single-pass model for predicting visual attributes of food during fluidized bed drying.

CONNECTING TEXT

In the previous chapter, we developed an improved model known as Unet-Xception. With this model, we have achieved the first and second milestones needed to improve the state-of-the-art in monitoring visual attributes of food, in real-time, during fluidized bed drying.

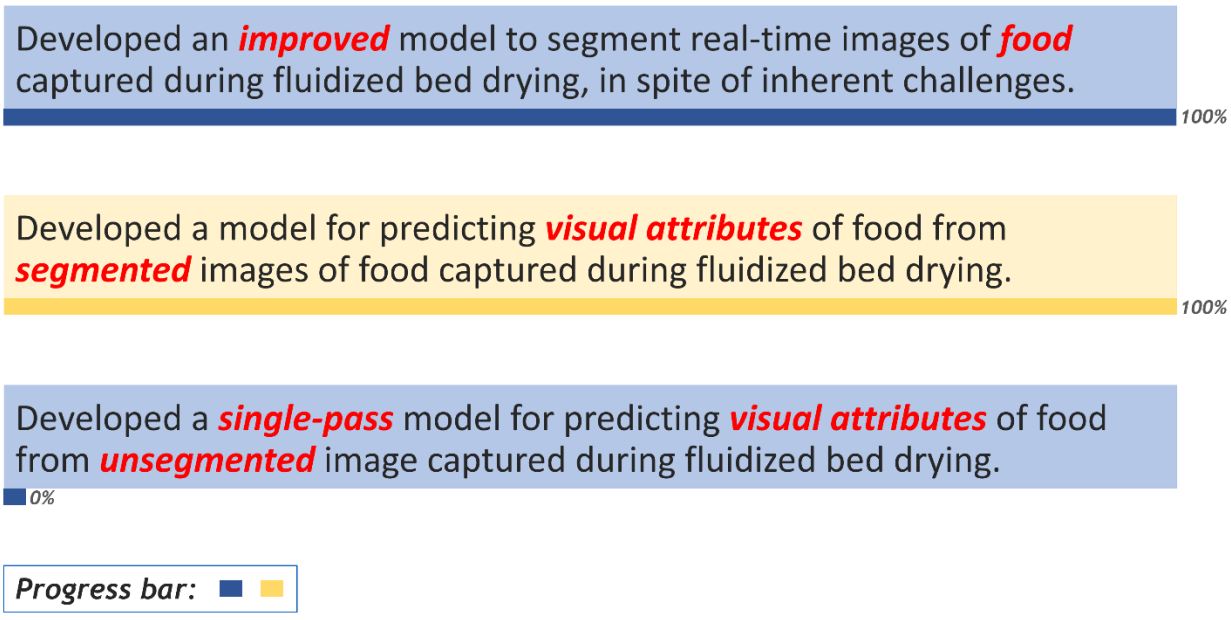


Figure 4.37 Milestones to improve the state-of-the-art

Unet-Xception significantly outperformed the classical CV pipeline in terms of image segmentation and visual attribute prediction. Hence, let us now develop the single-pass solution.

CHAPTER V

AN INTELLIGENT SINGLE-PASS SYSTEM FOR MONITORING VISUAL ATTRIBUTES OF FOOD DURING FLUIDIZED BED DRYING

5.1. ABSTRACT

In this study, we developed a single-pass deep-learning solution for monitoring visual attributes of green peas in real-time, during fluidized bed drying. Such an AI-driven solution could significantly reduce global food waste by ensuring consistent product quality during food drying. The deep learning model was trained to monitor color, texture, and size of peas. Color was monitored by predicting the CIE *L*, a* and b** indices. Texture was measured by predicting contrast, correlation, entropy, and homogeneity, while equivalent diameter, Feret diameter, filled area, perimeter, and roundness were used to monitor the product size. The single-pass model predicted visual attributes from real-time images of peas, without segmenting the image. The images were captured during fluidized bed drying runs conducted at 50, 55 and 60^{o}C. The performance of this single-pass model was compared to that of a classical computer vision model and a two-pass model that predicted visual attributes of peas from segmented outputs. The single-pass solution significantly outperformed the classical model, while attaining an almost similar coefficient of determination (R^2) as the two-pass model. The R^2 values achieved by each model are reported in the form [classical model, two-pass, single-pass]. a* and b* were the best indicators of color, with R^2 values of [0.7719, 0.8851, 0.8675] and [0.8694, 0.8981, 0.9064], respectively. Texture was best monitored by Homogeneity, with R^2 values of [-0.8660, 0.8017, 0.7845]. While the R^2 for the four product size parameters were [-1.0278, 0.7926, 0.8748] for equivalent diameter, [-1.6417, 0.8482, 0.8802] for ferret diameter, [-0.3663, 0.8391, 0.8901] for filled area, and [-1.5425, 0.825, 0.8421] for perimeter. The results of this study show that it is possible to monitor realtime changes in visual attributes of food, during fluidized bed drying, without segmenting the image of the food. Integrating this novel and adaptable AI-driven solution to food dryers could bring about consistent product quality, and by extension significantly reduce global food losses and waste.

5.2. INTRODUCTION

The color, texture, size, and shape of dried foods are important visual parameters that attract or dissuade consumers from purchasing dry products. If the product looks good, consumers pick it up from the shopping shelf and pay for it. On the other hand, if the dry food is unattractive, consumers walk away from it, and the rejected product adds to the enormous volume of food wasted annually. According to the Food and Agriculture Organization of the United Nations (FAO), 1/3 of the food produced annually is wasted or lost.[1-5,24,25] To reduce the volume of dry food that is wasted due to unattractive visual appearance, it is vital that we develop dryers that can control real-time changes in moisture content and visual appearance of the food. With such drying system, we would consistently end up with products that have long shelf life and are attractive to consumers.

Ironically, the control system for most food dryers is designed to achieve only the desired final moisture content. Hence, we are not in control of visual appearance of the final product, thereby leading to increased food loss and waste. To solve this problem, we need drying systems that control real-time changes in moisture content and visual attributes. However, such control systems cannot work in isolation. To regulate the drying conditions optimally, the control system would need real-time feedback on the moisture content and visual attributes of the product.

There have been a lot of breakthroughs in convective drying.[10,14-17,20,21,52-54] Researchers have developed various Computer Vision (CV) systems for monitoring the color, texture, size, and shape of food, in real-time, during convective drying. The CV system monitors visual attributes from real-time images of the food captured during the drying process. Some other researchers have even gone a step further to develop control systems that feed on real-time data generated by the CV systems.[11-13,85]

The CV system comprises broadly of two main parts – an image acquisition system and an image processing system. The first part captures illuminated images of the food in realtime, while the second part processes the image to obtain visual attributes of the food. The image processing step is usually conducted using a method we refer to as the *Two-pass* framework. As depicted in Figure 5.1, this framework comprises of two steps – image segmentation followed by visual attribute prediction. Once the image is captured by the acquisition system, it is fed to the image processing unit for image segmentation.

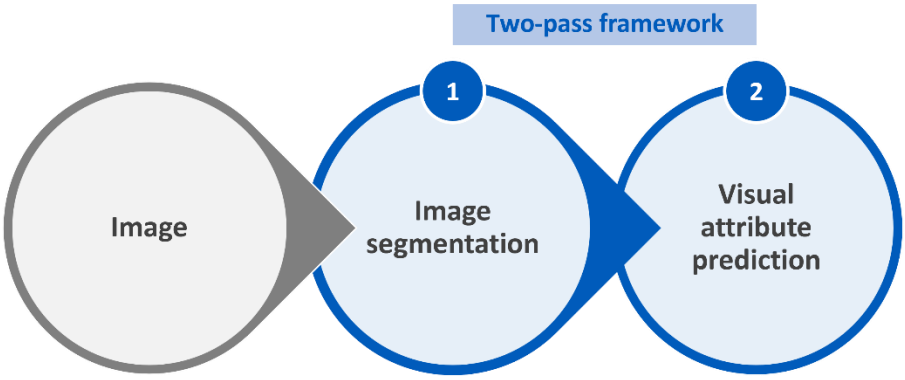


Figure 5.1 The Two-pass framework

The product of segmentation is a binary image containing 0s and 1s. The 1s represent the regions within the original image that are occupied by the food material, also known as the object of interest. While the 0s represent the background, that is, the regions within the original image that are not occupied by the food.

After segmentation, the RGB image and the binary mask are used to compute the average visual attributes of food materials within the original image. The visual attributes of the food in each image are generated by computing the color, texture, size, and shape of the objects of interest. In line with the two-pass framework, researchers have been able to monitor real-time changes in visual parameters of various foods during convective drying. Some of the investigated products include potato slices,[15,16] shrimps, [17,21,54] apple slices,[14,20] blueberry,[10] and bell pepper, [52].

To the best of our knowledge, up until 2019, there was no published work on monitoring visual attributes of food in real-time, during fluidized bed drying. Since then, Iheonye et al.[26], developed a novel classical CV system for monitoring real-time shrinkage of green peas during fluidized bed drying.[26] Their system was also based on the two-pass framework. The image acquisition system was designed to account for unequal distance of the individual peas facing the camera lens. The disparity in individual product distance is unique to images captured during fluidized bed drying. Since the peas were not at the same distance from the camera, Iheonye et al. built their hypothesis on the premise that visual attributes could be computed using a certain group of peas in the image referred to as ***"infocus peas"***.

According to the authors, during fluidization, as the peas randomly jumped up and down within the drying chamber, some peas that touch the wall of the dryer facing the camera lens. These

unique set of peas that touch the drying chamber were referred to as infocus peas, while the other peas behind infocus ones were called the ***defocused peas***. Figure 5.2, depicts a fluidized bed dryer containing infocus peas and defocused peas. Furthermore, during fluidization, as the peas jump up and down within the drying chamber, the peas would also be mixing, hence, each time a new image is captured, a unique set of infocus peas would be displayed on the image. Therefore, if numerous images were captured at a given instance during the drying process, the mean visual attribute of the infocus peas on all the images would be equal to the mean visual attribute of all the peas in the dryer, at that instance. If this process is repeated at short intervals throughout the drying process, we could monitor real-time changes in visual attributes of the peas.

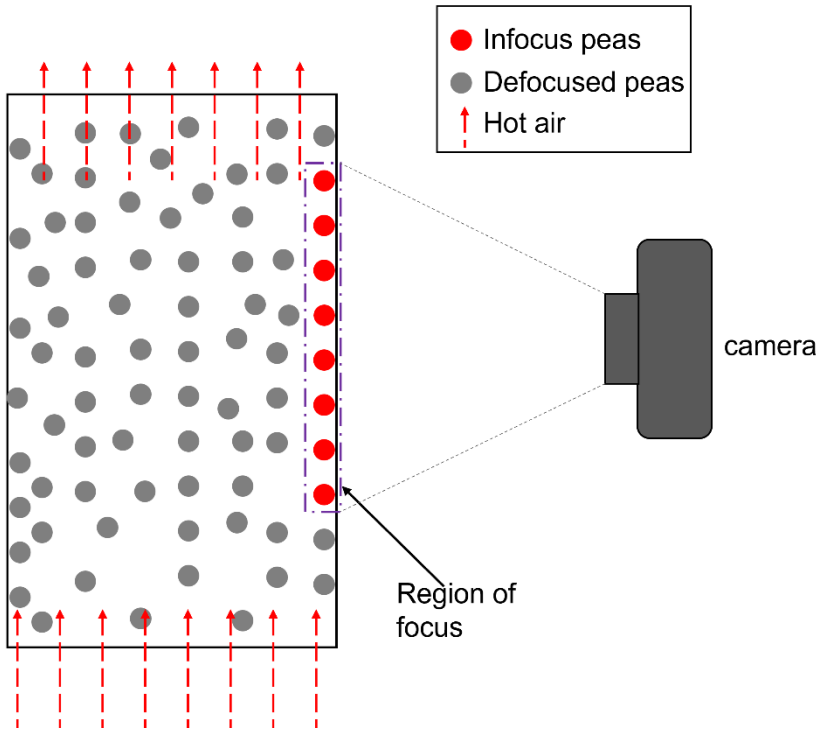
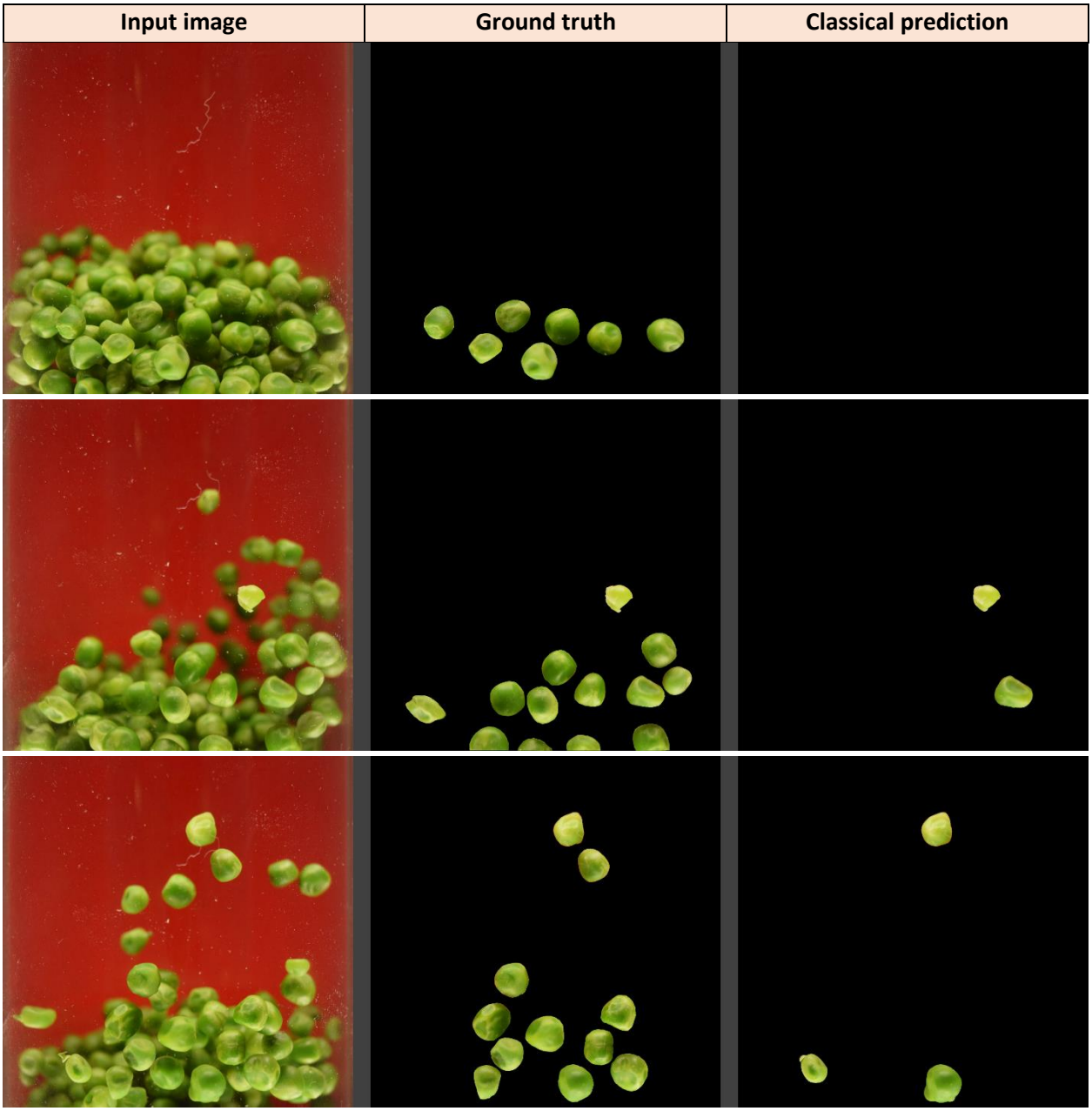


Figure 5.2 Image depicting infocus peas and defocused peas

Hence, to monitor real-time changes in visual attributes during fluidized bed drying, Iheonye et al. captured ten consecutive images every 15 minutes. During their research work, they conducted fluidized drying runs at 50, 55 and 60^{o}C. Using classical CV techniques, they developed a segmentation pipeline, designed to overcome challenges unique to images captured during fluidized bed drying. If these issues are not tackled, it becomes difficult and almost impossible to properly segment images captured during fluidized bed drying. Some of the issues include

overlapping objects of interest, clustering of uniformly coloured objects of interest, especially during the onset of drying, and non-uniform color and texture across the surface of individual objects of interest as drying progressed.

Their segmentation pipeline mostly succeeded in segmenting image regions where infocus peas were not clustered. However, it ran into problems when the infocus peas were overlapped, were clustered together during the initial stages of drying, or when the individual peas within the cluster peas had varying colour and texture across their surface. Figure 5.3 shows some examples.



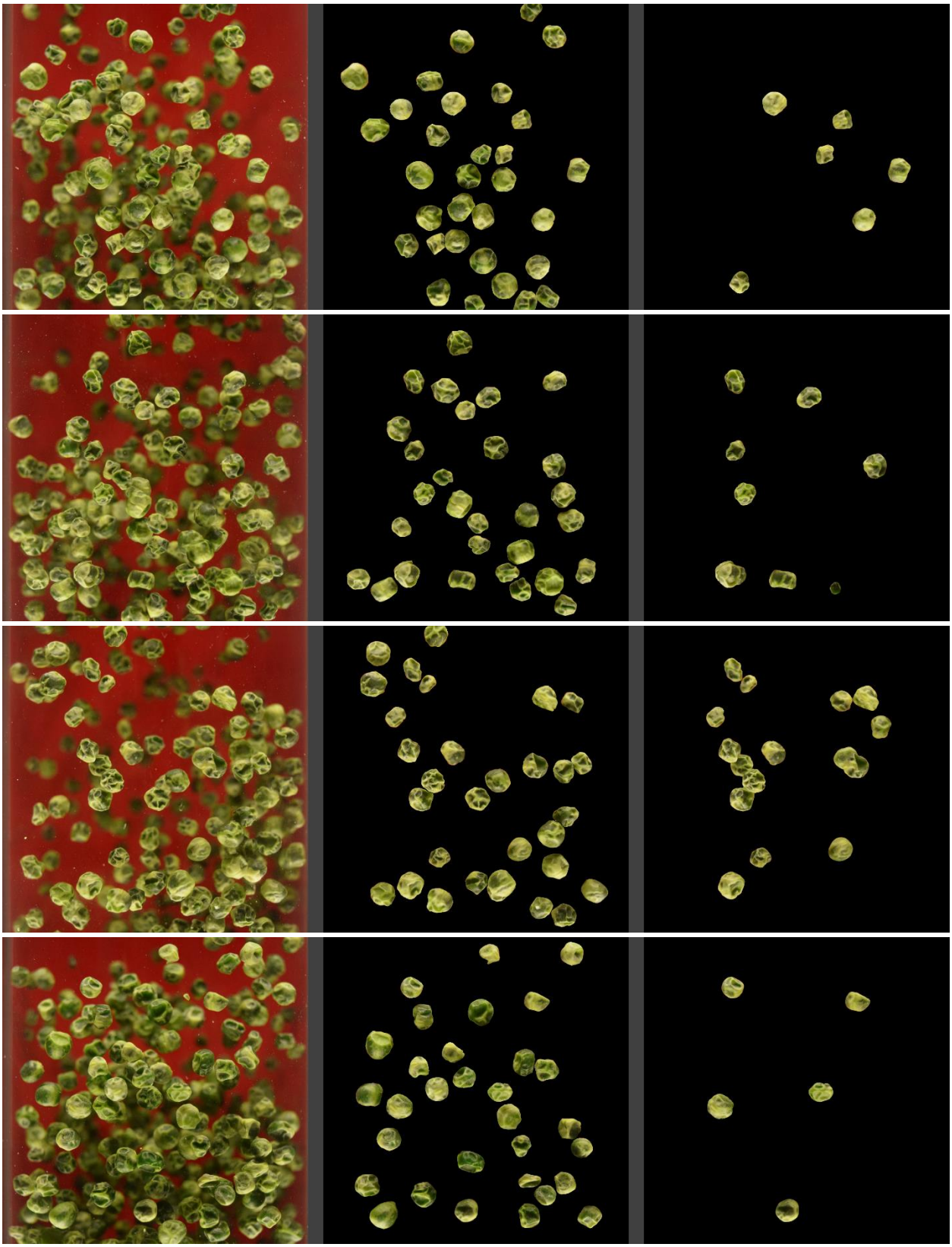


Figure 5.3 Original image, ground truth mask and segmented result using the classical CV pipeline.

Due to limitations of the Classical CV technique, the segmentation pipeline omitted some of the infocus peas, thus degrading the quality of information needed to compute visual attributes. The pipeline achieved a Mean Intersection-over-Union (MIoU) of 0.814. The coefficient of determination for predicting various attributes are shown in Table 5.1 below.

Table 5.1 R^2 values for classical, Xception, and Unet-Xception models in predicting twelve visual attributes

visual parameter		R^2
Color	L*	-0.8660
	a*	0.7719
	b*	0.8694
Texture	Contrast	-3.5691
	Correlation	-7.7618
	Entropy	-1.4961
	Homogeneity	0.7987
	Equivalent diameter	-1.0278
Size	Ferret diameter	-1.6417
	Filled area	-0.3663
	Perimeter	-1.5425
	Roundness	-8.0029

To increase the R^2 values, the segmentation result had to be improved. Hence, in Chapter 4, we developed an AI-driven solution that was also modelled along the Two-pass framework, comprising of image segmentation followed by visual attribute prediction.

Image segmentation was conducted using a trained U-Net model, while visual attributes were predicted by a deep learning model known as Xception. The combined model was referred to as "*Unet-Xception*". Once an image was fed to U-Net, instead of producing only a binary mask, as the classical pipeline did, Unet-Xception utilized the predicted binary mask from U-Net and the input image to generate a new RGB image containing only infocus peas. This image of infocus peas was then fed to the Xception model for visual attribute prediction.

Unet-Xception produced a higher MIoU of 0.9464, as well as improved coefficients of determination for predicting visual attributes. The R^2 values for the classical pipeline and Unet-Xception model are shown in Table 5.2.

Table 5.2 R^2 values for classical and Unet-Xception models in predicting eight visual attributes

visual parameter	classical	Unet-Xception
L*	-0.8660	0.8017
a*	0.7719	0.8851
b*	0.8694	0.8981
Homogeneity	0.7987	0.7947
Equivalent diameter	-1.0278	0.8082
Ferret diameter	-1.6417	0.8535
Filled area	-0.3663	0.8627
Perimeter	-1.5425	0.8152

Unet-Xception produced a smoother plot of real-time changes in visual attributes. By the end of the research, having developed a solution that significantly improved segmentation results and visual attribute prediction, a new research question sprung up:

"Is it possible to predict visual attributes of food, during fluidized bed drying, without segmenting the input image?"

Answering this question is the primary focus of this work. Hence, our research objective would be to develop a deep learning solution for monitoring visual attributes of green peas during fluidized bed drying, without segmenting the input images. This novel solution would monitor visual attributes in real-time as drying progresses. After developing this system, the prediction would be compared to those achieved by the classical pipeline and by Unet-Xception. That said, let us move on the next section where we would discuss our journey towards developing this ***Single-pass*** deep learning solution for monitoring visual attributes.

5.3. MATERIALS AND METHODS

Our discussion would be divided into three parts - image acquisition, image preprocessing and visual attribute prediction. When combined, these three sub-divisions make up the proposed Single-pass or end-to-end system for monitoring real-time changes in visual attributes during fluidized bed drying. Let us start with image acquisition.

5.3.1. Image Acquisition

Real-time images of green peas were captured using the same image acquisition system that was developed by Iheonye et al.[26] This was the same acquisition system that was used for developing

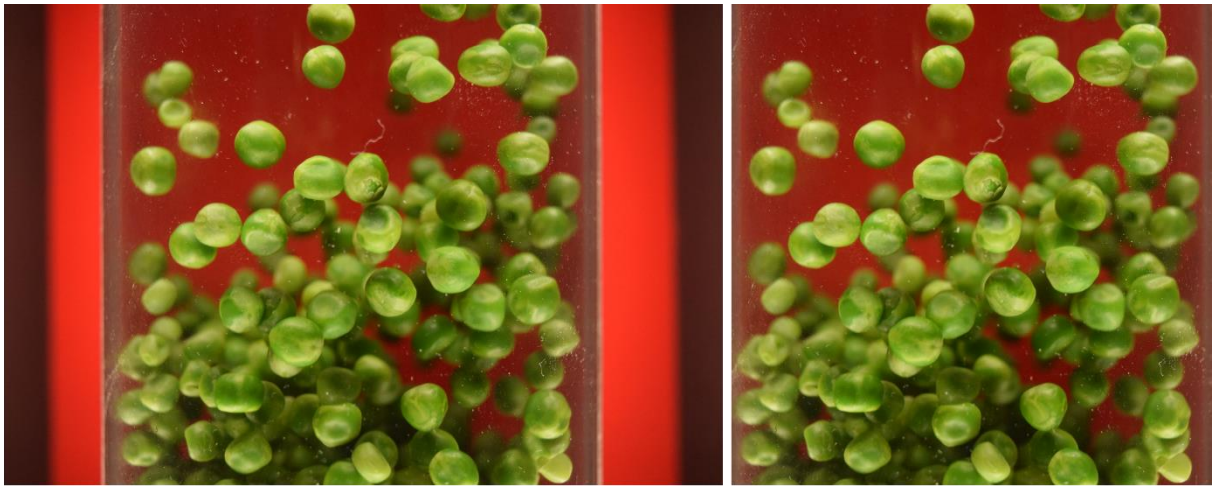
Unet-Xception, the two-pass deep learning solution that was described in section 4.3.4. A detailed description of the acquisition system can be found in section 4.3.1, as well as in the original article by Iheonye et al. [26]

Figure 5.4 Image acquisition system for fluidized bed drying[26]

The images that were used in this research work were the same images used to train Unet-Xception and develop the classical CV pipeline. These images of green peas were captured during fluidized bed drying experiment conducted at 60^{o}C. During the experiment, 10 images were captured every 15 minutes. By the end of the 75-minute drying run, 60 images were acquired. The original size of each image was 2000 x 2992 pixels. After acquiring the images, the next step was image preprocessing.

5.3.2. Image Preprocessing

The image preprocessing step comprised of five sub-steps, they include cropping, annotation, resizing, argumentation, and ground truth generation. Figure 5.5a below shows an example of the original 2000 x 2992 RGB image that was captured during the drying process.



a b

Figure 5.5 (a) Original image (b) cropped version

The region of the image that was important for this research was the area within the cylindrical drying chamber. Hence, the regions left and right of the cylinder were cropped out, as shown in Figure 5.5b. Since the width of the cylinder was 2000 pixels, cropping the regions outside the cylinder produced an equal sided image. An equal sided image would retain the shape and proportionality of the peas if the image had to be resized.

After cropping the images, the next step was to produce ground truth masks containing only infocus peas. The ground truth masks were produced manually using an image annotation tool known as LabelMe®- v5.0.1.[78] These masks were later used to compute the ground truth values of visual properties, for each real-time image. The values were latter used to obtain the cost per prediction when training the model.

Due to limitation in RAM size and the need to reshuffle the training set at the start of each epoch, it was not feasible to train the single-pass model using a 2000 x 2000 image, hence we moved on to the third preprocessing step - image resizing.

The images were resized to 1024 x 1024 pixels. Aside from system computing limitations, another reason why the image was resized was so that the model would be trained using the same size of image that Unet-Xception was trained on. Thereby, making it possible to compare the performance of both models. After resizing the images to 1024 x 1024, the dataset was split into training, validation, and testing sets.

Since the model was being trained to predict visual attributes, we ensured that when splitting the images, the visual attributes associated with the images were equally represented across the three datasets. Hence, the images were not split just based on when they were captured during the drying process, instead they were split using the stratified splitting method. The splitting parameter was the b* index.

Normally, after splitting the images into three datasets, the next step would have been to start training the model, however, 60 images would not be enough for training, validating, and testing a deep learning model. To generate sufficient realistic variants of the original images, we carried out a process known as image augmentation. During this process, realistic variants were generated by applying random transformations to each pair of image and corresponding mask. The transformations that were applied in this research include random crop, random up-down flip, random left-to-right flip, and random rotation. Since part of our objective was to monitor Color and texture, random transformations such as altering the brightness, saturation and contrast of the images were not part of the augmentation process. The augmentation pipeline was written in Python-v3.7. Some of the Python libraries that were used include Skimage-v0.18.0, Tensorflow-v2.7.0, and Panda-v1.40. Figure 5.6 displays a set of realistic variants that were obtained after augmentation.

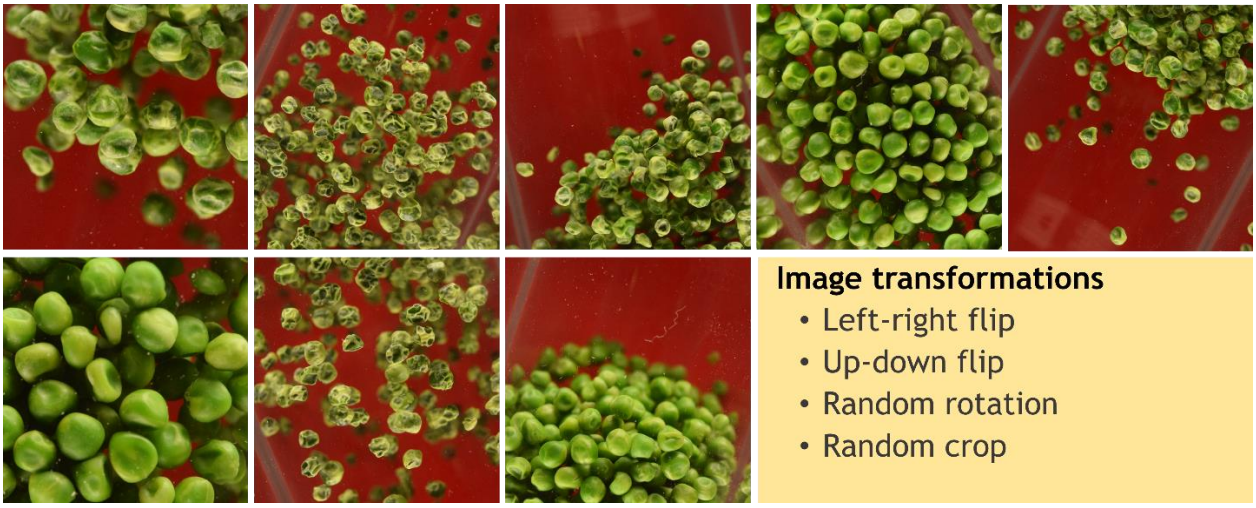


Figure 5.6 Realistic variants obtained after image augmentation

Augmentation was carried out separately on the training, validation, and test sets. The splitting was conducted before augmentation to ensure variants of an original image were not in more than one set. By so doing, the model would not be tested or validated using an image that shares significant similarity with the dataset the model was trained on. By the end of augmentation, there were 2974, 905 and 732 images in the training, validation, and test sets, respectively. Afterwards, the model was trained to predict visual attributes.

5.3.3. Visual Attribute Prediction

During this research work, the deep learning model was trained to monitor three broad visual properties - color, texture, and size. Color was monitored by predicting the CIE *L* a* b** color values. Texture was monitored by predicting the following statistical pixel intensity related measurements: entropy, energy, homogeneity, contrast, correlation, and uniformity. While size was monitored by predicting equivalent diameter, eccentricity, ferret diameter, filled area, perimeter, and roundness. A detailed description of each visual attribute, as well as the equation for generating the ground truth values are provided in section 4.3.5. Next, let us describe Xception, the deep learning model that was trained to predict visual attributes.

5.3.3.1. Xception

At this point in the paper, one may ask *"Why did the authors of this paper use Xception? Why did they not use EfficientNet, Inception-v4, GoogleNet, ResNet101 or VGG-16?"* Just like Xception, the other five networks are all renowned image classification models. So, why did we use Xception?

Similar to the thought process that guided our development of Unet-Xception, when considering which model to use for this research, we considered three factors, namely the number of model parameters to be trained, the computational demand, and the overall performance of the model. Amongst all the models considered, Xception was the most favourable.

The feature that makes Xception stand out from the rest was the unique form of convolutional layers that make up the model. This unique kind of convolution is called ***Depthwise Separable Convolution***. Although this form of convolution has been sparsely used in the past, it was brought back to prominence by Francis Chollet in 2016, when he developed the state-of-the-art image classifier known as Xception. A detailed description of depthwise separable convolution was

presented in section 4.3.6. Within that section, we also explained how this convolutional layer reduces the number of weights within Xception, as well as how it improves the effectiveness of the model in learning spatial and cross channels relationships between the input image and the expected output. So far, depthwise separable convolutional layers have been shown to outperform regular convolutional layers. Let us now discuss how Xception was trained to predict visual attributes.

5.3.3.2. Model hyperparameters and training

The Xception model was trained using images of size 1024 x 1024 pixels. Back-propagation was conducted using the Adam optimization algorithm. Momentum decay hyperparameters β_1 and β_2 values were set to 0.9 and 0.999, respectively. Epsilon was set to 1×10^{-7} for numerical stability.

The model loss was computed using Mean Squared Error (MSE) and Mean Absolute Error (MAE). Mean absolute error was used to calculate model loss in predicting correlation and contrast, due to the presence of outliers within their dataset. While the loss for other visual parameters were calculated using Mean Squared Error. At the end of each epoch, the model was evaluated on the validation set, by computing the MAE and Coefficient of Determination (R^2).

There were three rounds of training. During the first round, the model was trained to predict all the twelve visual attributes. The initial weights were those of an Xception model trained on the ImageNet dataset. These weights were obtained from the TensorFlow deep learning library. The Xception model was trained for 85 epochs, using an image size of 1024 x 1024, a batch size of one, 2994 steps per epoch, and learning rate that ranged from 3e-5 to 3e-7. Dropout and L2 regularization were applied to avoid overfitting. The dropout rate for the visual parameters ranged from 0.1 - 0.5, while the regularization factors ranged from 0.0001 - 0.1.

By the end of the first round of training the overall performance of the model was reduced due to the low coefficients of determination when predicting contrast, correlation, entropy, and roundness. Hence, for the second round of training, these four visual attributes were eliminated from the list of predicted outputs. The same hyperparameters were applied for first and second training rounds. After the second round of training, the original training set and validation set were combined to form a new set of training examples. This new dataset, as well as the original

hyperparameters were used to train the model for the last time. However, this time around, the model was evaluated on the test set.

The data input pipeline, as well as the model architecture were developed and trained using the Tensorflow-v2.7.0 deep-learning library. At this point in the research, we were done training Xception to predict eight visual attributes. Let us discuss how the model performed.

5.4. RESULTS AND DISCUSSION

To describe how well the Single-pass model performed, we would compare the performance to that of the classical model and the Unet-Xception model. Since the integral component of the single-pass solution was an Xception model, going forward, we would refer to the single-pass model as "Xception". Table 5.3 shows the coefficient of determination achieved using Xception when it was trained to predict twelve visual attributes. Just like Unet-Xception, after the first round of training, Xception also performed poorly in predicting contrast, correlation, entropy, and roundness, thus reducing the overall performance of the model. That is why the number of predicted outputs were reduced to eight during the second round of training.

Table 5.3 R^2 values for classical, Unet-Xception, and Xception models in predicting twelve visual attributes

Visual parameter	Classical	Unet-Xception	Xception
L	-0.8660	0.7731	0.7170
a	0.7719	0.8669	0.8072
b	0.8694	0.8521	0.8852
Contrast	-3.5691	0.3128	0.5075
Correlation	-7.7618	0.1231	0.2469
Entropy	-1.4961	-0.0139	0.3761
Homogeneity	0.7987	0.7234	0.8930
Equivalent diameter	-1.0278	0.7926	0.8748
Ferret diameter	-1.6417	0.8482	0.8802
Filled area	-0.3663	0.8391	0.8901
Perimeter	-1.5425	0.8250	0.8421
Roundness	-8.0029	0.3474	0.6447

Table 5.4 contains the improved R^2 values when the model was trained to predict eight visual attributes. Xception outperformed the classical model in predicting on all visual attributes.

Table 5.4 R^2 values for classical, Unet-Xception, and Xception models in predicting eight visual attributes

Visual parameter	Classical	Unet-Xception	Xception
L	-0.8660	0.8017	0.7845
a	0.7719	0.8851	0.8675
b	0.8694	0.8981	0.9064
homogeneity	0.7987	0.7947	0.8988
Equivalent diameter	-1.0278	0.8082	0.8606
Feret diameter	-1.6417	0.8535	0.8707
Filled area	-0.3663	0.8627	0.8837
perimeter	-1.5425	0.8152	0.8566

By observing the R^2 values alone, it seems like Xception slightly outperformed Unet-Xception in predicting all visual attributes, except L^* and a^* color indices. However, there is more to Xception's performance than R^2 reveals. To spot this important discovery, let us observe some figures in which the classical model, Unet-Xception and Xception were used to plot real-time changes in visual attributes, for peas dried at 50^{o}C. The figures of interest are Figures 5.7 – 5.9 for color, Figure 5.10 for texture, and Figures 5.11 – 5.14 for Size. While observing each of these figures, we would notice that Unet-Xception produced the smoothest trend of all the three models. The next in line was Xception, followed by the classical model.

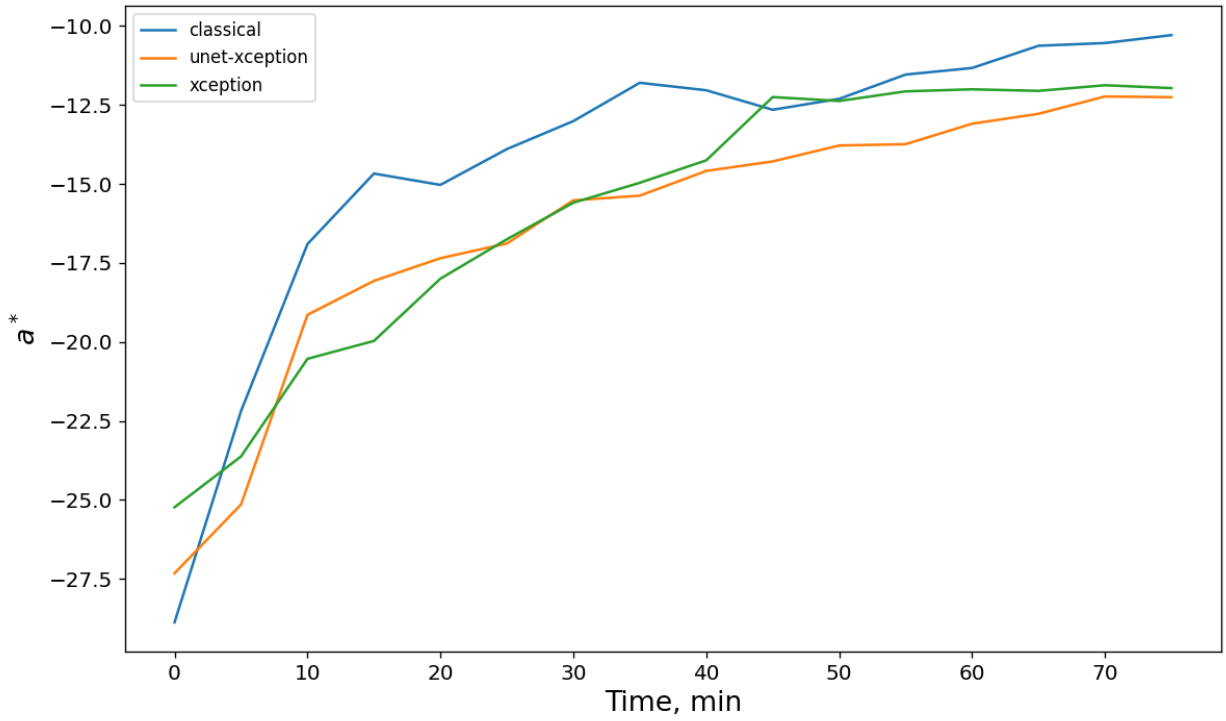


Figure 5.7 Real-time plot of a, produced by the classical model, Unet-Xception and Xception.*

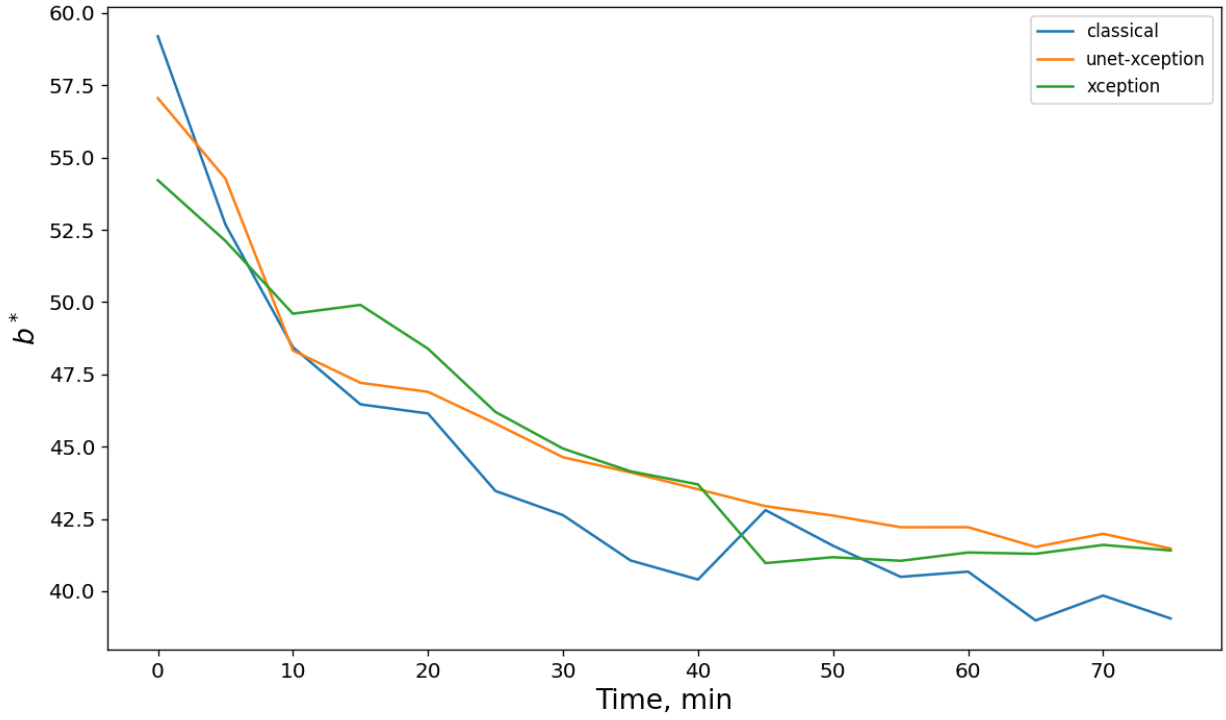


Figure 5.8 Real-time plot of b, produced by the classical model, Unet-Xception and Xception.*

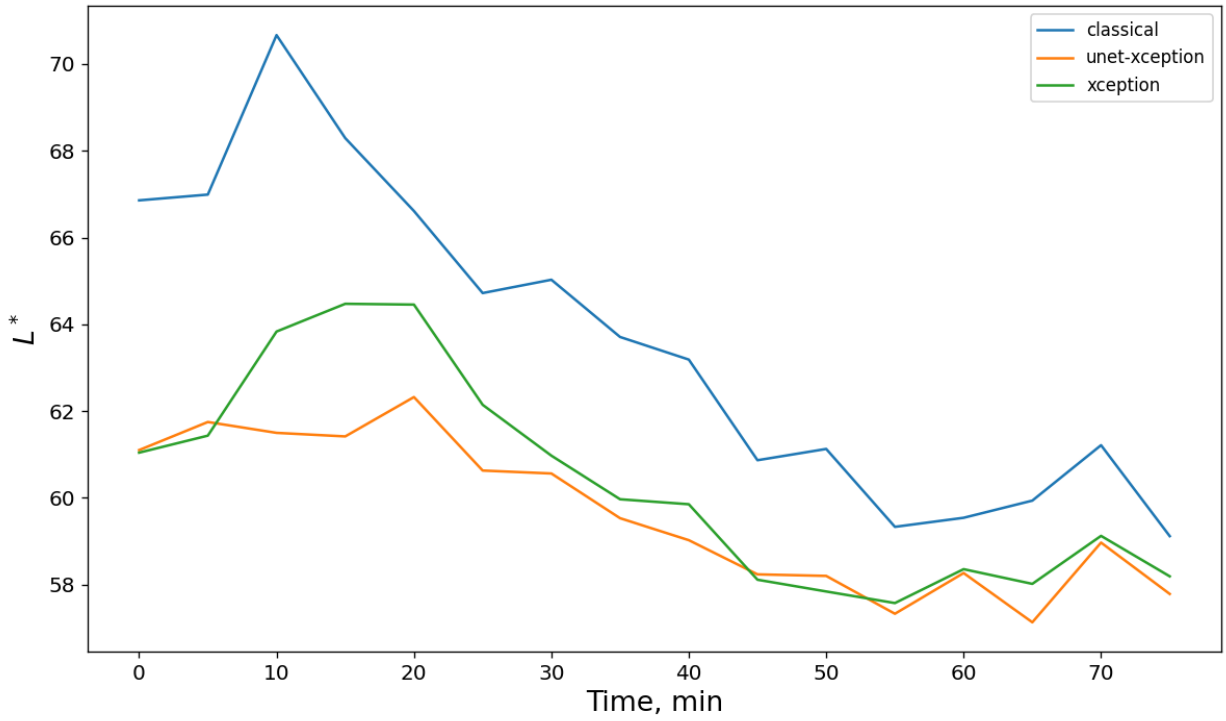


Figure 5.9 Real-time plot of L^*, produced by the classical model, Unet-Xception and Xception.

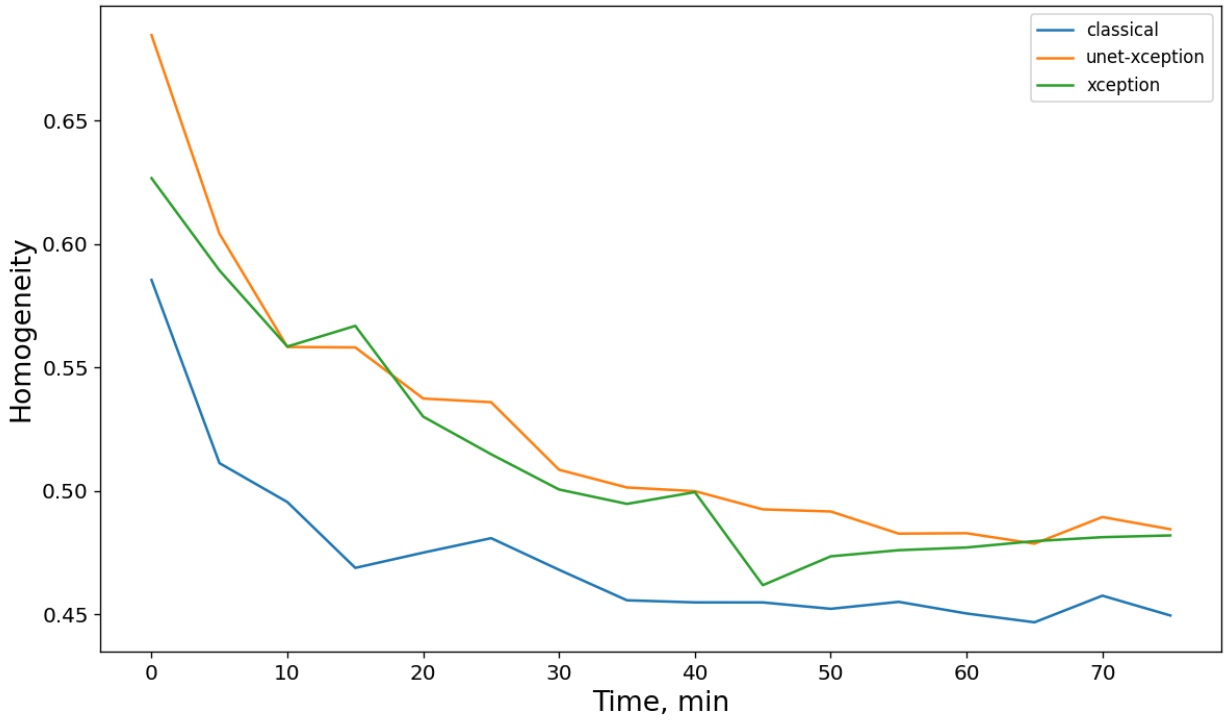


Figure 5.10 Real-time plot of homogeneity, produced by the classical model, Unet-Xception and Xception.

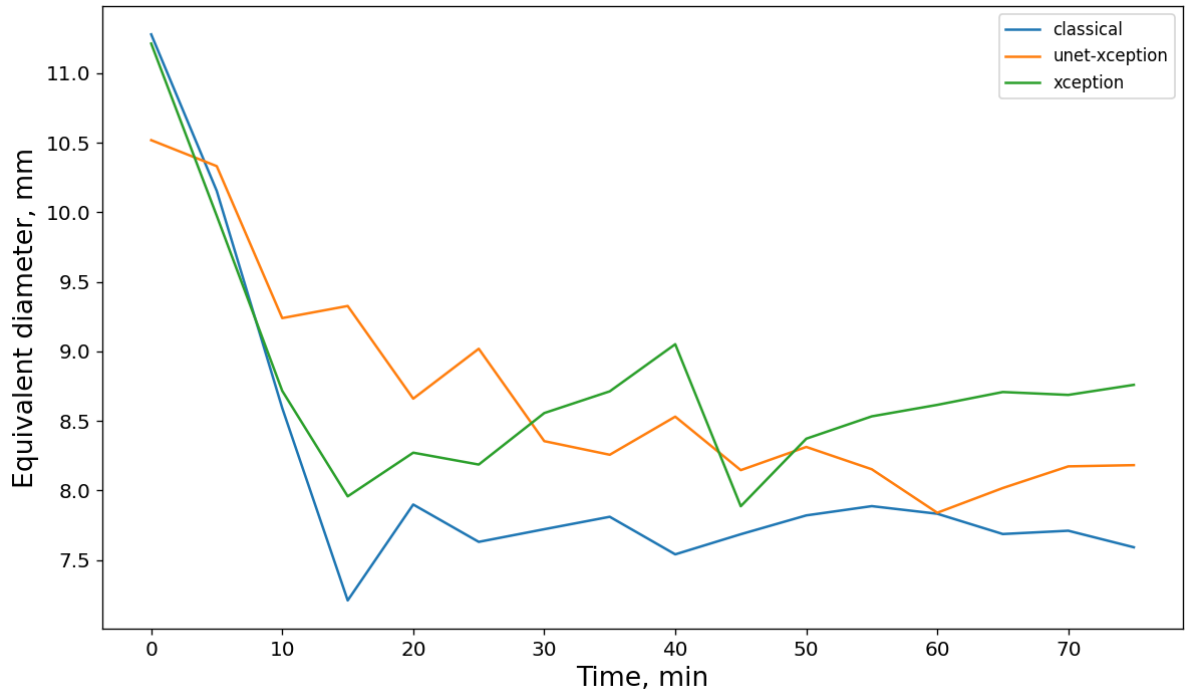


Figure 5.11 *Real-time plot of equivalent diameter, produced by the classical model, Unet-Xception and Xception.*

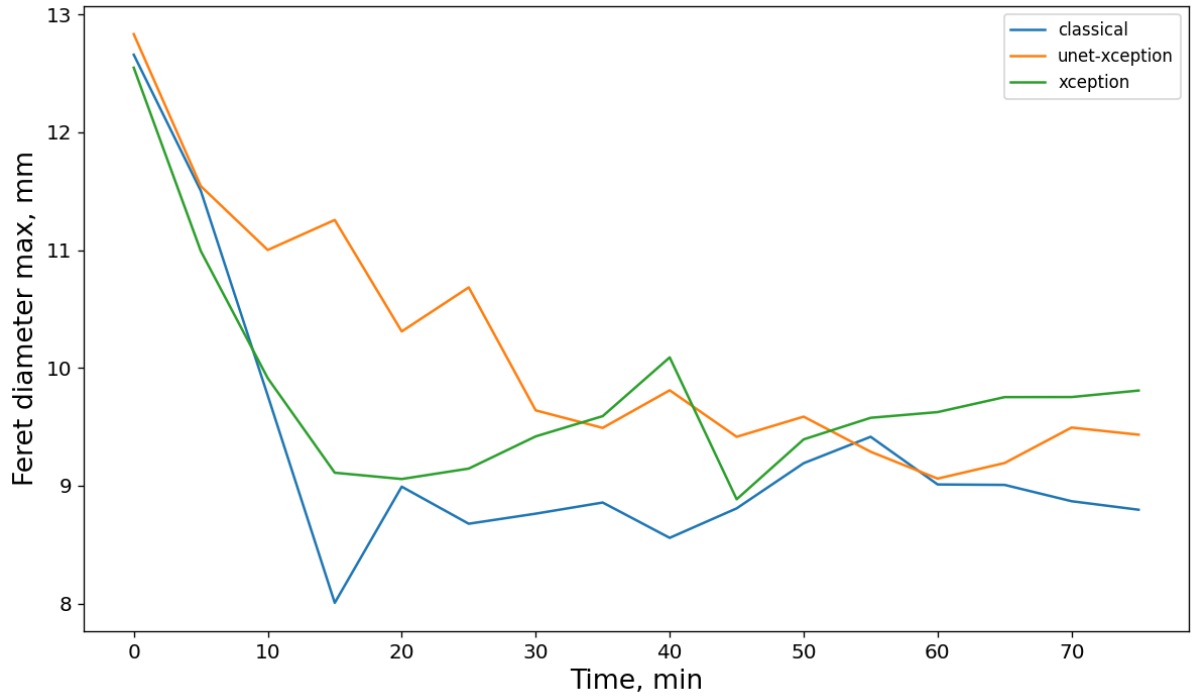


Figure 5.12 *Real-time plot of feret diameter, produced by the classical model, Unet-Xception and Xception.*

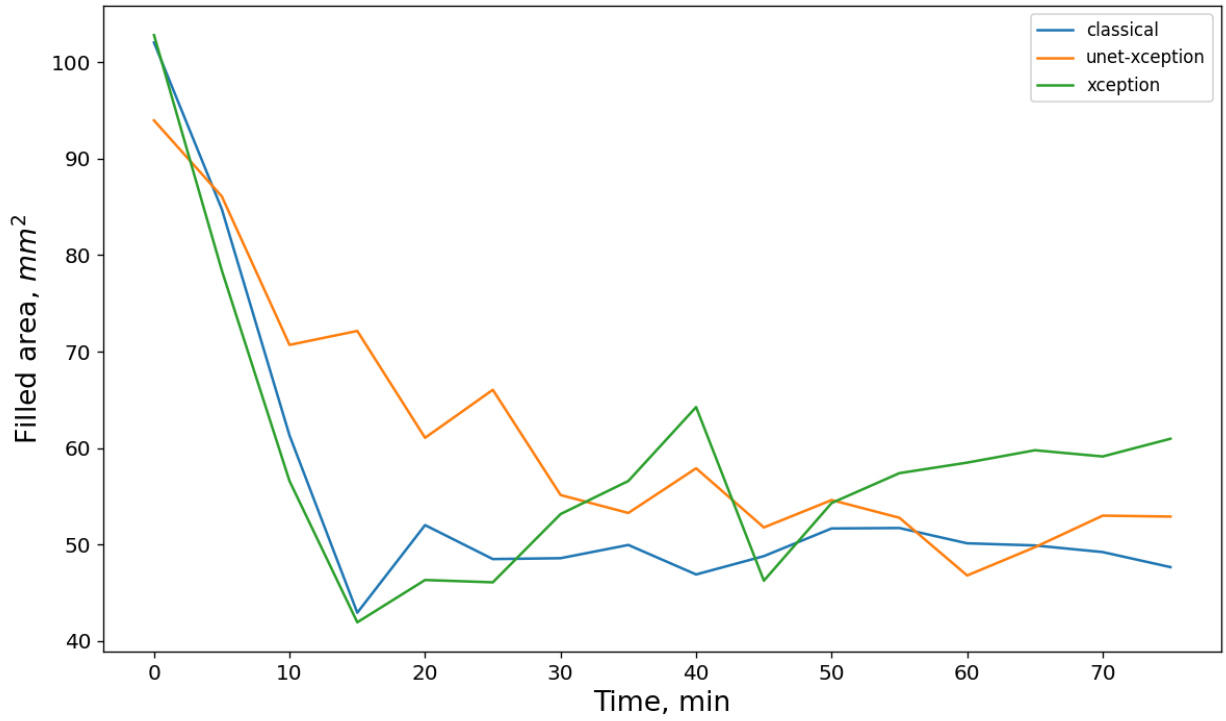


Figure 5.13 Real-time plot of filled area, produced by the classical model, Unet-Xception and Xception.

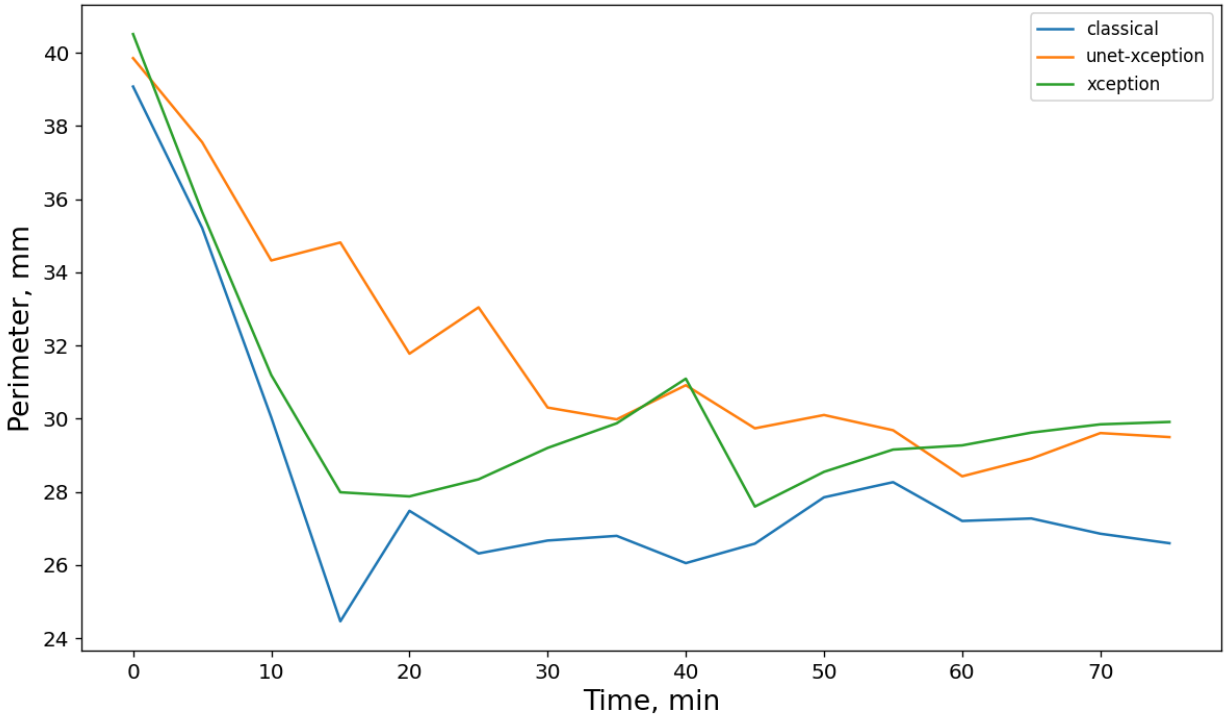


Figure 5.14 Real-time plot of perimeter produced by the classical model, Unet-Xception and Xception.

Having observed all the plots, we see that notwithstanding that Xception had higher R^2 values, the plots generated by Unet-Xception predicted real-time changes better. Why?

Our answer could be found in Figure 5.15 below. This figure contains two rows of images. Each row is made of three images each. The first image was a real-time image captured during fluidized bed drying. The second image is the ground truth containing only infocus peas, while the third image is the segmentation result generated by the U-Net model within Unet-Xception.

As we examine the U-Net prediction in Figure 5.15, we would notice that the model detected additional infocus peas that were missing from the ground truth mask. Since visual attributes were computed using only infocus peas, the Unet-prediction that was fed to the Xception model (that was part Unet-Xception) contained more accurate information for predicting visual attributes of peas within the original image. This explanation could be why the Unet-Xception model produced a smoother trend than the one-pass Xception model.

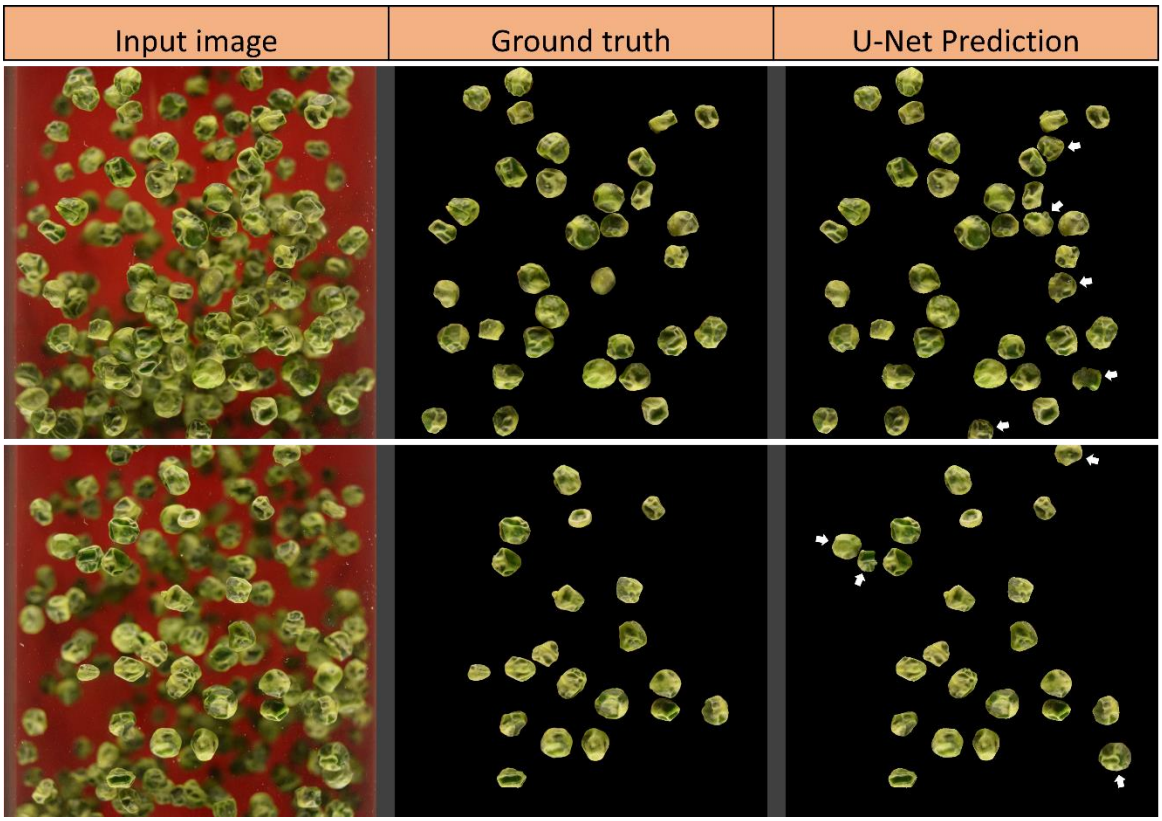


Figure 5.15 U-Net predictions showing additional infocus peas (not in the ground truth) that were detect by the model.

To throw more light on why Unet-Xception performed slightly better than Xception, we need to compare how both models were trained. Let us start with Unet-Xception. As depicted in Figure 5.16, Unet-Xception was an amalgamation of the Unet and Xception models.

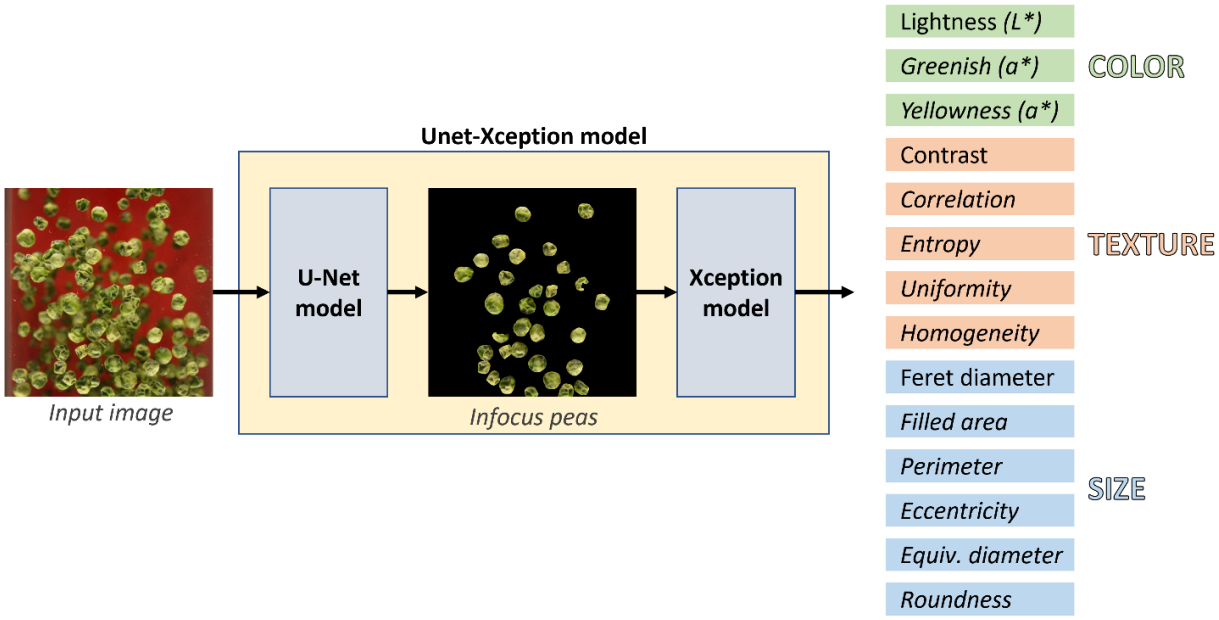


Figure 5.16 Image of the Unet-Xception model comprising of U-Net and Xception models.

Before merging them together to form Unet-Xception, each of the models were trained separately. U-Net was trained using real-time images captured during the drying process. These images contained infocus peas, defocused peas, and a red background. On the other hand, the Xception model that U-Net was merged to, was trained with RGB images containing only infocus peas.

Once training was over, U-Net was able to predict outputs that contain most of the infocus peas within the ground truth mask and much more, as we have seen. The presence of additional infocus peas within the predicted mask meant that in some cases, the ground truth mask was not a very accurate representation of all the infocus peas within the input image. However, unlike U-Net, there was an occasional discrepancy between the prediction and ground truth, that was non-existent when training the Xception model that was merged to U-Net. Xception was trained on RGB images containing only infocus peas. The ground truth values for visual attributes were generated from the same RGB images of infocus peas that Xception was trained on. Hence these ground truth values were an accurate presentation of the RGB image fed into the model. Once the Xception

model was fed an RGB image of infocus peas, the model produced the best possible value of visual attributes.

Table 5.5 contains a column labelled *"Xception-infocus"*. Xception-infocus refers to the Xception model within Unet-Xception. The R^2 values within this column were generated when Xception-infocus was evaluated using RGB images of infocus peas, within the test set.

Table 5.5 R^2 values for classical, Xception-infocus, Unet-Xception, and Xception models in predicting eight visual attributes

Visual parameter	Classical	Xception-infocus	Unet-Xception	Xception
L	-0.8660	0.8215	0.8017	0.7845
a	0.7719	0.9062	0.8851	0.8675
b	0.8694	0.8967	0.8981	0.9064
homogeneity	0.7987	0.9033	0.7947	0.8988
Equivalent diameter	-1.0278	0.9298	0.8082	0.8606
Ferret diameter	-1.6417	0.9241	0.8535	0.8707
Filled area	-0.3663	0.9460	0.8627	0.8837
perimeter	-1.5425	0.9095	0.8152	0.8566

We notice that as a stand-alone model, when Xception-infocus was used to predict visual properties of peas, it performed better than the other three models. This high performance meant that Unet-Xception was good at predicting visual attributes. The slightly reduced coefficient of determination was because in some cases, the ground truth values on visual attribute did not fully account for all the infocus peas within the original image. This reasoning provides insight on how the predictions from the single-pass model Xception could be improved. Achieving this goal could be possible by re-training Xception with ground truth values generated from updated RGB image of infocus peas, containing previously omitted peas.

The number of weights in Unet-Xception and Xception are about 56 and 21 million respectively. Improving the results of Xception as described above, would save a lot of computational resources during deployment, when using the single-pass solution. Notwithstanding the slight improvement of Unet-Xception over Xception, the real-time trend produced using Xception shows that it is possible to predict real-time changes in visual properties using a single-pass deep learning solution.

5.4.1. Monitoring visual attributes using Xception

In this section, we would discuss the results obtained when Xception (the single-pass model) was used to monitor visual attributes of peas, for fluidized bed drying experiments, conducted at 50, 55, 60^{o}C.

5.4.1.1. Color

Figure 5.17 shows the real-time trend in greenish index *a**. For all drying temperatures, *a** trended upwards throughout the drying duration. The increase in a* meant that as drying progressed, the color of the peas changed from green to dark green.

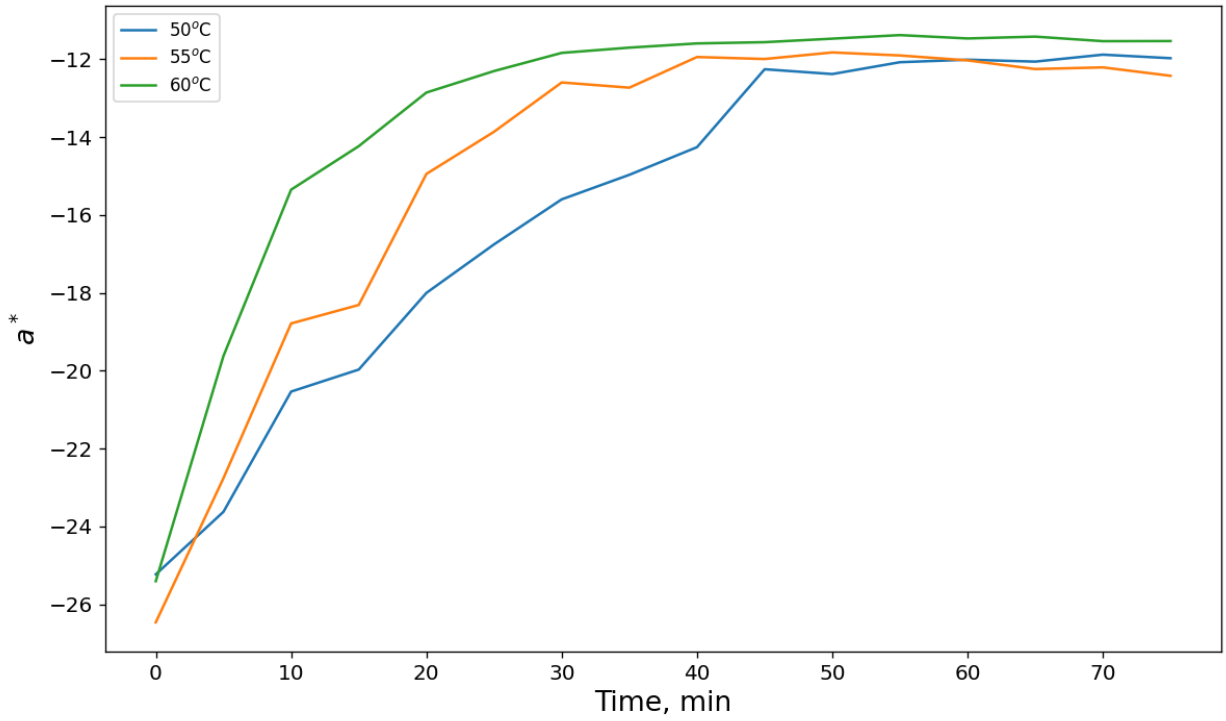


Figure 5.17 Real-time plot of a, predicted by Xception, for 50, 55 and 60^{o}C drying runs.*

From the graph, we notice that air temperature influenced the rate of color change. The fastest increase in a* occurred within the first 20 minutes for peas dried at 60 ^{o}C, followed by 40 minutes for peas dried at 55^{o}C, and 45 minutes for peas dried at 50 ^{o}C. The change in a* could be attributed to degradation in chlorophyll on the surface of the pea. The upward trend may also be linked to Maillard browning reaction, due to air temperature.[17,92-94] Kaur et al. [95] obtained similar findings when drying peas using a three-stage convective drying method.

Figure 5.18 displays real-time trend in yellowish index b*, for peas dried at 50, 55 and 60 ºC. Yellowish index decreased over time, for all drying temperatures. The reduction in b* index could have been caused by Maillard browning reaction. The change could also be linked to formation of brown pigments on the surface of the peas as drying progressed. Another reason for the downward trend could be due to increase in concentration of β-carotenoid caused by moisture loss and pea shrinkage.[96,97]

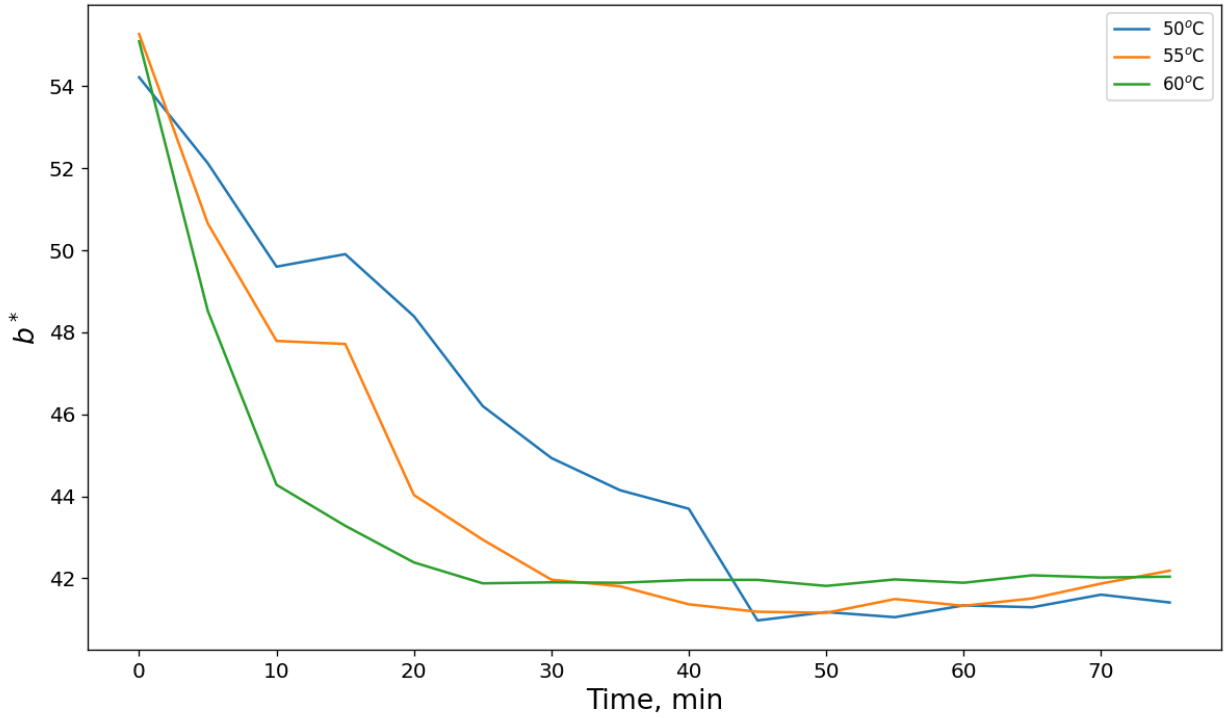


Figure 5.18 Real-time plot of b, predicted by Xception, for 50, 55 and 60°C drying runs.*

The real-time changes in Lightness index L* is presented in Figure 5.19. Lightness value describes the luminance or the ability of the peas to reflect light.[90,98] Overall, lightness reduced as the product lost moisture during the drying process. Meaning that the product appeared darker by the end of drying. Zelinska et al.[99] observed the same trend when drying green peas. The reduction in L* could be attributed to degradation of chlorophyll into gray-brown compounds such as pheophytin or Pheophorbide.[99]

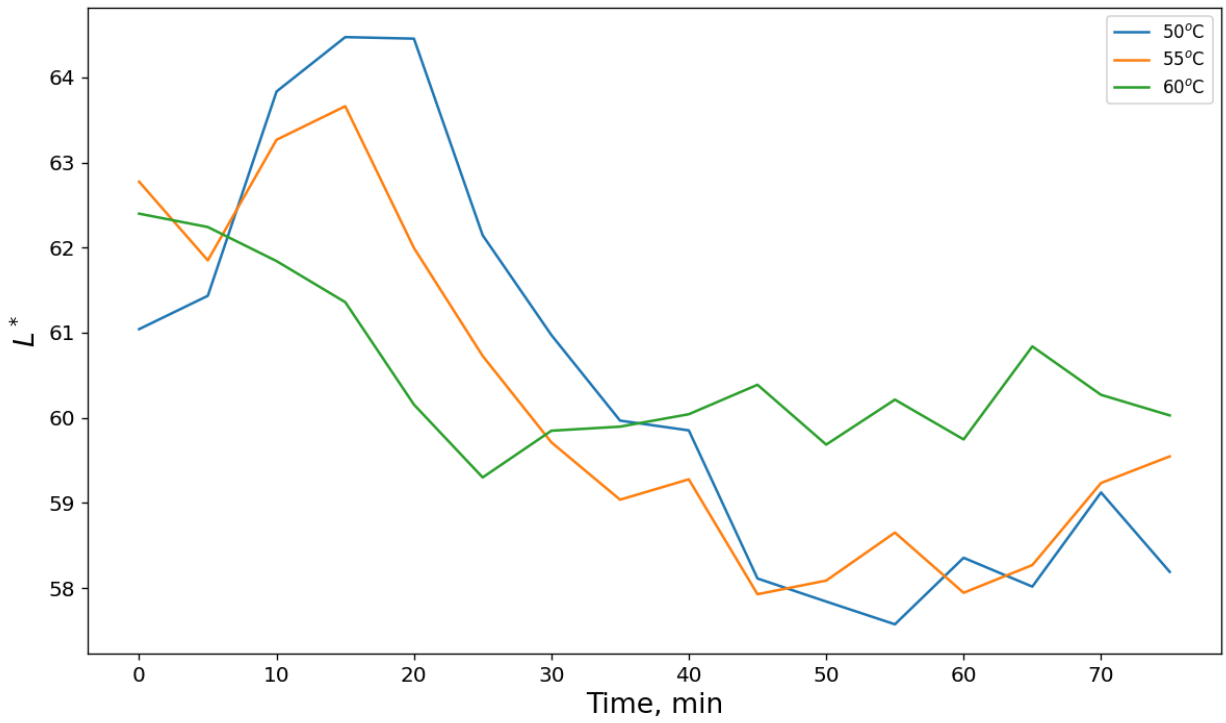


Figure 5.19 Real-time plot of L, predicted by Xception, for 50, 55 and 60°C drying runs.*

Observing the 50, and 55°C trends, we notice that between the 5th to 15th minute, both plots briefly trend upwards. Afterwards, they trend downwards for the rest of the drying duration. Th reason for both earlier bumps could be because of the ground truth data for images captured between the 5th to 15th minute which did not account for all the infocus peas within the original image.

The final *L** values for peas dried at 50 and 55°C were lower than those dried at 60°C. The reduced *L** values meant that, by the end of drying, the peas dried at 50 and 55°C appeared darker than those dried at 60°C. The reason for the variation in final *L** value could be due to high air temperature. This increased temperature of 60°C may have caused significant early moisture loss and case hardening, thus leading to a higher *L** by the end of drying.

As drying progressed, lighter and darker green regions developed on the pea surface. The appearance of varying intensities of green was observed for peas dried at the three temperatures. Another reason for the discrepancy in the final L* value could be explained by considering the combined effect of formation of light and dark regions on the peas, together with the rate of moisture loss, and uniformity in shrinkage. Peas dried at 60°C had a high rate of moisture loss at the earlier stages of drying, thus leading to case hardening. As the peas lost water, while case

hardened, they shrunk more uniformly on all sides. Hence, both their lighter and darker green regions were visible within the realtime image. So, when L* was predicted, the light green regions help maintain a high L* value for peas dried at 60°C. On the other hand, peas dried at 50 and 55°C gradually lost moisture as drying progressed. The gradual loss of moisture, together with non-uniform shrinkage across the surface of individual pea could be the reason for the lower *L** values. As the product shrank, more of the lighter regions collapsed inward. Hence the darker and harder regions contributed more towards the lower final L* value. In summary, with the one-step deep learning solution we can now monitor real-time changes in food color, during fluidized bed drying.

5.4.1.2. Texture

Amongst all textural parameters Xception was trained to predict, homogeneity turned out to be the best for monitoring visual texture. The parameter had the highest coefficient of determination. Also, the real-time trend was consistent for all the drying runs. Figure 5.20 displays real-time changes in homogeneity for peas dried at 50, 55, and 60°C. For the three drying temperatures, homogeneity trended downwards as the peas dried. Fernandez et al.[8] achieved similar downward trend when investigating the effect of drying conditions on the shape, texture, and color of apples.

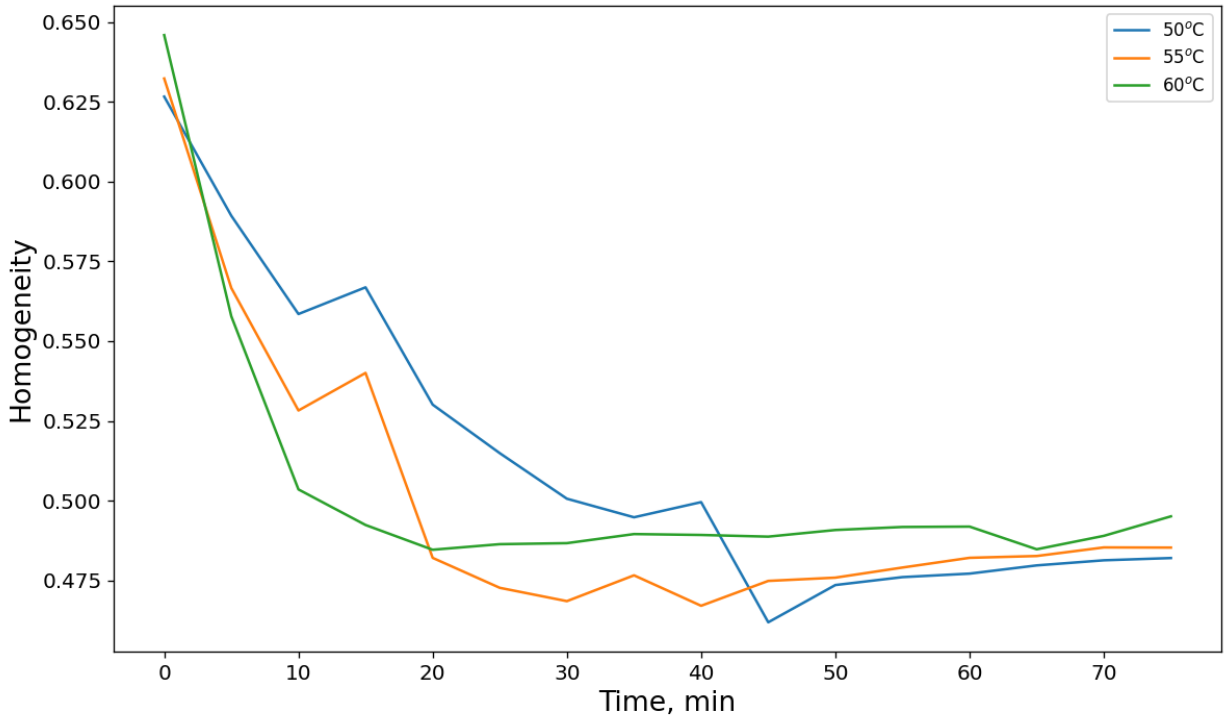


Figure 5.20 Real-time plot of homogeneity, predicted by Xception, for 50, 55 and 60°C drying runs.

The trend in textural homogeneity is an indication of how uniform or disorganized the pixel intensities on the pea surface evolves, as drying progresses. This evolution is depicted in Figure 5.21. The Figure is a series of real-time images, captured at 5 minute intervals, during a drying run, conducted at 50^{o}C. Each image was part of 15 consecutive images captured every 5 minutes, from 0 – 75 minutes.

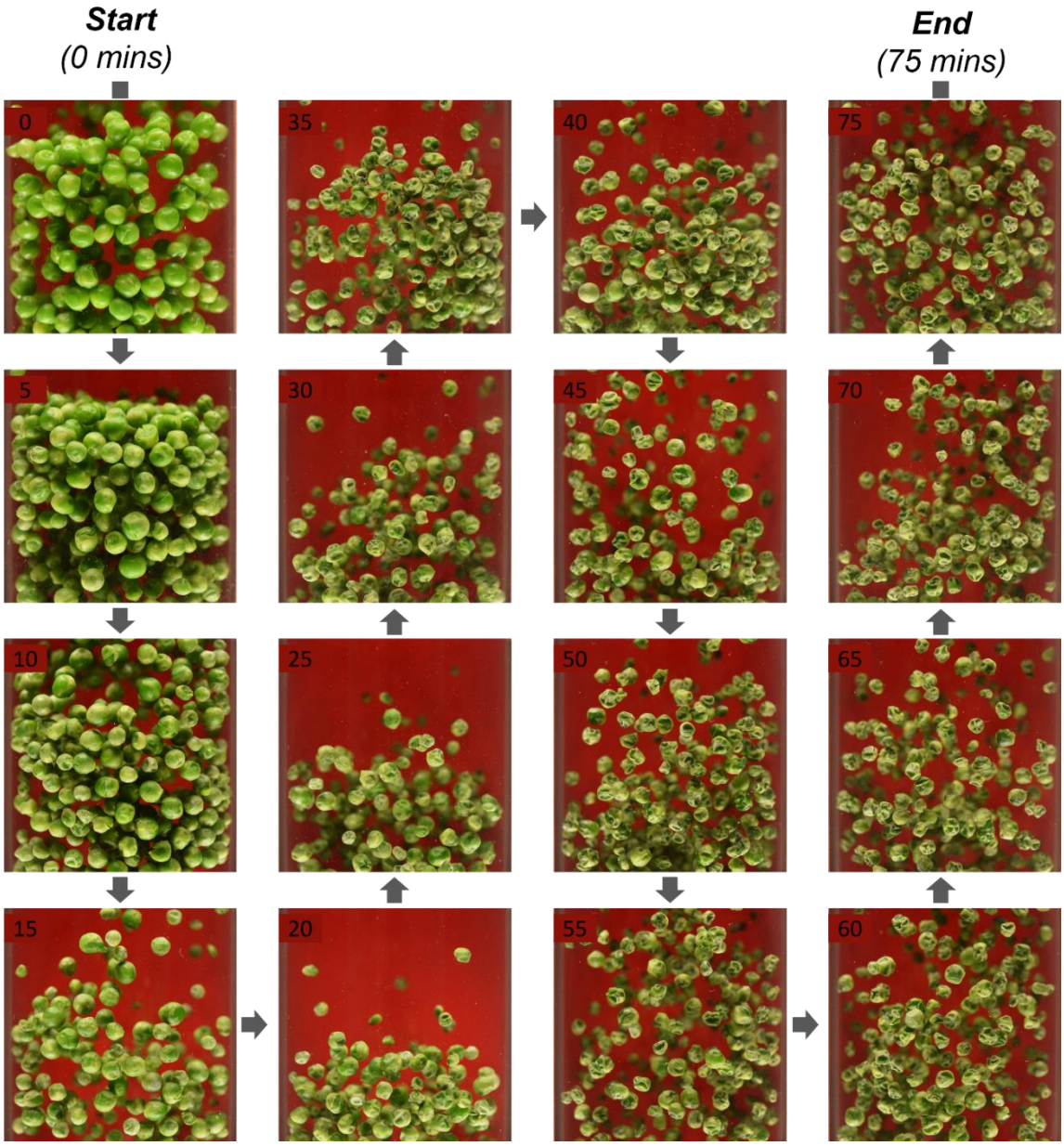


Figure 5.21 Real-time images captured at 5-minute interval, for peas dried at 50^{o}C.

At time 0, the color on the surface of individual peas appears uniform across the pea surfaces. As the pea loses water the color changes. The change is not uniform across the surface of individual peas. Some regions of each pea appear darker than the others. The variation in pixel intensity and the randomness in distribution of the intensities on the surface of the peas affect homogeneity.

As variability in pixel intensity across the surface increases, the rougher the surface of the peas appears, and the lower the homogeneity value becomes. The trend in homogeneity could help food processer know if the product is losing moisture uniformly across it surface. Such knowledge could help reduce early onset of case hardening.

5.4.1.3. Size

During the drying process, as the food loses moisture, it reduces in size. Assuming there were no sensors to monitor moisture content, such as Near infra-red spectrometers, a model would be trained to predict moisture using real-time data of product size. So, aside being used feedback for monitoring visual quality, product size could be used to monitor moisture content. Figure 5.22 – Figure 5.25 display plots of real-time changes in equivalent diameter, ferret diameter, filled area and perimeter. For all the drying temperatures, each of the size parameter trended downward over the course of drying duration.

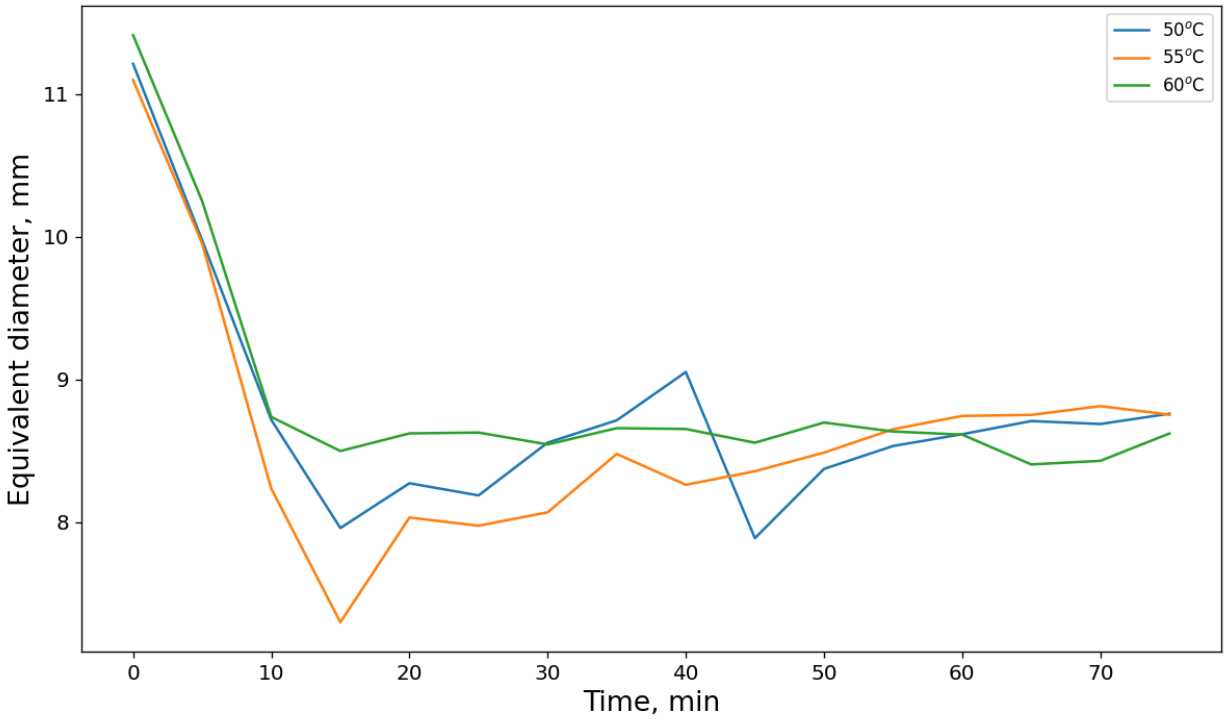


Figure 5.22 Real-time plot of equivalent diameter, predicted by Xception, for 50, 55 and 60°C drying runs.

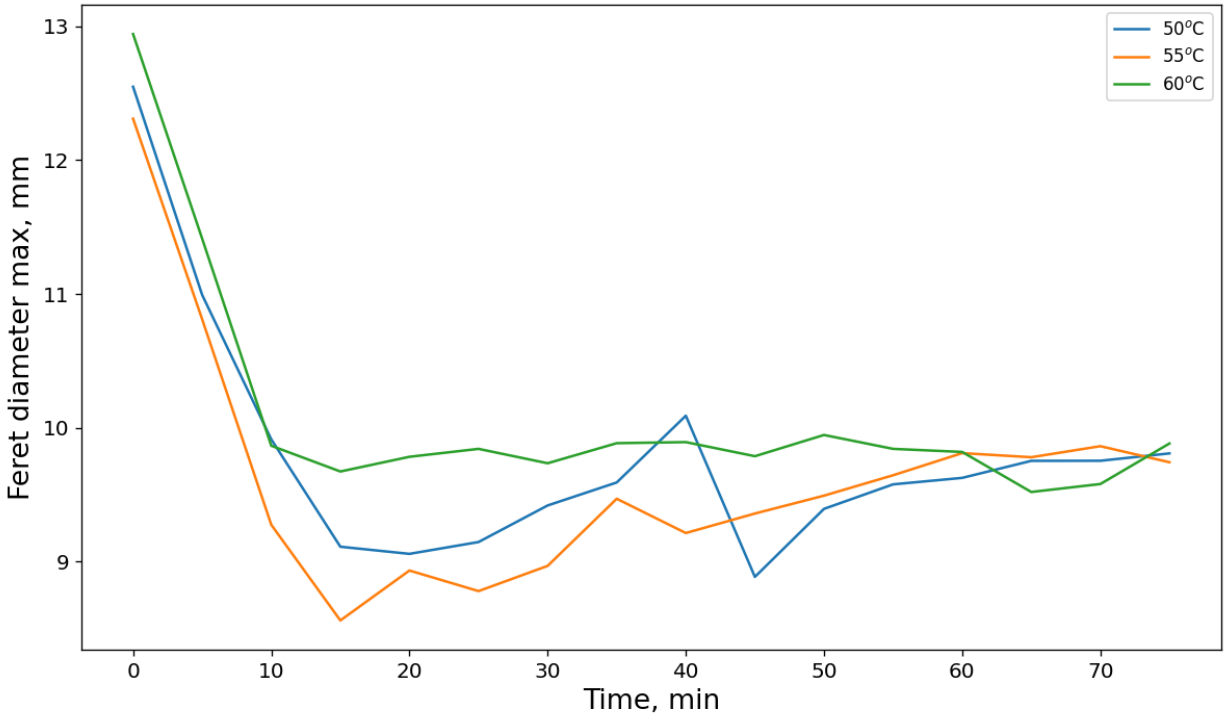


Figure 5.23 Real-time plot of feret diameter, predicted by Xception, for 50, 55 and 60°C drying runs.

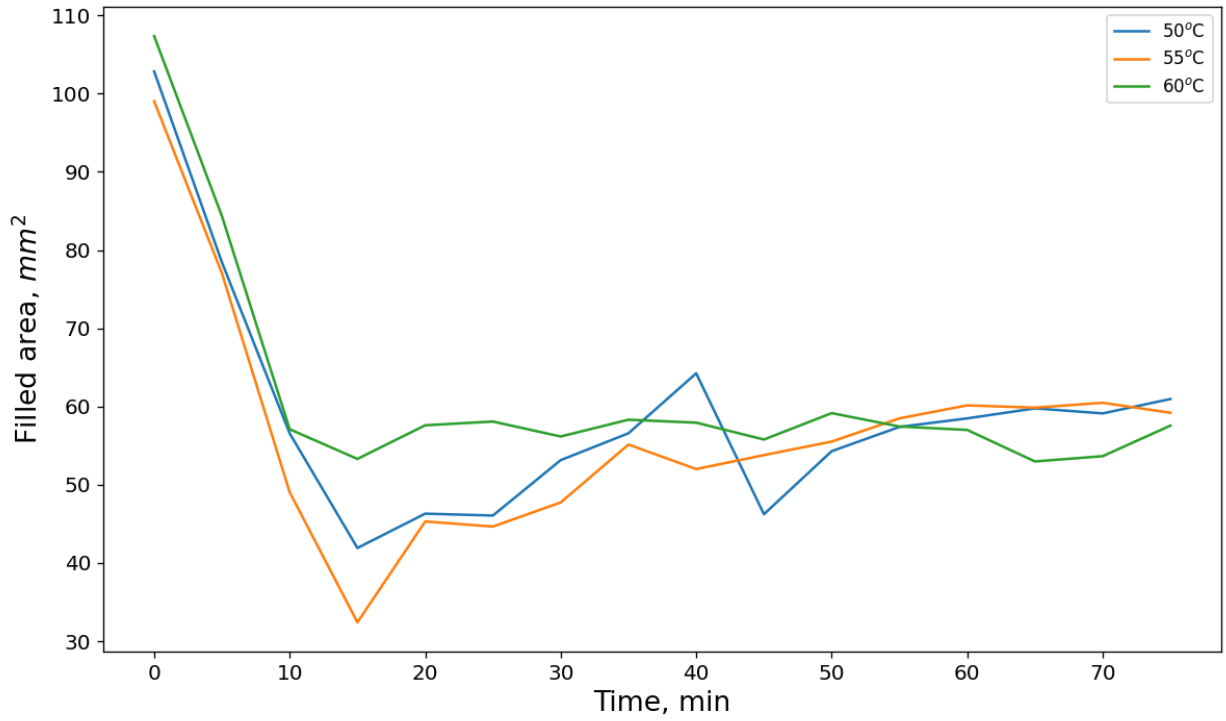


Figure 5.24 Real-time plot of filled area, predicted by Xception, for 50, 55 and 60°C drying runs.

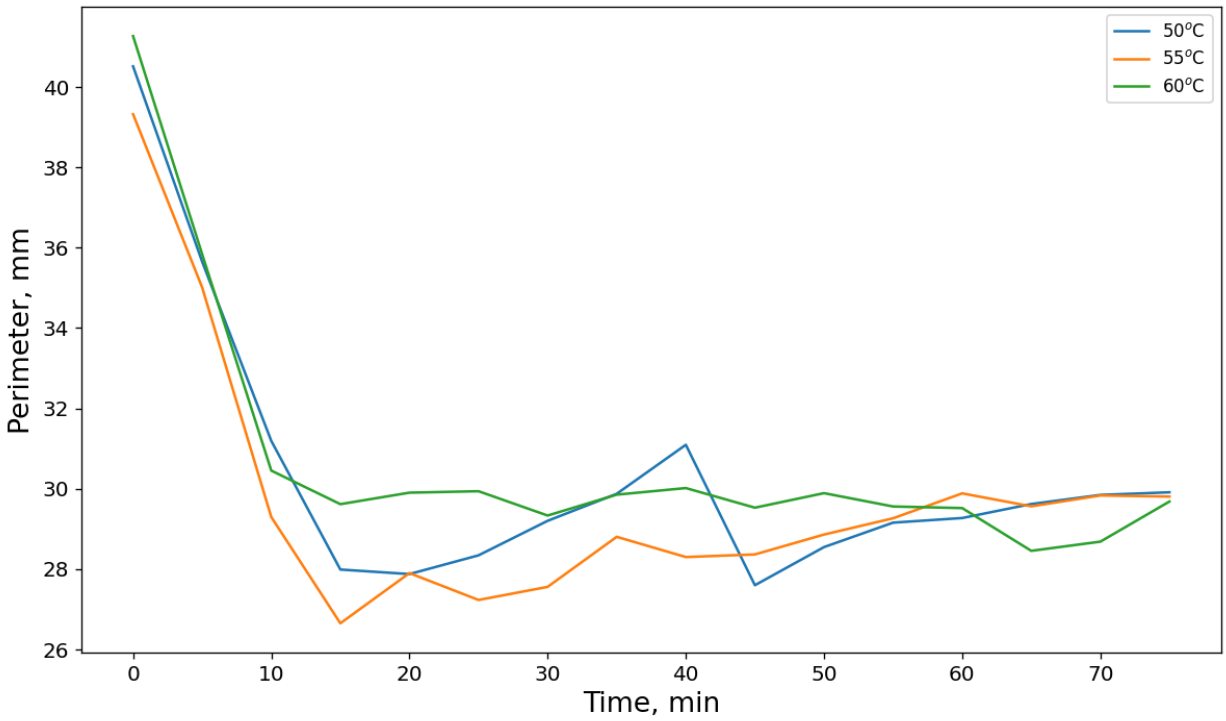


Figure 5.25 Real-time plot of perimeter, predicted by Xception, for 50, 55 and 60°C drying runs.

Having gone through all the real-time plots associated with each visual property, we can say that the single-pass solution could be an effective tool for monitoring real-time changes in visual attributes of food, during fluidized bed drying.

5.5. CONCLUSION

This research work was designed to answer one question ***"is it possible to monitor real-time changes in visual attributes of peas, during fluidized bed drying, without segmenting images of the peas, from which the prediction would be made?"*** To answer this question, we embarked on developing a single-pass deep learning solution. The backbone of this solution was the Xception model. Hence, we referred to the solution as "Xception". The solution was developed to monitor color, texture, and size. Xception monitored color by predicting L* a* and b* indices. Texture was monitored using homogeneity, while equivalent diameter, ferret diameter, filled area, perimeter, and roundness were predicted to monitor size. Xception was trained to predict visual attributes from RGB images of green peas that were captured during fluidized bed drying experiments. The drying run was conducted at 60^{o}C.

After training, the performance of Xception was compared to that of a classical model and a two-pass deep learning solution known as Unet-Xception. Both the classical model and Unet-Xception were built in line with the two-pass framework, involving image segmentation followed by visual attribute prediction. Xception and Unet-Xception significantly outperformed the classical model in predicting visual attributes. In terms of coefficient of determination, Xception slightly performed better than Unet-Xception in predicting all visual attributes, except L^* and a^* indices. When deployed to monitor new drying runs conducted at 50, 55 and 60^{o}C, the real-time plots produced by Xception and Unet-Xception had similar patterns and trend. Thus, proving that the single-pass solution could be an effective tool for monitoring visual attributes in real-time, during fluidized bed drying. Furthermore, since the Xception had fewer model weights that Unet-Xception, during deployment, the single-pass solution would require significantly lesser computing power than the two-pass solution.

Interestingly, amongst the three models, the real-time plots produced by Unet-Xception was the smoothest, followed by Xception, then the classical model. This meant that the performance of the single-pass model could be slightly improved by retraining it on ground truth data that accounted

for all the infocus peas within the real-time image. Finally, the results from this study show that during fluidized bed drying, visual attributes of food material could be monitored in real-time, without segmenting the captured image. In the future, the capabilities of this single-pass solution could be broadened to include monitoring real-time changes in moisture content and nutrition content.

Our results show that with the single-pass model, visual attributes of food can be monitored in real-time during fluidized bed drying. Since we can now monitor visual attributes in real-time, it also means that it is possible to control these attributes in real-time too. We know the drying parameters that affect them. Hence, our next study could be to develop a control system that learns how different operating parameters affect the color, texture, size, and moisture content of food in real-time, during fluidized bed drying. Once the control system is merged to the single-pass solution, and the union is then integrated with a drying system, we would end up with a dryer that consistently produces dried food that is visually appealing and has a longer shelf life. Such dryer could significantly reduce global food losses and waste.

CONNECTING TEXT

At the point in the research, we had developed a two-pass deep learning solution and a single-pass deep learning solution. With the Two-pass solution, we had accomplished the first two milestones, while the single-pass solution achieved the third milestone.

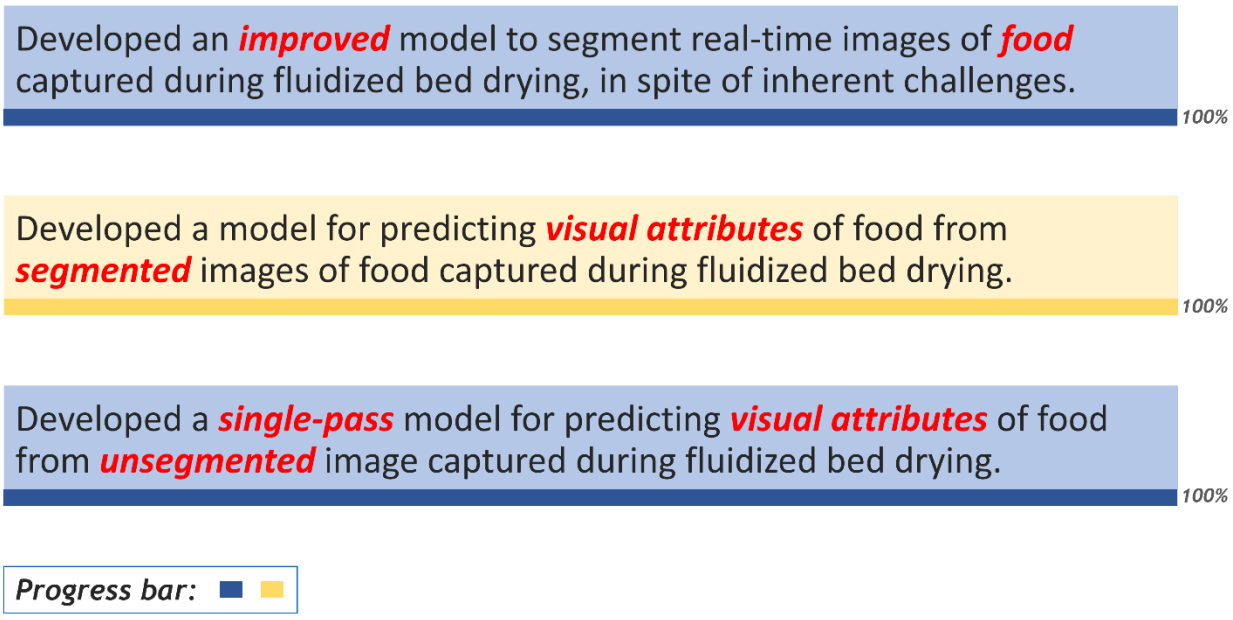


Figure 5.26 Milestones to improve the state-of-the-art

By achieving the three milestones, we have developed two effective systems for monitoring visual attributes of food, in real-time during fluidized bed drying. In the next chapter, we would present our overall conclusions; itemize our contributions to knowledge; and outline future research works.

CHAPTER VI

CONCLUSIONS, CONTRIBUTION TO KNOWLEDGE AND FUTURE WORK

6.1. CONCLUSIONS

This study was designed to answer one major research question - ***"Is it possible to monitor visual attributes of food, in real-time, during fluidized bed drying, using images of the food captured during the drying process?".*** If the answer to this question is "Yes, we can monitor visual attributes in real-time", then, it means that we can also control these attributes in real-time. By developing systems that can monitor and control visual attributes of food during fluidized bed drying, we are at the precipice of building intelligent drying systems that have the capability to consistently produce dry foods that are visually attractive and have long shelf life. Such consistency in product quality could significantly reduce global food losses and waste.

To answer our research question, this study was conducted using green peas, since it exhibits most of the image-related challenges that have made monitoring visual attributes of food during fluidized bed drying an unsolved problem. Due to the mixing action that occurs during fluidization, our hypothesis was that visual attributes could be predicted by extracting information from only those peas that were in contact with the dryer wall. These peas were referred to as *infocus peas*. During the study, we developed an image acquisition system, and built a classical CV pipeline and two deep learning models. The image acquisition system captured real-time and well illuminated images of green peas. The system prevented motion-related blur around the peas, and produced significant contrast amongst the infocus peas, the defocus peas, and the image background. The classical CV pipeline and the first deep learning solution (Unet-Xception) were modeled in line with the Two-pass framework. This framework involves two operational steps – image segmentation, followed by visual attribute prediction using the segmented output. On the other hand, the second deep learning model (Xception) was a Single-pass solution. Meaning that the model predicted visual attributes of peas, without segmenting the input image.

In terms of segmentation performance, the classical CV pipeline was overwhelmed by challenges that negatively affect the segmented output. Some of these challenges include overlapping and

clustered peas within the image, unequal distance of peas from the camera lens, and uneven color and texture across the surface of individual peas. Thus, the segmented result did not detect so many infocus peas (objects of interest). On the other hand, Unet-Xception overcame these challenges, achieving a Mean Intersection-over-Union of 0.9464, compared to 0.81901 by the classical CV pipeline. Unet-Xception even detected additional infocus peas that were missing in the ground truth mask. Hence, it was expected that this Two-pass deep learning model would produce smoother trends when predicting visual attributes. This expectation turns out to be true.

The classical CV pipeline and the two deep learning solutions monitored color, texture, and size of the peas in real-time. Color was monitored by predicting Lightness (L*), greenish index (a*), and yellowish index (b*). Visual texture was monitored in terms of contrast, correlation, entropy, homogeneity, and uniformity, while Feret diameter, filled area, perimeter, and equivalent diameter were used to monitor product size. Unet-Xception and Xception significantly out-performed the classical pipeline, in predicting visual attributes. When used to monitor fresh drying experiments, Unet-Xception produced the smoothest real-time trend. The coefficients of determination achieved by the three solutions are presented in the form [classical pipeline, Unet-Xception, Xception].

a* and b* indices were the best predictors of color, with R^2 values of [0.7719, 0.8851, 0.8675] and [0.8694, 0.8981, 0.9064], respectively. Homogeneity was the best parameter for monitoring texture, with R^2 values of [-0.8660, 0.8017, 0.7845]. Finally, the following R^2 values were obtained for size: [-1.0278, 0.7926, 0.8748] for equivalent diameter, [-1.6417, 0.8482, 0.8802] for ferret diameter, [-0.3663, 0.8391, 0.8901] for filled area, and [-1.5425, 0.825, 0.8421] for perimeter.

Since Unet-Xception identified additional objects of interest that were missing from the manually annotated ground truth masks, it was expected that the R^2 values would be slightly dampened. However, since Unet-Xception comprised of a highly performing Unet and a well-trained Xception, we expected Unet-Xception to produce the smoothest trend when used to monitor fresh drying runs. This expectation turned out to be true. All these findings provide an answer to our original research question, and the answer is ***"Yes, with either a Two-pass or a Single-pass deep learning solution, we can monitor visual attributes of food, in real-time, during fluidized bed drying, from images of the food captured during the drying process"***.

In future studies, we hope to broaden the capabilities of both deep learning solutions, e.g. by training them to also monitor moisture content and nutrient content of the food in realtime. Once

this goal is achieved, the improved monitoring system could be merged with a control system to ensure consistent product quality. Such an improved and intelligent drying system could contribute immensely toward reducing global food losses and waste.

6.2. CONTRIBUTIONS TO KNOWLEDGE

The following are some of the important contributions to knowledge that emanate from this study:

1. Various classical computer vision systems have been developed to monitor visual attributes of food, in real-time, during the drying process. However, none has been developed for fluidized bed dryers. This is the first study that developed a two-pass classical computer vision system to monitor visual attributes of food, in real-time, during fluidized bed drying. This study provides future researchers vital information on how to develop such CV systems. It also stipulates various challenges associated with chaotic-type images that the classical CV approach cannot solve.
2. Previous real-time computer vision systems computed visual attributes of food using static-type images. This is the first work where a CV system was developed to monitor visual attributes of food from chaotic-type images. Hence, this work informs the research community on how to develop real-time solutions to monitor visual attributes of food, even when the products are in random motion.
3. Previous works on monitoring visual attributes of food were conducted in drying systems where the individual food materials were all at the same distance from the camera. Hence, such research was geared towards solving a 2-dimensional problem. The information was needed to compute visual attributes using spatial data. However, in the case of fluidized bed drying, we are faced with a 3-dimensional problem. The individual products are at different distances from the camera lens, hence, to compute visual attribute, such as size, we would need both spatial and depth information. This is the first study that provides a roadmap to researchers on how to develop effective real-time solutions to tackle this problem.
4. This thesis is also the first work that developed an effective two-pass ***deep learning*** solution, to monitor visual attributes of food in real-time during fluidized bed. This system out-performed the classical CV system, both in image segmentation and visual attribute

prediction. Hence when coupled with a control system, the food dryer could consistently produce dry products that are visually attractive and have a longer shelf life.

5. All published research that monitored visual attributes of food during drying were modelled in line with the Two-pass framework. The framework predicts visual attributes using the segmented output of the captured image. This is the first study where a single-pass deep learning solution was developed to predict visual attributes of food in real-time, without segmenting the image of the food. This single-pass achieved similar (and in cases slightly better) predictions as the two-pass deep learning solution.
6. We cannot optimize visual attributes of food dried in a fluidized bed dryer, unless these attributes can be measured during the drying process. Our Two-pass and Single-pass deep learning solutions would be an indispensable tool for food drying researchers to generate real-time feedback on visual attributes of food during fluidized bed drying. This stream of data could then be used to optimize the drying process so as to end up with dried products that are visually appealing and dried to a safe moisture content.

6.3. FUTURE WORK

1. The two-pass and single-pass deep learning solutions were trained to predict visual attributes of green peas. To improve their versatility, the models should be trained on more food products.
2. Consumers are attracted to food products that are visually appealing, nutritious and have long shelf life. The two deep learning solutions we developed, can help food processors monitor visual attributes in real-time. With the integration of near infrared spectrometry, we could investigate the possibility of upgrading both AI-driven solutions to predict moisture content, visual attributes, and nutritional information all at the same time.
3. Our two-pass deep learning model known as *Unet-Xception* comprises of two sub-models, U-Net for image segmentation and Xception for predicting visual attributes using the segmented output from U-net. We could investigate how a different state-of-the-art semantic segmentation model, such as DeepLabV3, would perform when used in place of U-Net.

4. Investigate the possibility of adapting both deep learning solutions to work on other drying systems, such as convective dryers.
5. By coupling either the two-pass or single-pass solutions to a control system, we would end up with an intelligent system that would ensure consistent dry food quality. Hence, we hope to develop a control system using reinforcement learning.

CHAPTER VII

REFERENCES

[1] FAO. State of Food and Agriculture: Moving Forward on Food Loss and Waste Reduction. **2019**.

[2] FAO. Food Loss and Food Waste. *http://www.fao.org/food-loss-and-food-waste/en/ http://www.fao.org/food-loss-and-food-waste/en/* (accessed August 10, 2019).

[3] FAO. Key facts on food loss and waste you should know. *http://www.fao.org/save-food/resources/keyfindings/en/ http://www.fao.org/save-food/resources/keyfindings/en/*2019).

[4] FAO. Global Initiative on Food Loss and Waste Reduction. *http://www.fao.org/3/a-i4068e.pdf http://www.fao.org/3/a-i4068e.pdf* (accessed August 10, 2019).

[5] Sebastian Stępień; Dawid Dobrowolski Loss and waste in the food supply chain: an introduction to the problem. *Progress in Economic Sciences* **2017**DOI: 10.14595/PES/04/022.

[6] Oman Kadunc, N.; Šibanc, R.; Dreu, R.; Likar, B.; Tomaževič, D. In-line monitoring of pellet coating thickness growth by means of visual imaging. *International Journal of Pharmaceutics* **2014**, *470*, 8-14. DOI: 10.1016/j.ijpharm.2014.04.066.

[7] Beaudin, M. *The agri-food industry in Atlantic Canada: recognition and development;* Canadian Institute for Research on Regional Development: Moncton, NB, 2004.

[8] Fernández, L.; Castillero, C.; Aguilera, J. M. An application of image analysis to dehydration of apple discs. *Journal of Food Engineering* **2005**, *67*, 185-193. DOI: 10.1016/j.jfoodeng.2004.05.070.

[9] Paulus, I.; Schrevens, E. Evaluating and modelling the size classification of apples. *Journal of Agricultural Engineering Research* **1999**, *74*, 411-419. DOI: 10.1006/jaer.1999.0479.

[10] Chen, Y.; Martynenko, A. Computer vision for real-time measurements of shrinkage and color changes in blueberry convective drying. *Drying Technology* **2013**, *31*, 1114-1123. DOI: 10.1080/07373937.2013.775587.

[11] Martynenko, A. I.; Yang, S. X. An Intelligent Control System for Thermal Processing of Biomaterials. Presented at the Apr 2007.

[12] Martynenko, A. Computer vision for real-time control in drying. *Food Engineering Reviews* **2017**, *9*, 91-111. DOI: 10.1007/s12393-017-9159-5.

[13] Martynenko, A. I. Computer-vision system for control of drying processes. *Drying Technology* **2006**, *24*, 879-888. DOI: 10.1080/07373930600734067.

[14] Wang, D.; Martynenko, A.; Corscadden, K.; He, Q. Computer vision for bulk volume estimation of apple slices during drying. *Drying Technology* **2017**, *35*, 616-624. DOI: 10.1080/07373937.2016.1196700.

[15] Yadollahinia, A.; Jahangiri, M. Shrinkage of potato slice during drying. *Journal of Food Engineering* **2009**, *94*, 52-58. DOI: //doi.org/10.1016/j.jfoodeng.2009.02.028.

[16] Yadollahinia, A.; Latifi, A.; Mahdavi, R. New method for determination of potato slice shrinkage during drying. *Computers and Electronics in Agriculture* **2009**, *65*, 268-274. DOI: 10.1016/j.compag.2008.11.003.

[17] Hosseinpour, S.; Rafiee, S.; Mohtasebi, S. S.; Aghbashlo, M. Application of computer vision technique for on-line monitoring of shrimp color changes during drying. *Journal of Food Engineering* **2013**, *115*, 99-114. DOI: 10.1016/j.jfoodeng.2012.10.003.

[18] Sturm, B.; Hofacker, W. C.; Hensel, O. Optimizing the drying parameters for hot-air–dried apples. *Drying Technology* **2012**, *30*, 1570-1582. DOI: 10.1080/07373937.2012.698439.

[19] Mulet, A.; Garcia-Reverter, J.; Bon, J.; Berna, A. Effect of shape on potato and cauliflower shrinkage during drying. *Drying Technology* **2000**, *18*, 1201-1219. DOI: 10.1080/07373930008917772.

[20] Nadian, M. H.; Rafiee, S.; Aghbashlo, M.; Hosseinpour, S.; Mohtasebi, S. S. Continuous real-time monitoring and neural network modeling of apple slices color changes during hot air drying. *Food and Bioproducts Processing* **2015**, *94*, 263-274. DOI: 10.1016/j.fbp.2014.03.005.

[21] Hosseinpour, S.; Rafiee, S.; Mohtasebi, S. S. Application of image processing to analyze shrinkage and shape changes of shrimp batch during drying. *Drying Technology* **2011**, *29*, 1416-1438. DOI: 10.1080/07373937.2011.587620.

[22] Oman, N.; Možina, M.; Tomaževič, D.; Pernuš, F.; Likar, B. A study of a visual inspection technique for in-process monitoring of coating of pharmaceutical pellets. Presented at the 2011 IEEE International Conference on Industrial Technology, Auburn, AL, USA, 2011.

[23] Sampson, D. J.; Chang, Y. K.; Rupasinghe, H. P. V.; Zaman, Q. U. A dual-view computer-vision system for volume and image texture analysis in multiple apple slices drying. *Journal of Food Engineering* **2014**, *127*, 49-57. DOI: 10.1016/j.jfoodeng.2013.11.016.

[24] Environment and Climate Change Canada *Taking stock: Reducing food loss and waste in Canada;* Waste reduction and Management Division, Environment and Climate Change Canada: 2019.

[25] Julian, P.; Mark, B.; Sarah, M. Food waste within food supply chains: quantification and potential for change to 2050. *Philosophical Transactions of the Royal Society B: Biological Sciences* **2010**, *365*, 3065-3081. DOI: 10.1098/rstb.2010.0126.

[26] Iheonye, A.; Gariepy, Y.; Raghavan, V. Computer vision for real-time monitoring of shrinkage for peas dried in a fluidized bed dryer. *Drying Technology* **2019**, 1-17. DOI: 10.1080/07373937.2019.1649277.

[27] United Nations *World Population Prospects 2019: Highlights;* United Nations: New York, 2019.

[28] Willenbockel, D. In *Scenarios for global agriculture and food security towards 2050: A review of recent studies;* New Challenges to Food Security; Routledge: 2014; pp 55-76.

[29] Rachid Serraj; Prabhu Pingali *Agriculture & food systems to 2050;* World Scientific Publishing Co Pte Ltd: Singapore, 2019.

[30] FAO. The future of food and agriculture – Alternative pathways to 2050. *Food and Agriculture Organization of the United Nations (FAO)* **2018**, 224.

[31] Mayor, L.; Moreira, R.; Sereno, A. M. Shrinkage, density, porosity and shape changes during dehydration of pumpkin (Cucurbita pepo L.) fruits. *Journal of Food Engineering* **2011**, *103*, 29-37. DOI: 10.1016/j.jfoodeng.2010.08.031.

[32] Demirhan, E.; Ozbek, B. Color Change Kinetics of Microwave-Dried Basil. *Drying Technology* **2009**, *27*, 156-166. DOI: 10.1080/07373930802566101.

[33] Nicolas, J. J.; Richard-Forget, F. C.; Goupy, P. M.; Amiot, M. J.; Aubert, S. Y. Enzymatic browning reactions in apple and apple products. *Critical reviews in food science and nutrition* **1994**, *34*, 109-157. DOI: 10.1080/10408399409527653.

[34] Oliveira, S. M.; Brandão, T. R. S.; Silva, C. L. M. Influence of Drying Processes and Pretreatments on Nutritional and Bioactive Characteristics of Dried Vegetables: A Review. *Food engineering reviews* **2015**, *8*, 134-163. DOI: 10.1007/s12393-015-9124-0.

[35] Gunasekaran, S. Computer vision technology for food quality assurance. *Trends in food science & technology* **1996**, *7*, 245-256. DOI: 10.1016/0924-2244(96)10028-5.

[36] Du, C. J.; Sun, D. W. Recent developments in the applications of image processing techniques for food quality evaluation. *Trends in Food Science & Technology* **2004**, *15*, 230-249. DOI: 10.1016/j.tifs.2003.10.006.

[37] Chen, Y.; Sun, D.; Cheng, J.; Gao, W. Recent Advances for Rapid Identification of Chemical Information of Muscle Foods by Hyperspectral Imaging Analysis. *Food engineering reviews* **2016**, *8*, 336-350. DOI: 10.1007/s12393-016-9139-1.

[38] Du, C. J.; Sun, D. W. Learning techniques used in computer vision for food quality evaluation: a review. *Journal of Food Engineering* **2006**, *72*, 39-55. DOI: 10.1016/j.jfoodeng.2004.11.017.

[39] Brosnan, T.; Sun, D. Improving quality inspection of food products by computer vision – a review. *Journal of Food Engineering* **2004**, *61*, 3-16.

[40] Patel, K.; Kar, A. Heat pump assisted drying of agricultural produce—an overview. *Journal of food science and technology* **2012**, *49*, 142-160. DOI: 10.1007/s13197-011-0334-z.

[41] Cubero, S.; Aleixos, N.; Molto, E.; Gomez-Sanchis, J.; Blasco, J. Advances in Machine Vision Applications for Automatic Inspection and Quality Evaluation of Fruits and Vegetables. *FOOD AND BIOPROCESS TECHNOLOGY* **2011**, *4*, 487-504. DOI: 10.1007/s11947-010-0411-8.

[42] Zareiforoush, H.; Minaei, S.; Alizadeh, M. R.; Banakar, A. Potential Applications of Computer Vision in Quality Inspection of Rice: A Review. *Food engineering reviews* **2015**, *7*, 321-345. DOI: 10.1007/s12393-014-9101-z.

[43] Guine, R.; Barroca, M. J. Effect of drying treatments on texture and color of vegetables (pumpkin and green pepper). *Food and Bioproducts Processing* **2012**, *90*, 58-63. DOI: 10.1016/j.fbp.2011.01.003.

[44] Krokida, M. K.; Tsami, E.; Maroulis, Z. B. Kinetics on color changes during drying of some fruits and vegetables. *Drying Technology* **1998**, *16*, 667-685. DOI: 10.1080/07373939808917429.

[45] Zielinska, M.; Markowski, M. Color Characteristics of Carrots: Effect of Drying and Rehydration. *International journal of food properties* **2012**, *15*, 450-466. DOI: 10.1080/10942912.2010.489209.

[46] Aral, S.; Beşe, A. V. Convective drying of hawthorn fruit (Crataegus spp.): Effect of experimental parameters on drying kinetics, color, shrinkage, and rehydration capacity. *Food Chemistry* **2016**, *210*, 577-584.

[47] Huang, M.; Wang, Q.; Zhang, M.; Zhu, Q. Prediction of color and moisture content for vegetable soybean during drying using hyperspectral imaging technology. *Journal of Food Engineering* **2014**, *128*, 24-30. DOI: //doi.org/10.1016/j.jfoodeng.2013.12.008.

[48] Priddy, K. L.; Keller, P. E. *Artificial Neural Networks: An Introduction;* SPIE: .

[49] Nahimana, H.; Zhang, M. Shrinkage and color change during microwave vacuum drying of carrot. *Drying Technology* **2011**, *29*, 836-847. DOI: 10.1080/07373937.2011.573753.

[50] Fathi, M.; Fathi, M.; Mohebbi, M.; Mohebbi, M.; Razavi, S.; Razavi, S. Application of Image Analysis and Artificial Neural Network to Predict Mass Transfer Kinetics and Color Changes of Osmotically Dehydrated Kiwifruit. *Food and Bioprocess Technology* **2011**, *4*, 1357-1366. DOI: 10.1007/s11947-009-0222-y.

[51] Zenoozian, M. S.; Devahastin, S.; Razavi, M. A.; Shahidi, F.; Poreza, H. R. Use of artificial neural network and image analysis to predict physical properties of osmotically dehydrated pumpkin. *Drying Technology* **2008**, *26*, 132-144. DOI: 10.1080/07373930701781793.

[52] Romano, G.; Argyropoulos, D.; Nagle, M.; Khan, M. T.; Muller, J. Combination of digital images and laser light to predict moisture content and color of bell pepper simultaneously during drying. *Journal of Food Engineering* **2012**, *109*, 438-448. DOI: 10.1016/j.jfoodeng.2011.10.037.

[53] Martynenko, A. I.; Davidson, V. J.; Brown, R. B. Intelligent computer vision system (SAIF) for automated inspection of ginseng root quality. Presented at the CSAE/SCGR 2005 Meeting, Winnipeg, Manitoba, 2005.

[54] Hosseinpour, S.; Rafiee, S.; Aghbashlo, M.; Mohtasebi, S. S. A novel image processing approach for in-line monitoring of visual texture during shrimp drying. *Journal of Food Engineering* **2014**, *143*, 154-166. DOI: 10.1016/j.jfoodeng.2014.07.003.

[55] Nadian, M. H.; Nadian, M. H.; Abbaspour-Fard, M. H.; Sadrnia, H.; Golzarian, M. R.; Tabasizadeh, M.; Martynenko, A.; Martynenko, A. Improvement of kiwifruit drying using

computer vision system (CVS) and ALM clustering method. *Drying Technology* **2017**, *35*, 709-723. DOI: 10.1080/07373937.2016.1208665.

[56] Saldaña, E.; Siche, R.; Huamán, R.; Luján, M.; Castro, W.; Quevedo, R. Computer vision system in real-time for color determination on flat surface food. *Scientia Agropecuaria* **2013**, *4*, 55-63. DOI: 10.17268/sci.agropecu.2013.01.06.

[57] Gonzalez, R. C.; Woods, R. E. *Digital image processing;* Pearson: New York, NY, 2018.

[58] Gonzalez, R. C.; Woods, R. E.; Eddins, S. L. *Digital image processing using MATLAB;* Gatesmark Publ: Natick, Mass, 2009.

[59] Buishand, J. G.; Gabelman, W. H. Investigations on the inheritance of color and carotenoid content in phloem and xylem of carrot roots (Daucus carota L.). *Euphytica* **1979**, *28*, 611-632. DOI: 10.1007/BF00038928.

[60] Zenoozian, M. S.; Devahastin, S. Application of wavelet transform coupled with artificial neural network for predicting physicochemical properties of osmotically dehydrated pumpkin. *Journal of Food Engineering* **2009**, *90*, 219-227. DOI: 10.1016/j.jfoodeng.2008.06.033.

[61] Gao, X.; Tan, J. Analysis of expanded-food texture by image processing .1. Geometric properties. *Journal of Food Process Engineering* **1996**, *19*, 425-444. DOI: 10.1111/j.1745-4530.1996.tb00403.x.

[62] Možina, M.; Tomaževič, D.; Leben, S.; Pernuš, F.; Likar, B. Digital imaging as a process analytical technology tool for fluid-bed pellet coating process. *European Journal of Pharmaceutical Sciences* **2010**, *41*, 156-162. DOI: 10.1016/j.ejps.2010.06.001.

[63] Kucheryavski, S.; Esbensen, K. H.; Bogomolov, A. Monitoring of pellet coating process with image analysis—a feasibility study. *Journal of Chemometrics* **2010**, *24*, 472-480. DOI: 10.1002/cem.1292.

[64] Tousey, M. D. The granulation process 101: basic technologies for tablet making. *Pharmaceutical technology (2003)* **2002**, *26*, S8.

[65] Chopra, R.; Alderborn, G.; Podczeck, F.; Newton, J. M. The influence of pellet shape and surface properties on the drug release from uncoated and coated pellets. *International journal of pharmaceutics* **2002**, *239*, 171-178. DOI: 10.1016/S0378-5173(02)00104-7.

[66] Chan, L. W.; Chan, W. Y.; Heng, P. W. S. An improved method for the measurement of colour uniformity in pellet coating. *International journal of pharmaceutics* **2001**, *213*, 63-74. DOI: 10.1016/S0378-5173(00)00619-0.

[67] Watano, S.; Miyanami, K. Image processing for on-line monitoring of granule size distribution and shape in fluidized bed granulation. *Powder Technology* **1995**, *83*, 55-60. DOI: 10.1016/0032-5910(94)02944-J.

[68] Mozina, M.; Burmen, M.; Tomazevic, D.; Pernus, F.; Likar, B. Automatic visual inspection of pharmaceutical pellets in coating process. Presented at the 2009.

[69] Ronneberger, O.; Fischer, P.; Brox, T. In *U-Net: Convolutional Networks for Biomedical Image Segmentation;* Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015; Springer International Publishing: Cham, 2015; pp 234-241.

[70] Hao Chen; Xiaojuan Qi; Lequan Yu; Pheng-Ann Heng DCAN: Deep Contour-Aware Networks for Accurate Gland Segmentation. Presented at the 2016.

[71] Chen, H.; Qi, X.; Yu, L.; Dou, Q.; Qin, J.; Heng, P. DCAN: Deep contour-aware networks for object instance segmentation from histology images. *Medical Image Analysis* **2017**, *36*, 135-146. DOI: 10.1016/j.media.2016.11.004.

[72] Liang-Chieh Chen; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A. L. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2018**, *40*, 834-848. DOI: 10.1109/TPAMI.2017.2699184.

[73] Kaiming He; Xiangyu Zhang; Shaoqing Ren; Jian Sun Deep Residual Learning for Image Recognition. Presented at the 2016.

[74] Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. **2016**.

[75] Cornell University Library, arXiv.org. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv.org* **2015**.

[76] Szegedy, C.; Wei Liu; Yangqing Jia; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. Presented at the Jun 2015.

[77] Chollet, F. Xception: Deep Learning with Depthwise Separable Convolutions. Presented at the Jul 2017.

[78] Russell, B. C.; Torralba, A.; Murphy, K. P.; Freeman, W. T. LabelMe: A Database and Web-Based Tool for Image Annotation. *International journal of computer vision* **2007**, *77*, 157-173. DOI: 10.1007/s11263-007-0090-8.

[79] Watano, S. Direct control of wet granulation processes by image processing system. *Powder Technology* **2001**, *117*, 163-172. DOI: 10.1016/S0032-5910(01)00322-9.

[80] Madiouli, J.; Sghaier, J.; Orteu, J.; Robert, L.; Lecomte, D.; Sammouda, H. Non-contact measurement of the shrinkage and calculation of porosity during the drying of banana. *Drying Technology* **2011**, *29*, 1358-1364. DOI: 10.1080/07373937.2011.561460.

[81] Ramos, I. N.; Miranda, J. M. R.; Brandão, T. R. S.; Silva, C. L. M. Estimation of water diffusivity parameters on grape dynamic drying. *Journal of Food Engineering* **2010**, *97*, 519-525. DOI: //doi.org/10.1016/j.jfoodeng.2009.11.011.

[82] Campos-Mendiola, R.; Hernández-Sánchez, H.; Chanona-Pérez, J. J.; Alamilla-Beltrán, L.; Jiménez-Aparicio, A.; Fito, P.; Gutiérrez-López, G. F. Non-isotropic shrinkage and interfaces during convective drying of potato slabs within the frame of the systematic

approach to food engineering systems (SAFES) methodology. *Journal of Food Engineering* **2007**, *83*, 285-292. DOI: 10.1016/j.jfoodeng.2007.02.027.

[83] Saadevandi, B. A.; Turton, R. The application of computer-based imaging to the measurements of particle velocity and voidage profiles in a fluidized bed. *Powder Technology* **1998**, *98*, 183-189. DOI: 10.1016/S0032-5910(98)00056-4.

[84] Martynenko, A. I. Computer Vision System for Ginseng Drying: Remote Sensing, Control and Optimization of Quality in Food Thermal Processing. Presented at the ADM Verlag, Saarbrücken, 2008.

[85] Martynenko, A. Computer-vision control system for ginseng drying, PhD Thesis, ProQuest Dissertations Publishing, 2006.

[86] FAO. World Hunger Statistics. *http://www.foodaidfoundation.org/world-hunger-statistics.html http://www.foodaidfoundation.org/world-hunger-statistics.html*2019).

[87] Aurélien Géron *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow;* O'Reilly Media: Sebastopol, 2019.

[88] Kingma, D. P.; Ba, J. Adam: A Method for Stochastic Optimization. **2014**.

[89] Pabst, W.; Gregorová, E. *Characterization of particles and particle systems;* ICT: Prague, 2007.

[90] León, K.; Mery, D.; Pedreschi, F.; León, J. Color measurement in L ∗ a ∗ b ∗ units from RGB digital images. *Food research international* **2006**, *39*, 1084-1091. DOI: 10.1016/j.foodres.2006.03.006.

[91] Mann, P. S. *Introductory statistics;* Wiley: Hoboken, NJ, 2011.

[92] Chua, K. J.; Hawlader, M. N. A.; Chou, S. K.; Ho, J. C. On the study of time-varying temperature drying - Effect on drying kinetics and product quality. *Drying technology* **2002**, *20*, 1559-1577. DOI: 10.1081/DRT-120014052.

[93] Lopez, A. (Universidad Publica de Navarra, Pamplona,Spain.); Pique, M. T.; Boatella, J.; Romero, A.; Ferran, A.; Garcia, J. Influence drying conditions on the hazelnut quality. III. Browning. *Drying technology* **1997**, *15*, 989-1002. DOI: 10.1080/07373939708917273.

[94] Kaveh, M.; Jahanbakhshi, A.; Abbaspour-Gilandeh, Y.; Taghinezhad, E.; Moghimi, M. B. F. The effect of ultrasound pre-treatment on quality, drying, and thermodynamic attributes of almond kernel under convective dryer using ANNs and ANFIS network. *Journal of food process engineering* **2018**, *41*, e12868-n/a.

[95] Kaur, P.; Zalpouri, R.; Singh, M.; Verma, S. Process optimization for dehydration of shelled peas by osmosis and three-stage convective drying for enhanced quality. *Journal of food processing and preservation* **2020**, *44*, n/a. DOI: 10.1111/jfpp.14983.

[96] Contreras, C.; Martín-Esparza, M. E.; Chiralt, A.; Martínez-Navarrete, N. Influence of microwave application on convective drying: Effects on drying kinetics, and optical and mechanical properties of apple and strawberry. *Journal of food engineering* **2008**, *88*, 55-64. DOI: 10.1016/j.jfoodeng.2008.01.014.

[97] Niamnuy, C.; Devahastin, S.; Soponronnarit, S. Effects of Process Parameters on Quality Changes of Shrimp During Drying in a Jet-Spouted Bed Dryer. *Journal of food science* **2007**, *72*, E553-E563. DOI: 10.1111/j.1750-3841.2007.00516.x.

[98] Kwok, K. C.; MacDougall, D. B.; Niranjan, K. Reaction kinetics of heat-induced colour changes in soymilk. *Journal of food engineering* **1999**, *40*, 15-20. DOI: 10.1016/S0260-8774(99)00031-X.

[99] Zielinska, M.; Zapotoczny, P.; Alves-Filho, O.; Eikevik, T. M.; Blaszczak, W. A multi-stage combined heat pump and microwave vacuum drying of green peas. *Journal of Food Engineering* **2013**, *115*, 347-356. DOI: https://doi.org/10.1016/j.jfoodeng.2012.10.047.