Research Design II
Fruit Classification using Convolutional Neural Networks

Russell Cachia, Mario Mallia Milanes
Institute of Information & Communication Technology
Malta College of Arts, Science & Technology
Corradino Hill
Paola PLA 9032
{russell.cachia.g53991, mario.mallia.milanes, ictar}@mcast.edu.mt

I. Introduction, Positioning, Research Onion

A. Description of Theme and Topic Rational

In the field of computer vision and image processing, the theme of fruit classification is a very interesting area of research, with the increasing advancements in technology and its agricultural importance. Nowadays, the importance of fruit classification is present in multiple different fields, such as dietary monitoring, food processing, and agricultural automation.

With the advancements in Computer Vision and Deep Learning [1], Convolutional Neural Networks (CNNs) have emerged as a game changer in this context, offering a powerful and automated solution for fruit quality control, sorting, and also grading. By making use of the capabilities of CNNs, fruit industry stakeholders can streamline their operations, minimize manual labor, reduce post-harvest losses, and also ensure consistent product quality, ultimately leading to profitability and customer satisfaction.

In today's world, where the visual complexity and uniqueness of fruits present significant challenges, harnessing the power of Convolutional Neural Networks (CNNs) [2] emerges as a promising solution.

B. Positioning and Research Onion

This study positions itself and the researcher at the intersection of computer vision techniques and practical applications which in this case is fruit classification. With the industry ever-expanding, the aim is to improve fruit classification accuracy.

Considering the given nature of the research problem and objectives, a positivist philosophy is best suited. This is due to the study primarily being quantitative, relying on a dataset of 60,000 images for fruit classification. The large-scale quantitative approach aligns with positivism's emphasis on empirical evidence and the scientific method. By adopting a positivist stance, our study aims to uncover objective truths about fruit classification using Convolutional Neural Networks (CNNs), contributing to the empirical knowledge base in this field.

The research approach taken for this study was a deductive one, starting from established theories and then extending the knowledge to classify fruit using CNNs. This deductive approach allows the researcher to formulate hypotheses based on existing knowledge and test them using quantitative analysis of the dataset.

The research strategy employed was experimental research. By forming a hypothesis, experimental research enables the researcher to systematically test the hypothesis through controlled experiments. Experimental research allowed for the control of variables and conditions during the classification process. Since the study relies on quantitative data and aims to evaluate the performance of CNN models, the experimental research provides a suitable framework for conducting rigorous quantitative analysis.

For the time horizon of this study, a cross-sectional approach was adopted, collecting data from a sample of fruit images. This approach offers valuable insights into the performance and applicability of CNNs in real-world fruit classification scenarios.

In terms of data collection methodology, the researcher adhered to a mono method. The use of the mono method ensures consistency and reliability in data collection, maintaining uniformity throughout the study's execution. Employing such techniques enhances the credibility of the dataset, finally leading to strengthening the research outcomes without repeating information unnecessarily.

C. Background to this Research Theme

Diving into the background of Fruit Classification by using Convolutional Neural Networks is of utmost importance to highlight the evolution of computer vision techniques and their application in real-world situations. Fruit classification using Convolutional Neural Networks has gained significant attention in recent years due to its potential applications in agriculture, food processing, and computer vision. CNNs have proven to be effective in image classification tasks, leveraging their ability to automatically learn and extract meaningful features from raw image data. The architecture of CNNs, consisting of convolutional, pooling, and fully connected layers, allows them to capture spatial hierarchies and patterns present in fruit images [2]. Various studies have explored the use of CNNs for fruit classification, employing different methodologies and data sets. Previous studies have shown high levels of accuracy when it comes to testing these CNNs.

In this study the researcher intends to provide an overview of some existing studies on fruit classification using CNNs. By analysing the strengths and limitations of these studies, the researcher seeks to identify areas for further improvement and future directions in fruit classification using CNNs. These studies are as follows:

1) Xiong et al. developed an algorithm for fruit-picking robots using artificial illumination to recognize grapes. Their system's accuracy was 92% [3].
2) Another study conducted by Katarzyna and Pawel performed fruit classification for supermarket retail sales systems. They suggested using a 9-layered deep neural network to classify six different apple cultivars. They claimed 99.78% accuracy [4].
3) In another study, Kumari and V. Gomathy have classified fruits using color and texture characteristics. To extract the region of interest, they used HSV color space thresholding. After using a three-level discrete wavelet transform, they then retrieved color data from the hue and saturation channels and texture information from the luminance channel. They used an SVM classifier to categorize 10 fruit types from the supermarket produce data set. Their method was 95.3% accurate [5].
4) A CNN based classifier was developed by Sakib, Ashrafi and Siddique which was able to recognize 25 classes of fruits on Fruits 360 data set with 100% test and 99.79% train accuracies [6].

TABLE I
STATE OF THE ART RESULTS

#	Study	Citation	Metrics	Results
1	AI Robots recognizing grapes	[3]	Accuracy	92%
2	Fruit Classification in Retail Systems	[4]	Accuracy	99.78%
3	HSV Thresholding for Fruit Classification	[5]	Accuracy	95.3%
4	CNN based Classifier	[6]	Accuracy	99.79%

D. Hypothesis

The objective of this study is to explore the application of computer vision and machine learning techniques in fruit classification. The primary focus is on assessing the accuracy of fruit evaluation, with the choice of datasets and algorithms serving as independent variables to analyze their influence on the accuracy of classification. Therefore, the hypothesis implies that through the utilization of computer vision technologies, it is viable to achieve precise identification of various types of fruits.

E. Research Aim and Purpose Statement

The primary aim of the study is to explore the potential of computer vision techniques in improving fruit classification accuracy. Monitoring the impact of different datasets and algorithms on the performance of CNN models for fruit classification. By evaluating these variables the study will assess the feasibility and effectiveness of these computer vision techniques.

The feature extraction process is important in order to obtain information in data. This process is to determine the performance of the machine learning [7]. By focusing on the accuracy of fruit evaluation as the main dependent variable, the study aims to understand how the choice of datasets and algorithms influences classification performance. Through rigorous analysis, the study seeks to validate the hypothesis. This study addresses two specific research questions:

1) How does the choice of data set impact the performance of CNN models for fruit classification?
2) How much of a difference do layers create with regard to accuracy?

By addressing these research questions the researcher aims to contribute valuable insights to the field of automated fruit classification and help inform future research and projects in of this domain.

II. REVIEW OF RESEARCH METHODOLOGIES AND MAP

A. Short Literature Review about the Methodologies Used in Other Studies

Upon viewing different relevant literature to the topic at hand, various approaches can be used.

1) **CNN-Based Models**
Several studies have leveraged Convolutional Neural Networks (CNNs) for fruit classification. These models have shown high accuracy in categorizing fruits as "fresh" or "rotten," with the proposed CNN model achieving an accuracy of 98.23% [8].
2) **Transfer Learning Models**
Transfer learning models like AlexNet, LeNet-5, VGG-16, and VGG-19 have been compared with CNN-based models for fruit categorization. The results indicate that the proposed CNN model outperforms transfer learning approaches due to achieving a classification accuracy of 98.2% [8].
3) **Support Vector Machines**
This paper reviews various fruit identification methods, including the use of SVMs. In this study, a method that uses multi-feature fusion to identify five kinds of fruits was presented. The authors used global histogram, LBP, HOG, and GaborLBP for feature extraction and managed to achieve 81.35% accuracy using LibSVM [9].

B. Difference between Academic and Non-Academic Material

The difference between academic and non-academic material lies in their intended audience and purpose.

Academic material is aimed towards scholarly audiences, featuring formal language and a structured format, it is based on facts, evidence, and always contains citations [10]. Authored by experts in the field, these materials undergo a rigorous peer-review process to ensure accuracy and contribute to the field's knowledge base. Common forms include academic journals, textbooks, and conference papers.

In contrast, non-academic material is written for a general audience, employing a more informal, personal, and subjective tone and less structure. These materials, which can include news articles, blogs, and magazines, are typically written by journalists or enthusiasts without comprehensive citations or a peer-review process. Their primary purpose is to inform or entertain rather than to fulfill scholarly criteria [10].

C. Recommendation of 5 articles from Peer-Reviewed Journals

In order to dive deeper and grasp a better understanding of Fruit Classification, the researcher has selected five articles from peer-reviewed journals. These articles have been carefully chosen given their relevance and contribution to the subject. By understanding these sources, the researcher aims to enrich his understanding which eventually leads to a more accurate Fruit Classification model.

1) Fruit Maturity Classification Using Convolutional Neural Networks Method. [11]
2) Implementation of Fruits Recognition Classifier using Convolutional Neural Network Algorithm for Observation of Accuracies for Various Hidden Layers. [12]
3) Application of Convolutional Neural Network-Based Detection Methods in Fresh Fruit Production: A Comprehensive Review. [13]
4) Fruit Disease Classification using Convolutional Neural Network. [14]
5) Fruits Classification using Convolutional Neural Network. [15]

D. Contextualised Literature and Research Material with Critical Literature Arguments

In this section, the researcher provides a small overview of the existing literature and research material used for this study. As already mentioned above, fruit classification has a significant role in the world of agriculture and food processing, which as an industry, play a vital role for us humans. These applications can range from fruit quality assessment to automated production sorting. CNNs, which are one of the many different types of deep learning algorithms, have shown promising results in automating this process as they are very capable into efficiently extracting features from fruit images and classifying them into predefined categories.

Numerous studies have explored the application of CNNs for fruit classification, employing various architectures, training strategies, and dataset sources. For example, Xiong et al. [3] used the Chan-Vese algorithm with morphological processing methods to segment fruit photos. Both daytime and nighttime illumination conditions allowed for the extraction of color characteristics. To locate the fruit harvesting location after grape recognition, they employed the minimal bounding rectangle and Hough Line detection. Their system's accuracy was 92%. This study focuses on green grape detection and picking-point calculation specifically in a night-time natural environment using a CCD vision sensor with artificial illumination. However, the narrow scope of the study limits its applicability to other environments or lighting conditions. It would also be valuable to evaluate the system's performance in various lighting conditions and different types of grape varieties to assess its generalizability.

In another study, Kumari and V. Gomathy [5] have classified fruits using color and texture characteristics. To extract the region of interest, they used HSV color space thresholding. This study lacks sufficient information about the data set used for fruit classification. Apart from this, this study focuses on fruit classification using statistical features, but it does not provide a comprehensive explanation of the specific statistical features employed, such as mean, standard deviation, or any histogram-based features.

Our research aims to address these limitations by dealing with techniques to improve model performance across multiple different fruit categories. By contextualizing the study with the broader landscape of fruit classification research, the researcher aims to contribute valuable insights and methodologies that advance the state-of-the-art in this field.

E. Literature Map

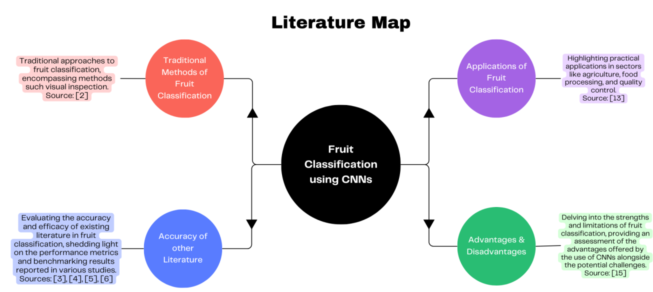

Fig. 1. Literature Map

Figure 1 helps give a better understanding of the literature relevant to this study, by mainly focusing on four aspects; The Traditional Methods of Fruit Classification, The Comparison of Accuracy and metrics of other literature, Applications of Fruit Classification in the field, and lastly The Advantages and Disadvantages of using CNNs to classify fruits.

III. Reflection of the Chosen Methodology

A. Definition of Research Questions

In this study 2 research questions were identified:

1) How does the choice of data set impact the performance of CNN models for fruit classification?
   This research question aims to investigate the influence of the dataset on the performance of Convolutional

Neural Network (CNN) models in the task of fruit classification. Specifically, it seeks to understand how variations in dataset characteristics such as size, diversity, and quality of images, and lighting scenarios all affect the accuracy of CNN models trained for fruit classification. By examining different datasets and their corresponding performance metrics, this question seeks to uncover insights into the optimal dataset choices for achieving high accuracy in fruit classification tasks using CNNs.

2) How much of a difference do layers create with regards to accuracy?
This research question focuses on exploring the impact of CNN architecture, particularly the number of layers, on the accuracy of CNN models for fruit classification. By systematically comparing CNN models with varying numbers of layers and analyzing their corresponding classification performance, this question aims to explain the relationship between network depth and classification accuracy in fruit classification tasks. Understanding the significance of network architecture can provide valuable insights for optimizing CNN models to achieve higher accuracy in fruit classification.

B. Definition of Objectives

The primary objective of this study is to investigate the factors that impact the performance of Convolutional Neural Network (CNN) models in fruit classification tasks. More specifically, this study aims to assess the influence of dataset selection and network architecture on the accuracy and generalization capabilities of CNN models for fruit classification. By systematically examining different studies with datasets that vary in characteristics and network architectures with different depths, this study seeks to identify the optimal combinations of dataset and network architecture that yield the highest accuracy in fruit classification.

Additionally, this study aims to provide insights into the practical implications of these findings for real-world applications of fruit classification. By understanding how dataset choice and network architecture affect classification accuracy, this study aims to offer guidance to researchers and practitioners in selecting appropriate datasets and designing optimal CNN architectures for fruit classification tasks. Ultimately, the objective is to contribute to the advancement of techniques for automated fruit classification, with potential applications in agriculture, food processing, and quality control.

C. Understanding of Research Philosophies, Approaches and Main Research Paradigms

In this study, the researcher adopts a pragmatist research philosophy, recognizing the importance of practical outcomes while still remaining open to different methodologies and perspectives. By embracing pragmatism, the researcher aims to strike a balance between theory and practical applications in fruit classification accuracy.

This study holds both a quantitative and qualitative stance in order to ensure a comprehensive analysis of fruit characteristics and classification algorithms. The quantitative side of this study is due to gathering large amounts of data in the shape of images of different fruits. As shown above, the dataset used in the study was close to around 60,000 images of fruits. On the other hand, the qualitative side of this study involves the interpretation of visual patterns and textures in these fruit images, contributing to a holistic understanding of fruit classification.

The study adopts a positivist research paradigm, emphasizing empirical observation and scientific inquiry. Through the systematic collection and analysis of data, the researcher aims to uncover objective patterns and correlations that facilitate accurate fruit classification.

By integrating these research philosophies, approaches, and paradigms, the study seeks to advance the field of fruit classification using CNNs, contributing to both theoretical knowledge and practical applications in agricultural automation and food industry quality control.

D. Methodology and its Explanation

This section outlines the methodology employed in this literature review, which aims to explore the classification of fruits using Convolutional Neural Networks (CNNs). The increasing demand for automated fruit classification systems has led to the emergence of machine learning techniques as viable solutions. In this study, the primary objective of the researcher is to investigate the effectiveness of CNNs in accurately classifying fruits based on their visual features. The main problem addressed in this study revolves around achieving high accuracy and precision in fruit classification using Convolutional Neural Networks (CNNs), while also considering the size and diversity of the data set. Accurate and efficient fruit classification plays a crucial role in various domains such as agriculture, food processing, and supply chain management.

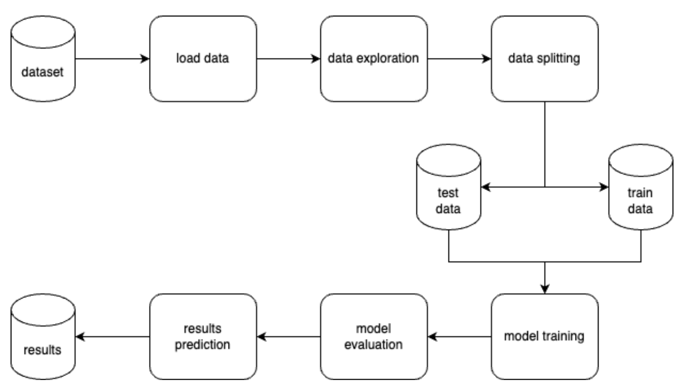

Fig. 2. Research Pipeline

Fig 2 displays the chosen research methodology employed in this study which shall be explained in detail below.

1) **Dataset**

The first step in the research pipeline involves gathering a suitable data set for fruit classification. To complete this, the researcher started from the *Fruits360* data set and then modified it for the research project in hand. Table II exhibits the size and structure of the data set being used in this study, outlining each category of fruit being classified alongside the number of images being used for this CNN. A grand total of 58,380 images were used.

TABLE II
CUSTOM DATASET USED

#	Fruit	Images
1	Apple	6,404
2	Apricot	492
3	Avocado	918
4	Banana	1,430
5	Beetroot	450
6	Blueberry	952
7	Cactus Fruit	490
8	Cauliflower	702
9	Cherry	3,444
10	Chestnut	1,104
11	Clementine	490
12	Coconut	490
13	Corn	912
14	Cucumber	860
15	Dates	490
16	Eggplant	468
17	Fig	702
18	Ginger	297
19	Grape	3,419
20	Grapefruit	982
21	Guava	490
22	Hazelnut	464
23	Kaki	490
24	Kiwi	466
25	Lemon	1,472
26	Mango	916
27	Melon	738
28	Mulberry	492
29	Nectarine	972
30	Onion	1,360
31	Orange	1,459
32	Papaya	492
33	Passion Fruit	490
34	Peach	1,722
35	Pear	5,037
36	Pepper	1,838
37	Pineapple	983
38	Pitahaya Red	490
39	Plum	1,767
40	Pomegranate	492
41	Potato	1,776
42	Raspberry	490
43	Redcurrant	492
44	Strawberry	1,230
45	Tomato	5,103
46	Walnut	735
47	Watermelon	475
	Total	58,380

2) **Load Data**
Once the data set was collected and finalized, the researcher proceeded to load the data into the system. This involves implementing mechanisms to efficiently load and store the images along with their corresponding labels. Proper data organization and management are crucial to ensure smooth processing throughout the pipeline.

3) **Data Exploration**
In this stage, the researcher conducted exploratory data analysis to gain insights into the data set. The researcher analyzed the distribution of fruit categories, identified any class imbalances or biases, and assessed the quality and consistency of the images. This exploration aids in understanding the characteristics and challenges associated with the data set.

4) **Data Splitting**
In this stage, the data was split into 2 parts; 80% of the data was used to train the model, while the other 20% of the data was used for validation. The training set is used to train the model, while the validation set is reserved for evaluating the model's performance. Proper partitioning techniques, such as random sampling and stratified sampling, are employed to ensure representative samples from each fruit category in both subsets. This section was finalized by defining the CNN model as a sequence of respective layers of the CNN and also compiling them for multi-class classification. Certain techniques such as *'Flatten', 'Dropout'* and *'Dense'* were employed before compiling the model.

5) **Model Training**
In this stage, the researcher trained the fruit classification model using a CNN architecture. The training process involves feeding the training data to the model, adjusting the model's internal parameters through an optimization algorithm, and iteratively refining the model's performance. Techniques like transfer learning and fine-tuning were applied to leverage pre-trained models and enhance training efficiency. The model was tested frequently, having a final run with 100 epochs, measuring metrics such as loss, accuracy, validation loss, and validation accuracy for each epoch repetitively.

6) **Model Evaluation**
Once the model was successfully trained, the researcher proceeded to evaluate its performance using the testing data set. The trained model is applied to classify the fruit images in the testing set, and the predictions are compared against the ground truth labels. Evaluation metrics such as loss history (*for validation & training*), accuracy (*for validation & training*), precision, and recall are calculated to assess the model's effectiveness in fruit classification. To enhance the model's robustness and generalization capability, the researcher employed data augmentation techniques during the evaluation process. Data augmentation involves applying various transformations to the training images; such as rotation, scaling, flipping, and cropping, to increase the diversity and variability of the training data. By incorporating augmented data during evaluation, the researcher can obtain more reliable and representative performance estimates of the model's classification ability. Furthermore, the researcher assessed the impact of data augmentation

on the model's performance by comparing the results obtained with and without augmentation. This analysis allows us to determine the effectiveness of data augmentation in mitigating overfitting and improving the model's ability to handle variations and variations in fruit images. The model evaluation stage provides crucial insights into the model's classification accuracy, precision, and overall performance, considering both the original and augmented data sets. By considering these evaluation metrics and analyzing the impact of data augmentation, the researcher gained a comprehensive understanding of the model's capabilities and limitations in fruit classification tasks. Finally, a confusion matrix is drawn up to help visualize the performance of this classification model.

7) **Results Prediction**
With the trained and evaluated model, the researcher can use it to predict the fruit category of unseen or future fruit images. New images were fed into the model, and the model's predictions were obtained. This step showcases the practical applicability of the trained model in real-world scenarios.

8) **Results**
Finally, the researcher analyzed and presented the results obtained from the evaluation and prediction stages. The researcher reported the performance metrics achieved by the model, discussing the accuracy, precision, and other relevant indicators. The results are summarized and visualized in a comprehensive manner into graphs, allowing for a clear understanding of the model's classification capabilities.

E. Reflections on Validity, Reliability & Generalizability/Transferability

To ensure validity in this study involves validating the effectiveness of the classification model in accurately identifying the fruit images. To enhance validity, multiple data preprocessing techniques were employed to ensure the best quality of the images.

To enhance reliability, the researcher implemented standardized training procedures and regularization techniques to mitigate overfitting and improve model generalization.

In the context of fruit classification, achieving generalizability involves demonstrating the efficacy of the classification model across diverse fruit varieties and cultivation practices. To enhance generalizability, the researcher employed data augmentation techniques to augment the training dataset with variations in lighting, orientation, and occlusion, simulating real-world conditions. Additionally, the researcher conducted transfer learning experiments to assess the adaptability of the model to new fruit categories with limited training data. By prioritizing generalizability/transferability, the researcher seeks to maximize the utility and impact of the research beyond the confines of the laboratory, facilitating its adoption in real-world applications such as agricultural automation and food industry quality control.

F. Ethical Considerations for this Study

Different ethical considerations were taken into mind for this study.

1) **Data Privacy**
The dataset used was taken from an online available dataset named Fruits360 and reconstructed for the best-fitting purpose of this study. All the data usage terms of the dataset were respected

2) **Data Integrity and Bias Mitigation**
As just mentioned, the dataset was viewed and reconstructed for this study, redundant and duplicate images were removed, and even categories that weren't relevant were deducted from the dataset. This was done in order to have a more efficient model training process and a dataset which is representative of the target subject and without any bias.

IV. Results, Analysis and Discussion

A. Presentation of Results using Established Metrics

In this section, the researcher will proceed to present the results of the study conducted.

As explained in Section III, during the **Model Evaluation** stage, evaluation metrics such as loss history, accuracy, precision, and recall are calculated and displayed in graph form.

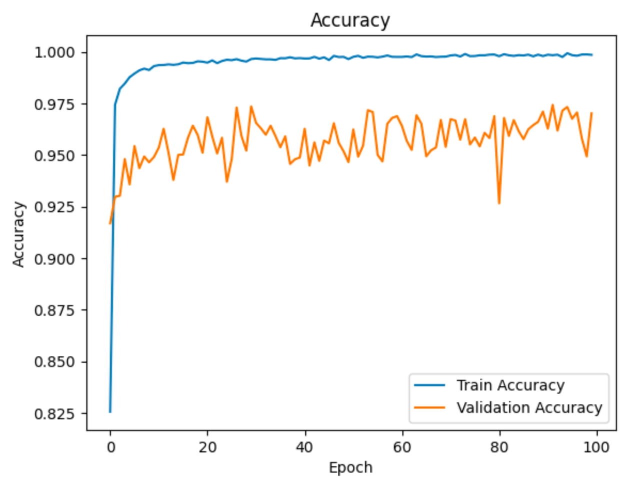

Fig. 3. Accuracy Graph

B. Analysis and Interpretation of Results Presented

The accuracy metric in Figure 3 provides a measure of how well the CNN model has classified the independent fruit categories correctly. The graph reveals a consistent increase in accuracy as the model learned from the training data. It started at a relatively low accuracy level and gradually improved with each epoch, eventually converging to an impressive overall accuracy of 98% on the train data set. This indicates that the model correctly classified 98% of the fruit images, showcasing its ability to classify between different fruit types. The validation data set had a less accurate experience, averaging around 95% accuracy. This upward trend highlights the model's ability

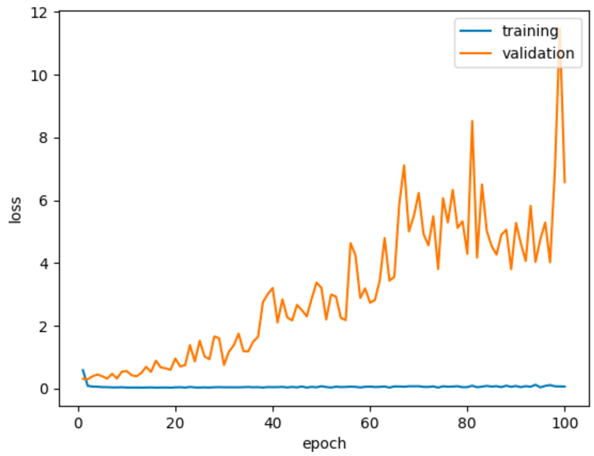

Fig. 4. Loss History Graph

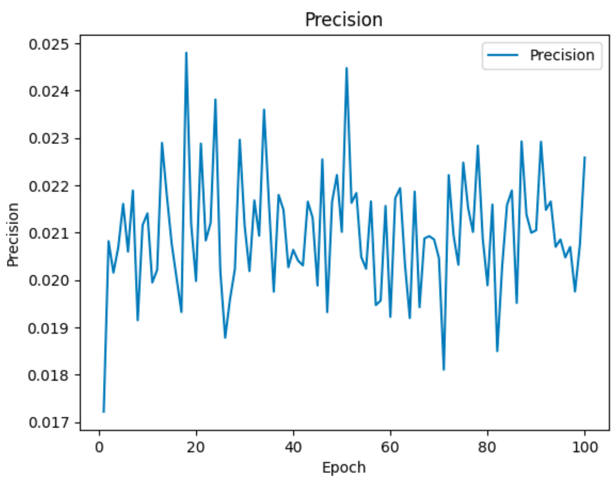

Fig. 5. Precision Graph

to generalize from the training data and accurately classify previously unseen fruit images.

Monitoring the loss history's progress is crucial to understanding how well the CNN model is learning during the training phase. Figure 4 illustrates the loss history, depicting the change in the loss value as the model underwent successive training epochs. The initial loss was relatively high, indicating a significant discrepancy between the predicted and actual fruit labels. However, as the training progressed, the training loss steadily decreased, indicating that the model successfully learned to minimize the divergence between predicted and actual labels. The decline in the loss function demonstrates the model's ability to capture important features and patterns within the fruit images, leading to more accurate classifications.

Figure5 demonstrates the precision rates across different fruit categories. The precision in this context, denotes the accuracy of correct predictions made by the CNN model. After every 20 epochs, the graph represents the precision score

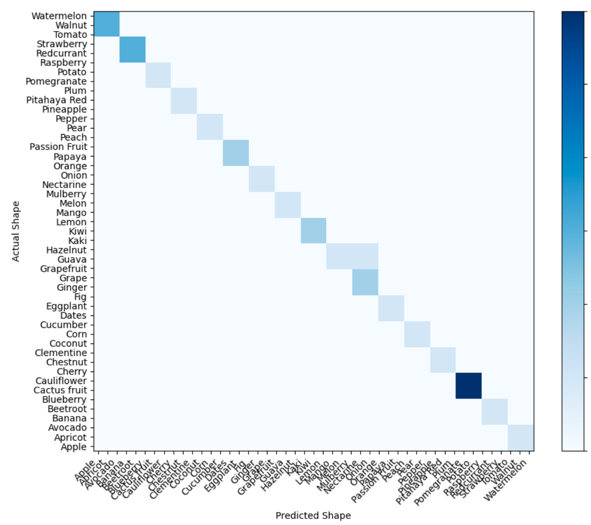

Fig. 6. Confusion Matrix

```
Class Apple: 2.05%
Class Apricot: 2.05%
Class Avocado: 2.05%
Class Banana: 2.05%
Class Beetroot: 2.05%
Class Blueberry: 2.05%
Class Cactus fruit: 2.05%
Class Cauliflower: 2.05%
Class Cherry: 2.05%
Class Chestnut: 2.05%
Class Clementine: 2.05%
Class Coconut: 2.05%
Class Corn: 2.05%
Class Cucumber: 2.05%
Class Dates: 2.05%
Class Eggplant: 2.05%
Class Fig: 2.05%
Class Ginger: 2.05%
Class Grape: 2.05%
Class Grapefruit: 2.05%
Class Guava: 2.05%
Class Hazelnut: 2.05%
Class Kaki: 2.05%
Class Kiwi: 5.58%
Class Lemon: 2.05%
...
Class Tomato: 2.05%
Class Walnut: 2.05%
Class Watermelon: 2.05%
This image most likely belongs to Kiwi with a 5.580 percent confidence.
```

Fig. 7. Confidence List

for specific fruits, illustrating the effectiveness of the model between different fruit categories. The higher the precision score indicates that the model is adept at correctly identifying a particular fruit from its image.

To further evaluate the performance of the fruit classification model, the researcher constructed a confusion matrix. The confusion matrix provides a comprehensive representation of the model's classification results, showing the number of correct and incorrect predictions for each fruit class. Figure 6 presents the confusion matrix for the model's predictions on the test data set. It allows the researcher to assess the specific areas where the model excelled or struggled in fruit classification, providing insights into potential areas for improvement and fine-tuning.

In addition to loss and accuracy, it is crucial to examine the confidence level of the model's predictions. The model's confidence refers to how certain it is about its classifications. Figure 7 displays the confidence distribution of the model's predictions on the test data set. The researcher observed that the model consistently demonstrated a 2.05% confidence for the images that were not present, with a 5.58% confidence on the correct predictions. Understanding the model's confidence can show us that when the correct fruit was classified, the model itself was more than double in confidence for the correct fruit than it was for the incorrect fruit.

C. Comparing & Contrasting with Different Results

In this study on fruit classification using CNNs, the researcher achieved a high accuracy rate of 98% during the training phase and 95% accuracy during the validation phase. These results demonstrate the effectiveness of the model in accurately classifying fruits, alongside its confidence.

When comparing these findings to study [6], who reported implementing a fruit recognition classifier using a convolutional neural network algorithm. Their study achieved remarkable results, with a test accuracy of 100% and a training accuracy of 99.79%. These findings highlight the outstanding performance of their model in accurately classifying fruits.

Although this study's accuracy rates are slightly lower compared to study [6], achieving accuracies of 98% in training and 95% in validation is still indicative of a strong performance in fruit classification. It is important to consider that variations in data sets, model architectures, training configurations, and preprocessing techniques can influence the reported accuracies across different studies.

Some notable differences between the two studies where:

1) The sizes of the data sets used. Both studies started from the *Fruits360* data set and modified it in their own way. This study had a total of 53,800 images while study [6] had a total of 18,000 images.
2) The model being used in study [6] is not heavily specified, but it is mentioned that CNN and deep learning are used, which leaves us to believe that the difference in results, apart from being affected by the data set size, was due to the added layers and filtering which in the end provide a higher accuracy percentage.

D. Discussion of Results in Relation to Original Hypothesis

As stated above in Section I, the original hypothesis of this study was "on assessing the accuracy of fruit evaluation, with the choice of datasets and algorithms serving as independent variables to analyze their influence on the accuracy of classification". When comparing the accuracy percentage alongside the size of the dataset used in this study with other studies, one can confirm that these independent variables do have an impact on the outcome of a study. After comparison with other studies, it became evident that the choice of dataset significantly influences the robustness of the classification model.

V. Conclusion

A. Main Conclusion

B. Addressing of Research Questions and Objectives

C. Identification of Shortcomings

D. Suggestions for Further Research

References

[1] M. Hassaballah and K. M. Hosny, "Recent advances in computer vision," *Studies in computational intelligence*, vol. 804, pp. 1–84, 2019.

[2] D. Bhatt, C. Patel, H. Talsania, J. Patel, R. Vaghela, S. Pandya, K. Modi, and H. Ghayvat, "Cnn variants for computer vision: history, architecture, application, challenges and future scope," *Electronics*, vol. 10, no. 20, p. 2470, 2021.

[3] J. Xiong, Z. Liu, R. Lin, R. Bu, Z. He, Z. Yang, and C. Liang, "Green grape detection and picking-point calculation in a night-time natural environment using a charge-coupled device (ccd) vision sensor with artificial illumination," *Sensors*, vol. 18, no. 4, p. 969, 2018.

[4] R. Katarzyna and M. Paweł, "A vision-based method utilizing deep convolutional neural networks for fruit variety classification in uncertainty conditions of retail sales," *Applied Sciences*, vol. 9, no. 19, p. 3971, 2019.

[5] R. S. S. Kumari and V. Gomathy, "Fruit classification using statistical features in svm classifier," in *2018 4th International Conference on Electrical Energy Systems (ICEES)*. IEEE, 2018, pp. 526–529.

[6] S. Sakib, Z. Ashrafi, M. Siddique, and A. Bakr, "Implementation of fruits recognition classifier using convolutional neural network algorithm for observation of accuracies for various hidden layers," *arXiv preprint arXiv:1904.00783*, 2019.

[7] S. I. C. Myongkyoon Yang, "Fruit classification using convolutional neural network(cnn)," *Precision Agriculture Science and Technology*, pp. 1–8, 2021.

[8] R. Pathak, "Classification of fruits using convolutional neural network and transfer learning models," *Management Information and Decision Sciences*, vol. 21, no. 3, 2021.

[9] F. Liu, L. Snetkov, and D. Lima, "Summary on fruit identification methods: A literature review," *Proceedings of the 2017 3rd International Conference on Economics, Social Science, Arts, Education and Management Engineering (ESSAEME 2017)*, vol. 119, 2017.

[10] D. Vaghani, "Difference between academic and non-academic writing," Mar 2022.

[11] S. Anatya, V. Christanti Mawardi, and J. Hendryli, "Fruit maturity classification using convolutional neural networks method," *IOP Conference Series: Materials Science and Engineering*, vol. 1007, no. 1, p. 012149, Dec 2020.

[12] S. Sakib, Z. Ashrafi, and A. Bakr Sidique, "Implementation of fruits recognition classifier using convolutional neural network algorithm for observation of accuracies for various hidden layers," 2022.

[13] C. Wang, S. Liu, Y. Wang, J. Xiong, Z. Zhang, B. Zhao, L. Luo, G. Lin, and P. He, "Application of convolutional neural network-based detection methods in fresh fruit production: A comprehensive review," *Frontiers in Plant Science*, vol. 13, May 2022.

[14] N. Pradheep, G. K. R. Praveen, M. N. C. Purna, N. Kalaivani, and V. Nandalal, "Fruit disease classification using convolutional neural network," *IEEE Xplore*, 2022.

[15] M. Khatun, F. Ali, N. Aman Turzo, J. Nine, and P. Sarker, "Fruits classification using convolutional neural network," *GRD Journals- Global Research and Development Journal for Engineering*, vol. 5, no. 8, Jul 2020.