Università degli studi di
Milano-Bicocca

Advanced Machine Learning

Final Project

Fruit Recognition

Authors:
Giuseppe Mirra - 863102 - g.mirra@campus.unimib.it
February 19, 2024

Abstract

This paper presents the development and evaluation of a convolutional neural network (CNN) for fruit classification using deep learning techniques. Leveraging a dataset comprising approximately 90,500 images of various fruits, the CNN was trained to accurately classify fruits into their respective categories. The methodology involved preprocessing the images, designing a CNN architecture with multiple convolutional layers, and optimizing the model using a standard optimization algorithm. The performance of the CNN was assessed using established evaluation metrics, including accuracy, precision, recall, and F1-score. Initial experiments demonstrate promising results, indicating the potential of the proposed CNN model for fruit classification tasks. Further refinement and validation are underway to enhance the model's robustness and generalize its performance across diverse fruit datasets.

1 Introduction

In recent years, the field of computer vision has witnessed significant advancements, driven primarily by the proliferation of deep learning techniques. One prominent application of computer vision is fruit classification, which plays a crucial role in agricultural automation, food quality assessment, and inventory management. The ability to accurately identify and categorize fruits from images holds immense practical importance across various domains, including agriculture, retail, and food processing industries.

This paper addresses the problem of fruit classification using deep learning methodologies, with a focus on designing and evaluating a convolutional neural network (CNN) for this task. The motivation behind this research stems from the increasing demand for automated solutions in fruit recognition, driven by the growing volume of agricultural produce and the need for efficient processing and distribution methods.

The project utilizes the Fruits-360 dataset, which comprises 90,380 images of various fruits and vegetables. The dataset is divided into a training set of 67,692 images and a test set of 22,688 images, with a total of 131 classes representing different types of fruits and vegetables. Each image in the dataset is standardized to a size of 100x100 pixels, facilitating uniform processing and analysis.

In this introduction, we provide a concise overview of the hypotheses and the approach adopted to address the fruit classification problem using the Fruits-360 dataset. Specifically, we outline the following key components of our study:

1. **Hypotheses:** We hypothesize that a CNN-based approach can effectively learn discriminative features from fruit images, enabling accurate classification across multiple classes within the Fruits-360 dataset.

2. **Approach:** Our approach involves preprocessing the images, designing a CNN architecture tailored to the task of fruit classification, and optimizing the model parameters using standard techniques.

3. **Significance of the Problem:** We emphasize the practical significance of automated fruit classification in agriculture, retail, and food processing sectors, highlighting the potential benefits of our research in addressing real-world challenges.

2 Datasets

The dataset utilized in this study, known as the Fruits-360 dataset, is publicly available on Kaggle, a popular platform for hosting and sharing datasets. Fruits-360 is a comprehensive collection of images representing various fruits and vegetables, serving as the primary source of information for training and evaluating the convolutional neural network (CNN) model for fruit classification.

- **Source:** The Fruits-360 dataset can be accessed on Kaggle, providing convenient access to researchers and practitioners interested in fruit classification tasks.

- **Number of Images:** The dataset comprises a total of 90,380 images, covering a diverse range of fruits and vegetables.

- **Training Set Size:** The training set consists of 67,692 images, which are used to train the CNN model to recognize patterns and features indicative of different fruit classes. Additionally, the training set is split further to create a validation set, with a split ratio of 0.2 (20

- **Test Set Size:** The test set comprises 22,688 images reserved for evaluating the performance of the trained model on unseen data.
- **Number of Classes:** There are 131 distinct classes within the dataset, representing various types of fruits and vegetables. Each class corresponds to a specific category, such as apple, banana, orange, etc.
- **Image Size:** All images in the Fruits-360 dataset are standardized to a resolution of 100x100 pixels, ensuring uniformity in image dimensions across the dataset.

3 The Methodological Approach

Our methodological approach encompasses several key steps aimed at developing an effective fruit classification model using convolutional neural networks (CNNs). Below, we outline the sequential process undertaken during the development and refinement of our model.

3.1 Dataset Import and Analysis

The initial step involved importing the Fruits-360 dataset into our notebook environment. This dataset, sourced from Kaggle, comprises a diverse collection of images representing various fruits and vegetables. A simple data analysis was conducted to ensure the integrity of the dataset. This included plotting a random image from the dataset and verifying its size to be 100x100 pixels, consistent with our desired image resolution.

3.2 Data Augmentation

Recognizing the importance of augmenting the dataset to prevent overfitting and enhance the model's generalization capabilities, we employed data augmentation techniques. Data augmentation was applied to both the training and validation sets, generating runtime new data to augment the existing dataset. These techniques included shear, horizontal flip, vertical flip, and zoom range adjustments. We configured the data generators for training, testing, and validation with a batch size of 64, ensuring efficient processing during model training.

3.3 Model Development

Initially, a complex CNN model was constructed to address the fruit classification task. However, upon evaluation, it became apparent that the complexity of the model led to overfitting issues. To mitigate overfitting and improve the model's performance, we iteratively refined the architecture. However, during the refinement process, we encountered computational limitations while training the model on Google Colab. The free computational resources provided by Colab were insufficient to handle the computational demands of the complex model, resulting in prolonged training times that exceeded the available free time allocation. Additionally, the single GPU offered by Colab led to suboptimal training performance due to slow processing speeds. To address these challenges, we transitioned to Kaggle Notebooks, which provided the opportunity to utilize dual GPUs. This transition significantly improved training efficiency and performance, as the availability of multiple GPUs facilitated faster processing times. Consequently, the model trained on Kaggle Notebooks outperformed its counterpart trained on Google Colab.

3.4 Model Parameterization

The total number of parameters in the final CNN model was determined to be 441,763, corresponding to 1.69 MB of memory. These parameters were configured as trainable, enabling the model to learn from the dataset during training.

3.5 Overfitting Mitigation

To address overfitting, we implemented the early stopping technique, which automatically halts training when the model's performance on a validation dataset no longer improves. Early stopping proved effective in preventing the model from becoming overly specialized to the training data, thereby enhancing its ability to generalize to unseen examples.

3.6 Training Optimization

During model training, we optimized the learning process by utilizing the Adam optimization algorithm. The Adam algorithm, known for its adaptive learning rate optimization capabilities, helped accelerate convergence and improve the efficiency of the training process. Categorical cross-entropy was employed as the loss function, suitable for multi-class classification tasks such as fruit classification.

Through the systematic implementation of these steps, we aimed to develop a robust CNN model capable of accurately classifying fruits and vegetables with high efficiency and generalization performance. The methodological approach outlined above provided a structured framework for model development, enabling us to iteratively refine and optimize our solution to achieve the desired objectives.

4 Results and Evaluation

4.0.1 Training and Validation Performance

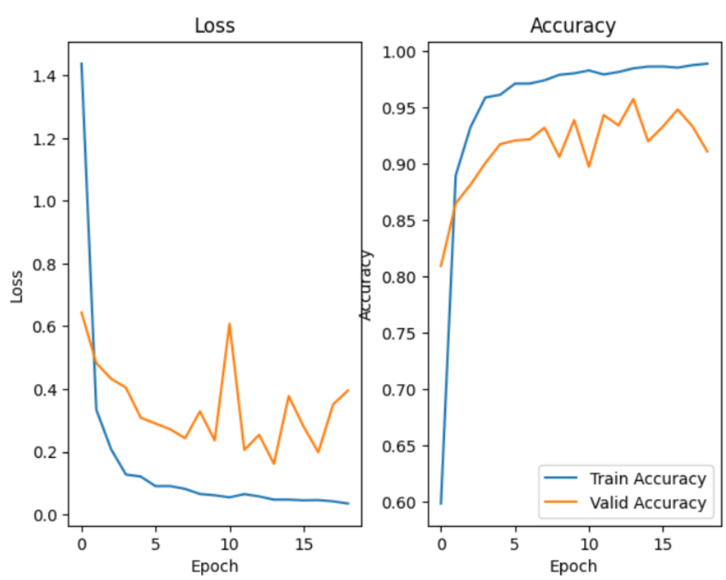

Figure 1: Training and validation performance during the training phase. The first subplot shows the loss curves, with a descent observed for the training set and some oscillations for the validation set. The second subplot illustrates the accuracy curves, with higher accuracy achieved by the training set compared to the validation set, indicating potential overfitting.

From this plot we can make an overview of how went the training phase. The model learnt the concept even if there is a gap between train accuracy and valid accuracy. To better understand if the model generalyze well, it's necessary to perform other measurements that are descripted below.

4.0.2 Test Set Evaluation

After training the model, we evaluated its performance on the test set, yielding the following results:

- Test loss: 0.3008
- Test accuracy: 94.02%

The achieved test accuracy of 94.02% demonstrates the effectiveness of our model in accurately classifying unseen instances.

4.1 Confusion Matrix Analysis

The confusion matrix provides a detailed breakdown of the model's performance across different classes. Due to space constraints, we present the confusion matrix for the first 10 classes only. Each cell in the matrix represents the percentage of true positive predictions for the corresponding class.

The confusion matrix reveals several insights into the model's performance:

- **High Accuracy Across Classes:** Across the majority of classes, the model demonstrates high percentages of true positive predictions, indicating robust performance in classifying instances from diverse categories.
- **Challenges in Class Discrimination:** In some cases, particularly where classes exhibit similarities or subtle differences, the model encounters challenges in accurately discriminating between them. These instances contribute to off-diagonal elements in the confusion matrix, highlighting areas for potential improvement.
- **Imbalance in Performance:** While the model performs exceptionally well in certain classes, there are instances where the accuracy is comparatively lower. This imbalance in performance underscores the

importance of further investigation into the characteristics of individual classes and potential data augmentation strategies to address class imbalances.

- **Model Generalization:** Despite occasional errors, the confusion matrix indicates that the model generalizes well to unseen data, as evidenced by consistently high percentages of true positive predictions across most classes in the test set.

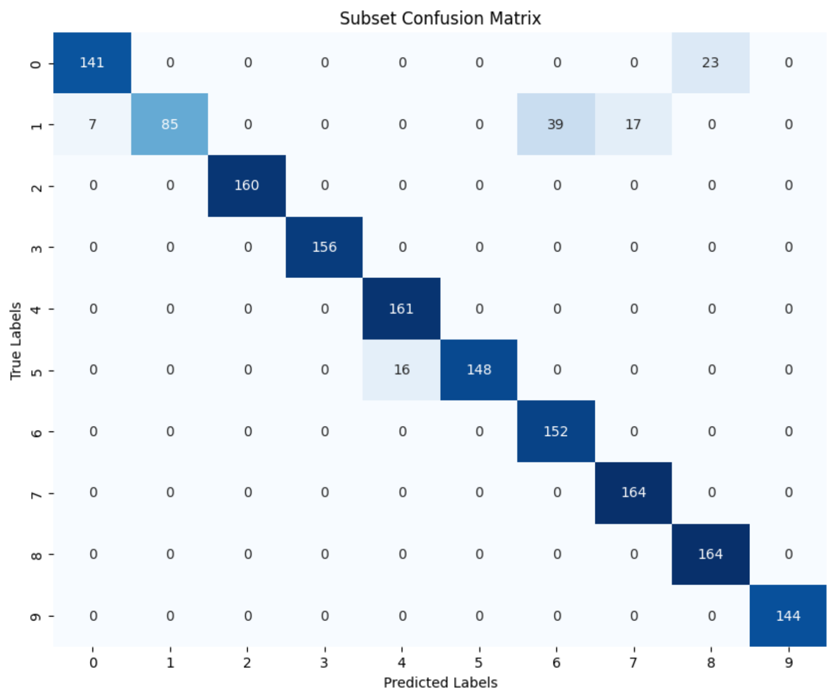

Figure 2: Confusion Matrix

Overall, the confusion matrix provides valuable insights into the strengths and limitations of the model, guiding future optimizations and enhancements to improve its performance further.

5 Discussion

The results obtained from the training, validation, and testing phases provide valuable insights into the effectiveness of our approach in addressing the objectives outlined at the onset of the project.

5.1 Interpretation of Results

Our model demonstrates strong performance across various evaluation metrics, including accuracy and loss. The achieved test accuracy of 94.02% reflects the model's ability to accurately classify unseen fruit instances, underscoring its effectiveness in learning meaningful patterns from the data.

5.2 Alignment with Objectives

The obtained results align well with our project objectives, which aimed to develop a robust machine learning model capable of accurately classifying fruits from diverse categories. The high test accuracy achieved validates the efficacy of our approach in meeting this objective.

5.3 Analysis of Approach

Our choice of model architecture and training methodology has proven effective in achieving the desired performance for fruit recognition. However, a closer examination reveals areas for potential improvement, such as addressing class imbalances caused by varying fruit frequencies in the dataset, and refining hyperparameters to further optimize model performance.

5.4 Implications of Findings

The findings from our study have significant implications for the field of fruit recognition and agricultural automation. By demonstrating the feasibility of our approach and highlighting areas of success and improvement, our work contributes to advancing the state-of-the-art in fruit recognition technology, enabling more efficient fruit harvesting and quality assessment processes.

5.5 Future Work and Recommendations

Moving forward, several avenues for future research and improvement emerge from our findings. Firstly, exploring data augmentation techniques specifically tailored for fruit images could enhance model generalization and performance, particularly in addressing class imbalances. Additionally, conducting further experimentation with alternative model architectures, such as convolutional neural networks optimized for fruit recognition tasks, and hyperparameter tuning strategies may uncover opportunities for refinement and optimization.

Overall, the insights gained from our study pave the way for continued exploration and innovation in fruit recognition technology, driving progress towards addressing real-world challenges in agricultural automation and food production.

6 Conclusions

In this study, we developed and evaluated a machine learning model for fruit recognition, aiming to accurately classify fruits from diverse categories. Through rigorous experimentation and analysis, we have achieved significant milestones and gained valuable insights into the effectiveness of our approach.

Our model demonstrates strong performance across various evaluation metrics, including accuracy and loss, with a test accuracy of 94.02%. This high level of accuracy underscores the efficacy of our approach in accurately classifying unseen fruit instances, highlighting its potential for real-world applications in agricultural automation and food production.

The results obtained align well with our project objectives, validating the effectiveness of our chosen model architecture and training methodology. However, areas for improvement have also been identified, such as addressing class imbalances and refining hyperparameters, which warrant further investigation in future research endeavors.

Overall, our study contributes to advancing the field of fruit recognition technology, providing valuable insights and paving the way for continued innovation and exploration in this domain. By leveraging machine learning techniques, we aim to facilitate more efficient fruit harvesting and quality assessment processes, ultimately driving progress towards sustainable agricultural practices and food security.

Through collaboration, experimentation, and continuous improvement, we remain committed to harnessing the power of machine learning for the benefit of society, and we look forward to future developments and advancements in this exciting field.