|  |  |
| --- | --- |
| Module Code | CN7023 |
| Module Title | Artificial Intelligence and Machine Vision |
| LSBF Coursework Instructor | Mr. Alvin Yeo |
| Coursework Title | Neural Network & Image Processing Classification Project (3000 words) Certificates related to AI and Machine Vision Group presentation based on approach to the main project. |
| Weighting | 100% |
| Handout Date | 20 JUNE 2024 |
| Coursework Submission Date | 02 SEPTEMBER 2024 |
| Students’ Name & Students’ ID | |  |  | | --- | --- | | Student Name(s) | Student ID Number(s) | | Kumarage Nimsha Erin | u2781408 | | Yang Yi | u2781423 | | Tomiris Mukhamedgaliyeva | u2781411 | |
| Documents uploaded in LSBF Canvas: | ⚪ Assignment Report  ⚪ Presentation Report  ⚪ Source Code, Scripts & Output  ⚪ Turnitin Report |
| UEL Turnitin Submission Similarity Index % | 13% |
|  | |

A document with signature on it

Description automatically generated

NEURAL NETWORK & IMAGE PROCESSING CLASSIFICATION FOR GRAPES

Kumarage Nimsha Erin: u2781408

Yang Yi: u2781423

Tomiris Mukhamedgaliyeva: u2781411

**Abstract**

This report presents a system designed to accurately distinguish different types of grapes, such as Concord (blue grape), Thompson Seedless (white grape), and Red Globe (pink grape) using Convolutional Neural Networks (CNNs). The AlexNet architecture was utilized to analyse and classify images. Through rigorous training and testing, the model achieved an overall classification accuracy of 97**%**.

1. **Introduction**

The objective of this report is to accurately distinguish between different types of grapes, categorizing them into 3 groups (blue grape, white grape, pink grape) using advanced image classification techniques. This report is structured into five sections: introduction, methodology, simulations, results, and conclusion. The methodology section will discuss different neural network architectures explored and used for classification, examining reasons why some proved more suitable than others. Ultimately, AlexNet was chosen. The simulations section will delve into the details of how AlexNet was applied. In the end, we will evaluate the results obtained and discuss the conclusions drawn from the analysis.

**2. Methodology**

In this study, we observe the advancements of Convolutional Neural Network (CNN) architectures (AlexNet, VGG, Inception, ResNet and EfficientNet). Various research articles were reviewed to identify the best architecture for classification system and following insights were obtained:

1) *AlexNet*

In vegetables image classification study, AlexNet achieved an accuracy of 92.1%, showcasing its robustness in handling image classification tasks (Zhu et al., 2018). Additionally, another study on fruit freshness categorization showed that AlexNet achieved rising accuracy rates of 98.2%, 99.8%, and 99.3% across three datasets (Amin et al., 2023).

2) *VGG*

VGG was considered due to its straightforward and uniform architecture. Its deeper layers can potentially lead to higher classification accuracy compared to AlexNet (Xu et al., 2024). Moreover, VGG have been shown to perform well in various image classification tasks. A study based on Vegetable Recognition and Classification demonstrated that VGG achieved high accuracy rate of 96% (Li et al., 2020).

3) *Inception (GoogleNet)*

GoogleNet has been prone to overfitting when the training dataset is limited, and its large number of parameters makes it challenging to apply in environments with limited resources (Xu et al., 2024). Inception has demonstrated high performance in various classification tasks. In a study related to fruit and vegetable classification, Inception achieved testing accuracy of 96.00%, underscoring its capability in handling image classification tasks (Yuesheng et al., 2021). However, despite its successes, it was not chosen due to its tendency to overfit.

4) *ResNet*

Although ResNet's residual learning approach makes it highly effective for very deep networks, we determined not to go with ResNet because study on fruit freshness classification showed that simpler models like AlexNet provided satisfactory accuracy, making ResNet's complexity unnecessary for our task (Amin et al., 2023). Even with its high performance in studies, ResNet is impractical for our specific use case.

5) *EfficientNet*

A study involving the classification of tomato fruit on the vine into ripe, immature, and damaged categories demonstrated the effectiveness of one of the newest models, EfficientNet, specifically EfficientNet-B0. The same study highlighted EfficientNet-B0's ability to deliver a high accuracy (97%) with reasonable resource usage, indicating its potential for high-accuracy image classification tasks under resource constraints (Phan et al., 2023). EfficientNet-B0 seems to be a better choice over other variants because it offers a good balance between performance and computational efficiency. While larger EfficientNet variants might offer slightly better accuracy, their increased computational cost and complexity are not justified for our case.

After thorough evaluation, it was decided to implement AlexNet, VGG (16 and 19), and EfficientNet-B0:

1) VGG: Despite its potential for high accuracy, VGG was ultimately not chosen due to its long training time and higher computational demands. Both VGG-16 and VGG-19 were attempted but found less practical in comparison to others.

2) EfficientNet-B0: Selected for its balance between accuracy and computational efficiency, and as one of the newest models, it was expected to perform well. However, practical implementation showed it took a substantial amount of time, and the computer overheated, leading to shutdowns.

3) AlexNet: AlexNet provided slightly better results. Chosen for its speed and simplicity, as it demonstrated a better performance (~0.5% higher accuracy) than EfficientNet-B0 over 10 epochs. Given the high accuracy (90+%) achieved by both architectures, AlexNet was preferred due to its faster training times and lower resource consumption.

As a result, after evaluating several architectures, we have determined that AlexNet is the most suitable.

AlexNet has been utilized for both classification and combined classification and regression tasks. When used for regression, the same architecture struggles because it is not designed to predict continuous values from high-dimensional image data. Consequently, AlexNet configured for classification performs significantly better on our dataset than when it is adapted for regression. It is essential to select neural network architectures that align with the specific nature of the task. AlexNet shines in classification tasks due to its ability to handle discrete categories, such as different types of grapes. Therefore, this report will focus solely on exploring classification using AlexNet, without incorporating regression.

**3. Simulations**

This section will provide a comprehensive explanation of selected dataset, AlexNet architecture, process of training, validation, and testing, learning algorithm used, and data encoding.

**3.1. Dataset description**

The grapes dataset consists of 1968 images, focusing on three distinct types of grapes. Each category is carefully organized into separate folders, with a clear and consistent naming convention that ensures easy access and efficient categorization. This structured approach facilitates seamless data handling during the model training and evaluation process.

The images are standardized with a resolution of 100 x 100 pixels. This resolution is particularly suitable for image classification tasks, enabling the model to discern key features without being overwhelmed by unnecessary details.

The dataset captures various visual characteristics of grapes, including differences in colour and texture, which are crucial for accurate classification. Some examples can be visualized in Figure 1.

|  |
| --- |
|  |
| Figure 1. Blue, Pink and White Grapes |

The structured organization, combined with the high-quality images, ensures that the dataset provides a strong foundation for training machine learning models, allowing for precise and reliable classification of the different grape types.

**3.2. AlexNet Network Architecture**

AlexNet architecture is designed to efficiently capture intricate features of images through a series of convolutional, pooling, and fully connected layers (Fu’adah et al., 2021). These layers apply multiple filters to images, capturing various patterns and edges essential for distinguishing between different classes. The architecture can be visualized in Figure 2.

|  |
| --- |
|  |
| Figure 2. AlexNet Architecture Visualisation |

**3.3. Training, validation and testing**

This section details processes of training, validating, and testing the AlexNet model for classifying data. Our approach ensures the model is well-trained and performs accurately on unseen data.

**Step 1. Data preparation**

Loading Data:

Training and testing datasets are loaded and divided into training (70%), validation (15%), and test (15%) sets, ensuring the necessary data is available to train and evaluate the model.

Data Enhancement:

To improve generalization ability of model, data augmentation is performed on training data. Techniques such as random horizontal flipping, cropping, and rotation are applied to create varied training samples.

Data Normalization:

Image data is scaled to the range [0, 1] and mean normalization is performed to standardize input data, which helps in faster and more stable convergence during training.

**Step 2. Model Definition**

Pre-trained AlexNet model is used, with last layer modified to match the number of dataset categories. AlexNet consists of multiple convolutional layers, pooling layers, and fully connected layers, which are essential for extracting image features and performing classification.

**Step 3. Compile the Model**

The model is compiled by selecting appropriate optimizer, loss function, and evaluation metric. Adam optimizer is chosen for its efficiency in handling large datasets and its adaptive learning rate capabilities. The learningrate is set to 0.0001 to ensure a stable and gradual learning process. The categorical cross-entropy loss function is used, as it is suitable for multi-class classification tasks, measuring performance of model by comparing predicted class probabilities with actual class labels. Evaluation metric selected is accuracy, which provides straightforward measure of how often model's predictions are correct.

**Step 4. Train the Model**

Using the fit method, the model is trained on training data over 10 epochs. During this phase, the model learns to recognize patterns and features in the data. The accuracy of each epoch remains above 90% for most of the time.

**Step 5. Evaluate the Model**

To evaluate the model's performance, the saved optimal model is loaded, the validation data is pre-processed, and the model.evaluate method is used. This evaluation on the validation dataset returns a performance metric: an accuracy of 96.47%.

**3.4. Machine learning algorithm**

AlexNet model used for classifying dataset employs a supervised machine learning (ML) algorithm. Supervised learning is a type of ML where the model is trained on labelled dataset. Each training example consists of an input paired with correct output, allowing model to learn mapping from inputs to outputs that can be used to predict output for new, unseen data. The dataset used for training includes both input data (e.g., images) and corresponding labels (e.g., categories). For our dataset, each directory contains images of a specific type of photos and is labelled accordingly. This organization allows model to learn the relationship between the inputs and correct outputs, which is crucial for making accurate predictions on unseen data (Raschka, S. & Mirajalili, V.,2017). Supervised learning is fundamental to AlexNet model's ability to classify images in chosen dataset. By learning from labelled training data, model can develop a robust understanding of features that distinguish different classes, leading to accurate predictions.

**3.5. Data encoding**

Image Data Generator from Keras was utilized for data preparation. This method was advantageous because it seamlessly integrated image preprocessing and label encoding into a single, streamlined process. It ensured consistency and minimized the risk of errors that could arise from manual data transformation. This not only optimized workflow but also allowed us to focus more on fine-tuning the model, ultimately contributing to its accuracy and robustness.

**4. Results**

In this section, we present a detailed analysis of the model's performance on images classification task. The results are described in three different ways (percentage, accuracy curve, confusion matrix) to provide a comprehensive understanding of model's accuracy.

AlexNet architecture achieved impressive results on the dataset, with the test set reaching an accuracy of 97%. This high accuracy indicates that architecture is highly effective in correctly classifying images of various fruits, vegetables, and nuts, demonstrating its strong generalization ability to new, unseen data. The performance metric underscores the robustness of architecture, validating effectiveness of the supervised learning approach and thoroughness of data preparation, training, and evaluation processes.

During the testing phase of model, we achieved an impressive accuracy rate of approximately 97% which can be visualised on accuracy curve in Figure 3. This high level of accuracy indicates that the model is performing well in classifying images into their respective classes.

Moreover, we utilized confusion matrix (Figure 4) to evaluate the performance of the model. Model had a bit more images of white and blue grapes, which contributed to better performance on these classes, as the model had more data to learn from. To simulate real-world testing scenarios and enhance the model's robustness, we applied image augmentation techniques. This involved generating variations of the original images, such as zooming in and making slight colour changes, making it appear as though the model was encountering new images it had never seen before.

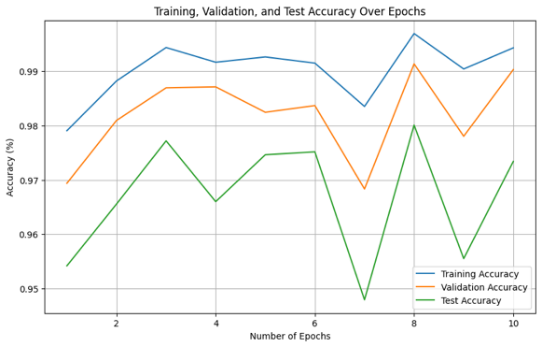


Figure 3. Accuracy curve

A blue squares with white text

Description automatically generated

Figure 4. Confusion Matrix

**5. Conclusion**

The objective of this project was to explore and develop effective methods for distinguishing between 3 types of grapes using machine learning techniques. This objective was successfully achieved by applying AlexNet for the classification task, resulting in a test accuracy of 97%.

By leveraging a pre-trained AlexNet model, we were able to achieve high accuracy in a shorter amount of time, demonstrating the efficiency and effectiveness of using transfer learning in this context. While our model's performance was strong, there is still room for optimization to potentially improve accuracy even further.

The success of our model suggests that deep learning models like AlexNet can be effectively employed in real-world applications, such as automated sorting systems in agricultural industries and quality control in food processing facilities.

Finally, deploying the model in a real-world setting and evaluating its performance under practical conditions would be valuable next steps.

**References**

Amin, U., Shahzad, M.I., Shahzad, A., Shahzad, M., Khan, U., and Mahmood, Z. (2023) 'Automatic fruits freshness classification using CNN and transfer learning', Applied Sciences, 13(14), p. 8087. [Pakistan].

Fu’adah, Y.N., Wijayanto, I., Pratiwi, N.K.C., Taliningsih, F.F., Rizal, S., and Pramudito, M.A. (2021) 'Automated classification of Alzheimer’s disease based on MRI image processing using convolutional neural network (CNN) with AlexNet architecture', Journal of Physics: Conference Series, 1844(1), p. 12020. Bristol: IOP Publishing. [Indonesia].

Li, Z., Li, F., Zhu, L. & Yue, J. (2020) 'Vegetable Recognition and Classification Based on Improved VGG Deep Learning Network Model', International Journal of Computational Intelligence Systems, vol. 13, no. 1, pp. 559–564.[China].

Phan, Q.-H., Nguyen, V.-T., Lien, C.-H., Duong, T.-P., Hou, M. T.-K., and Le, N.-B. (2023) 'Classification of tomato fruit using Yolov5 and convolutional neural network models', Plants, 12(4), p. 790.[Vietnam].

Raschka, S. and Mirajalili, V. (2017) Python machine learning: machine learning and deep learning with Python, scikit-learn, and TensorFlow. 2nd edn, fully revised and updated. Birmingham, England: Packt.[UK].

Xu, L., Chen, X., Yang, X., and Parida, P. (2024) 'Tourism image classification based on convolutional neural network SqueezeNet: Taking Slender West Lake as an example', PloS one, 19(1), p. e0295439. United States: Public Library of Science.[India].

Yuesheng, F., Jian, S., Fuxiang, X., Yang, B., Xiang, Z., Peng, G., Zhengtao, W. & Shengqiao, X. (2021) 'Circular Fruit and Vegetable Classification Based on Optimized GoogLeNet', IEEE Access.[USA].

Zhu, L., Li, Z.B., Li, C., Wu, J., and Yue, J., 2018. High performance vegetable classification from images based on AlexNet deep learning model. International Journal of Agricultural and Biological Engineering, 11(4), pp.217-223.[China].