Convolutional Neural Network Models for Image Classification

Group 4
Mohd Darish Khan (2101MC29)
Suryansh Jaiswal(2101MC41)
Hardik Singh(2101MC19)

September 18, 2024

1 Introduction

In this report, we describe the design, training, and evaluation of three convolutional neural network (CNN) models for image classification. Each model varies in terms of its architecture and complexity. The models were implemented using PyTorch and evaluated on a dataset of images with the goal of achieving high classification accuracy. The models were trained on a dataset of three classes and evaluated using accuracy and loss metrics.

2 Model Architectures

The following three models were created and trained:

- **Model 1**: A two-layer CNN with 8 and 16 filters respectively.
- **Model 2**: A three-layer CNN with 8, 16, and 32 filters respectively.
- **Model 3**: A four-layer CNN with 8, 16, 32, and 64 filters respectively.

These models were trained using a cross-entropy loss function and optimized with the Adam optimizer. Each model used max-pooling layers after every convolutional block, followed by fully connected layers and softmax activation to output class probabilities.

3 Training and Validation

The models were trained for a maximum of 100 epochs, with early stopping implemented to halt training when the validation loss did not improve for 20 consecutive epochs. The training loss and accuracy for each epoch were recorded.

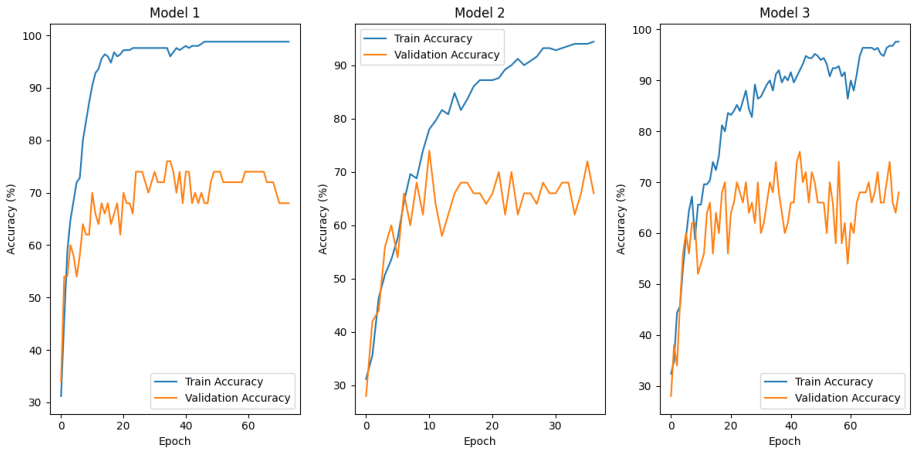


Figure 1: Training and validation accuracy over epochs

4 Test Results

After training, the models were evaluated on the test set. The test accuracies for the three models are as follows:

- **Model 1**: 73%
- **Model 2**: 65%
- **Model 3**: 71%

5 Confusion Matrix

The confusion matrix for Model 1 was computed to visualize the performance of the model on the test set.

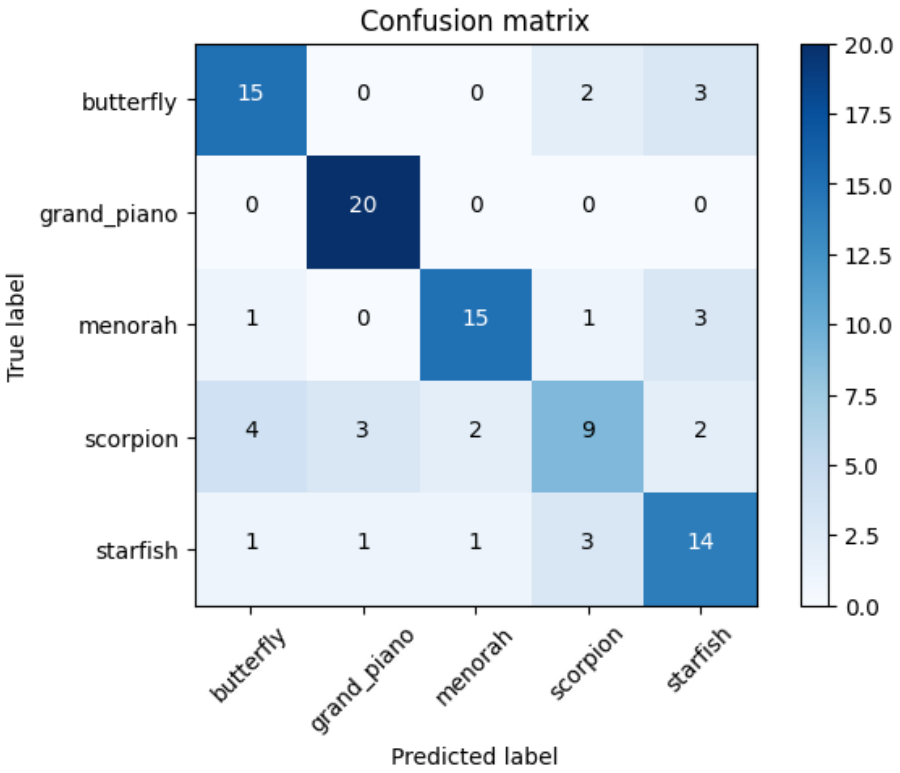
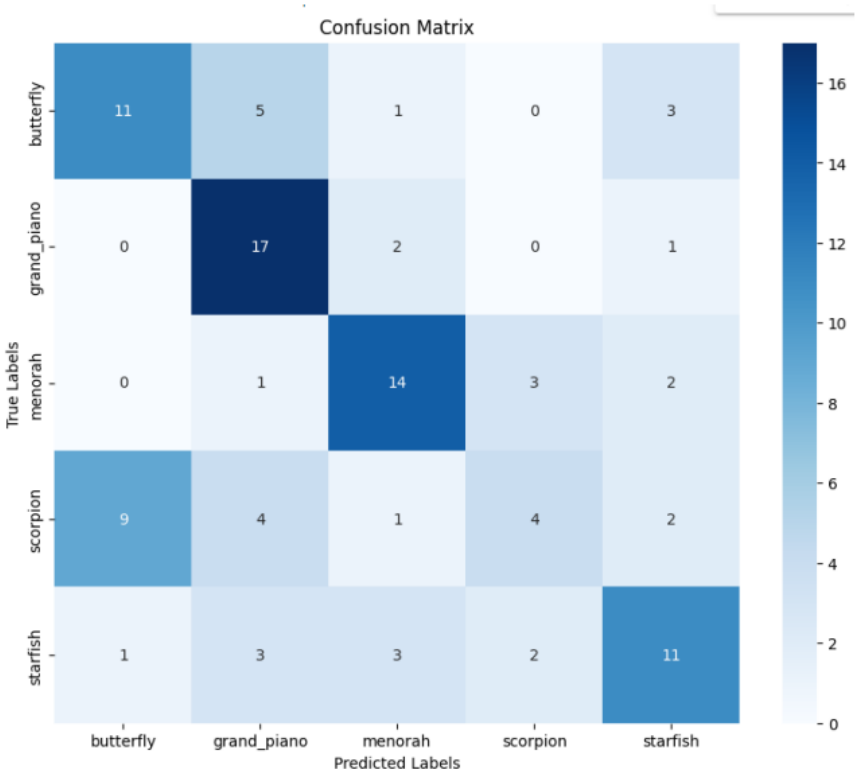


Figure 2

Comparison with FCNN without regularisation



6 Feature Maps

The feature maps generated by the first and second convolutional layers for an input image are shown below. These visualizations provide insight into the learned filters and the hierarchical representations formed by the CNN.

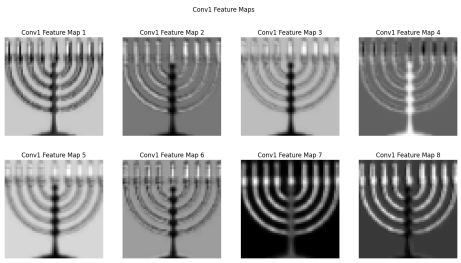


Figure 3: Feature maps from the first convolutional layer of Model 1

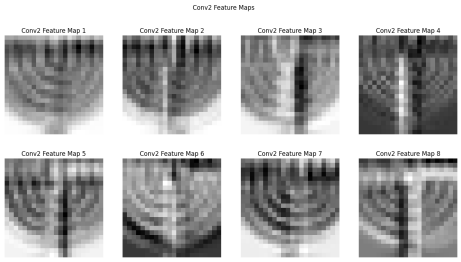


Figure 4: Feature maps from the first convolutional layer of Model 1

7 Maximally Activating Neuron Analysis

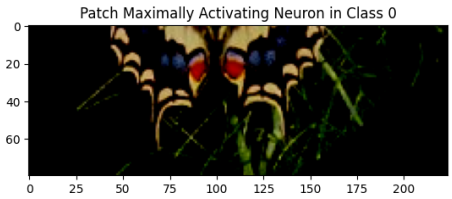
In this section, we examine one image from the training set for each class and pass each image through the CNN with the best-performing architecture (Model 1). The goal is to identify the neuron in the last convolutional layer that is maximally activated for each image. We then trace back to the patch in the image that causes the neuron to fire and visualize these patches.

7.1 Methodology

For each class, an image from the training set is selected and passed through the CNN. The output of the last convolutional layer is analyzed to find the neuron with the maximum activation value. We then trace back from this neuron to the corresponding patch in the original input image that caused the neuron to fire. The size and position of the patch are determined based on the receptive field of the network at that layer.

7.2 Results

The patches in each image that maximally activate the selected neuron are visualized below:



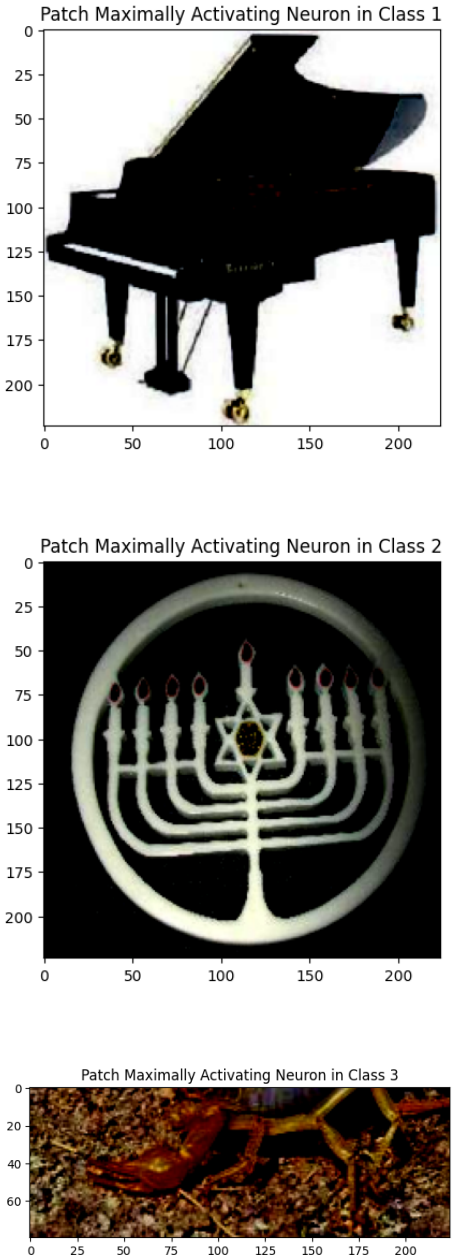


Figure 5: Enter Caption

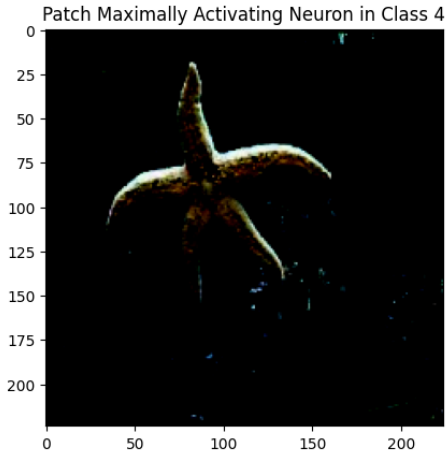


Figure 6: Enter Caption

8 Conclusion

In this report, we implemented and evaluated three CNN models for image classification. Model 1 achieved the highest test accuracy of 73%. While Model 2 underperformed, Model 3 performed comparably to Model 1 but with a more complex architecture.