CSC-345: Object Recognition Report

Introduction

With the widespread accessibility of cameras, image data has grown exponentially, which has driven a need for image classification algorithms for purposes such as search and automatic moderation[1]. In this document, we approach image classification using the CIFAR-100 dataset(from Canvas) with the goal of producing accurate and precise machine-learning models.

The focus of our investigation involves implementing and comparing the performance of two powerful techniques: Neural Network (NN) and Convolutional Neural Network (CNN). NNs are versatile models capable of learning complex relationships in data, while CNN excels at developing internal representations of two-dimensional images, making them well-suited for image-related tasks typically outperforming random forest approaches[2], [3].

This report presents a comprehensive analysis of our chosen methods, detailing the feature extraction process, training, and evaluation of the provided dataset. Our goal is to achieve accurate classification results and critically assess the strengths and limitations of each approach.

Method

For both classification methods similar models will be applied to both label types with evaluation being performed using both results.

Neural Network (NN):
The NN approach involves the following key steps:

1. **Feature Extraction:**
    a. Utilizing the Histogram of Oriented Gradients (HOG) to represent image features as vectors. Which has been show to outperform other feature extraction techniques[4].
    b. Normalizing the feature vectors to ensure consistent scales across different features.
2. **Model Construction:**
    a. Implementing a neural network model using the Keras library.
    b. Designing a simple model with densely connected layers.
    c. Configuring the model with an appropriate loss function, optimizer, and evaluation metrics.
3. **Training and Evaluation:**
    a. Training the NN on the training datasets, monitoring performance on a validation set.
    b. Evaluating the trained model on the testing dataset to assess capabilities.
    c. Analysing classification results, generating a confusion matrix, and calculating accuracy.
    d. Adjust model according to evaluation until satisfactory results are achieved.

Convolutional Neural Network (CNN):
The CNN approach involves the following key steps:

1. **Data Preprocessing:**
    a. Reshaping the input images to match the expected input shape of the CNN.
    b. Normalizing pixel values
2. **Model Construction:**
    a. Creating a model using convolutional, pooling, and densely connected layers.

b. Configuring the model with an appropriate loss function, optimizer, and evaluation metrics.

3. **Training and Evaluation:**
   a. Training the CNN on the training dataset, , monitoring performance on a validation set.
   b. Evaluating the CNN's performance on the testing dataset.
   c. Analysing classification results, generating a confusion matrix, and calculating accuracy.
   d. Adjust model according to evaluation until satisfactory results are achieved.

These methods aim to apply neural networks to exploit the repetitiveness of features in images, while being applied in a more flexible way to allow for adjustments to avoid overfitting to the dataset. To mitigate the occurrence of over fitting a early stopping monitor is applied to monitor the validation loss. We then also drop nodes between layers to try and stop the methods from learning the dataset[5], [6].

Results

During early testing phase it became apparent that validation loss would be a major issue with the model with as both models were had too many nodes in each layer. At that point accuracy was at 29% for coarse NN and 50 for coarse CNN, however identifying that loss was an issue, it was elected that the parameters per layer would be reduced, nodes would also be dropped from each layer. This brought down validation loss similarly to figure1.

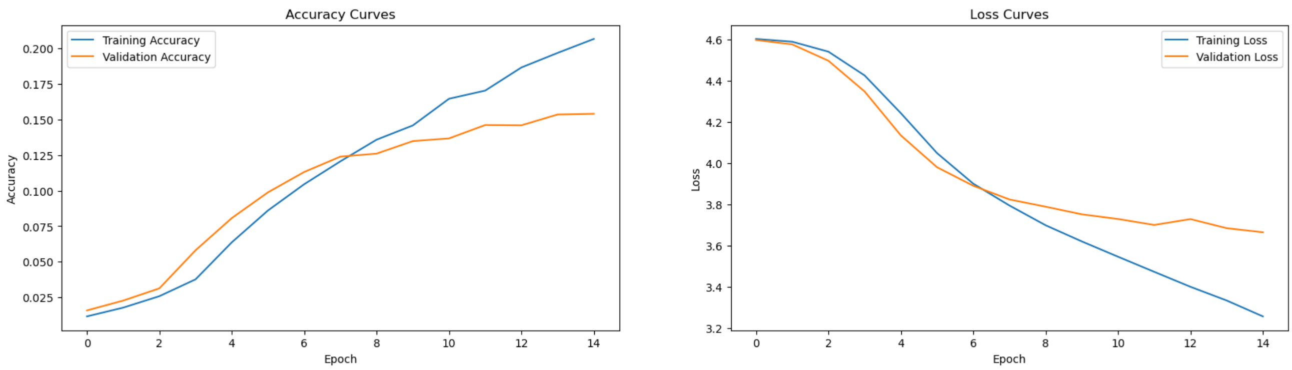

Figure 1: Fine labelled NN model

Model	Neural Network (NN)		Convolutional Neural Network (CNN)	
Labelling	Coarse	Fine	Coarse	Fine
Accuracy	27.50%	19.25%	47.75%	36.07%
Validation Loss	2.3933	3.5470	1.7003	2.5953

Confusion Matrices

Due to the size of the matrices for the models reducing their size to fir the document would making readability difficult and quickly exceed the page limit. Only the coarse labelled model from the CNN is given however the matrices for the other models can be found in the attached notebook.

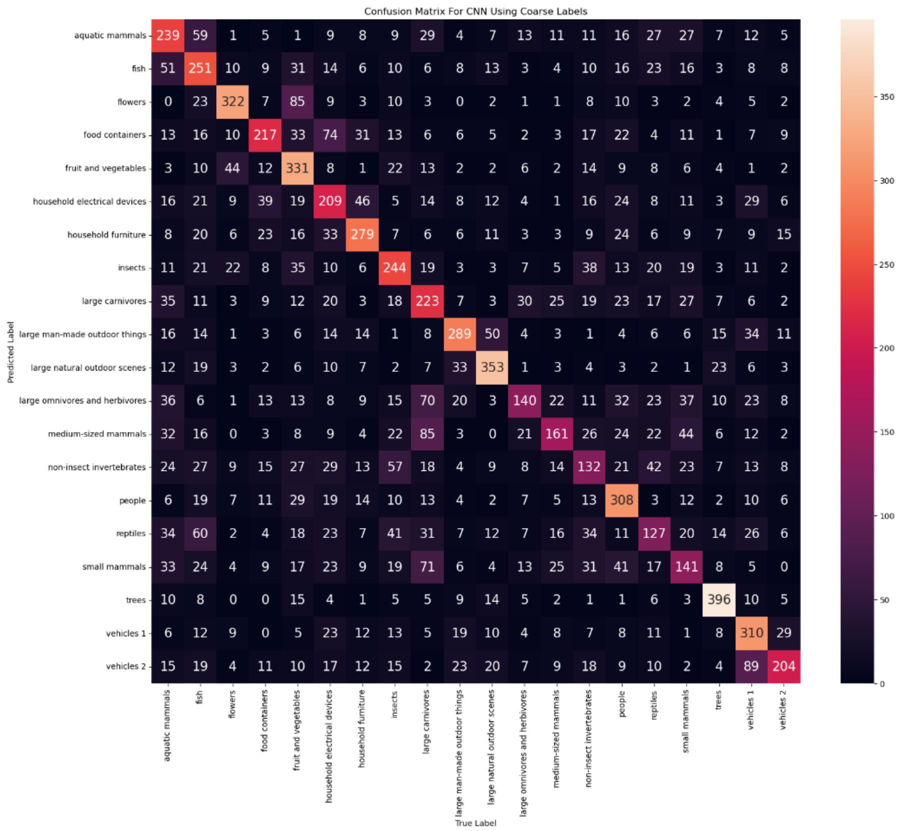

Figure 2: Confusion Matrix for coarse labelled CNN

Conclusion

The results of the experiments reveal interesting insights into the performance of Neural Network (NN) and Convolutional Neural Network (CNN) models on the provided image dataset.

Neural Network (NN):

The NN model exhibited limitations in capturing the complexity of certain labels with a very low. The achieved accuracy of approximately 27.50% for both fine and coarse labels suggests that the model struggled to generalize well to the testing dataset. The relatively low accuracy indicates that the model failed to discern the subtle features necessary for precise object classification.

Convolutional Neural Network (CNN):

In contrast, the CNN model demonstrated improved performance, especially in classifying coarse labels. The accuracy for acceptable labels reached 36.07%, while the accuracy for coarse labels surpassed expectations at 47.75%. The CNN's ability to discern more nuanced details from image allows it to classify the finer labelled data to more consistently.

Comparative Analysis:

- **Accuracy Improvement:**
- The CNN outperformed the NN in terms of accuracy, indicating the effectiveness of convolutional layers in image classification tasks.
- **Label Granularity Impact:**

- Notably, the CNN exhibited a significant accuracy improvement for coarse labels compared to fine labels. This suggests that the CNN architecture excelled in recognizing broader object categories.

In conclusion, while the CNN demonstrated superior performance, both models have room for enhancement. The exploration of advanced architectures, regularization techniques, and data augmentation strategies could contribute to more accurate and robust object classification.

References

[1] F. Kou, J. Du, Y. He, and L. Ye, ‘Social network search based on semantic analysis and learning’, *CAAI Transactions on Intelligence Technology*, vol. 1, no. 4, pp. 293–302, 2016, doi: https://doi.org/10.1016/j.trit.2016.12.001.

[2] S. Y. Chaganti, I. Nanda, K. R. Pandi, T. G. N. R. S. N. Prudhvith, and N. Kumar, ‘Image Classification using SVM and CNN’, in *2020 International Conference on Computer Science, Engineering and Applications (ICCSEA)*, 2020, pp. 1–5. doi: 10.1109/ICCSEA49143.2020.9132851.

[3] ‘What is a Convolutional Neural Network?’, NVIDIA Data Science Glossary. Accessed: Dec. 13, 2023. [Online]. Available: https://www.nvidia.com/en-us/glossary/data-science/convolutional-neural-network/

[4] N. Dalal and B. Triggs, ‘Histograms of oriented gradients for human detection’, in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05)*, 2005, pp. 886–893 vol. 1. doi: 10.1109/CVPR.2005.177.

[5] C. Sitaula and N. Ghimire, ‘An Analysis of Early Stopping and Dropout Regularization in Deep Learning’, vol. 5, no. 1.

[6] moodayday™, ‘Dropout & Early stopping’, AI³ | Theory, Practice, Business. Accessed: Dec. 13, 2023. [Online]. Available: https://medium.com/ai%C2%B3-theory-practice-business/dropout-early-stopping-188a23beb2f9