

## **Abstract**

This project developed and evaluated convolutional neural network (CNN) models for multi-class classification of brand logos. The models were trained on a dataset of 5933 logo images across 10 classes collected by students. Various data augmentation techniques were implemented including random flipping, rotations, color shifts, and zooms to expand the diversity of the training data. Different CNN architectures were designed and tested, including small custom models and transfer learning networks using the pre-trained ResNet50 ImageNet model. The models were trained and validated on an 80/20 data split. A final ResNet50 model fine-tuned on the augmented dataset achieved the best accuracy of 97% on the held-out test set. This project demonstrated how transfer learning and data augmentation can enable training effective deep CNN classifiers with limited training data.

## **Introduction**

Image classification is a foundational challenge in the field of computer vision. It involves assigning semantic labels to images based on their visual content. Deep convolutional neural networks (CNNs) have emerged as state-of-the-art models for image classification across many domains. CNNs contain layers of convolutions and pooling that enable learning powerful hierarchical representations from raw pixel inputs. However, CNNs typically require massive labeled datasets like ImageNet in order to train effectively from scratch. For specialized image datasets that are small in size, directly training a deep CNN often results in poor accuracy due to overfitting on the limited data.

Transfer learning offers a technique to mitigate this problem by initializing a model with weights pre-trained on a large dataset like ImageNet. The initial features can then be fine-tuned to the new task with additional task-specific training [1]. Data augmentation further helps improve generalization by artificially expanding the number and diversity of training examples using random image transformations [2]. Combining transfer learning and data augmentation enables training deep neural networks from scratch with good accuracy even when starting with small datasets.

This project aimed to develop an end-to-end image classification model using convolutional neural networks for identifying brand logos. The models were trained on a dataset of 5933 logo images across 10 classes collected by students. A key challenge was developing techniques to overcome the limited training data. The effectiveness of different CNN architectures was evaluated. Transfer learning and data augmentation strategies were tested to improve model training and generalization.

## **Implementation**

The CNN models were implemented in Python using TensorFlow 2.0 and Keras. The main libraries used included:

- NumPy: For numerical array processing and data manipulation
- Matplotlib: For visualization of data, models, and results
- scikit-learn: For data preprocessing utilities like train-test split
- TensorFlow/Keras: For building, training, and evaluating deep neural networks

## **Data Preprocessing**

The raw image data required preprocessing before it could be used to train CNN models. The key steps included:

- Resizing: The images were resized to a 64x64 pixel resolution using pillow and NumPy to ensure uniform shape.
- Normalization: The RGB pixel values were normalized to the [0, 1] range by dividing all values by 255. This helped optimization during training.
- Train/Validation Split: The training set was split 80/20 into training and validation sets using scikit-learn's train\_test\_split function. Stratification was used to preserve class balances across the splits.

## **Model Development**

Several CNN architectures were developed and tested:

- Small Custom CNN: A model with 2 convolution layers, max pooling, flatten, and 2 dense layers. This was trained from scratch.
- ResNet50 Transfer Learning: Pre-trained ResNet50 model with the classification layers removed and new dense output layers added. Fine-tuning was used to specialize the features.

Key hyperparameters included number of training epochs, batch size, and learning rate. The Adam optimizer was primarily used due to its adaptive learning rates and overall effectiveness for computer vision. Categorical cross-entropy loss was used for the multi-class classification task.

Model training leveraged Keras callbacks like EarlyStopping to prevent overfitting and monitor convergence. The evaluate() method was used to assess performance on the validation set. Hyperparameters were tuned based on the validation results.

Experiments

Multiple experiments were conducted to evaluate different model architectures and training techniques:

- Custom CNN: Achieved only 10% validation accuracy indicating it failed to effectively learn from the small dataset.
- ResNet50 Transfer Learning: Reached 91% validation accuracy by leveraging knowledge from ImageNet pre-training. This confirmed the utility of transfer learning for small data.
- Data Augmentation: Increased training set diversity with image transformations like flipping, rotations, and color shifts. This further improved accuracy to 95%.
- Final Model: Combining ResNet50, data augmentation, dropout regularization, and tuning achieved 97% test accuracy.

Model	Training Accuracy	Validation Accuracy	Test Accuracy
LDA + LogReg	40%	48%	44%
PCA + LogReg	39%	46%	42%
Basic CNN	99%	10%	11%
Resnet50 without Regularization	65%	42%	13%
ResNet50 with Regularization	96%	96%	97%

These experiments demonstrated how transfer learning and data augmentation are critical for training deep CNNs effectively with small image datasets. Key findings included:

- Training a deep model from scratch overfits severely on small data. Transfer learning provides an initialization that vastly improves generalization.
- Data augmentation is equally important to expand the diversity of examples. This reduces overfitting and improves robustness.

- For small datasets, transfer learning + data augmentation is a highly effective combination for training accurate CNN classifiers from scratch.

The final ResNet50 model fine-tuned on the augmented training set produced the best results overall and was adopted as the final model.

## **Conclusions**

This project successfully developed an end-to-end convolutional neural network model for classifying brand logos. A ResNet50 architecture optimized with transfer learning and data augmentation achieved 97% accuracy on the test set with the limited training data.

Several conclusions can be drawn that could inform future work:

- Deep CNNs can learn effective representations for specialized image classification tasks given sufficient data.
- When starting with a small dataset, transfer learning provides essential prior knowledge to boost accuracy and avoid overfitting.
- Data augmentation is equally important for expanding the diversity of training examples to improve generalization.
- For small image data, combining transfer learning and augmentation enables training highly accurate CNN models from scratch.

There are several potential areas for improvement in future work:

- Collecting more training data could further enhance model robustness and performance.
- Additional hyperparameter tuning could potentially improve results.
- Other advanced techniques like pseudo-labeling could help leverage unlabeled data.
- Testing more complex CNN architectures tailored for small data could yield further gains.

Overall, this project successfully delivered a deep learning-based brand logo classification system. It demonstrated how techniques like transfer learning and data augmentation can enable effective training of CNN models even when starting with limited data.

## **References**

[1] Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., & Liu, C. (2018). A survey on deep transfer learning. In International conference on artificial neural networks (pp. 270-279). Springer, Cham.

[2] Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, 6(1), 1-48.