Project Report: Image Classification Using CNNs

Project Overview

This project focuses on using convolutional neural networks (CNNs) for image classification tasks. Several CNN models were developed and tested using the CIFAR-10 dataset and a custom dataset consisting of food images. The main objective was to evaluate the model's performance in terms of accuracy and understand the behavior of models under different configurations and training conditions.

Methodology

Data Preparation

- **CIFAR-10 Dataset:** Used for initial training and testing of the CNN models. It consists of 60,000 32x32 color images in 10 classes.
- **Custom Food Image Dataset:** Comprises images categorized into three classes: Cakes, Pasta, and Pizza. This dataset was used to test the model's ability to generalize to new, unseen data.

CNN Architecture

- The CNN architecture included layers designed to extract features (convolutional layers) and layers to reduce the dimensionality of the feature maps (pooling layers). These were followed by fully connected layers that performed the classification based on the features extracted by the convolutional layers.

Training

- The model was trained using stochastic gradient descent (SGD) with a momentum of 0.9 and a learning rate of 0.001. The models were trained for 30 epochs on the CIFAR-10 dataset.

Testing and Evaluation

- The models were evaluated based on their classification accuracy on the test set. Additionally, confusion matrices were generated to analyze the model's performance across different classes.

Results

CIFAR-10 Model Performance

- The CNN trained on the CIFAR-10 dataset achieved a peak accuracy of approximately 59% on the test images.
- The accuracy per class varied, with some classes like 'car' achieving higher accuracies compared to others like 'cat'.

Custom Food Image Dataset

- When the CIFAR-10 trained model was tested on the custom food image dataset, the accuracy dropped, indicating challenges in generalizing to new data types.

Fine-Tuning and Transfer Learning

- Fine-tuning the pre-trained CIFAR-10 model on the food image dataset slightly improved accuracy, demonstrating the benefits of transfer learning in adapting to new domains.

Dropout Implementation

- Incorporating dropout in the network architecture helped in managing overfitting, which was evident from the comparative stability in loss across training epochs.

Feature Extraction and Machine Learning Classifiers

- Features extracted from the CNN were used to train SVM and AdaBoost classifiers. The SVM classifier outperformed AdaBoost, suggesting differences in how these models handle feature spaces derived from deep networks.

Discussion

- The confusion matrices revealed that certain classes are more challenging to distinguish than others, possibly due to similarities in color and texture patterns between classes.
- The experiments with dropout highlighted its effectiveness in enhancing model generalization by reducing overfitting.
- Fine-tuning demonstrated that transfer learning could be a potent strategy for adapting pre-trained models to new tasks with limited data availability.

Conclusion

This project illustrated the capability of CNNs in handling image classification tasks with varying degrees of complexity and dataset characteristics. Future work could explore more sophisticated CNN architectures, deeper networks, and advanced training strategies like learning rate annealing to further enhance model performance.