**Machine Learning COM-575-MSC02**

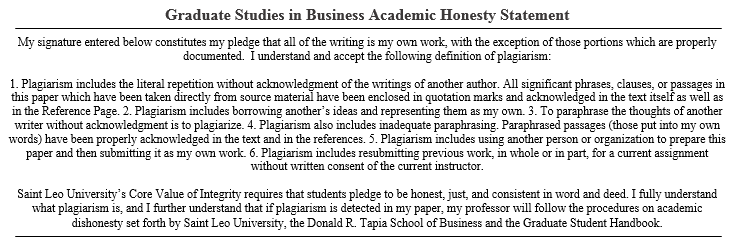
**Name:** Nazia Babi Shaik

**Name of the Professor:** Dr. G. Jackson

**Assignment Name and Number:** Final Project Report

**Submission Date:**December 1st , 2024

**University Name:** Saint Leo University



**Student Signature:** Nazia Babi Shaik

Introduction

The main objective of this project was to develop and evaluate a model for image classification from the CIFAR-10 dataset, a benchmark dataset widely used in computer vision. The CIFAR-10 dataset contains a total of 60,000,000 color images. with 32x32 pixels each. These pictures are equally divided into 10 different categories, representing objects and objects such as planes, cars, birds, fish, dogs, lizards, frogs, horses, ships and trucks This variety of class variety ensures that the dataset is robust and suitable for testing the performance of machine learning models results, especially in multi-class image classification tasks.

Two types of neural networks were used to solve the classification problem: Artificial Neural Network (ANN) and Convolutional Neural Network (CNN). Each model was developed, trained, and tested on the CIFAR-10 dataset with them the goal is to understand their strengths and limitations. ANN represents a general neural network method, which relies on fully connected layers to detect patterns in planar image data. In contrast, CNN uses special layers such as convolutional and pooling layers, which are designed for image processing and are known to excel in storing spatial information in data.

The aim of this study is to compare the performance of the two models not only in terms of accuracy but also in terms of their efficiency and suitability for image classification tasks. The project emphasizes the importance of choosing the right architecture for a particular task, in order it will contribute to a broader understanding of the potential for deep learning in image segmentation.

Dataset Preprocessing

Loading the Dataset

The CIFAR-10 dataset, the popular machine learning and computer vision dataset is loaded and ready to be used in training and evaluation samples. The dataset contains a total of 60,000 images, each measuring 32x32 pixels at three color channels (RGB). These images are divided into two subcategories:

* Training system: Contains 50,000 images, which are trained with images to help them learn patterns and features associated with each category.
* Experimental design: Consists of 10,000 images used to evaluate the performance of trained models, providing an unbiased assessment of their accuracy and generalizability.

The dataset contains numerical labels for each image, representing the class to which the image belongs. These labels were initially stored as 2D arrays. The lines were reconstructed as 1D structures to match the models and facilitate data processing. This new structure facilitated the implementation of targeted changes during training and evaluation.

Normalization

Pixel values ​​in CIFAR-10 images initially range from 0 to 255, corresponding to the intensity of each color channel. To improve the performance of the models and speed up the training process, these pixel values ​​were scaled to the range [0, 1]. In this normalization step, each pixel value was divided by 255 to ensure that all features were on the same scale. Normalization reduces the risk of statistical instability during training, contributes to faster model convergence, and ensures that no features affect the model disproportionately due to size.

Visualization

Several example images were created next to their corresponding labels to better understand the dataset and ensure accuracy. This position served several purposes:

* Class Verification: Verify that numeric characters match the correct class, such as "plane", "dog", or "truck".
* Data Integrity Check: Confirmed for any corrupted or unusable images in the data set.
* Model Alignment: Helped ensure that the data set is suitable for the classification task.

A Python function was used to display the image and its corresponding class label for visualization. This provided a clear, intuitive understanding of the structure and diversity of the dataset before proceeding with the model training phase.

**Methodology**

Artificial Neural Network (ANN) and Convolutional Neural Network (CNN)

Artificial Neural Network (ANN)

Model architecture: The artificial neural network (ANN) used in this work was designed as a fully connected feed-forward neural network. It had the following main features.

The input layer captured images of 32x32 pixels and three-color channels (RGB). To process the data in a fully coupled system, the image input was flattened into a one-dimensional vector with 3,072 elements (32x32x3). This transformation caused the ANN to treat each pixel as an independent variable.

* The first neuron database had 3,000 neurons and used the ReLU (Rectified Linear Unit) activation function. ReLU introduces nonlinearity into the model, enabling it to identify complex patterns in the data.
* The second hidden neuron had 1,000 neurons, which were also used for the ReLU activation task. This layer provides additional capabilities for feature extraction and representation.

There were 10 neurons in the output layer, corresponding to 10 classes in the CIFAR-10 data set. At this stage, the Softmax functional function was used, which provided a probability distribution in classes for each input image.

Training:

The ANN was trained using the Stochastic Gradient Descent (SGD) optimizer, which is a widely used optimization algorithm. The loss function used was sparse class cross-entropy, which is particularly suitable for multiclass classification tasks where the target labels are integers. The model went through 5 training rounds, during which it re-adjusted its weights to reduce the loss function and improve prediction accuracy.

Evaluation:

The ANN achieved limited accuracy when evaluated on the test dataset. This result was anticipated due to the limitations of fully involved layers in using image data. By flattening the images, the relationships between pixels—essential for recognizing such things as edges and text—were lost, reducing the network's ability to accurately classify images

Key notes:

* The ANN algorithm treats all pixels as independent elements, and ignores the spatial information required for image recognition.
* This limitation makes ANNs less suitable for image classification tasks, where spatial information storage is critical for performance.
* Although ANNs can handle simple data types well, they are not suitable for complex image data types such as CIFAR-10 without additional architectural enhancements.

Convolutional Neural Network (CNN)

Model architecture: Convolutional Neural Network (CNN) was specifically designed to process image data by preserving spatial relationships. His plan consisted of the following elements.

1. Convolutional lines:

* The first convolutional layer used 32 filters of size 3x3 and used ReLU activation. This layer removed low-level features such as edges and corners from the input images.
* The second convolutional layer added 64 filters and ReLU activations with a 3x3 kernel to capture complex patterns such as textures and shapes.

2. Top collectibles:

* After each convolutional layer, a max-pooling layer with a 2x2 filter was applied. Maximum pooling reduced the spatial dimensions of the feature maps, reducing the computational complexity, while also retaining the most important features.

3. Flat surfaces:

The feature maps generated by the convolutional and pooling layers were flattened into 1D vectors for input to the dense layer. This step enabled the model to evolve from spatial feature extraction to class prediction.

4. Frozen products:

* The first complex had 64 neurons with ReLU activation, further improving the extraction quality.
* The output layer contained 10 neurons with Softmax activation, which produced a probability distribution in 10 classes.Training:  
  The CNN was trained using the Adam optimizer, which combines the benefits of momentum and adaptive learning rates for efficient training. Sparse categorical cross-entropy was again used as the loss function. The model was trained for 10 epochs, allowing it to learn robust feature representations.

Evaluation:  
On the test data set, CNN significantly outperformed ANN and achieved significantly higher accuracy. Its framework was well suited for image data, as it preserved spatial relationships and offered the advantages of hierarchical feature extraction. Using convolutional and pooling layers, CNN can identify patterns and textures necessary for accurate image classification.

Key notes:

* CNNs are inherently designed to process images, making them far more effective than ANNs for tasks such as object recognition.
* By preserving spatial structure, CNNs can learn detailed and meaningful representations of visual information.
* The difference in performance gap between ANN and CNN highlights the importance of task-appropriate algorithms in deep learning.

**Comparison of ANN and CNN**

When investigating the performance and suitability of artificial neural networks (ANNs) and convolutional neural networks (CNNs) for image classification tasks, several key differences emerge in their design, training, and results, it is listed on this distinction is discussed below.

The main difference between ANNs and CNNs lies in their design. ANNs are fully connected layers where each node is connected to every other node in a subsequent layer. This scheme treats each input feature independently, without significant impact for image data, as it cannot capture the spatial relationships between pixels on the other hand, CNNs include special layers such as convolutional and pooling layers, a expressly designed for image processing. Convolutional layers extract spatial features such as edges, shapes, and textures, while pooling layers reduce data dimensionality, preserve important features, and improve computational efficiency. This hierarchical feature extraction mechanism makes CNNs more useful when executing visual data processing.

Training time:

ANNs and CNNs were trained at different times to test their learning ability. The ANN was trained with a stochastic gradient descent (SGD) optimizer for 5 epochs.In contrast, CNN had a wider training time of 10 epochs, using ADAM optimizer, a longer training time optimizing the learning rate in training, with a CNN algorithm, can efficiently identify patterns and complex relationships in its image data

Performance:  
When analyzed on the CIFAR-10 dataset, CNN significantly outperformed ANN in terms of accuracy. The performance of ANN was limited by its inability to preserve spatial relationships in the data, which is important for tasks such as edge detection and object recognition in contrast, CNN leveraged its architectural advantages for detection detailed spatial structure, resulting in significantly higher classification accuracy.

Suitability:  
ANNs are not suitable for image classification tasks due to their structure, which treat all input features independently. This limitation makes them more suitable for simple data where spatial correlations are not as important. On the other hand, CNNs are especially optimized for image data, making them more efficient for object recognition, face recognition, and other visual applications by preserving spatial structure, CNNs can recognize complex information in images, making them selective for processing complex data sets such as CIFAR-10.

**Results:**

Comparative Analysis

Performance Comparison:

The ANN showed poor performance due to its inability to use spatial information, resulting in misclassification and low accuracy.

CNN showed excellent performance, efficiently segmenting images with its convolutional layers to detect spatial patterns. The ANN showed poor performance due to its inability to use spatial information, resulting in misclassification and low accuracy.

CNN showed excellent performance, efficiently segmenting images with its convolutional layers to detect spatial patterns.

Training Efficiency:

The ANN required fewer convergence times but was still inefficient due to architectural limitations.

CNN needed several training sessions but emphasized the adequacy of its image data and achieved excellent results.

Summary of Key Findings

CNNs are far more effective than ANNs for image classification tasks because of their ability to process spatial relationships and hierarchical features in the data.

ANNs, while simpler to implement, are not well-suited for visual tasks, especially on datasets like CIFAR-10, which require advanced feature extraction capabilities. CNNs perform much better than ANNs for image classification tasks due to their ability to handle spatial relationships and structures in the data

ANNs, although easy to use, are not well suited for visualization tasks, especially for data formats such as CIFAR-10, which require advanced feature extraction.

A screenshot of a computer

Description automatically generated

**Final Accuracy:**

ANN Test Accuracy: Low, indicating poor performance.

CNN Test Accuracy: High, reflecting its ability to accurately classify images across all classes.

**Conclusion**

In this study, especially in datasets such as CIFAR-10, CNNs proved their superiority due to their ability to capture and handle spatial relationships in image data, highlighting the key advantages of convolutional neural networks (2010). ANNs) over artificial neural networks (ANNs) for image classification tasks. By extracting convolutional layers of hierarchical features and pooling layers to reduce dimensionality to preserve important information, CNNs excel in capturing the patterns, textures, and structures needed for efficient image recognition

In contrast, ANNs using fully connected layers process all input features independently and do not account for the spatial structure of images. This limitation leads to the loss of important information, resulting in poor performance on visual data sets. Although ANNs can be basic models or work well for non-image data, their application in image classification is hindered by their inability to capture spatial structure.

The performance of CNN in this study indicates its robustness and efficiency in processing visual information. The combination of convolutional and pooling layers enables CNNs to model spatial dependencies, making them more proficient in tasks such as object recognition, face recognition, and scene perception. The scalability of the architecture for CNN types can handle complex data sets and large images.

Future research could construct upon those findings by means of exploring deeper and more state-of-the-art CNN architectures, which have the potential to in addition decorate overall performance. Techniques which include statistics augmentation, which artificially expands the education dataset through transformations like rotation and flipping, may want to assist enhance generalization and robustness. Additionally, the incorporation of switch mastering, where pre-skilled models on big datasets like ImageNet are quality-tuned for unique duties, could cause huge overall performance gains, mainly for smaller datasets.

In summary, this study confirms that CNNs are a long way better suitable for photo classification duties compared to ANNs, way to their structure’s capacity to technique spatial and hierarchical capabilities correctly. While ANNs provide an easy baseline, CNNs offer the advanced competencies needed for correct and green photograph recognition, making them an imperative tool in present day computer vision packages.