Jeffery Dirden
Bradley Johnson
Professor McManus
November 30, 2024

Image Recognition Report

For our midterm object detection challenge we decided to use the CIFAR-10 dataset using ResNet50 which was challenging, but rewarding in the knowledge gained.

Setting Up and Customizing Kaggle, the kaggle.json file was set up to authenticate and access Kaggle datasets, and the required kaggle Python package was installed. Using Kaggle's API, the CIFAR-10 dataset was downloaded, and the files were then unzipped. The train that was extracted.The py7zr package was used to decompress the 7z file. Dataset structure and data inspection. The /train/ directory contains images in.png format. TrainLabels.csv is a CSV file that contains the labels. Distribution of Labels: The distribution of categories was determined by analyzing the label column. Preprocessing Images; To train the model, images were transformed into numpy arrays. In order to facilitate categorical processing, labels were converted to integers. Array Conversion to NumPy: Every image was read, converted to numpy arrays, and resized if necessary. Information and Labels: X: Features (NumPy arrays of images) Y: Labels (integers that are mapped).

Train-Test Split: The dataset was divided into subsets for testing (20%) and training (80%). Development of Models using a Basic Neural Network Model. A simple feedforward neural network that includes: Layer of input: Flattened (32, 32, 3) Dense (64 neurons, ReLU activation) hidden layer Dense (10 neurons, softmax activation) output layer. Advanced ResNet-50 Model

ResNet-50 pre-trained on ImageNet is used in a more reliable architecture. For fine-tuning, custom layers were added to the ResNet-50 model, which served as a convolutional base. Training and validation datasets were used to assess evaluation metrics. Model performance on test data is known as accuracy. Visualization: Validation loss versus training. Accuracy in training versus validation. Findings: Basic Model Accuracy: Only moderate accuracy was attained, suggesting the need for more sophisticated architectures. The ResNet-50 Performance: With dropout regularization and advanced features, accuracy is noticeably higher. Conclusions: In image classification tasks, sophisticated architectures such as ResNet-50 perform better than simple feedforward neural networks. To improve accuracy, proper preprocessing, architecture fine-tuning, and data augmentation are essential. Future research could involve: Trying out more pre-trained models. applying cutting-edge data augmentation methods adjusting hyper parameters such as optimizer type and learning rate.

Possible Hazards and a Plan for Mitigation working with deep learning models and the CIFAR-10 dataset carries a number of potential risks. An examination of these risks, divided into problems pertaining to data, computation, and models, can be data issues with data quality and integrity. Computational constraints with hardware limitations depending on the environment that you are using. Overfitting, under fitting, and hyper parameter tuning. Ethical and General risks such as bias in data, and misuse of the model.

Jeffery Dirden's reflection; My knowledge of how to efficiently handle and preprocess big datasets has increased as a result of working on this project. There were difficulties with the CIFAR-10 dataset, including managing a compressed format and deriving significant patterns from labeled data. The rewarding task of converting images into numerical representations while preserving data integrity brought to light the significance of exploratory data analysis. In order to avoid biased results, I also learned how important it is to balance the dataset. Aspects like managing unbalanced classes and enhancing data to enhance model generalization were especially instructive. My conviction that strong data preprocessing is essential as the cornerstone of any machine learning task has been strengthened by this project. For me, using data augmentation techniques to replicate a bigger and more varied dataset was the most exciting aspect. I'm more comfortable working with real-world datasets now, and I'm excited to use these abilities in upcoming projects.

Bradley Johnson's reflection; Understanding the trade-offs between simplicity and complexity in model design was the most important lesson learned from this project. To begin, we used a simple neural network to categorize CIFAR-10 pictures. Although this method gave a solid basis for understanding, it soon became clear that more sophisticated architectures, such as ResNet-50, were required to attain performance at the cutting edge. For me, transfer learning was the high point since it showed how to use pretrained models to increase accuracy while drastically cutting down on computational expenses. Additionally, I discovered how minor adjustments could have a significant impact on model performance by adjusting hyperparameters like learning rate and dropout values. I learned the value of a smooth transition between data pipelines and model training from working with the data preprocessing team. This experience reaffirmed that when data quality, computational efficiency, and model architecture are taken into account collectively, machine learning projects have the best chance of success. To push the limits of what we can accomplish with constrained data and computational resources, I intend to investigate additional sophisticated models and frameworks in the future.

Any project needs innovation to move forward, but machine learning is a field where innovation frequently results in breakthroughs. The creative concepts and methods listed below have the potential to improve the project and create new research opportunities. Data augmentation and synthesis, custom loss function for class specific accuracy, automated hyper parameter optimization, model interpretability and visualization, ensemble models, and semi supervised learning are a few examples of innovation that will play a big part in the future of completing this type of work.