Image Classification Report

Abstract

This report presents the approach taken to solve the vehicle image classification task. The dataset was analyzed and manually cleaned due to significant quality issues and imbalance. A deep learning model was trained using TensorFlow, experimenting with different architectures, including a custom CNN and ResNet. Due to the dataset's poor quality and imbalance, overfitting was a major challenge. Various techniques such as augmentation, dropout, and regularization were applied to mitigate overfitting. The best-performing model was exported to ONNX format and evaluated on a test set.

Data Analysis and Cleaning

During dataset analysis, it was observed that the dataset contained numerous misclassified images and an imbalance in class distribution. To address this, manual data cleaning was performed by checking each folder and removing incorrect or low-quality images. Unlike automated tools such as `fastdup` or `cleanvision`, this manual process ensured a thorough review of the dataset.

Data Preprocessing

- Images were resized to **180x180** pixels to ensure uniformity and compatibility with the model architecture.
- Data augmentation techniques such as rotation, flipping, and brightness adjustments were applied to artificially increase dataset variability and reduce overfitting.
- Normalization was performed by scaling pixel values between 0 and 1.

Model Architecture

Experiments were conducted with both a custom CNN model and ResNet. However, due to dataset limitations, the models faced overfitting challenges. The final chosen architecture was **ResNet50**, pre-trained on ImageNet, with fine-tuning on the dataset to leverage transfer learning benefits.

Training and Experimentation

- **Loss Function**: Categorical Cross-Entropy
- **Optimizer**: Adam
- **Learning Rate**: 0.001 (with learning rate decay)
- **Epochs**: 50 (with early stopping to prevent overfitting)
- **Techniques to Reduce Overfitting**:
    - Data Augmentation
    - Dropout Layers (0.5 probability)
    - L2 Regularization
    - Fine-tuning only specific layers of ResNet instead of the entire model

Results and Key Findings

- The initial custom CNN model overfitted due to the dataset issues.
- ResNet50 performed better but still faced challenges due to data imbalance.
- Training accuracy was significantly higher than validation accuracy, indicating overfitting.
- Despite augmentation and regularization, the poor dataset quality limited the model's generalization ability.
- **Best Model Performance (on Validation Set):**
    - Accuracy: **~78%**
    - Precision, Recall, and F1-score: Reported per class, showing lower performance on minority classes.
    - Confusion Matrix: Showed misclassification trends, especially for underrepresented categories.

Future Work

- Obtain a **better quality and balanced dataset** to improve model performance.
- Experiment with other architectures such as EfficientNet or Vision Transformers.
- Implement active learning strategies to enhance dataset quality progressively.

Note

The dataset provided was of **poor quality and highly imbalanced**, which significantly affected model performance. Even after **manual cleaning** of data by checking each folder, the model still struggled due to the inherent data limitations. If a well-balanced dataset with high-quality images were available, the training process would yield much better results.

Deliverables

1. notebooks for model training, evaluation and tonnx export.
2. Trained **ONNX model** with inference details.
3. Performance metrics and confusion matrix.
4. `README.md` with instructions on running the model.