Report: Age Detection Using Fine-Tuning VGG16 Model

Introduction

Age detection is an essential application in computer vision with numerous practical use cases, including security systems, targeted marketing, and healthcare. This report presents the process and outcomes of fine-tuning a pre-trained VGG16 model for age detection using the UTKFace dataset, a benchmark dataset for age and demographic analysis.

Background

The VGG16 model, developed by the Visual Geometry Group (VGG) at Oxford University, is widely recognized for its success in image classification tasks. By leveraging transfer learning, the model can be fine-tuned for specific applications, such as age estimation, which involves regression rather than classification. The UTKFace dataset, containing over 20,000 facial images labeled with ages ranging from 0 to 116 years, served as the foundation for training and validation.

Learning Objectives

The primary objectives of this project were:

1. To understand and implement transfer learning techniques using the VGG16 model.
2. To fine-tune the model for regression tasks specific to age prediction.
3. To evaluate the model's performance using key metrics such as test loss and mean absolute error (MAE).
4. To identify and address challenges in training and improving model accuracy.

Activities and Tasks

1. **Dataset Preparation:**
   - Collected and preprocessed the UTKFace dataset by resizing images to 224x224 pixels, normalizing pixel values, and splitting the dataset into training, validation, and test sets.
2. **Model Architecture:**
   - Loaded the pre-trained VGG16 model, excluding its top classification layers.
   - Added custom fully connected layers suitable for regression, culminating in a single neuron output for predicting age.
3. **Fine-Tuning:**
   - Unfroze the last few convolutional layers of the VGG16 model for domain-specific feature learning.
   - Used the Adam optimizer and mean squared error loss function for training.

4. **Training and Evaluation:**
    - Trained the model using early stopping and learning rate scheduling to prevent overfitting.
    - Evaluated performance on the test set, obtaining a test loss of 223.39 and a mean absolute error (MAE) of 10.61 years.

Skills and Competencies

The project developed and demonstrated the following skills and competencies:

- Proficiency in Python and deep learning frameworks such as TensorFlow/Keras.
- Understanding of transfer learning and fine-tuning techniques.
- Experience in data preprocessing and augmentation for robust model training.
- Ability to evaluate model performance using regression metrics.
- Problem-solving skills to address model overfitting and underfitting.

Feedback and Evidence

The model's performance, as indicated by an MAE of 10.61 years, is a reasonable baseline for age detection using a pre-trained architecture. Visualizations of predicted vs. actual ages and error distributions provided insights into the model's strengths and limitations. Feedback from peers suggested exploring alternative architectures and ensembling methods to improve accuracy.

Challenges and Solutions

1. **Image Preprocessing**:
    - **Challenge**: Raw images varied in size, quality, and lighting conditions, impacting model performance.
    - **Solution**: Applied preprocessing techniques such as resizing, normalization to enhance model generalization.

2. **Overfitting:**
    - Challenge: The model began overfitting after a few epochs.
    - Solution: Used only 10 epochs so that it won't overfit.

3. **High Error Margin:**
    - Challenge: Achieving lower MAE required significant experimentation.
    - Solution: Fine-tuned hyperparameters, including the learning rate and batch size, and experimented with different optimizers.

Outcomes and Impact

The project achieved the following outcomes:

- Developed an age detection model with a baseline MAE of 10.61 years.
- Gained hands-on experience with fine-tuning a pre-trained convolutional neural network (CNN).
- Identified areas for future improvement, such as incorporating additional datasets and exploring ensemble methods.

The impact of this work includes contributing to the broader application of age detection in real-world scenarios and providing a foundation for further research and optimization.

Conclusion

This project demonstrated the feasibility of using a fine-tuned VGG16 model for age detection with the UTKFace dataset. While the model achieved reasonable accuracy, there is scope for improvement through advanced techniques and additional data. The experience underscored the importance of transfer learning and regression techniques in computer vision, providing valuable insights and practical skills for future endeavors.