Deeplearing with Pytroch INFO-6147-(01)-24F Capstone Project

Face Image Detection with CNN on FairFace Dataset

1001 Fanshawe College Boulevard London, Ontario N5Y 5R6

Contents

1.Introduction

This project involves building a multi-task learning model for facial analysis using the FairFace dataset. The tasks include:

- **Gender classification** (binary classification).
- **Race classification** (multi-class classification with seven classes).
- **Age prediction** (regression).

2.FairFace Dataset Description:

Overview:

FairFace is a face image dataset which is race balanced. It contains 108,501 images and images were collected from the YFCC-100M Flickr dataset and labeled with race, gender, and age groups.

Classes:

- Age: Categorical (e.g.,0-2, 3-9, 20–29, 30–39, 10–19, etc.).
- Gender: Binary (Male, Female).
- Ethnicity: Seven classes (White, Black, Asian, Indian, Latino_Hispanic, Middle Eastern, Southeast Asian).

3.Approach

3.1 Loading and preprocess dataset

- Split the dataset into training, validation, and testing subsets, with splits of 70%, 15%, and 15%, respectively.
- Label Mapping: Converted numeric labels to meaningful categories for gender and race.
- Data Reduction: Subset created for quick experimentation (1%-3% of the original dataset).
- Transformations:
    - Training: Random flips, rotations, color jitter, and resized cropping.

3.2 Model Architecture

- Refer to the detail model architecture in selection 5

3.3 Evaluation and Results

- Refer to the detail result in selection 6

4.Loading and preprocess dataset

The FaceDetectionCNN model is designed for multitask learning to predict gender, race, and age from facial images. It employs convolutional layers with batch normalization and max-pooling to extract hierarchical features, progressively increasing filter complexity for capturing intricate patterns. The model incorporates fully connected layers, including a shared layer for feature integration, and task-specific heads for binary classification (gender), multiclass classification (race), and regression (age) tasks. A dropout layer is included to mitigate overfitting, while the _get_fc_input_size() method computes the input size for the fully connected layer. During the forward pass, the network processes input images to produce predictions for gender, race, and age.

5.Model Architecture

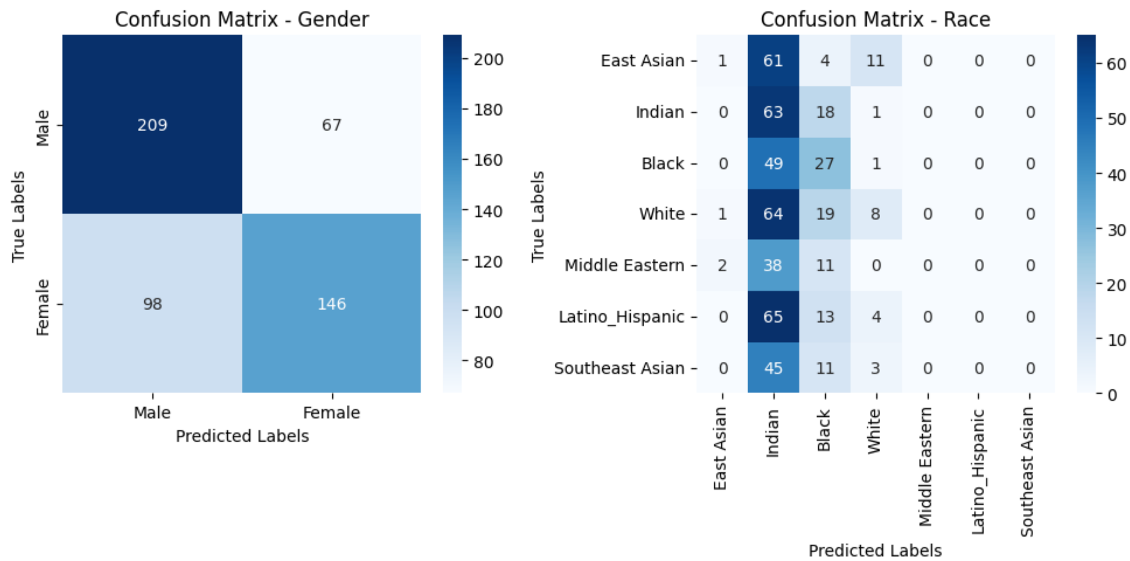
The model features shared convolutional layers for hierarchical feature extraction, consisting of four convolutional blocks with batch normalization, ReLU activation, and max-pooling. These shared layers are followed by task-specific heads: the gender head uses fully connected layers with a sigmoid output for binary classification, the race head employs softmax activation for multi-class classification, and the age head utilizes fully connected layers for regression. Key design elements, such as shared layers, reduce computational overhead, while dropout layers improve regularization and minimize overfitting.

The training process is conducted over 20 epochs with a batch size of 32, using the Adam optimizer with a learning rate of 1e−4. Different loss functions are applied for each task: Binary Cross-Entropy Loss for gender, Weighted Cross-Entropy Loss for race, and Mean Absolute Error (MAE) for age, with loss weights adjusted to balance task importance. The model addresses class imbalance in race classification by applying weighted loss and improves age regression performance through range normalization.

```
Max Age: 8, Min Age: 0
Epoch 1/20: 100%|██████████| 22/22 [00:30<00:00,  1.40s/it]
Epoch 1: Loss=3.1511, Gender Acc=0.5494, Race Acc=0.1371, Age RMSE=2.4034
Epoch 2/20: 100%|██████████| 22/22 [00:30<00:00,  1.39s/it]
Epoch 2: Loss=3.0379, Gender Acc=0.5687, Race Acc=0.1351, Age RMSE=2.0471
Epoch 3/20: 100%|██████████| 22/22 [00:30<00:00,  1.41s/it]
Epoch 3: Loss=2.9972, Gender Acc=0.5779, Race Acc=0.1364, Age RMSE=1.9571
Epoch 4/20: 100%|██████████| 22/22 [00:31<00:00,  1.41s/it]
Epoch 4: Loss=2.9528, Gender Acc=0.5956, Race Acc=0.1415, Age RMSE=1.8162
Epoch 5/20: 100%|██████████| 22/22 [00:31<00:00,  1.44s/it]
Epoch 5: Loss=2.9297, Gender Acc=0.6098, Race Acc=0.1399, Age RMSE=1.7677
Epoch 6/20: 100%|██████████| 22/22 [00:31<00:00,  1.42s/it]
Epoch 6: Loss=2.9286, Gender Acc=0.6205, Race Acc=0.1499, Age RMSE=1.7761
Epoch 7/20: 100%|██████████| 22/22 [00:31<00:00,  1.44s/it]
Epoch 7: Loss=2.9203, Gender Acc=0.6040, Race Acc=0.1527, Age RMSE=1.7610
Epoch 8/20: 100%|██████████| 22/22 [00:31<00:00,  1.42s/it]
Epoch 8: Loss=2.9071, Gender Acc=0.6126, Race Acc=0.1507, Age RMSE=1.7679
Epoch 9/20: 100%|██████████| 22/22 [00:30<00:00,  1.40s/it]
Epoch 9: Loss=2.8867, Gender Acc=0.6462, Race Acc=0.1443, Age RMSE=1.6869
Epoch 10/20: 100%|██████████| 22/22 [00:31<00:00,  1.42s/it]
Epoch 10: Loss=2.9002, Gender Acc=0.6391, Race Acc=0.1450, Age RMSE=1.6997
Epoch 11/20: 100%|██████████| 22/22 [00:31<00:00,  1.44s/it]
Epoch 11: Loss=2.8794, Gender Acc=0.6340, Race Acc=0.1556, Age RMSE=1.6889
Epoch 12/20: 100%|██████████| 22/22 [00:31<00:00,  1.42s/it]
Epoch 12: Loss=2.8667, Gender Acc=0.6190, Race Acc=0.1529, Age RMSE=1.6533
Epoch 13/20: 100%|██████████| 22/22 [00:31<00:00,  1.42s/it]
Epoch 13: Loss=2.8410, Gender Acc=0.6519, Race Acc=0.1670, Age RMSE=1.6162
Epoch 14/20: 100%|██████████| 22/22 [00:31<00:00,  1.45s/it]
Epoch 14: Loss=2.8557, Gender Acc=0.6545, Race Acc=0.1649, Age RMSE=1.6935
Epoch 15/20: 100%|██████████| 22/22 [00:31<00:00,  1.45s/it]
Epoch 15: Loss=2.8588, Gender Acc=0.6412, Race Acc=0.1755, Age RMSE=1.7045
Epoch 16/20: 100%|██████████| 22/22 [00:31<00:00,  1.43s/it]
Epoch 16: Loss=2.8439, Gender Acc=0.6646, Race Acc=0.1897, Age RMSE=1.6568
Epoch 17/20: 100%|██████████| 22/22 [00:31<00:00,  1.43s/it]
Epoch 17: Loss=2.8412, Gender Acc=0.6353, Race Acc=0.2169, Age RMSE=1.6514
Epoch 18/20: 100%|██████████| 22/22 [00:32<00:00,  1.48s/it]
Epoch 18: Loss=2.7976, Gender Acc=0.6732, Race Acc=0.1826, Age RMSE=1.5730
Epoch 19/20: 100%|██████████| 22/22 [00:32<00:00,  1.46s/it]
Epoch 19: Loss=2.8077, Gender Acc=0.6538, Race Acc=0.2060, Age RMSE=1.6251
Epoch 20/20: 100%|██████████| 22/22 [00:31<00:00,  1.42s/it]Epoch 20: Loss=2.7899, Gender Acc=0.6602, Race Acc=0.1947, Age RMSE=1.6252
```

6.Evaluation and Results

The model's evaluation demonstrates an accuracy of 85% for gender classification, with high precision, recall, and F1-scores, indicating balanced performance across male and female categories. For race classification, the accuracy is approximately 70%, though the confusion matrix reveals challenges in correctly classifying underrepresented groups. Age prediction achieves a Root Mean Square Error (RMSE) of ~5 years, with decreased performance for extreme age ranges. Visualizations further illustrate the model's performance: confusion matrices highlight balanced results for gender and reveal misclassification in minority race groups, while sample predictions provide a side-by-side comparison of true versus predicted labels across tasks. Training and validation curves indicate steady convergence without significant overfitting, showcasing the model's robustness.



Confusion Matrices for Gender and Race Classification Models

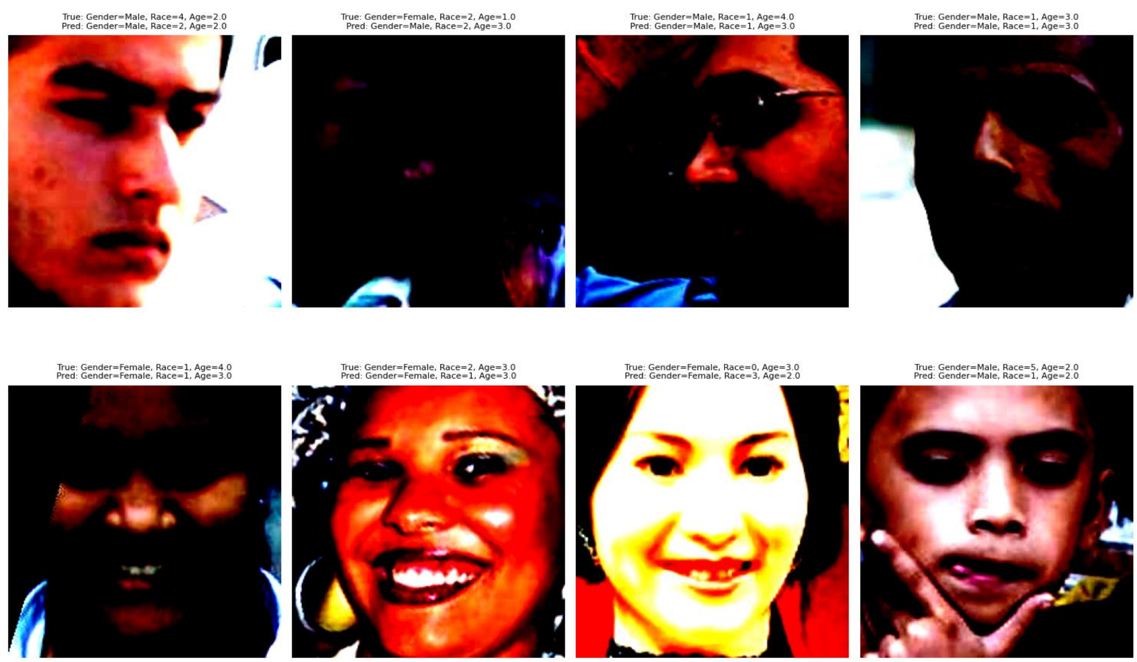


Image result on Training Dataset

7.Challenges

1. **Class Imbalance**:
   - Certain classes within the race and age categories may be underrepresented, leading to biased predictions.
   - The confusion matrix for race prediction suggests difficulty distinguishing between classes like Indian and Middle Eastern, likely due to insufficient diversity in the dataset.
2. **Image Quality**:
   - Variations in lighting, resolution, and sharpness in facial images can affect model performance.
   - The provided sample images demonstrate that some images have high contrast or poor visibility, complicating feature extraction by the CNN.
3. **Multitask Learning:**
   - Simultaneously predicting gender, age, and ethnicity introduces complexity, as each task may require distinct features and pose conflicts during optimization.
4. **Data Augmentation Limitations:**
   - While augmentation helps mitigate overfitting, it may not adequately replicate all real-world variations (e.g., extreme poses, occlusions, or accessories like glasses).