Efecan Kasapoğlu

31038

CS412

HW4

Colab Notebook Link: https://drive.google.com/file/d/1u5kjWpq7b0QCEFUfVVu0K-DustqKiyGz/view?usp=sharing

1. Introduction

In this project, I classified the gender of individuals using the CelebA30k dataset and developed a deep learning model. I performed the learning process with VGG16. Four different experiments were implemented with two different learning rates and two different fine-tuning strategies to examine the factors affecting the performance of the model. The obtained results were analyzed with accuracy, loss graphs and complexity matrix.

2. Methods

I split the dataset (30000) into 80% (24000) training, 10% (3000) validation, and 10% (3000) test sets. After that I used four different training configurations:

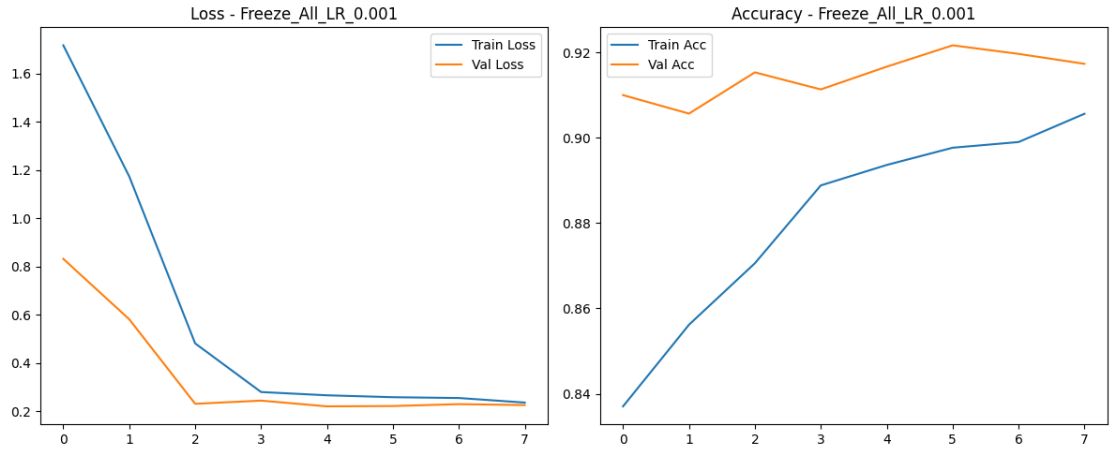
- Freezing all convolutional layers and 0.001 learning rate
- Freezing all convolutional layers and 0.0001 learning rate
- Freezing all weights, but fine-tuning the last convolutional block along with the classifier head and 0.001 learning rate
- Freezing all weights, but fine-tuning the last convolutional block along with the classifier head and 0.0001 learning rate

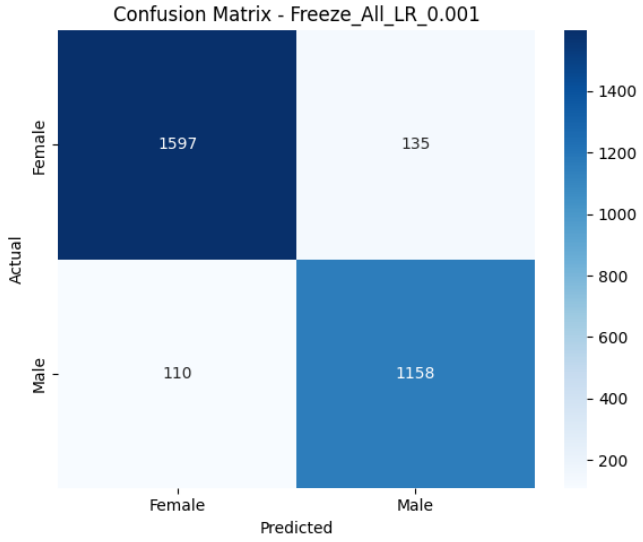
The number of epochs is fixed to 10.

3. Results

a. Freezing all convolutional layers and 0.001 learning rate

This experiment is done by 0.001 learning rate and freezing all convolutional layers and training only the classifier head.

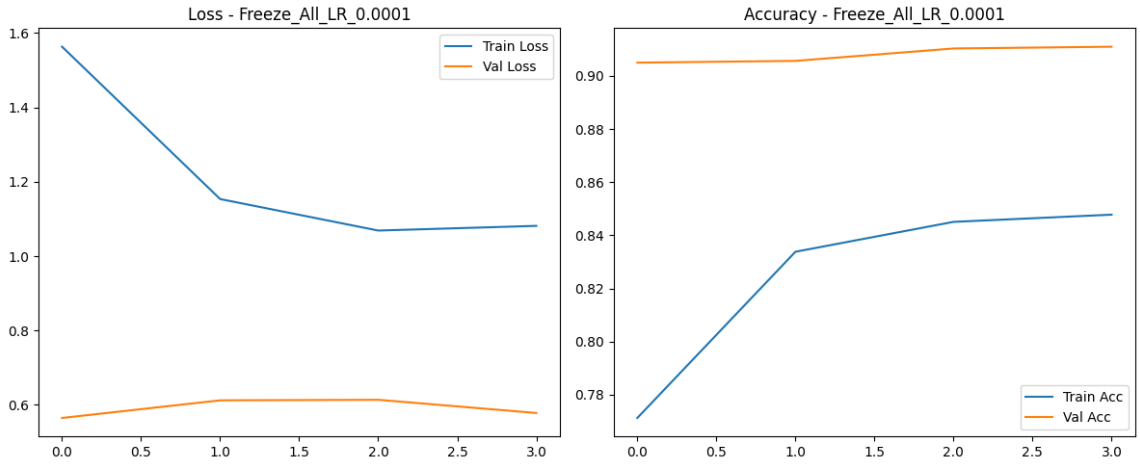


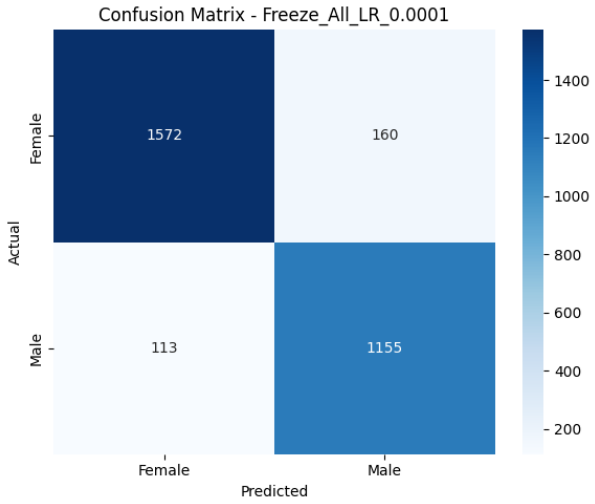


The test accuracy is 91.38%. Female precision is 0.94, recall is 0.92, f1 score is 0.93. Male precision is 0.9, recall is 0.91 and f1 score is 0.90. True answer for female is 1597 and wrong answer is 135. True answer for male is 1158 and wrong answer is 110.

b. Freezing all convolutional layers and 0.0001 learning rate

This experiment is done by 0.0001 learning rate and freezing all convolutional layers and training only the classifier head.

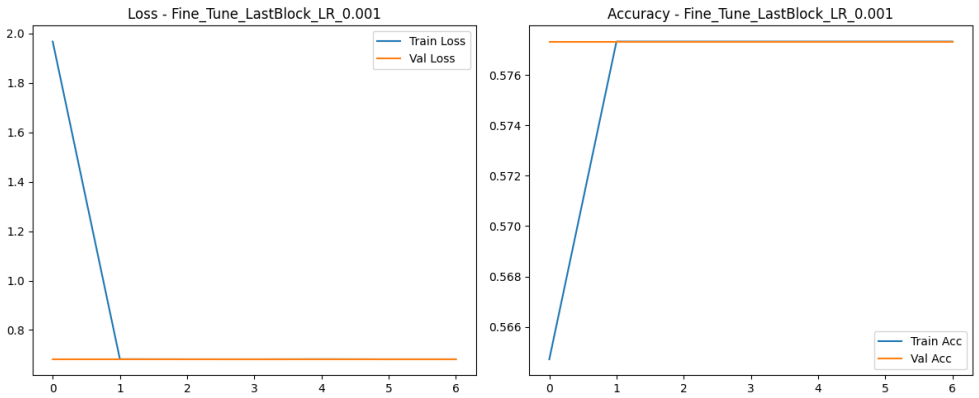


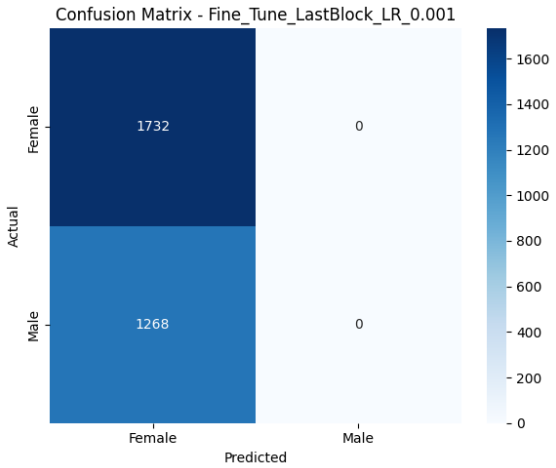


The test accuracy is 90.90%. Female precision is 0.93, recall is 0.91, f1 score is 0.92. Male precision is 0.88, recall is 0.91 and f1 score is 0.89. True answer for female is 1572 and wrong answer is 160. True answer for male is 1155 and wrong answer is 113.

c. Freezing all weights, but fine-tuning the last convolutional block along with the classifier head and 0.001 learning rate

This experiment is done by 0.001 learning rate and freezing all weights; however, fine-tuning the last convolutional block along with the classifier head.

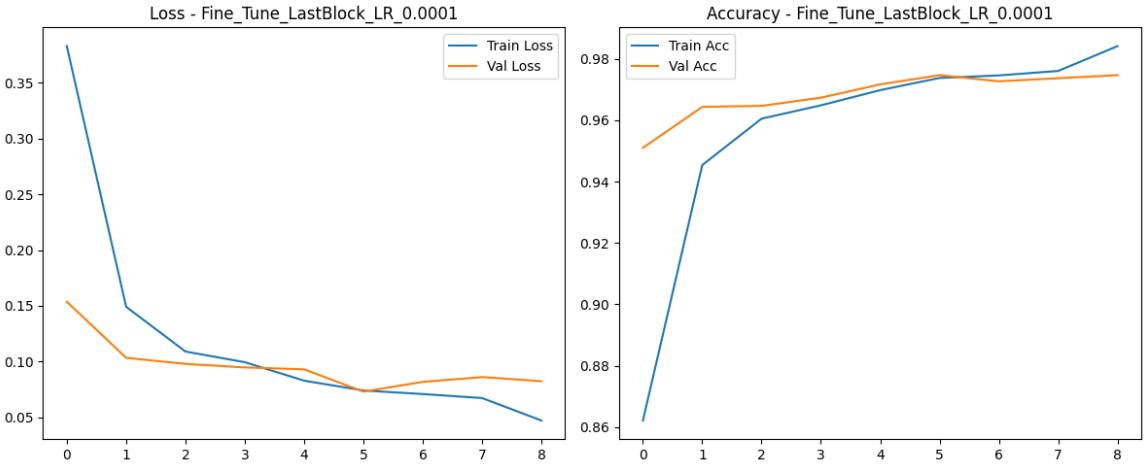


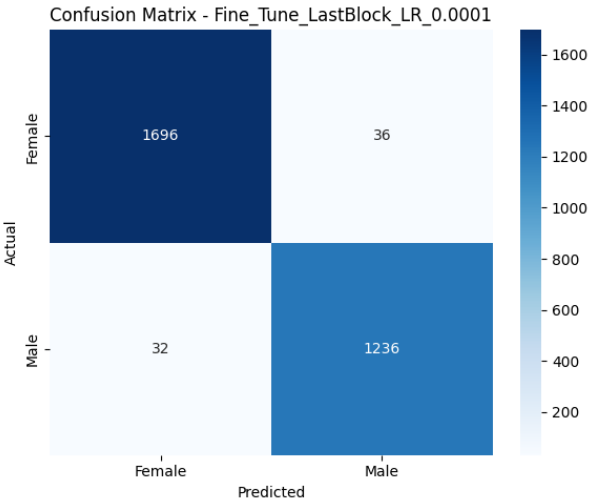


The test accuracy is 57.73%. Female precision is 0.58, recall is 1.00, f1 score is 0.73. Male precision is 0.00, recall is 0.00 and f1 score is 0.00. True answer for female is 1732 and wrong answer is 0. True answer for male is 0 and wrong answer is 1268.

d. Freezing all weights, but fine-tuning the last convolutional block along with the classifier head and 0.0001 learning rate

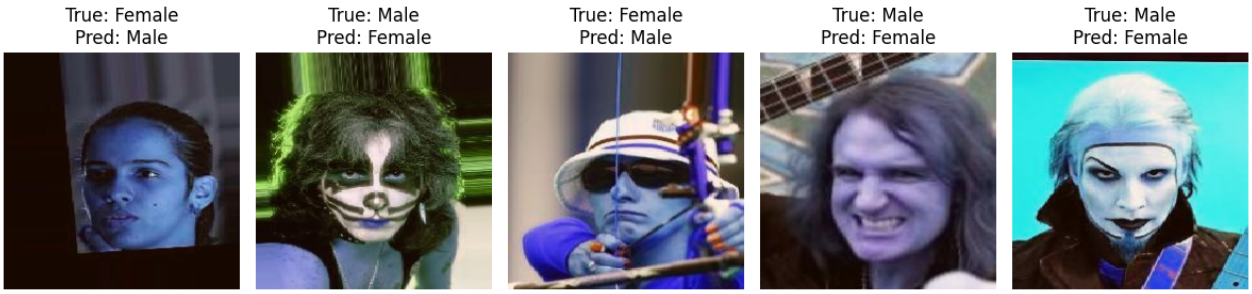
This experiment is done by 0.0001 learning rate and freezing all weights; however, fine-tuning the last convolutional block along with the classifier head.





The test accuracy is 97.73%. Female precision is 0.98, recall is 0.98, f1 score is 0.98. Male precision is 0.97, recall is 0.97and f1 score is 0.97. True answer for female is 1696 and wrong answer is 36. True answer for male is 1236 and wrong answer is 32.

Below are five randomly selected misclassified examples from the test set using the freezing all weights, but fine-tuning the last convolutional block along with the classifier head and 0.0001 learning rate model.



4. Discussion

Among the four different experiments I conducted, the best performance was with “Freezing all weights, but fine-tuning the last convolutional block along with the classifier head and 0.0001 learning rate”. It was the best result with a test accuracy of 97.73%. However; the worst performance was with “Freezing all weights, but fine-tuning the last convolutional block along with the classifier head and 0.001 learning rate”. It was the worst performing model with 57.73%.