



https://github.com/duringbug/ai_project_5.git

```
Epoch 1/10
100/100 [=====] - 21s 196ms/step -
loss: 2.6067 - accuracy: 0.5572 - val_loss: 0.9260 - val_accuracy: 0.5750
Epoch 2/10
100/100 [=====] - 17s 171ms/step -
loss: 0.7613 - accuracy: 0.6722 - val_loss: 0.9643 - val_accuracy: 0.6087
Epoch 3/10
100/100 [=====] - 18s 179ms/step -
loss: 0.3717 - accuracy: 0.8788 - val_loss: 1.0058 - val_accuracy: 0.6212
Epoch 4/10
100/100 [=====] - 18s 176ms/step -
loss: 0.1527 - accuracy: 0.9619 - val_loss: 1.2137 - val_accuracy: 0.6100
Epoch 5/10
100/100 [=====] - 18s 184ms/step -
loss: 0.1005 - accuracy: 0.9797 - val_loss: 1.2899 - val_accuracy: 0.6175
Epoch 6/10
100/100 [=====] - 18s 181ms/step -
loss: 0.0693 - accuracy: 0.9866 - val_loss: 1.3996 - val_accuracy: 0.6250
Epoch 7/10
100/100 [=====] - 18s 184ms/step -
loss: 0.0663 - accuracy: 0.9906 - val_loss: 1.4958 - val_accuracy: 0.6112
Epoch 8/10
100/100 [=====] - 19s 188ms/step -
loss: 0.0500 - accuracy: 0.9919 - val_loss: 1.4828 - val_accuracy: 0.6313
Epoch 9/10
100/100 [=====] - 19s 190ms/step -
loss: 0.0547 - accuracy: 0.9912 - val_loss: 1.4053 - val_accuracy: 0.6425
Epoch 10/10
100/100 [=====] - 20s 196ms/step -
loss: 0.0384 - accuracy: 0.9922 - val_loss: 1.5821 - val_accuracy: 0.6313
```

出现过拟合现象

解决方法:增加Dropout层，减小维度以减小复杂度

使用EarlyStopping：使用EarlyStopping回调函数，在验证集准确率不再提高时停止训练。

多模态情感分析下的正确率约为64%

设计考虑和亮点：

1. **多模态输入**：该模型采用了多模态输入，即图像和文本。这样的设计可以捕捉来自不同数据类型的信息，提高模型对数据的理解和表达能力。
2. **卷积神经网络（CNN）和长短时记忆网络（LSTM）结合**：图像部分使用了卷积神经网络，而文本部分使用了嵌入层和长短时记忆网络。这种结合可以有效地处理图像和文本数据，充分利用它们的特征。
3. **空间和时间上的 Dropout**：模型中使用了 Dropout 层，既有空间上的 Dropout 用于图像部分，也有时间上的 Dropout 用于文本部分。这有助于防止过拟合，提高模型的泛化能力。
4. **合并层**：通过使用 `concatenate` 合并图像和文本处理结果，模型能够融合两个不同来源的信息。这种融合可能使模型更全面地理解数据，提高预测的准确性。
5. **使用了 EarlyStopping 回调**：在模型训练过程中，采用了 EarlyStopping 回调函数，当验证集上的损失不再改善时，可以提前停止训练，防止过拟合。
6. **模型的保存和加载**：模型在训练后被保存到文件（`model.h5`），然后通过 `load_model` 函数重新加载，以便在未来的时间点进行预测。
7. **处理未标记数据**：通过加载未标记的数据，进行预测，并输出预测结果，实现了对新数据的分类。

消融实验结果

只有img:

```
(env) tangjianfeng@tangjianfengdeMacBook-Air project_5 % python -u "/Volumes/TJF_YINGPAN/"
Epoch 1/4
200/200 [=====] - 51s 255ms/step -
loss: 0.9372 - accuracy: 0.5903 - val_loss: 0.9438 - val_accuracy: 0.5925
Epoch 2/4
200/200 [=====] - 54s 272ms/step -
loss: 0.9063 - accuracy: 0.5984 - val_loss: 0.9283 - val_accuracy: 0.5875
Epoch 3/4
200/200 [=====] - 55s 277ms/step -
loss: 0.8996 - accuracy: 0.5991 - val_loss: 0.9082 - val_accuracy: 0.5763
Epoch 4/4
200/200 [=====] - 56s 280ms/step -
loss: 0.8803 - accuracy: 0.6025 - val_loss: 0.9777 - val_accuracy: 0.5813
```

没有过拟合，但准确率低（图片信息的维度大）

只有text

```
(env) tangjianfeng@tangjianfengdeMacBook-Air project_5 % python -u "/Volumes/TJF_YINGPAN/"
Epoch 1/4
200/200 [=====] - 12s 56ms/step -
loss: 0.9295 - accuracy: 0.5938 - val_loss: 0.9016 - val_accuracy: 0.5875
Epoch 2/4
200/200 [=====] - 11s 54ms/step -
loss: 0.7737 - accuracy: 0.6747 - val_loss: 0.8880 - val_accuracy: 0.6400
Epoch 3/4
200/200 [=====] - 11s 54ms/step -
loss: 0.4582 - accuracy: 0.8238 - val_loss: 1.0216 - val_accuracy: 0.6325
Epoch 4/4
200/200 [=====] - 11s 54ms/step -
loss: 0.3117 - accuracy: 0.8791 - val_loss: 1.1428 - val_accuracy: 0.6375
```

发生过拟合，但准确率高（文字信息的维度较小）

问题:

可能是我combined = concatenate([x, y])这个地方出错了，没有很好的利用二者有利的信息，把

噪声也加在一起了

我有个想法训练两个model，一个是img(x)的，一个是text(y)的，输出是0，1，2对应的得分，将这六个数据再进行一次model的训练,结果如下,我查询了其他多模态情感识别的网页，发现正确率基本都是65%~69%之间

多模态融合模型在验证集上的结果

```
(env) tangjianfeng@tangjianfengdeMacBook-Air project_5 % python -u "/Volumes/TJF_YINGPAN/
Epoch 1/4
200/200 [=====] - 69s 339ms/step -
loss: 0.9558 - accuracy: 0.5878 - val_loss: 0.9036 - val_accuracy: 0.5875
Epoch 2/4
200/200 [=====] - 69s 343ms/step -
loss: 0.7803 - accuracy: 0.6712 - val_loss: 0.8155 - val_accuracy: 0.6425
Epoch 3/4
200/200 [=====] - 69s 344ms/step -
loss: 0.3456 - accuracy: 0.8737 - val_loss: 1.0050 - val_accuracy: 0.6350
Epoch 4/4
200/200 [=====] - 70s 348ms/step -
loss: 0.1340 - accuracy: 0.9563 - val_loss: 1.1758 - val_accuracy: 0.6363
```