Let Me Know You

Sentiment analysis and age gender classfication
with neural network

Kun Liang, Harry Qi, Bohan Wu, Naiqiao Ye

04/29/2022

CONTENT

Introduction

Problem Statement

JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

- Sentiment Analysis
- (In progress.) Age and Gender Prediction
- Convolutional Neural Network!
- Multi-Classification

Motivation

- Problem: Can machine learn human faces and expressions from images?
- Emotions affect decision
  - “Emotions provide embodied information about the costs and benefits of anticipated action …” [1]
- Age affects decision processing
  - “Older people rely more on emotions and experience and less on reason than do younger people”[2]

Interests

JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

- Interested in related research of how emotion/age/gender impact decision making
- Connecting machine learning to sociology
- Experience in-depth computer vision model design and tunning

Datasets — Sentiment Analysis

JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

FER 2013 (Facial Expression Recognition 2013)

- Training set: 28709 examples
- Validation set: 3589 examples
- Test set: 3589 examples

Training set

Angry：3995
Disgust：436
Fear：4097
Happy：7215
Sad：4830
Surprise：3171
Neutral：4965

Datasets — Sentiment Analysis

JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

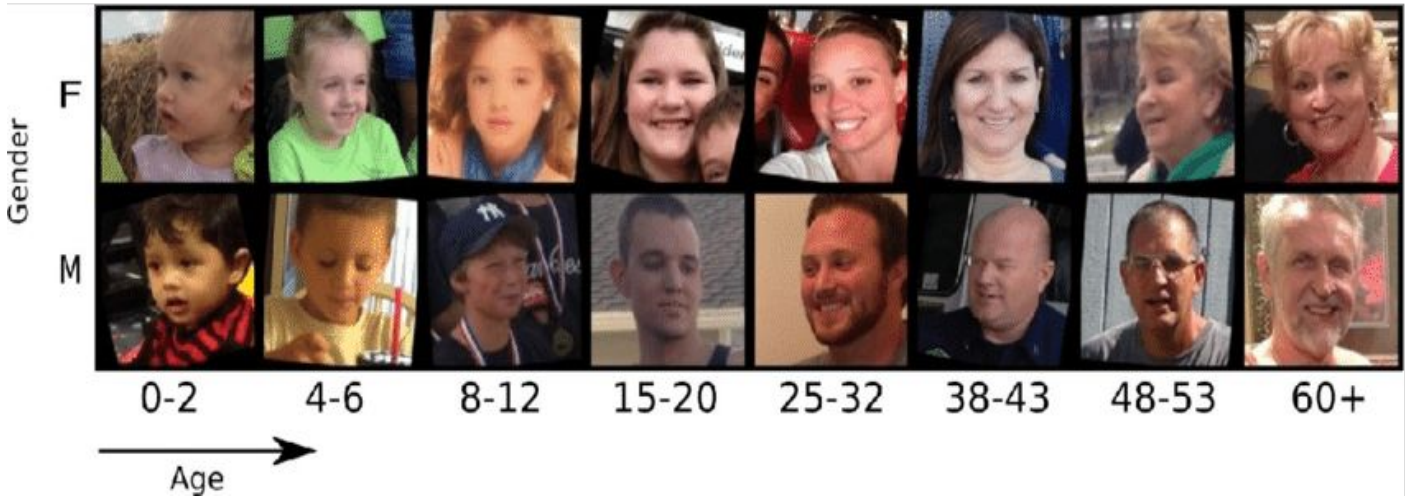
FER 2013 (Facial Expression Recognition 2013)

- 0=Angry 1=Disgust
  2=Fear 3=Happy
  4=Sad 5=Surprise
  6=Neutral
- Resolution: 48 × 48 × 3

Datasets — Age and Gender

Age And Gender Classification on Adience Gender

- Total number of pictures:
  - 19370 pictures
- Resolution:
  - 600 × 600 × 3
- Gender:
  - female: 9372
  - male: 8120



Picture website:
https://www.researchgate.net/figure/Adience-sample-Sample-of-each-age-group-and-gender-from-the-fourth-fold-of-the-Adience_fig1_318084198

Datasets — Age and Gender

Age And Gender Classification on Adience Gender

- Age groups:
  - (0-2),(4-6),(8-12),(15-20),

    (25-32),(38-43),(48-53)

    (60+)
- Re-define age groups:
  - attach labels to different age groups
- Two models:
  - age prediction
  - gender prediction

Picture website:
https://www.researchgate.net/figure/Adience-sample-Sample-of-each-age-group-and-gender-from-the-fourth-fold-of-the-Adience_fig1_318084198

Datasets — FER2013 & Adience

Data Preprocessing

- Data Normalization
  - Rescaling
- Data Augmentation
  - Rotation
  - Zoom
  - Horizontal Flip

Deliverables

Deliverables

- Must Accomplish (before presentation):
  - Data Processing for FER2013 & Adience(Interim Goal)
  - Baseline Model for sentiment analysis(Interim Goal)
  - ResNet Model for sentiment analysis
  - Hyper-tuned Baseline Model for sentiment analysis
  - Research previous work on FER2013 and implement one model as comparison (VGG16)

Deliverables

JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

- Must Accomplish (before presentation):
  - Data Processing for FER2013 & Adience(Interim Goal)
  - Baseline Model for sentiment analysis(Interim Goal)
  - ResNet Model for sentiment analysis
  - Hyper-tuned Baseline Model for sentiment analysis
  - Research previous work on FER2013 and implement one model as comparison (VGG16)

All Accomplished!

Deliverables

- Expect to Accomplish (before code submission):
  - Create baseline model for Adience dataset
  - 60%+ Validation Accuracy on our best model for sentiment analysis (best effort)
  - Finish implementing model for Adience
  - Fine tune Adience datasets and achieve maximum accuracy (85%+)

Methods

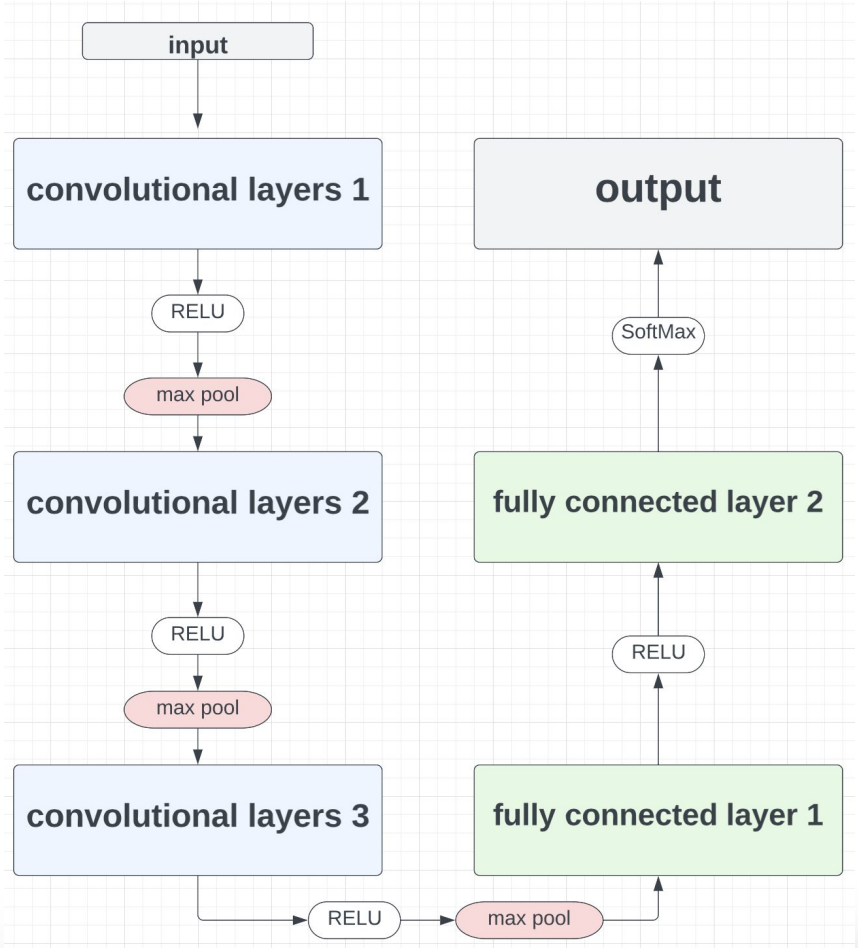
Methods

FER2013 for sentiment analysis

- Baseline Model
- Baseline CNN Model With Mega Tunning
- VGG16
- ResNet
- VGG16 + Attention
- Evaluation based on (Cross Entropy) Loss and Accuracy

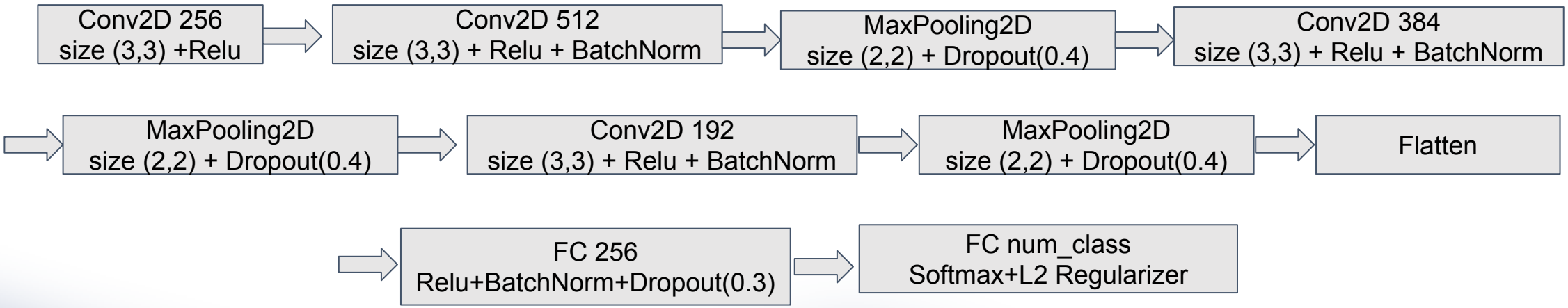
Baseline Model

- A Simple Convolutional Neural Network
- 3 convolution layers
- 2 fully connected linear layers
- Maxpool as pooling layer
- ReLU as activation functions
- Softmax as output layer activation function



Baseline CNN Model With Mega Tunning

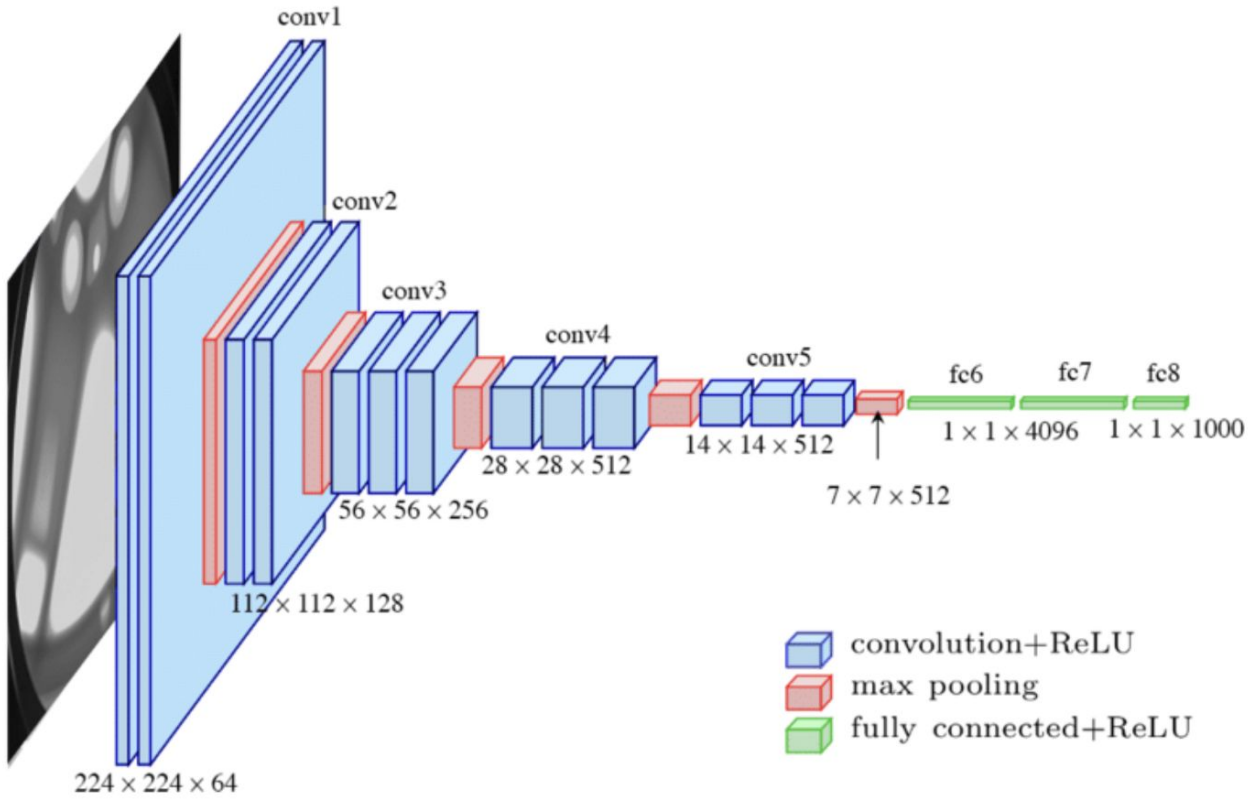
- One more convolution layer
  - Filter size tunning
- BatchNorm layer
- Dropout layer
- Flatten
- Regularizer



VGG16 Model

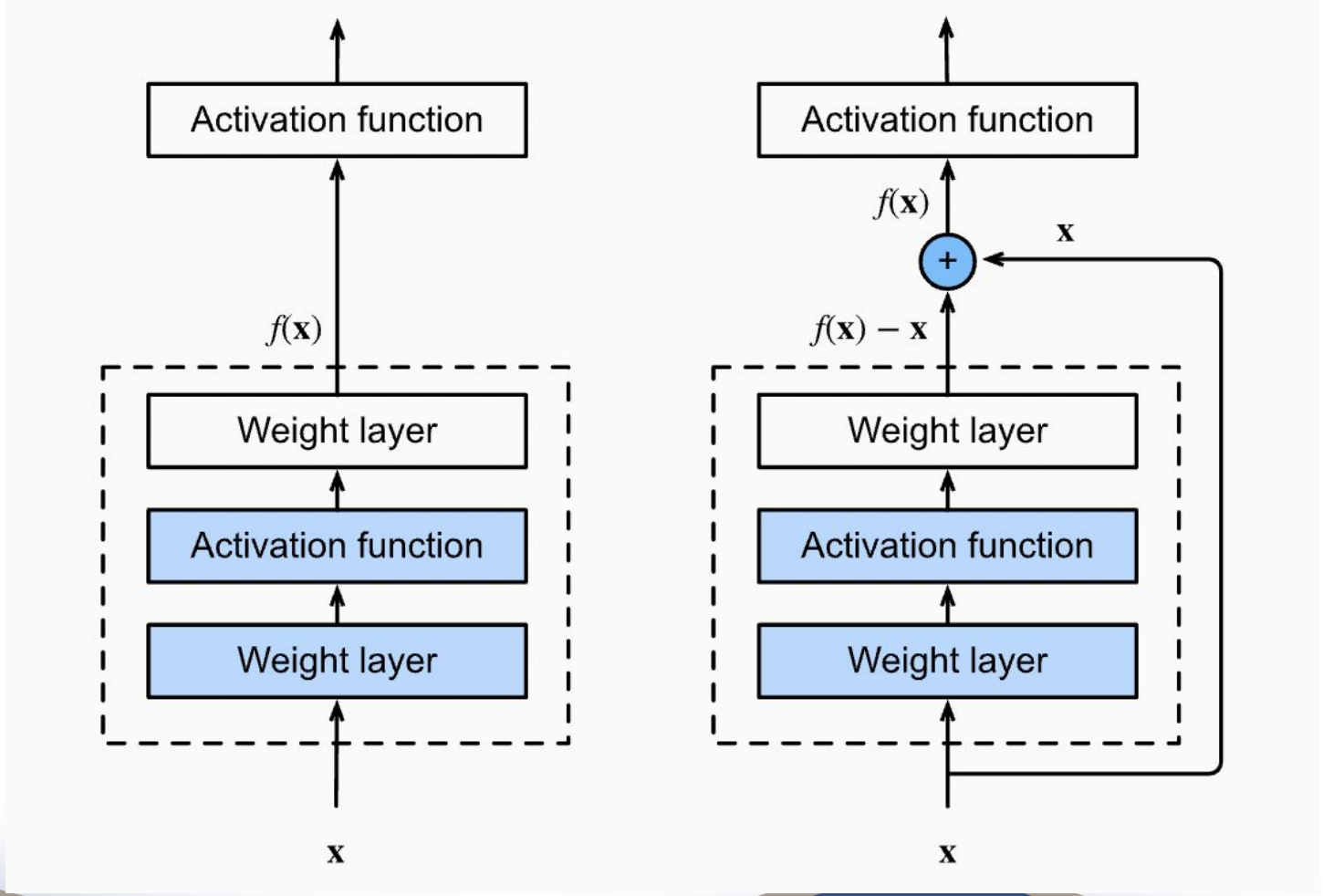
- VGG16: (Visual Geometry Group)
  - 13 convolutional layers + RELU
    - layers with parameters, W
  - 5 max pooling layers
    - no parameters
  - 3 fully connected layers
    - layers with parameters, W
- **16** layers with tunable parameters:
  - 13 convolutional layers
  - 3 fully connected layers

- Advantages / Disadvantages
  - simple structures
  - huge amount of parameters



picture from https://medium.com/mlearning-ai/an-overview-of-vgg16-and-nin-models-96e4bf398484
the paper link: https://arxiv.org/pdf/1409.1556.pdf [4]

ResNet Model



A regular block (Left)

A residual block (right)

ResNet Model

- Residual Neural Network (ResNet)
- *ResNet* models are implemented with double- or triple- layer skips
  - Nonlinearities (e.g. ReLU) and BatchNorm in between
  - An additional weight matrix may be used to learn the skip weights
- Why skip?

More at https://doi.org/10.48550/arXiv.1512.03385

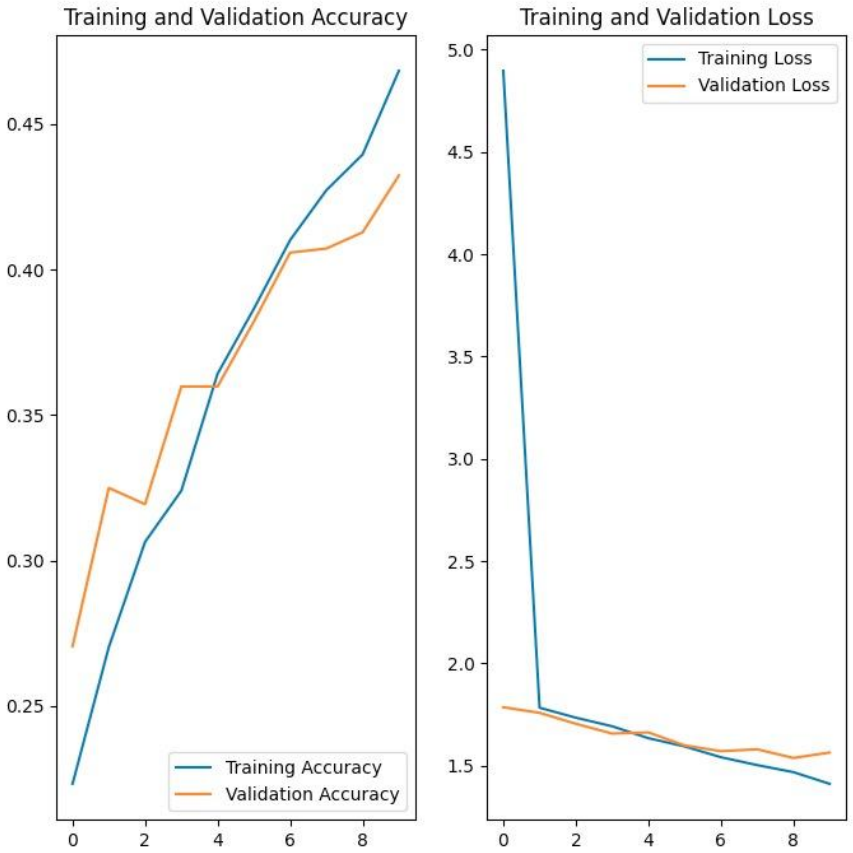
Results

Results

- Baseline Model
  - Highest Train/Valid Accuracy: 0.48/0.42
- Baseline Model with mega-tuning
  - Highest Train/Valid Accuracy: 0.73/0.65
- VGG16
  - Highest Train/Valid Accuracy: 0.84/0.51
- ResNet
  - Highest Train/Valid Accuracy: 0.71/0.61

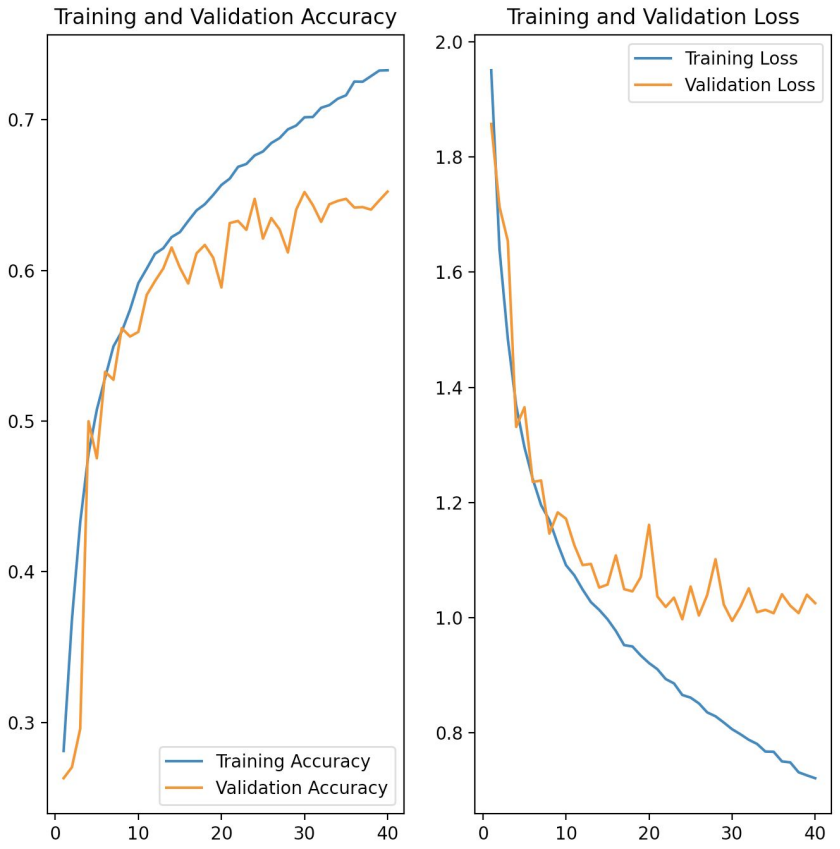
Results

- Baseline Model (Regular CNN)
  - 10 epchos, batch size 512
  - Train_Acc: 0.48
  - Val_Acc: 0.42
  - Action: Use dropout layer and Regularizer (L2) in tunning the model



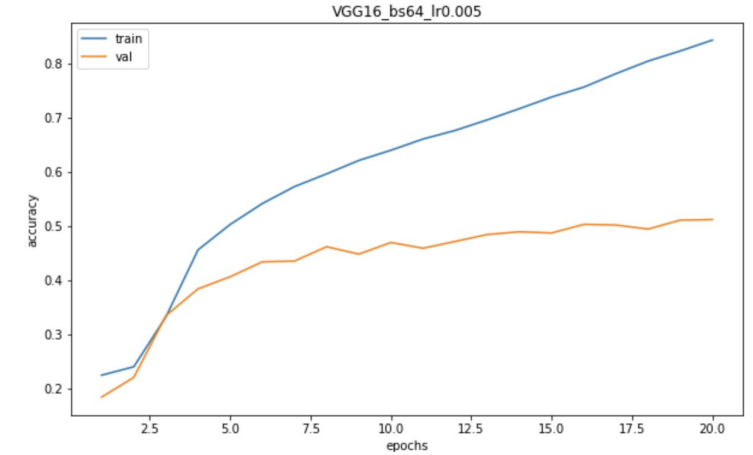
Results

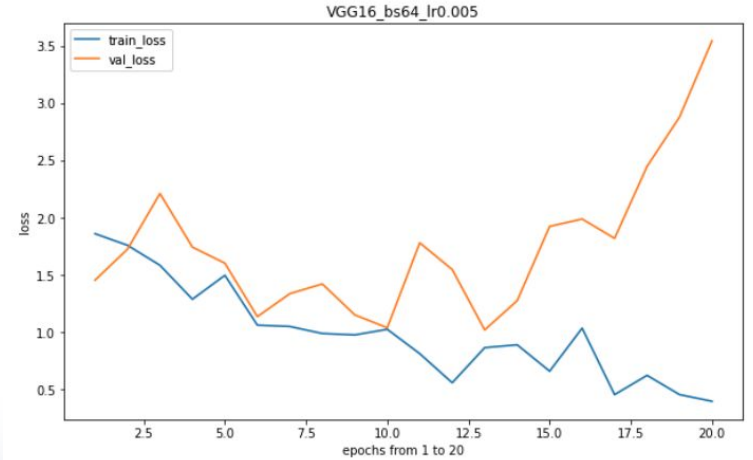
- Baseline Model With Mega Tunning
  - 40 epchos, batch size 128
  - Train_Acc: 0.73
  - Val_Acc: 0.65



Results

- VGG16
  - 20 epchos, batch size 64
  - Learning rate 0.0005
  - No drop-out stages
  - Adding drop-out stages reduces overall performance (Train ACC:0.67 / Val ACC: 0.43)
  - VGG16 improves the accuracy comparing to baseline, but suffers from over-fitting

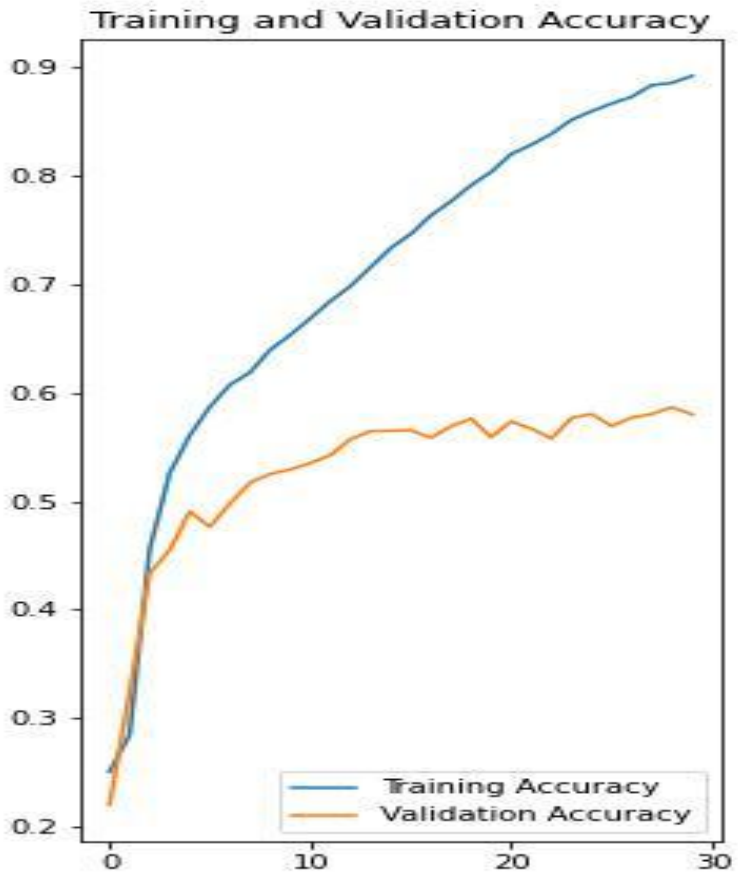




<Figure size 432x288 with 0 Axes>

Results

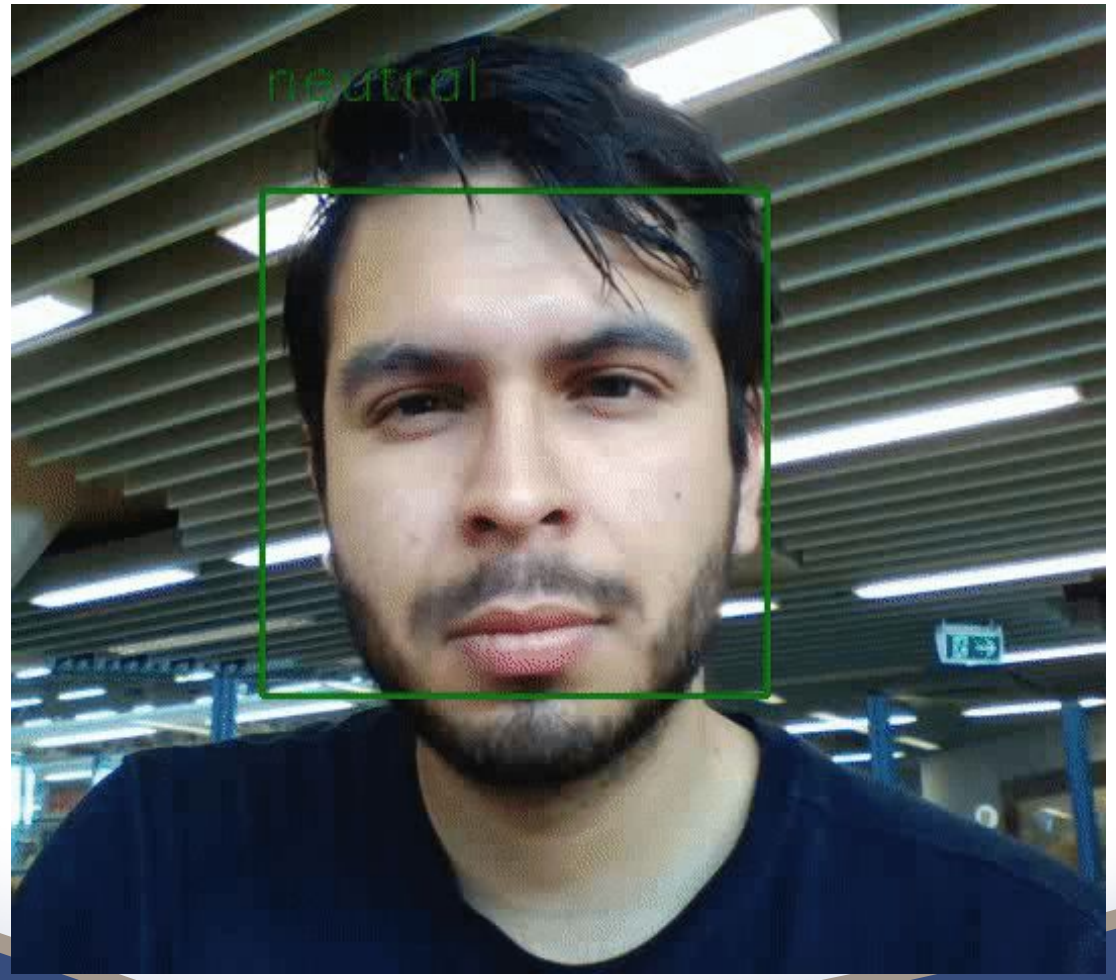
- ResNet
  - 20 epchos, batch size 64
  - Exceeds baseline performance
  - Suffers from over-fitting
  - Adding drop-out stages reduces overall performance



Future Deliverables

Deliverables

- Would like to accomplish (future plans)
  - Implement real-time video-feed prediction
  - Merge all three models
    - sentiment, gender, and age prediction in real-time
  - Face alignment in real-time video
  - …



Gain

- What Surprised us?
  - Pytorch vs. Tensorflow
  - Hyperparameter/Mega Tuning
- Takeway
  - Stick with one framework
  - Research

Reference

[1] Zadra, Jonathan R, and Gerald L Clore. “Emotion and perception: the role of affective information.” Wiley interdisciplinary reviews. Cognitive science vol. 2,6 (2011): 676-685. doi:10.1002/wcs.147
[2] Peters E, Hess TM, Västfjäll D, Auman C. Adult age differences in dual information processes: Implications for the role of affective and deliberative processes in older adults’ decision making. Perspectives on Psychological Science. 2007;2(1):1–23. doi: 10.1111/j.1745-6916.2007.00025.x.
[3] Goodfellow, Ian J., et al. “Challenges in Representation Learning: A Report on Three Machine Learning Contests.” Neural Networks, vol. 64, 2015, pp. 59–63., https://doi.org/10.1016/j.neunet.2014.09.005.
[4] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014).

THANKS !

Kun Liang Yanjie Qi Bohan Wu Naiqiao Ye