Convolutional Neural Networks: Selfie Generalization

Tom Youngblood

Main Objective

Main Objective: **Research Question**

Primary Research Question: explore how well a CNN, trained on easily generated user-generated data like **selfies**, performs on standard, **generalized** image recognition tasks.

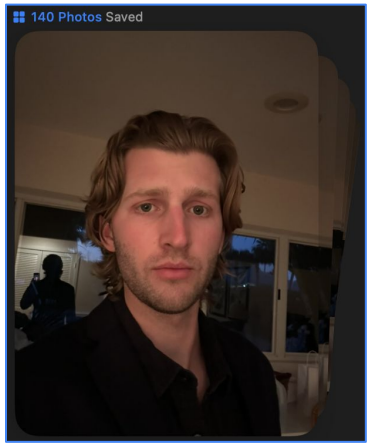
Main Objective: Process

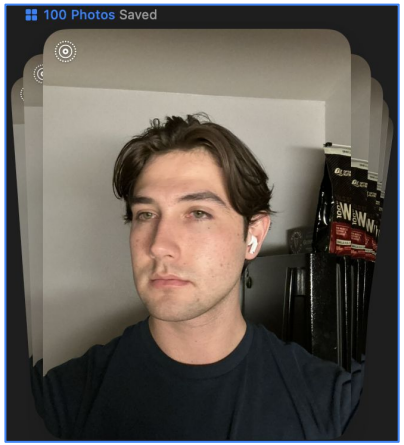
1. **Manual Data Collection**
2. **Data Cleaning**
3. **Initial Model Building**
    a. Making Predictions on Selfie Images
    b. Making Predictions on Non-Selfie Images
4. **Iterative Data-Tuning and Model-Tuning**
5. **Conclusion**

Data Collection Process

Data Collection Process

Collected ~400 Photos from Close Friends



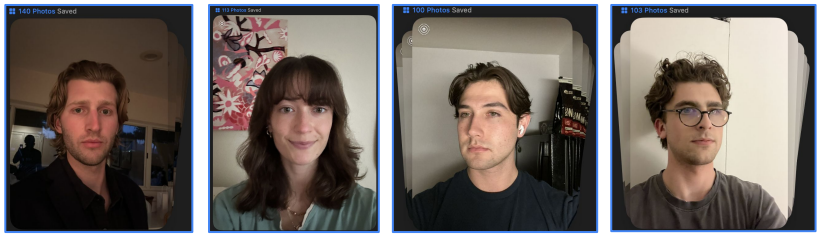


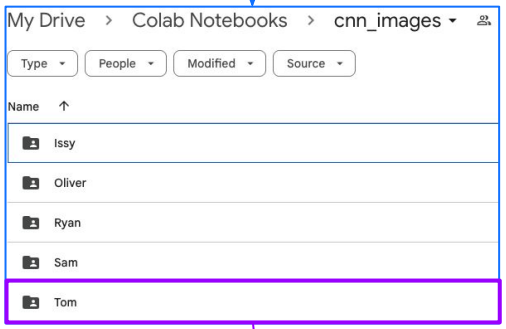
\+ 100 selfies of my own

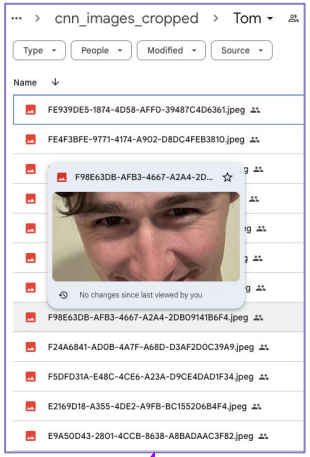
Data Cleaning Process

Data Cleaning Process

Aggregated Data in TF/Keras Generator Readable Format: Google Drive







Initial Model Building: Model 1

Model 1: Selfie Performance

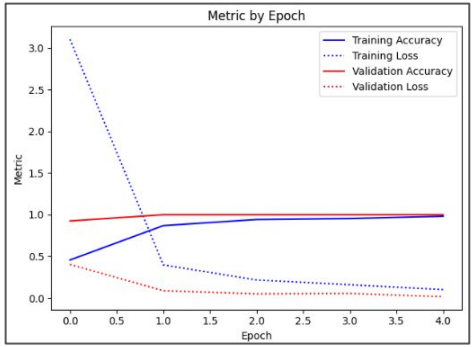
Initial Class Structure

```
Class: Issy: 113 images
Class: Oliver: 141 images
Class: Ryan: 100 images
Class: Sam: 104 images
Class: Tom: 86 images
```

VGG16 Model: Selfie Performance

```
Epoch 1/10
14/14 ━━━━━━━━━━━━━━━━━━━━ 41s 3s/step - accuracy: 0.3013 - loss: 4.7760 - val_accuracy: 0.9252 - val_loss: 0.4029
Epoch 2/10
14/14 ━━━━━━━━━━━━━━━━━━━━ 33s 2s/step - accuracy: 0.8407 - loss: 0.4826 - val_accuracy: 1.0000 - val_loss: 0.0877
Epoch 3/10
14/14 ━━━━━━━━━━━━━━━━━━━━ 34s 2s/step - accuracy: 0.9314 - loss: 0.2447 - val_accuracy: 1.0000 - val_loss: 0.0494
Epoch 4/10
14/14 ━━━━━━━━━━━━━━━━━━━━ 32s 2s/step - accuracy: 0.9669 - loss: 0.1510 - val_accuracy: 1.0000 - val_loss: 0.0544
Epoch 5/10
14/14 ━━━━━━━━━━━━━━━━━━━━ 37s 3s/step - accuracy: 0.9792 - loss: 0.0993 - val_accuracy: 1.0000 - val_loss: 0.0157
```

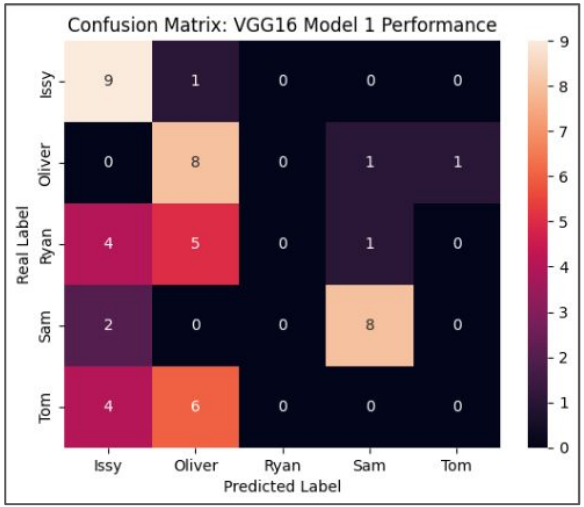
Selfie Performance: Graphical View



VGG16 Model: Example Predictions

VGG16 Model: CNN Designed for Image Classification Tasks

Model 1: Generalization to Non-Selfie Images



- Issy: 19 predictions. 47% accuracy.
- Oliver: 20 predictions. 40% accuracy.
- Ryan: 0 predictions. 0% accuracy.
- Sam: 10 predictions. 80% accuracy.
- Tom: 1 prediction: 0% accuracy
- The model is heavily biased against predicting Ryan (0 predictions) and Tom (1 prediction).

- **Overall Accuracy: 50%**
- **Heavy Predictive Bias**
  - Class imbalance?
- Note: File format issues

Next Steps: Further Optimization

Overview: Next Steps Taken

Step Taken	Reasoning	Selfie Accuracy	Non-Selfie Accuracy
Class Imbalance, Data Augmentation	Large predictive bias on first model	85.4%	40%
Unfreezing 10 Pre-Trained Layers	Possible overfitting to pretrained ImageNet weights	28%	-
Scheduled Learning Rate	Make ImageNet weights more effective	26%	-
Implementing CV2 for Facial Cropping	Data optimization to reduce image background impact on model performance	94%	30%

Deeper Look: Class Imbalance Solution

Step Taken	Selfie Accuracy	Non-Selfie Accuracy
Implementing CV2 for Facial Cropping	94%	30%

Original Class Balance

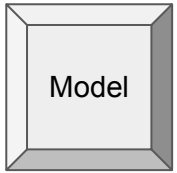
```
Class: Issy: 113 images
Class: Oliver: 141 images
Class: Ryan: 100 images
Class: Sam: 104 images
Class: Tom: 86 images
```

New Class Balance

```
Class: Issy: 100 images
Class: Oliver: 100 images
Class: Ryan: 100 images
Class: Sam: 100 images
Class: Tom: 100 images
```

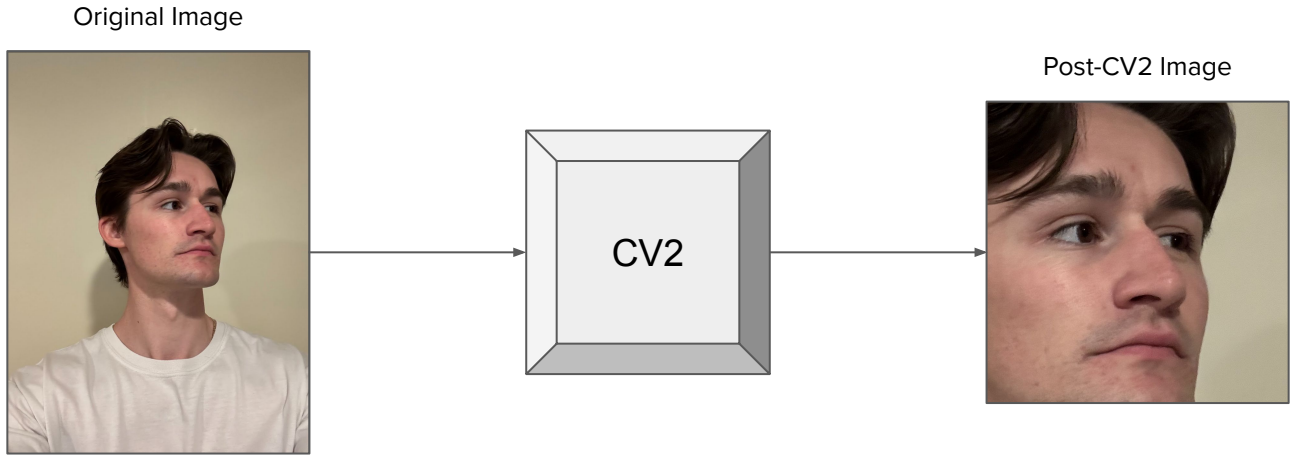
Increased Augmentation + Brightness Scaling

```
# New Train Datagen
train_datagen_2 = ImageDataGenerator(rescale=1./255,
                                     validation_split=0.2,
                                     rotation_range=45,
                                     width_shift_range=0.3,
                                     height_shift_range=0.3,
                                     brightness_range=[0.9, 1.1],
                                     shear_range=0.4,
                                     zoom_range=0.3,
                                     vertical_flip=True,
                                     horizontal_flip=True,
                                     fill_mode="nearest"
                                    )
```



Deeper Look: CV2 Cropping Solution

Step Taken	Selfie Accuracy	Non-Selfie Accuracy
Implementing CV2 for Facial Cropping	94%	30%



Conclusions

Conclusions

- Base VGG16 model with a simple custom classifier saw best performance
  - Achieved 98% accuracy on selfies
  - Achieved 50% accuracy on real-world images
  - No architectural, hyperparameter, or data-tuning changes improved accuracy
- Data cleanup and augmentation were important but could only push accuracy improvement so far with 400 training images.
- CNNs trained on controlled (selfie) datasets can generalize effectively to real world applications feasibly