NEW YORK INSTITUTE OF TECHNOLOGY

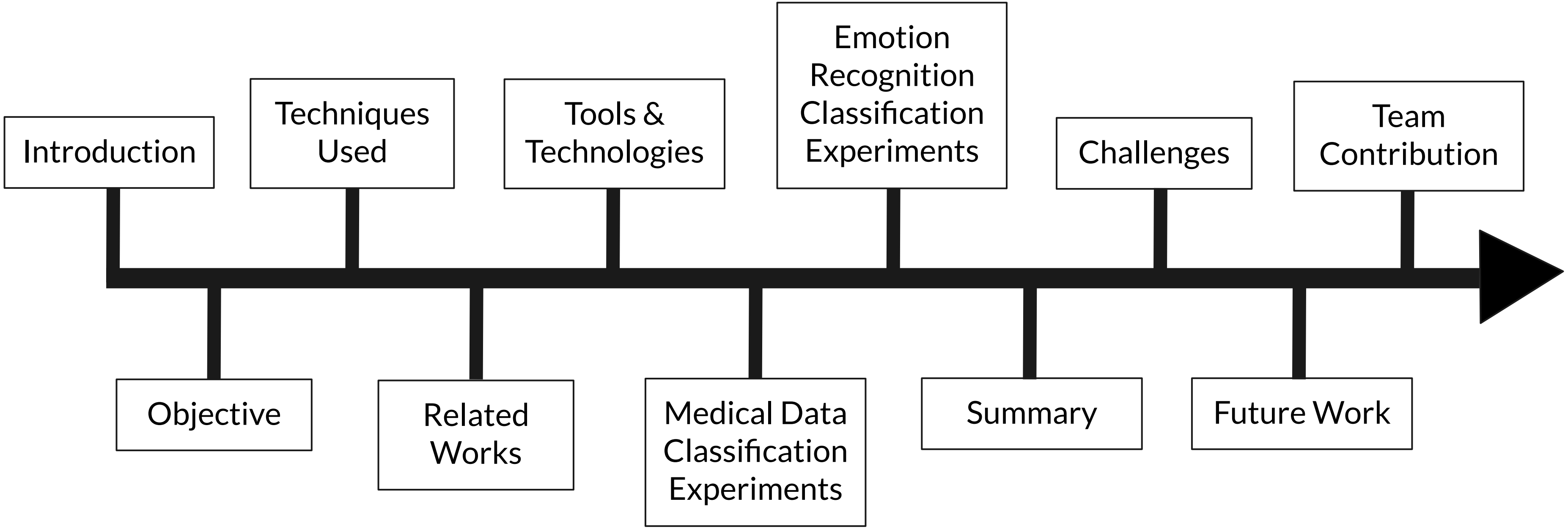
Image Classification: Feature Selection, Data Augmentation, and Transferred Learning

DTSC 870 / Spring 2022

Advisor: Dr. Cao

Team: Michael Trzaskoma, Hui (Henry) Chen

Presentation Overview



Introduction

Computer Vision

- The ability for computers to extract key details and information from visual inputs and has the ability to perform actions based on this information such as classification.

Source: https://towardsdatascience.com/everything-you-ever-wanted-to-know-about-computer-vision-heres-a-look-why-it-s-so-awesome-e8a58dfb641e

Introduction

<u>Image Classification</u>

- What is it?
  - Supervised learning task where computers extract features from images used to distinguish images of different classes
- Importance
  - Computers identifying features not conceived through the human eye to enhance classification accuracy
  - Improved accuracy can be life saving depending on domain
    - Ex: Medical application with disease classification
- Challenges
  - Requires a lot of labeled image data
  - Usable image data is very limited due to:
    - Unlabeled
    - Privacy concerns
    - Industry regulation

Objective

To explore the techniques of feature selection, data augmentation, and transfer learning seeking to improve Machine/Deep Learning model (SVM and CNN) image classification accuracy

Related Works

1. <u>Feature and Feature Selection:</u>
    a. Local Binary Patterns (LBP)
        i. Compare each pixel's neighbors to determine the visual descriptor of the image based on the gray levels co-occurrence matrix and the result as a binary number.

        ii. Paper utilizing LBP features:
            1. Face Expression Recognition using LBP by Ravi et al. [1]
                a. Utilized LBP to extract facial features: 76.23%

    b. Histogram of Oriented Gradients (HoG)
        i. Extract image features by computing an histogram of oriented gradients of an image squared cells.

        ii. Paper utilizing HoG:
            1. Facial Emotion Classification Using Texture Features by Kalsum et al. [2]
                a. Utilized Local Intensity Order Pattern (LIOP) and HoG to extract the facial details: 63% accuracy

    c. Principal Component Analysis (PCA)
        i. Reduce the number of features of a data set while preserving as much information as possible.

        ii. Paper utilizing PCA:
            1. Feature Extraction, Reduction and Classification Analysis by Sachdeva et al. [3]
                a. Utilized PCA in medical data to increased the accuracy from 80.8% to 89%

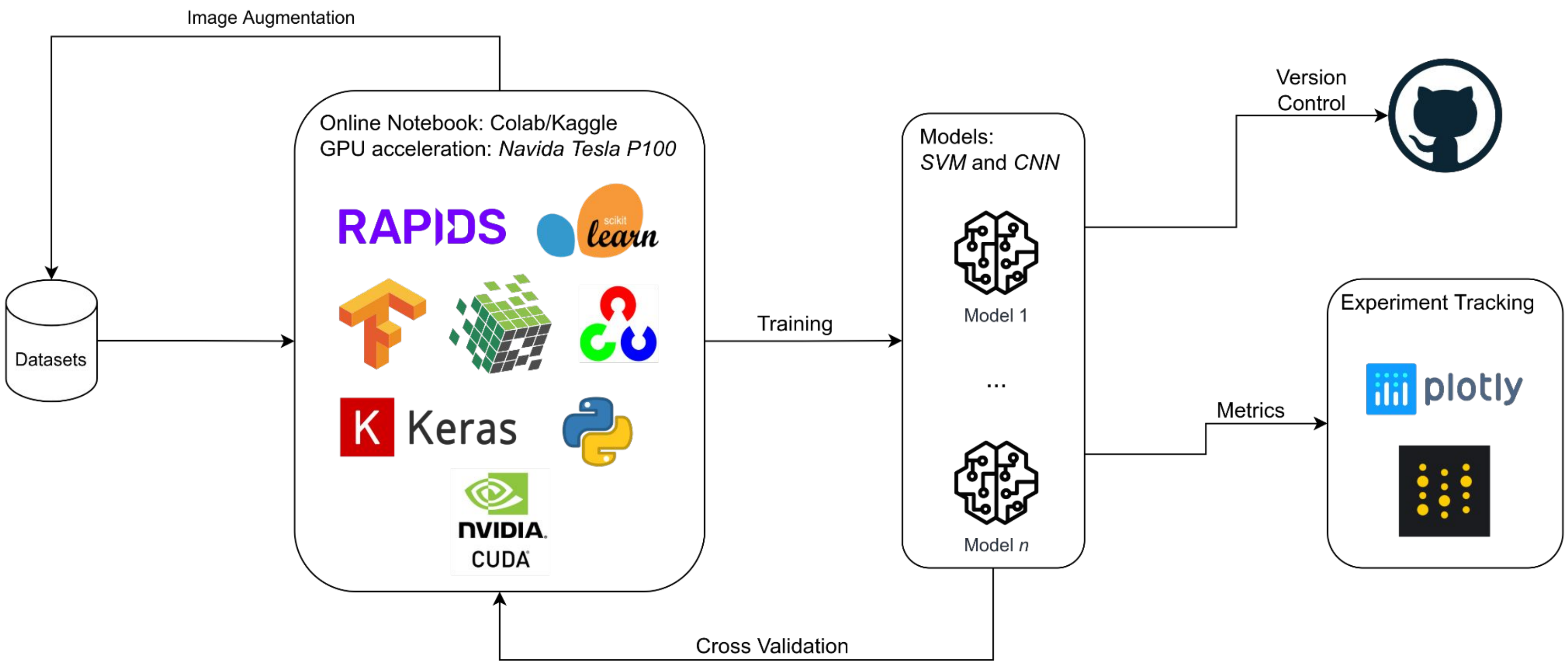
Related Works

2. Data Augmentation - method of performing visual modifications to a preexisting image
    a. Allows for the creation of more image data from an initial dataset
    b. Papers utilizing transfer learning:
        i. Enhancing Medical Image Classification via Augmentation by Tang et al. [4]
            1. Utilized augmentation in medical data with goal to improve image classification accuracy
            2. Accuracy improvement from 87.51% to 91.82%
            3. Observed that not all types of image of augmentation can be beneficial

Related Works

3. Transfer Learning - method for CNNs to learn from pretrained models who have been trained on millions of images in various domains
    a. Requires less data to be trained
    b. Can benefit classification tasks with small data sizes
    c. Papers utilizing transfer learning:
        i. Diabetic Retinopathy Severity Classification by Thota and Reddy [5]
            1. Search to improve medical data image classification utilizing transfer learning
            2. Improved previous standard classification accuracy from 54.31% to 74% accuracy utilizing transfer learning of VGG-16 model
        ii. Egg Incubation Image Classification by Junaidi et al. [6]
            1. When utilizing a customized CNN: 87%
            2. Utilizing VGG-19 model for transfer learning: 92%

Tools and Technologies



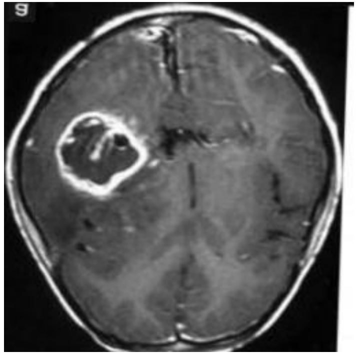
Scope

Classification tasks focused on two applications:

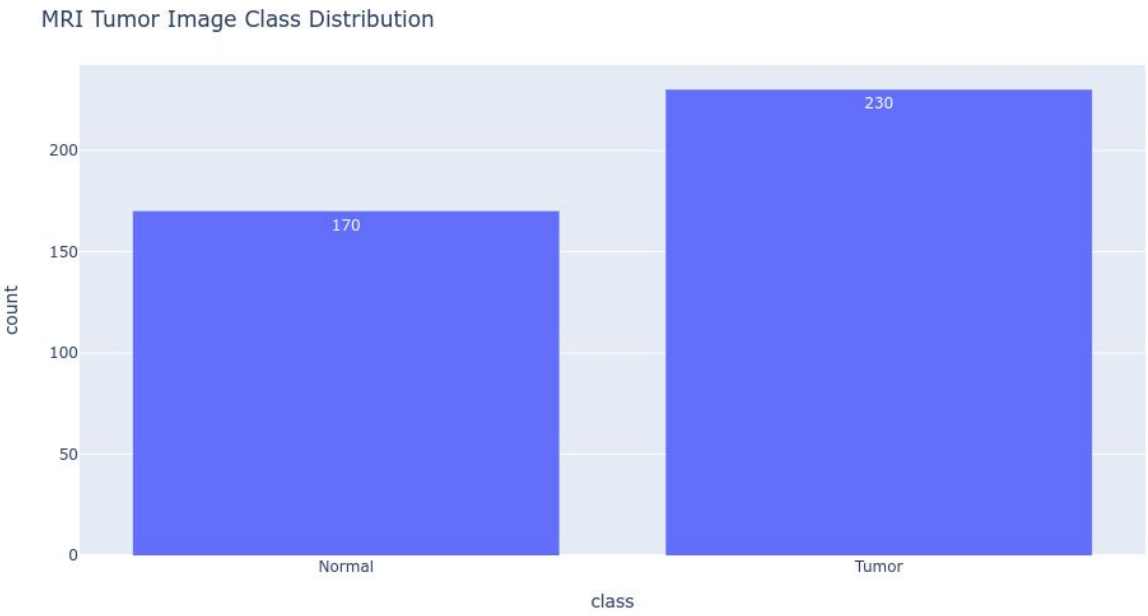
- Medical Data: MRI Brain Tumor Dataset
- Emotion Recognition: Facial Emotion Recognition 2013 Dataset

MRI - Dataset

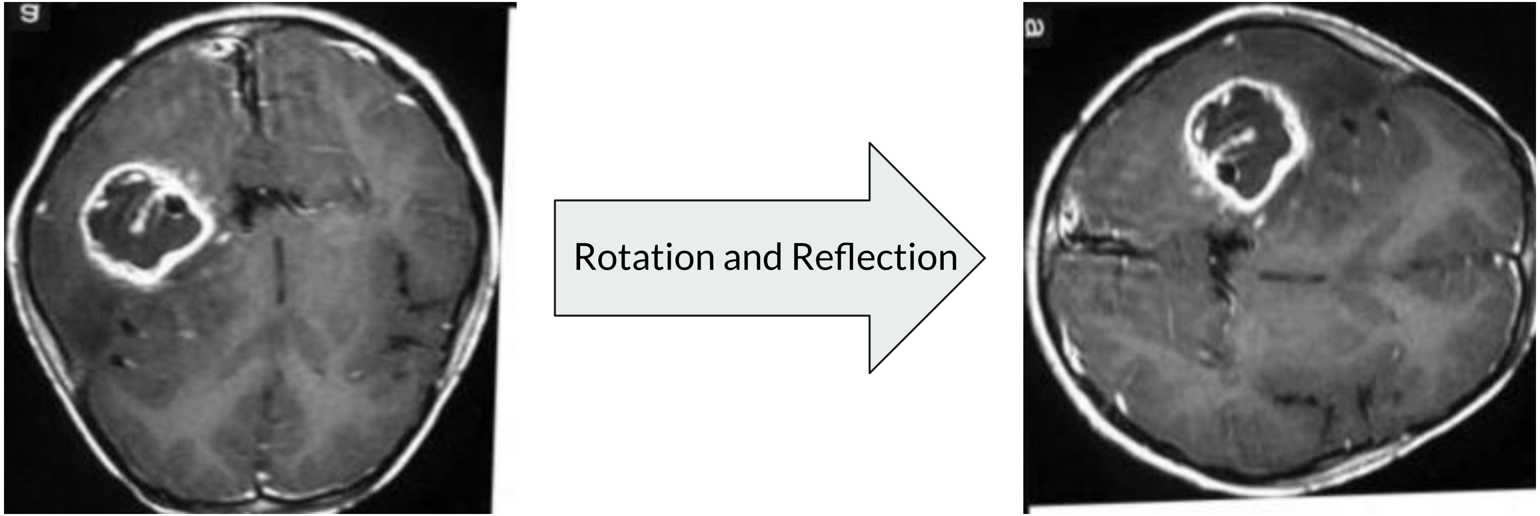
- MRI Brain Tumor Image
  - Available on Kaggle [7]
  - Classes: Normal and Tumor
  - Total size: 400 images
  - Image size: 256 * 256 * 3 (pixels)
  - Training set: 70% (280 images)
  - Testing set: 30% (120 images)



(Sample Image)



MRI: Image Augmentation

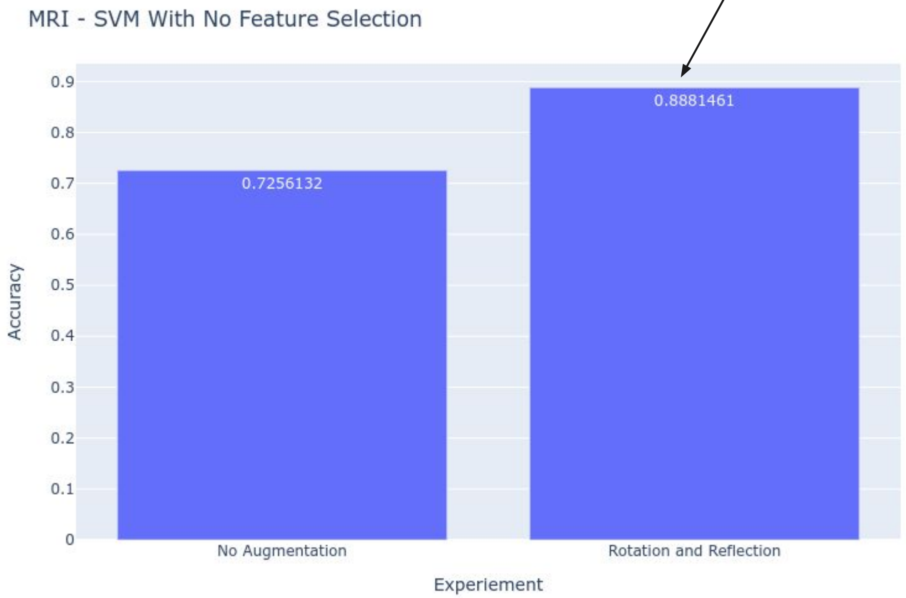


MRI SVM Experiments

MRI - SVM - No Feature Selection

Best Performance: **88.81% Accuracy**

1. No feature selection
   a. Image raw pixel data
      i. 256 * 256 * 3 = 196608 features

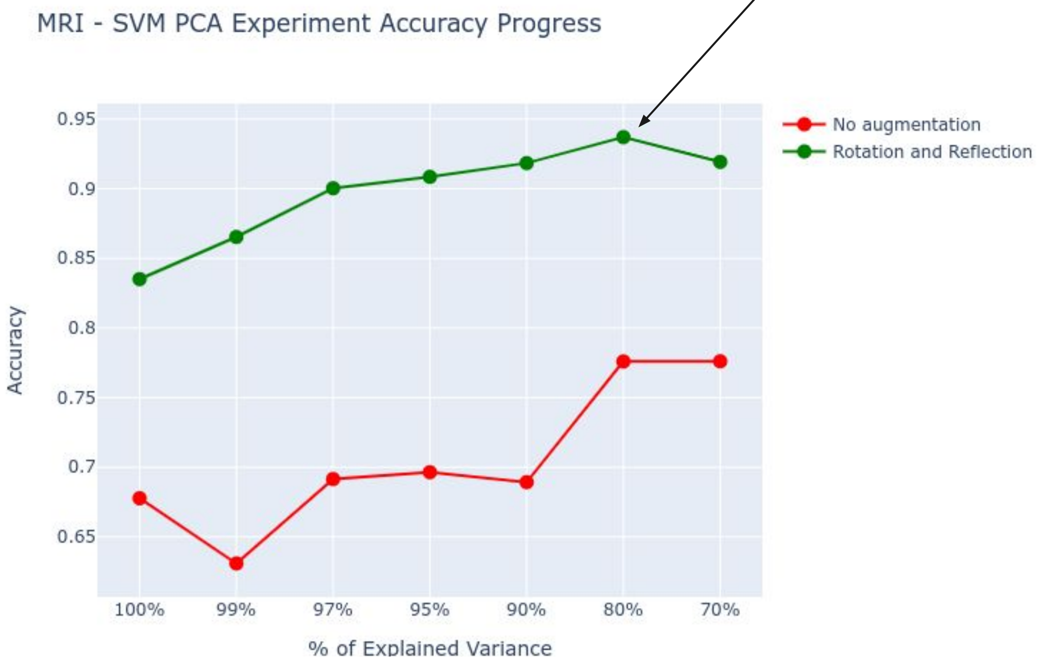


MRI - SVM - PCA

2. Principal Component Analysis (PCA)

Raw Pixel Features (Without PCA)	PCA Variance	# of Features
196608	100%	279
	99%	245
	97%	206
	95%	179
	90%	131
	80%	77
	70%	46

Best Performance: **93.70% Accuracy**



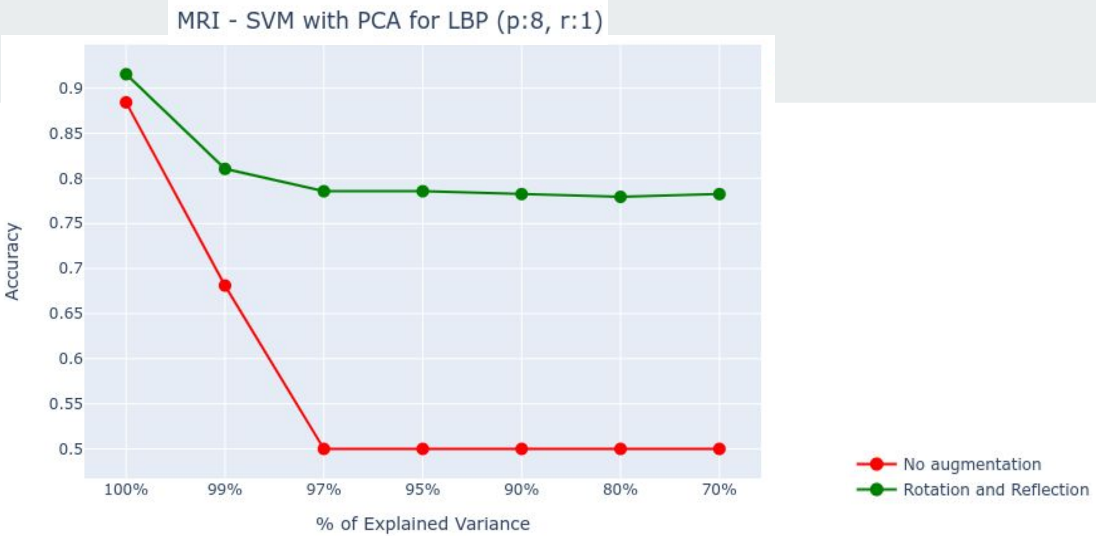
MRI - SVM - LBP + PCA

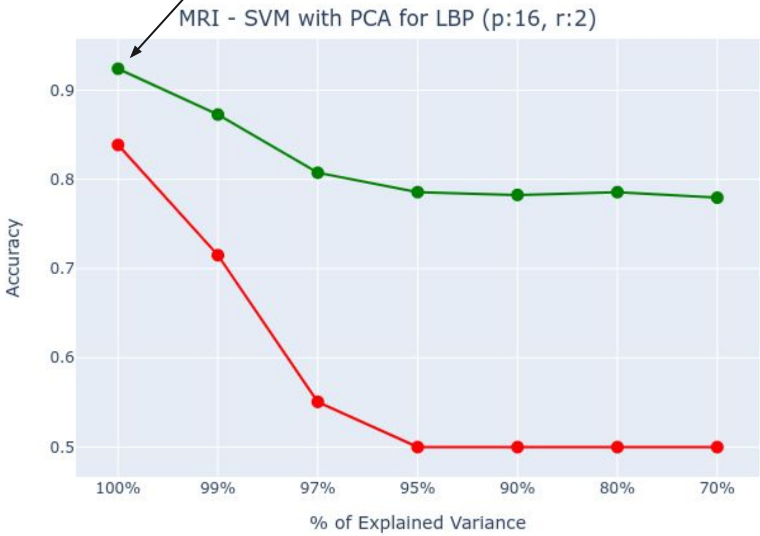
3. Local Binary Patterns (LBP) + PCA
    a. Tune the number of neighbors *P* within a radius of *R*.

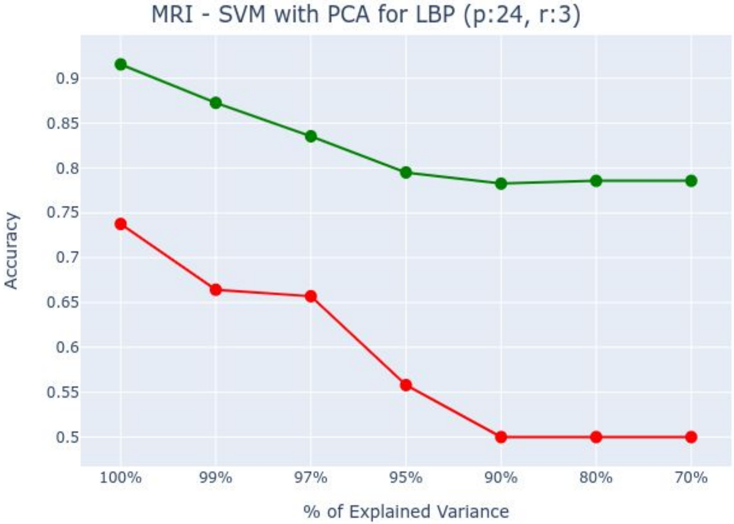
Raw Pixel Features (without PCA)	PCA Variance	Number of Features Without LBP	LBP (*P*: 8, *R*: 1)	LBP (*P*: 16, *R*: 2)	LBP (*P*: 24, *R*: 3)
196608	100%	279	279	279	279
	99%	245	272	273	273
	97%	206	262	264	265
	95%	179	253	255	257
	90%	131	232	235	239
	80%	77	193	200	205
	70%	46	158	166	173

MRI - SVM - LBP + PCA

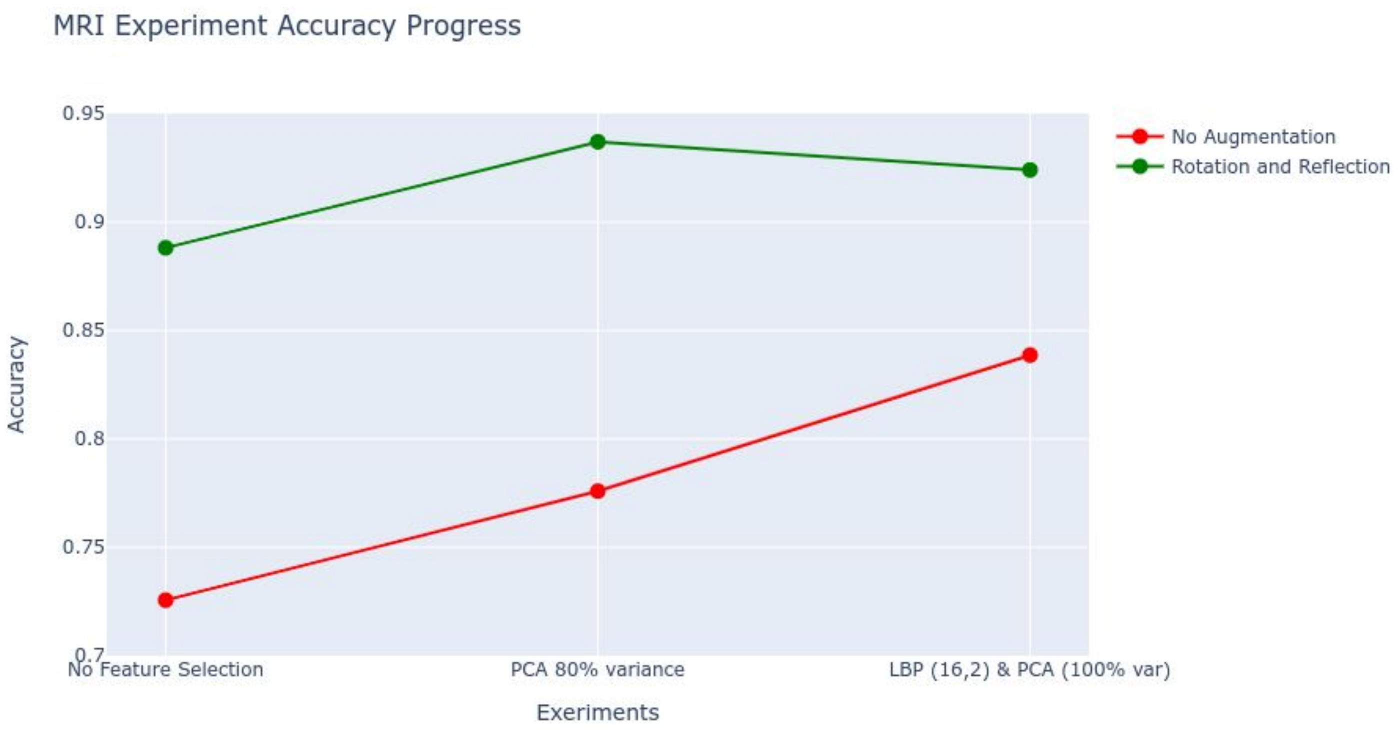
Best Performance: **92.40% Accuracy**







MRI - SVM - Results



MRI - SVM - Best Performance

PCA:

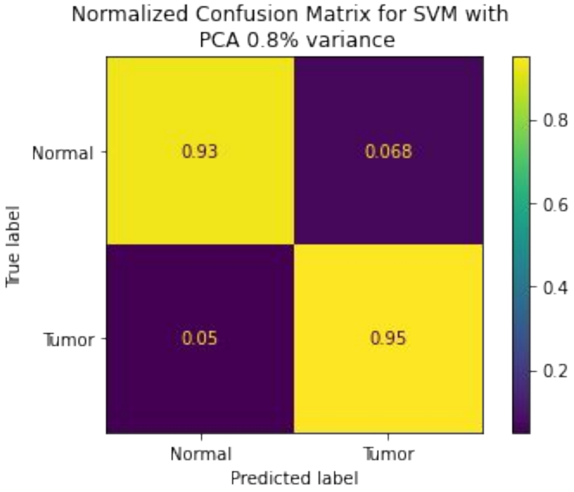
- 80% variance: 77 features

Accuracy: 93.70%

Parameters:

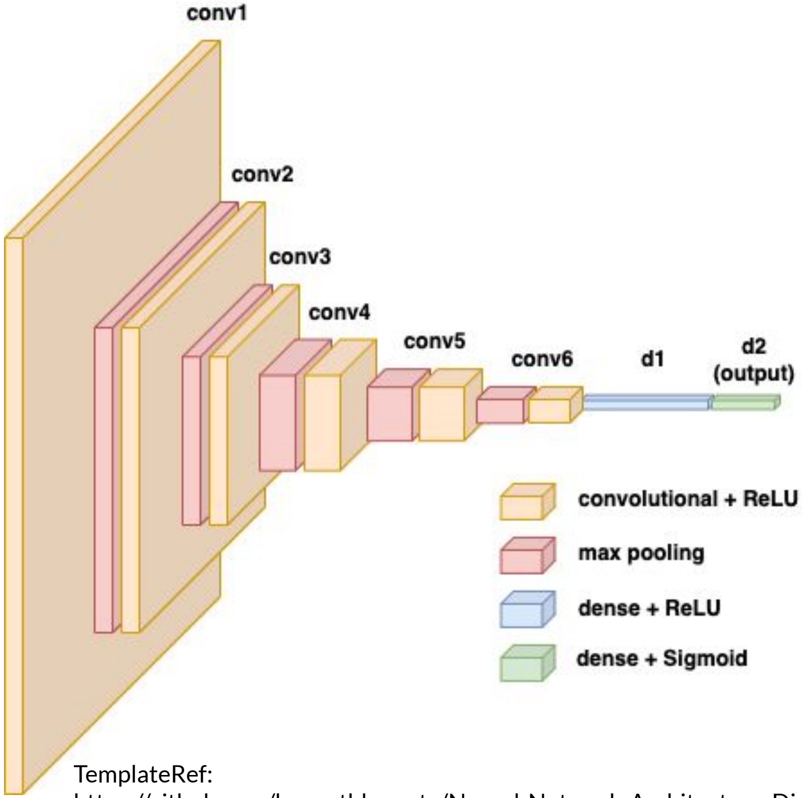
- C: 3.0
- Gamma: 0.003636364
- Kernel: RBF

Post Augment

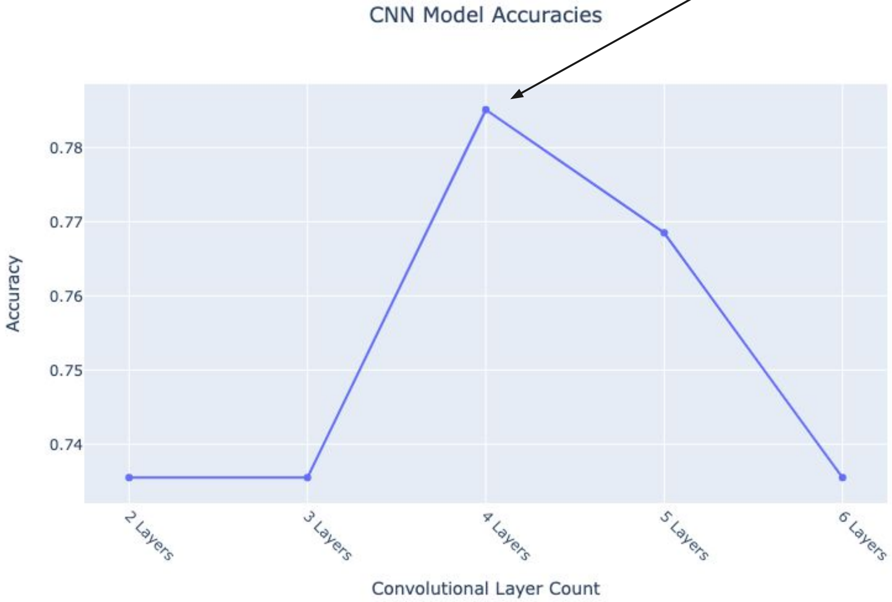


MRI CNN Experiments

MRI - CNN - Initial Structure



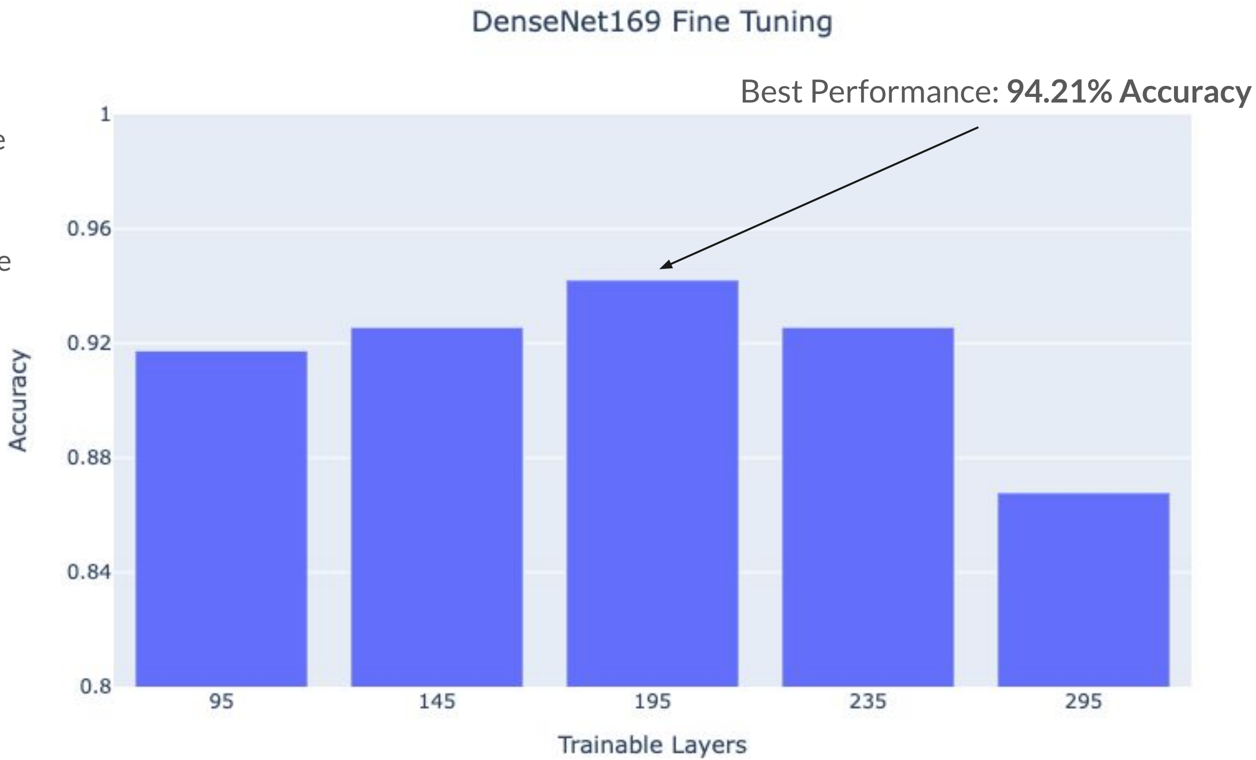
Best Performance: **78.51% Accuracy**



TemplateRef:
https://github.com/kennethleungty/Neural-Network-Architecture-Diagrams

MRI - CNN - Transfer Learning

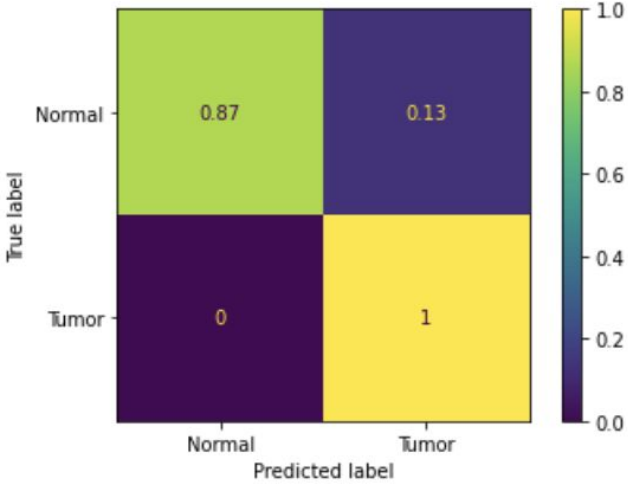
2. Transfer Learning:
   a. DenseNet169
      i. 695 possible trainable layers
   b. Freeze top layers and set x amount of bottom layers to be trained



MRI - CNN - Augmentation

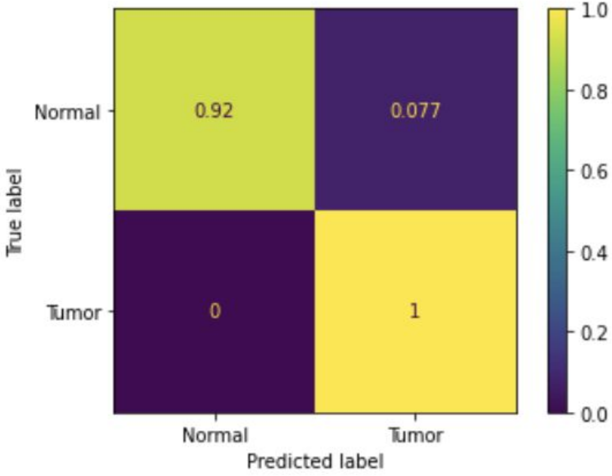
3. Data Augmentation:
    a. Utilizing DenseNet169
        i. Set 195 trainable layers

No Augment



Accuracy: 94.21%

With Rotation and Reflection Augment



Accuracy: 96.69%

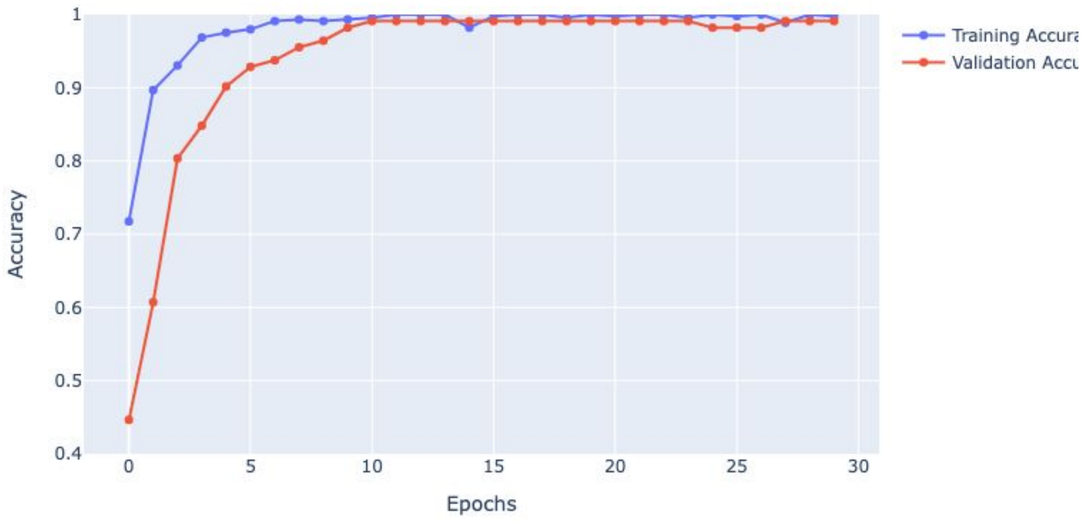
MRI - CNN - Results

Overview:

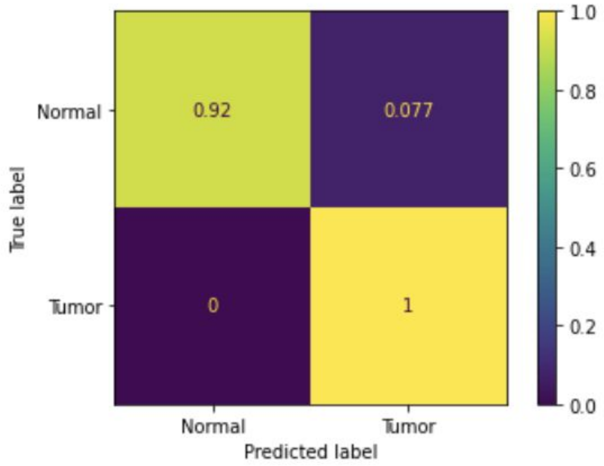
MRI - CNN - Best Performing Model

- Best Performing:
  - DenseNet169 transfer learning
  - 195 trainable layers
  - Rotation and Reflection Data Augmentation

Training Plot



Accuracy: 96.69%



MRI - Discussion

1. SVM utilizing no feature selection outperformed CNN utilizing no techniques
   a. SVM performance: 88.81% accuracy
   b. CNN performance: 78.51% accuracy

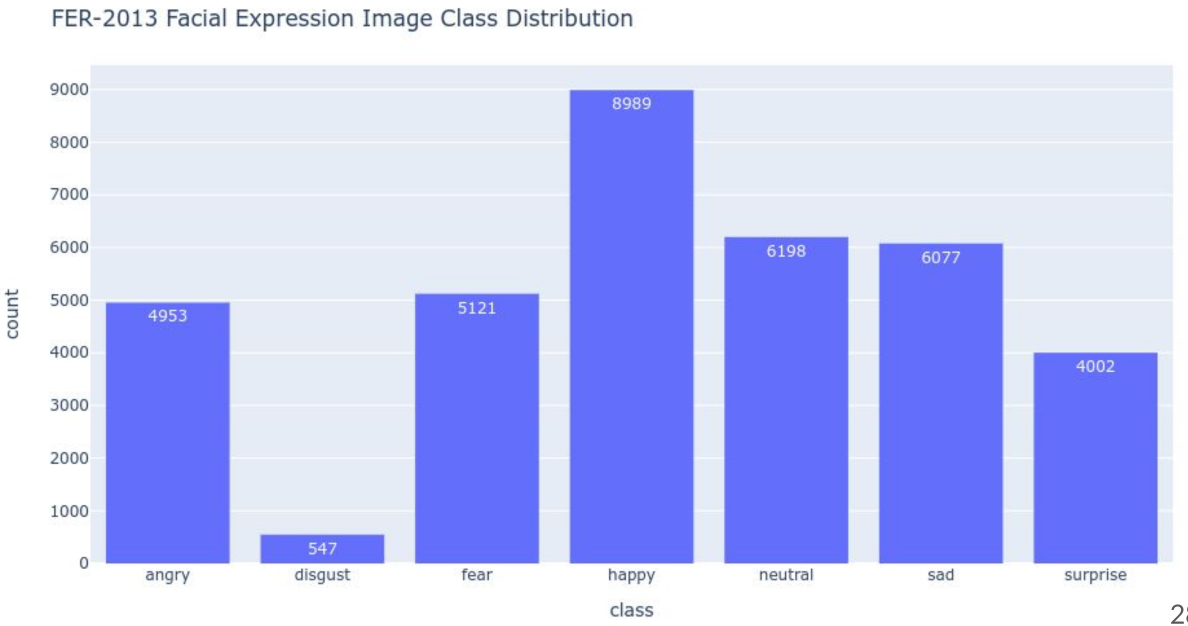
2. Once transfer learning and image augmentation was applied to the CNN, the CNN's performance surpassed the SVM model utilizing PCA selection and augmentation
   a. CNN with DenseNet169 transfer learning: 96.69% accuracy
   b. SVM with PCA on 80% variance: 93.70% accuracy

FER SVM Experiments

FER - Dataset

- FER-2013 Facial Expression Image
  - Available on Kaggle [8]
  - Classes: angry, disgust, fear, happy, neutral, sad, surprise
  - Total size: 32298 images
  - Image size: 48 * 48 (pixels)
  - Training set: 70% (22609 images)
  - Testing set: 30% (9689 images)

(Sample Image)



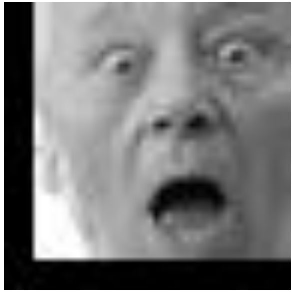
FER: Image Augmentations

Method 1

Rotation and
Reflection

Method 2

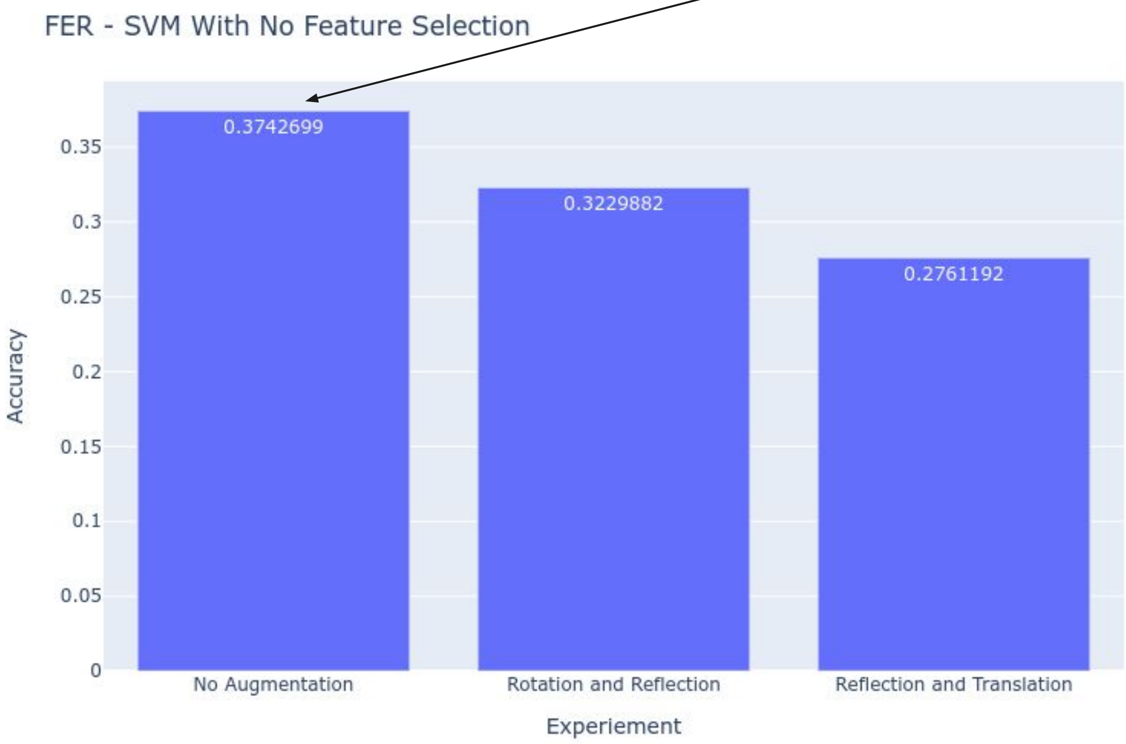
Reflection and
Translation



FER - SVM - No Feature Selection

Best Performance: **37.43% Accuracy**

1. No feature selection
   a. Image raw pixel data
      i. 48 * 48 = 2304 features

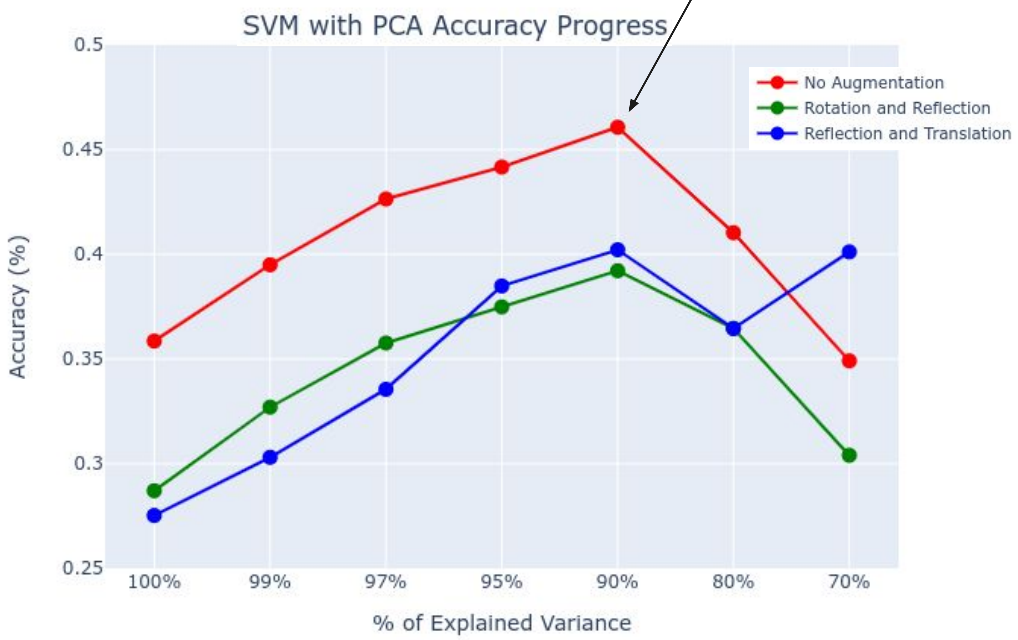


FER - SVM - PCA

2. Principal Component Analysis (PCA)

Raw Pixel Features (without PCA)	PCA Variance	# of Features
2304	100%	2304
	99%	904
	97%	425
	95%	256
	90%	104
	80%	32
	70%	13

Best Performance: **46.07% Accuracy**

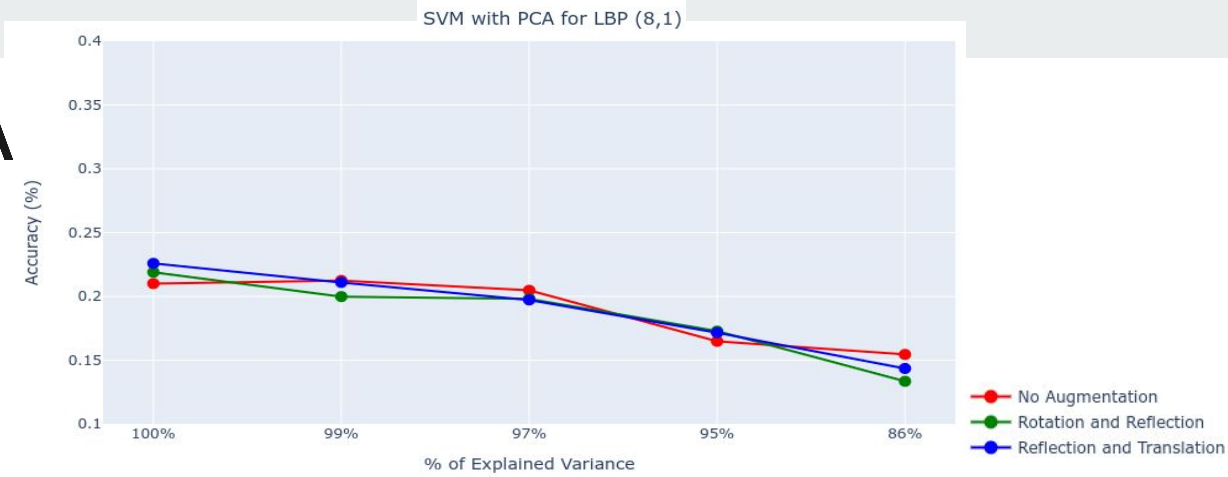


FER - SVM - LBP + PCA

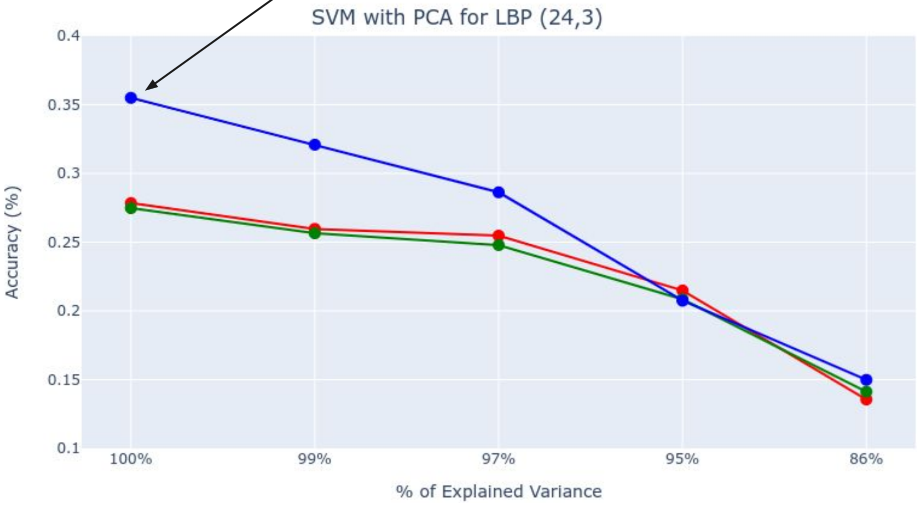
3. Local Binary Patterns (LBP) + PCA
   a. Tune the number of neighbors *P* within a radius of *R*.

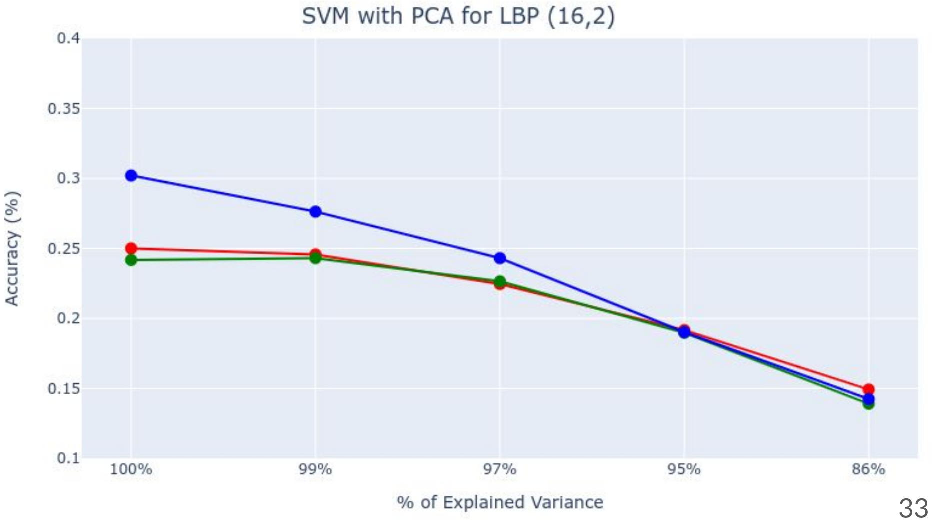
Raw Pixel Features (without PCA)	PCA Variance	LBP (*P*: 8, *R*: 1)	LBP (*P*: 16, *R*: 2)	LBP (*P*: 24, *R*: 3)
2304	100%	10	18	26
	99%	8	13	18
	97%	7	10	13
	95%	4	6	7
	86%	2	2	2

FER - SVM - LBP + PCA



Best Performance: **35.50% Accuracy**

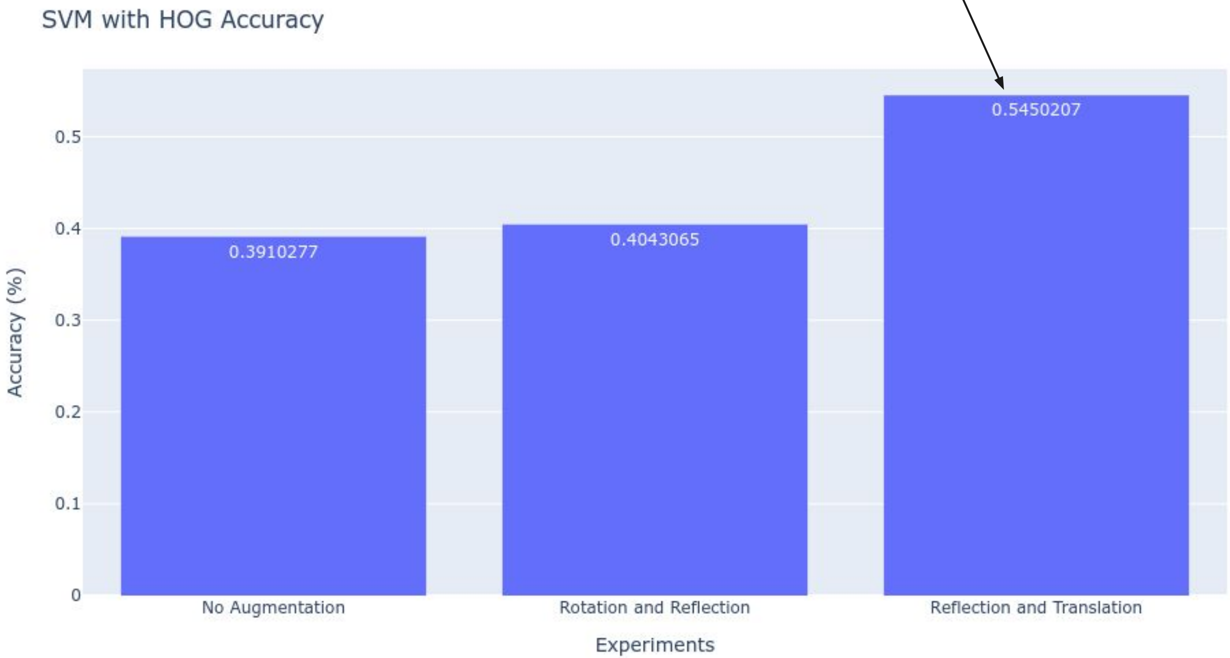




FER - SVM - HOG

Best Performance: **54.50% Accuracy**

4. Histogram of Oriented Gradients (HOG)
    a. Pixels per cell: 16 * 16
    b. Orientation: 9



FER - SVM - Results

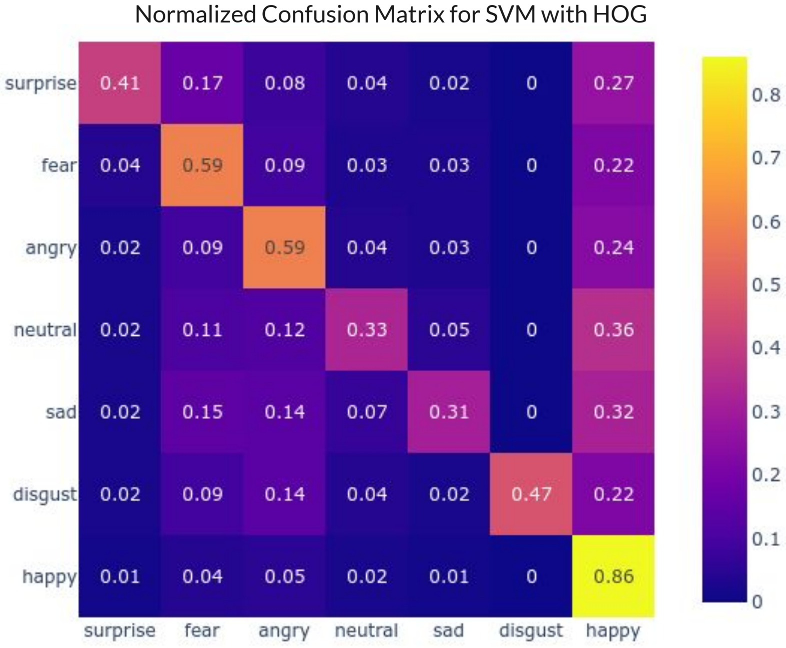
FER - SVM - Best Performance

HOG with Reflection and Translation

Accuracies: 54.50%

Parameters:

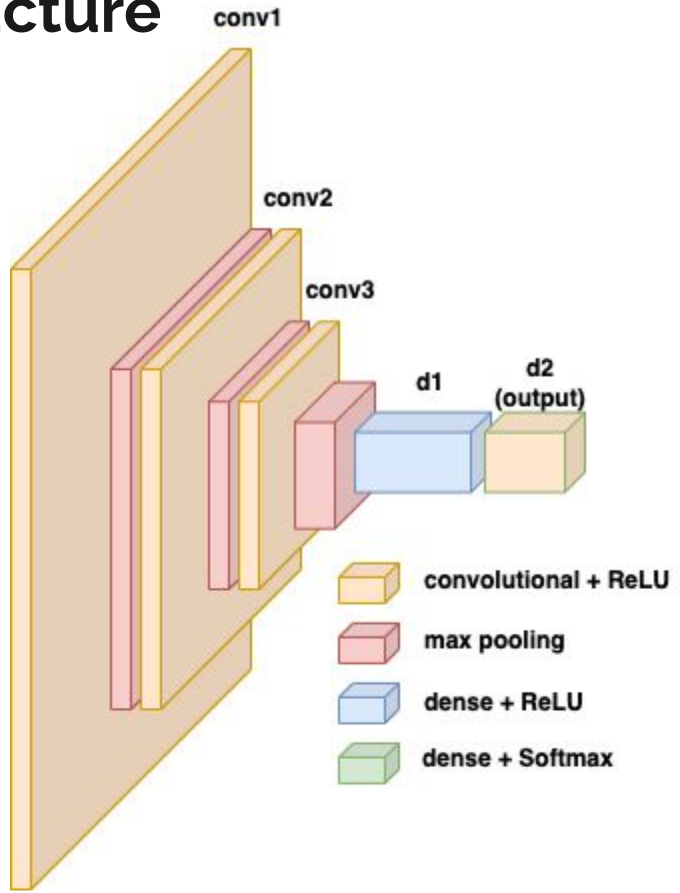
- C: 1.0
- Kernel: Polynomial
- Degree: 5
- Decision Function: one-vs-rest
- Applied class weights



FER CNN Experiments

FER - CNN - Initial Structure

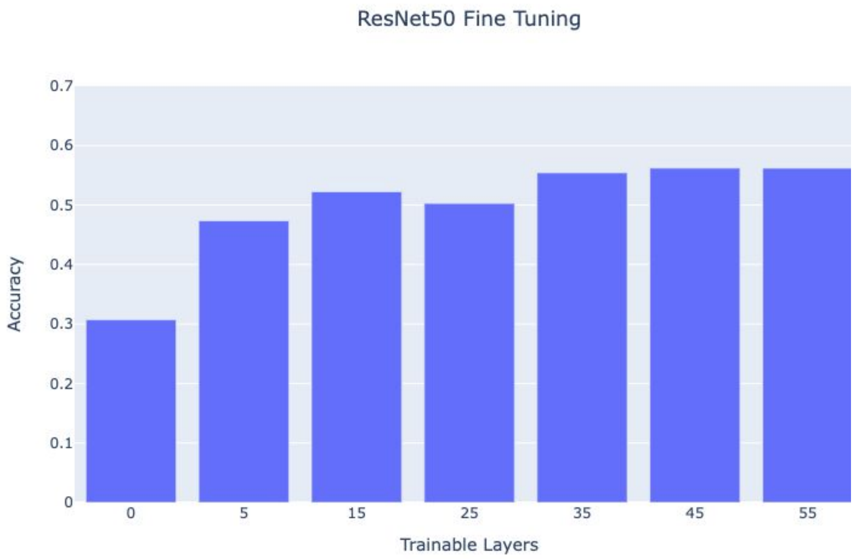
1. Initial CNN network:
    a. Performance:
        i. **Accuracy: 54.35%**

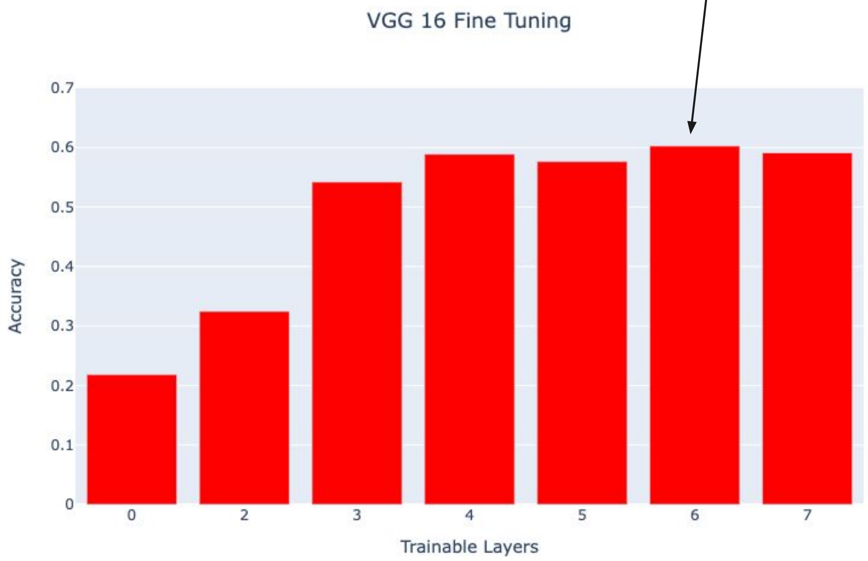


FER - CNN - Transfer Learning

2. Transfer Learning
   a. ResNet50 - 175 trainable layers total
   b. VGG16 - 19 trainable layers total

Best Performance: **60.28% Accuracy**

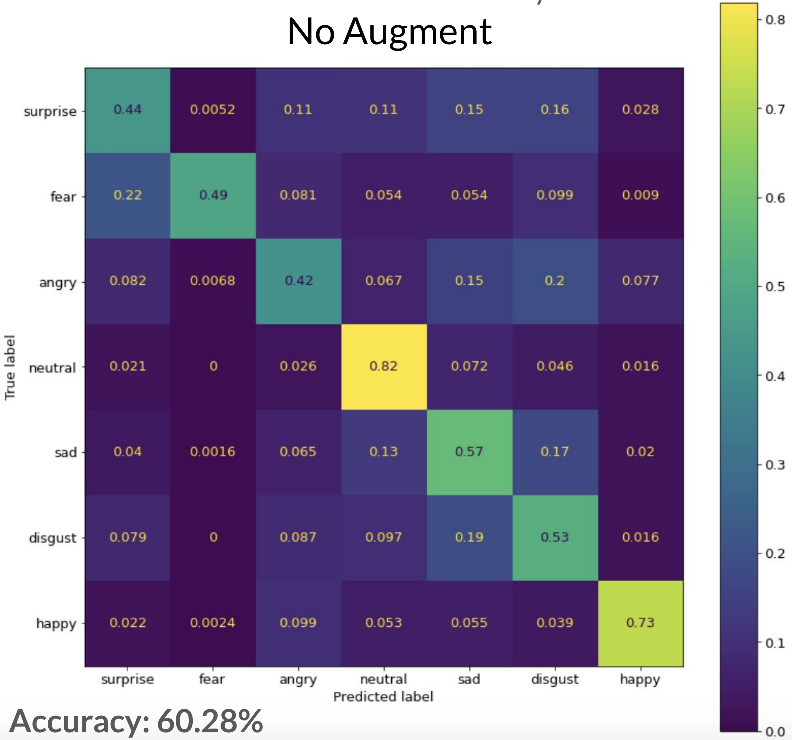




FER - CNN - Data Augmentation

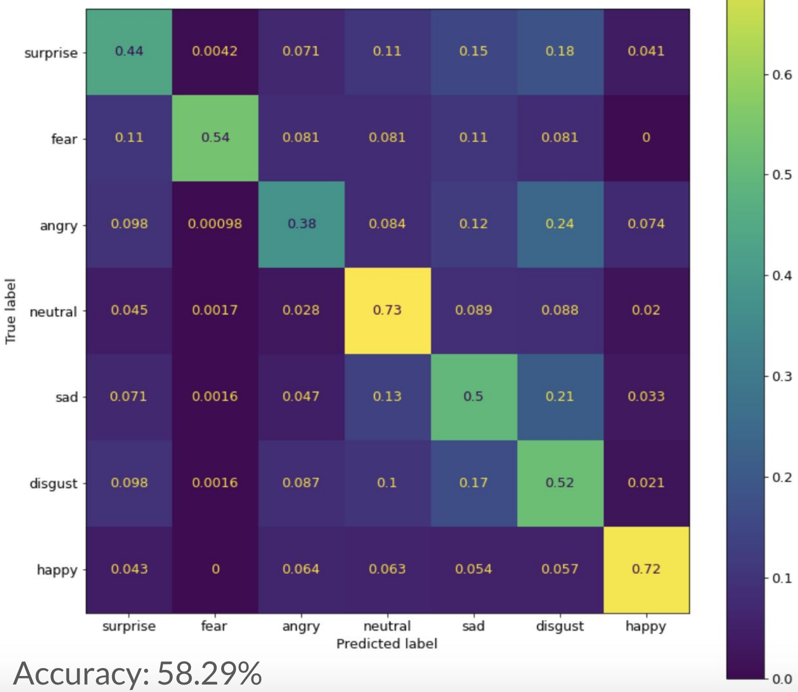
3. Utilization of Data Augmentation
    a. VGG 16 - 6 Trainable Layers

No Augment



Accuracy: 60.28%

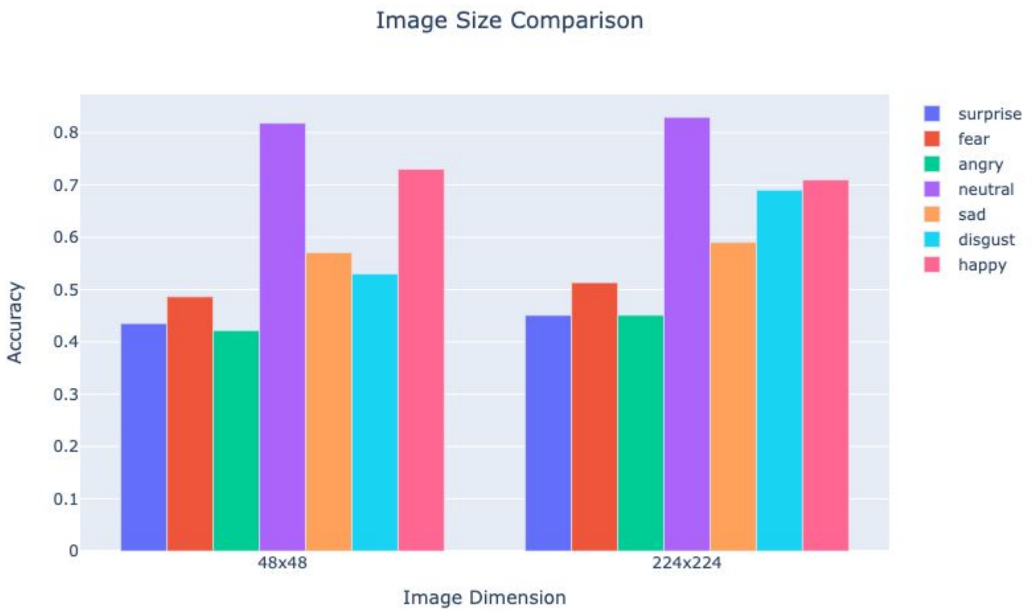
With Rotation and Reflection Augment



Accuracy: 58.29%

FER - CNN - Image Size Adjustment

4. Resizing Images
   a. Initial Image size could be too small for CNN to identify features
   b. VGG16 worked with larger image sizes (224x224)
   c. Accuracy Results:
      i. 48x48 Accuracy: 60.28%
      ii. 2**24x224 Accuracy: 64.11%**



FER - CNN - New Data Augmentation

5. Second attempt at Augmentation
    a. No Augmentation
        i. **Accuracy: 64.11%**
    b. Initial Augmentation - Reflection and Rotation
        i. **Accuracy: 62.52%**
    c. New Augmentation - Reflection and Translation
        i. **Accuracy: 64.43%**

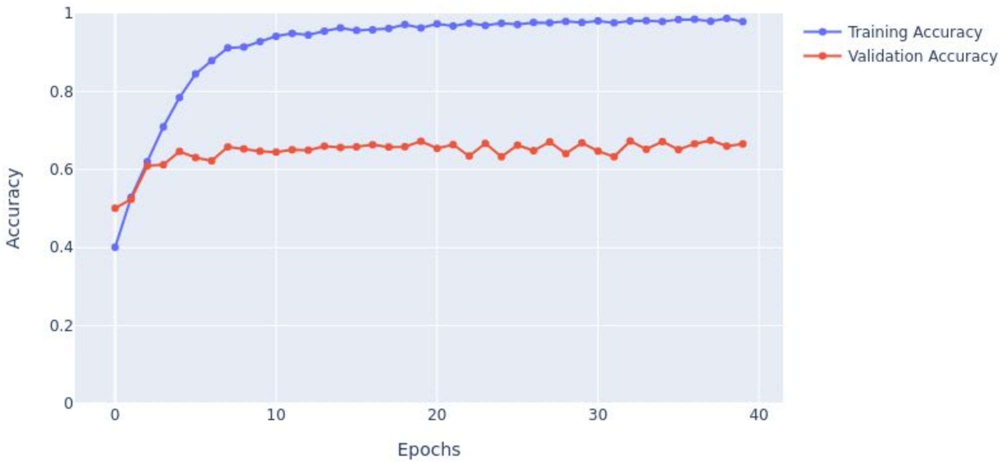
FER - CNN - Results

Overview:

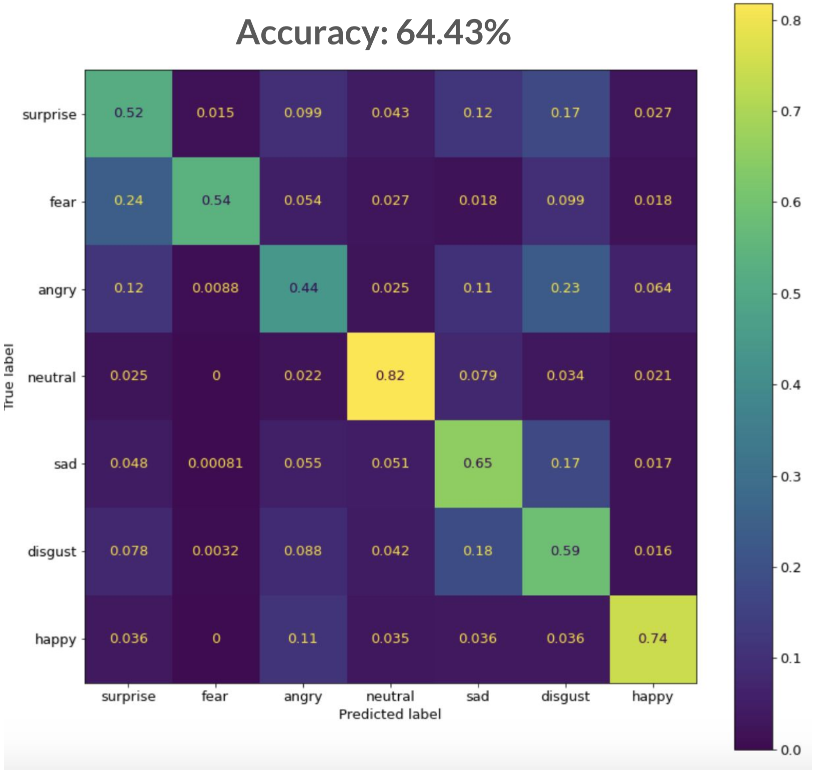
FER - CNN - Best Performing Model

- Best Performing:
    - VGG 16 transfer learning
    - 6 trainable layers
    - 224x224 Resized Images
    - Reflection and Translation Data Augmentation

Training Plot



Accuracy: 64.43%



FER - Discussion

1. CNN utilizing no techniques outperforms SVM without any feature selection
   a. CNN Performance: 54.35% accuracy
   b. SVM Performance: 37.43% accuracy

2. Once CNN applies transfer learning from the VGG network, it outperforms the SVM model utilizing HOG feature selection
   a. CNN with VGG 16 transfer learning: 64.43% accuracy
   b. SVM with HOG: 54.50% accuracy

Summary

1. CNN greatly benefits from transfer learning, being able to outperform SVM utilizing feature selections and data augmentation.
2. PCA works well on raw pixel data, but is less effective on handcrafted features such as LBP, and this is true for both medical and emotional data.
3. Applying image augmentation is a way to enrich the training data and boost the model performance.
4. A single methodology is not always applicable to every domain the same way.
    a. Image augmentation with FER 2013
        i. Augmentation method applied on MRI dataset had no effect on FER 2013
        ii. Required different augmentation methodology to be impactful

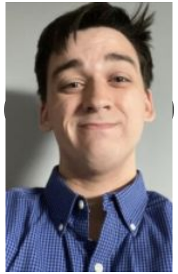
Challenges

1. Computation Resources on FER-2013 dataset
    a. Heavy computation resource is need.
    b. Kaggle notebook and Nvidia Rapids
        i. Reduced the amount of time of PCA to process on average from 8 mins → 2 mins 12 sec.
        ii. Reduced time per epoch when training CNNs on average from 5 mins → 40 sec.
2. Nvidia Rapids environment setup
    a. Credit to Ashwin Srinath, Software Engineer at Nvidia Rapids Team, for fixing a bug that we reported.
3. Large quantity of factors to account for when optimizing classifier accuracy
    a. For CNN: Trainable parameters, optimizers, learning rate scheduler, layer structure, regularization
    b. For SVM: Trainable parameters, regularization

Future Works

1. Look into the Oversampling, Downsampling, SMOTE, Generative Adversarial Networks (GAN) methods to balance the datasets.
2. Analysis and draw a comparison between how the model performs before and after balancing the datasets.
3. Explore classification with video data which incorporates visual and sound information.

Team Contribution



Michael Trzaskoma

Data Augmentation

CNN Experimentation

Hui (Henry) Chen

Data Augmentation

SVM Experimentation

References

[1] R. Ravi, S. V. Yadhukrishna and R. prithviraj, "A Face Expression Recognition Using CNN & LBP," 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), 2020, pp. 684-689, doi: 10.1109/ICCMC48092.2020.ICCMC-000127.

[2] Kalsum, T., Mehmood, Z., Kulsoom, F., Chaudhry, H., Khan, A., Rashid, M. and Saba, T., 2021. Localization and classification of human facial emotions using local intensity order pattern and shape-based texture features. Journal of Intelligent & Fuzzy Systems, 40(5), pp.9311-9331.

[3] Sachdeva, J., Kumar, V., Gupta, I., Khandelwal, N. and Ahuja, C., 2016. A package-SFERCB-"Segmentation, feature extraction, reduction and classification analysis by both SVM and ANN for brain tumors". Applied Soft Computing, 47, pp.151-167.

[4] X. Tang, C. Zhou, L. Chen and Y. Wen, "Enhancing Medical Image Classification via Augmentation-based Pre-training," 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2021, pp. 1538-1541, doi: 10.1109/BIBM52615.2021.9669817.

[5] N. B. Thota and D. Umma Reddy, "Improving the Accuracy of Diabetic Retinopathy Severity Classification with Transfer Learning," 2020 IEEE 63rd International Midwest Symposium on Circuits and Systems (MWSCAS), 2020, pp. 1003-1006, doi: 10.1109/MWSCAS48704.2020.9184473.

[6] A. Junaidi, J. Lasama, F. D. Adhinata and A. R. Iskandar, "Image Classification for Egg Incubator using Transfer Learning of VGG16 and VGG19," 2021 IEEE International Conference on Communication, Networks and Satellite (COMNETSAT), 2021, pp. 324-328, doi: 10.1109/COMNETSAT53002.2021.9530826.

[7] A. M. Hashan, "MRI based brain tumor images," Kaggle, 04-Apr-2021. Available: https://www.kaggle.com/mhantor/mri-based-brain-tumor-images.

[8] M. Sambare, "Fer-2013," Kaggle, 19-Jul-2020. Available: https://www.kaggle.com/msambare/fer2013.