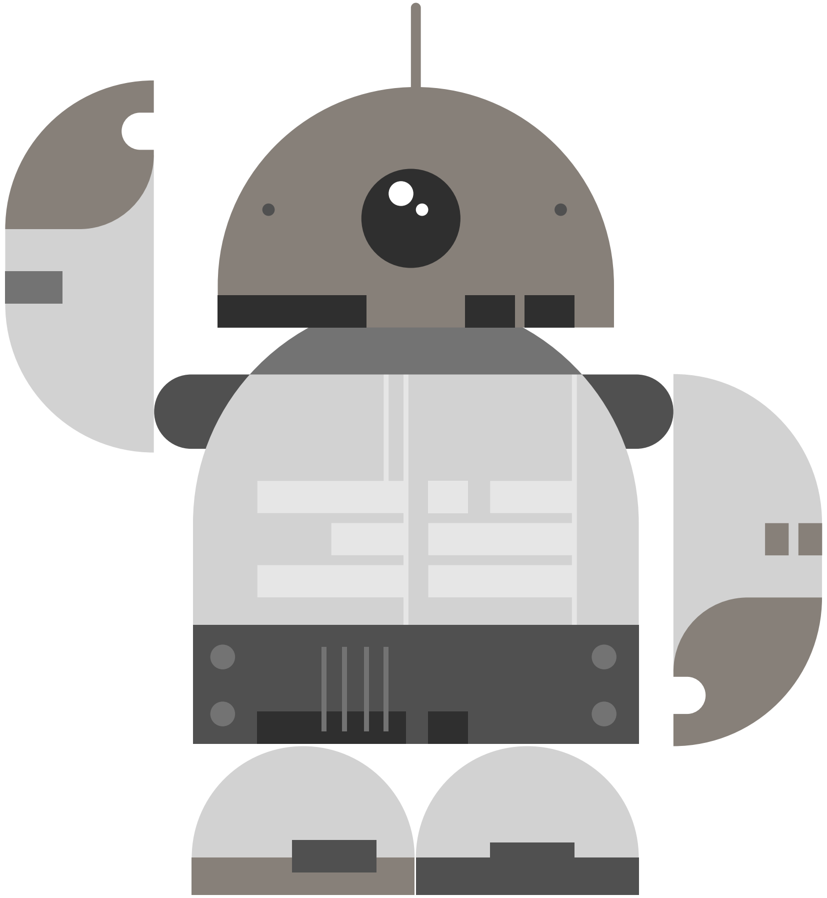
FACIAL EMOTION DETECTION

Ian Edmundson



PROBLEM DEFINITION

Facial expressions make up as much as 55% of human communication, conveying emotions and intent. Understanding the relationships between facial expressions and emotions can deepen our understanding of human behavior, leading to advancements in multiple fields.

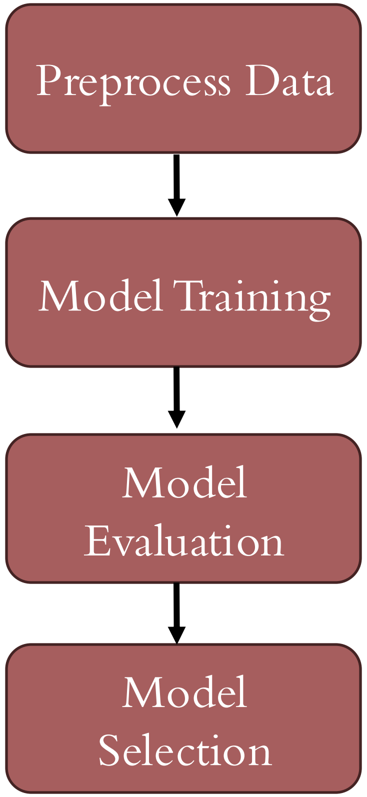
Why is this important?

By utilizing the power of machine learning, accurate detection of emotions can achieve impacting effects on fields such as:

- Psychology
- Healthcare
- Robotics and AI
- Interpersonal communications

APPROACH AND SOLUTION DESIGN

- Objective: Identify best performing model for emotion classification base on provided images.
- Approach:
  1. Model Design
     - Custom Convolutional models
       - Color and Grayscale images
     - Pre trained models
       - VGG16, ResNetV2, and EfficientNetV2
  2. Training and Evaluation
     - Train all models using the same training set to ensure consistency
     - Evaluate performance on same testing data to ensure fair comparison
  3. Performance comparison
     - Compare models based on evaluation metrics
     - Select best performing model



KEY FINDINGS AND INSIGHTS

- To ensure a clean and reliable dataset for model training, testing, and validation, thorough data preprocessing was performed.
  - Utilized the MTCNN model, trained to detect human faces, to help identify problematic images.
  - Removed images from training, testing, and validation datasets
- Example of problematic images are shown below:

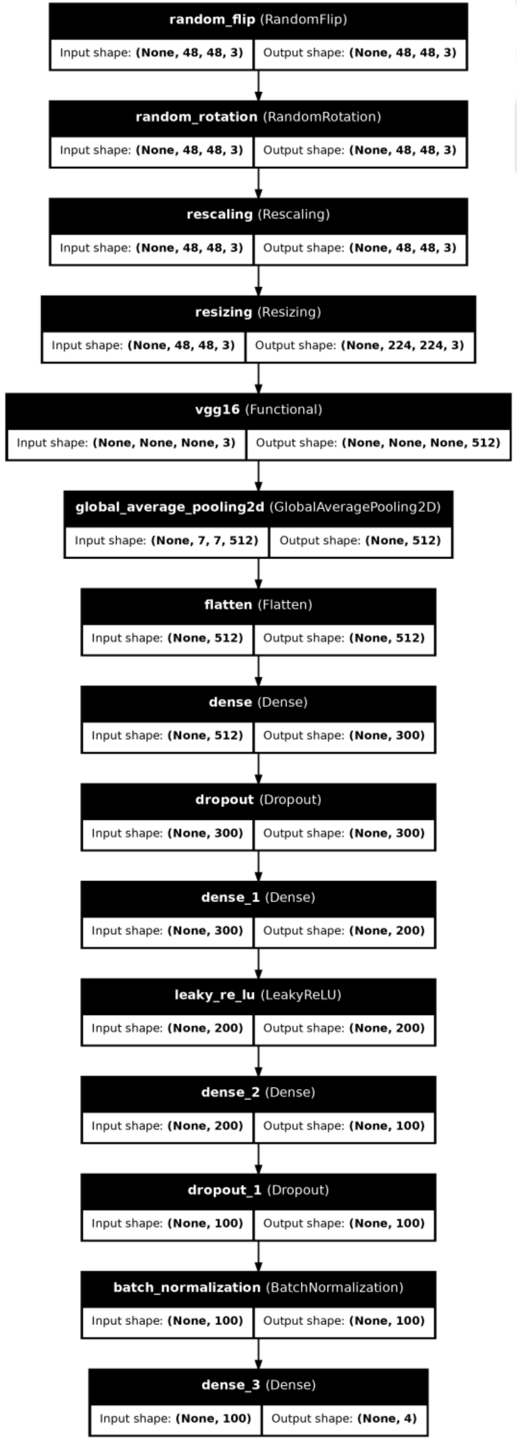



KEY FINDINGS AND INSIGHTS

Model	Training Accuracy	Validation Accuracy	Test Accuracy
VGG16 (4 Dense layers, Color)	80.05%	75.27%	79.69%
Custom 3 Convolutional layers 2 Dense layers (Color)	69.99%	71.59%	75.78%
ResNet V2 (4 Dense layers, Color)	76.37%	73.58%	74.22%
Custom 3 Convolutional layers 2 Dense layers (Grayscale)	68.44%	71.71%	73.43%
Custom 5 Convolutional layers 3 Dense layers (Grayscale)	67.86%	72.01%	71.87%
Custom 5 Convolutional layer 3 Dense layers (Color)	66.53%	72.03%	70.31%
EfficientNetV2 (4 Dense layers, Color)	25.03%	36.68%	25.00%

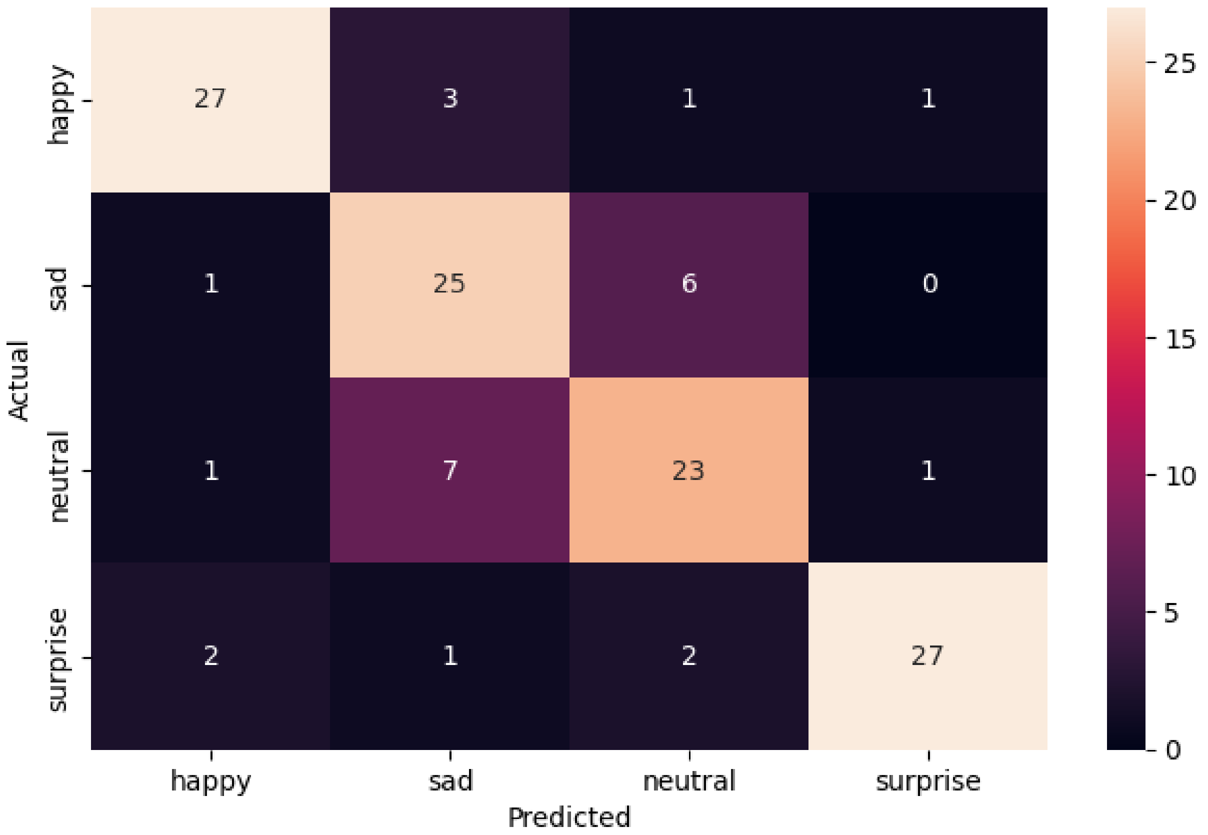
KEY FINDINGS AND INSIGHTS

- VGG16 achieved the highest test accuracy at 79.69%, making it the best performing model.
- The custom model with 3 Dense layers and 5 Convolutional layers showed a good performance but fell short compared to VGG16.
- EfficientNetV2 had the worst performance out of all models.



KEY FINDINGS AND INSIGHTS

- The confusion matrix shows how well the model predicts each emotion category
- This matrix shows that the categories happy and surprise predict with high accuracy:
  - **Happy** was correctly predicted 84.37%
  - **Surprise** was correctly predicted 90%
- The categories sad and neutral had a hard time distinguishing each other:
  - **Sad** misclassified as neutral 6 times but had a 70% success
  - **Neutral** misclassified as sad 7 times but had a 71.85% success



RECOMMENDATIONS AND NEXT STEPS

Recommendations	Explanation
Increase the amount of training data	• As data increases so does the ability for the model to generalize over the space • More data reduces the risk of overfitting • Increased data for sad and neutral could improve the predictions for the model
Extensive data cleaning and preprocessing	• Identify and remove bad images such as blank images or images lacking facial features • Address ambiguous images which caused issues between sad and neutral
Increased computing power	• Allows for more training iterations, finer tuning, and better optimization • Faster compute times
Hyperparameter tuning	• Methods like Hyperband and Bayesian optimization to efficiently explore the search space

CONCLUSION

- VGG16 demonstrated the best overall performance, showing a good balance between the training, validation, and test accuracy
- Despite the VGG16 models performance, the model faced challenges distinguishing between the sad and neutral categories.
    - This could mean that the images have similar facial features or lack of diverse training data