Facial Emotion Detection Deep Learning

Capstone Project

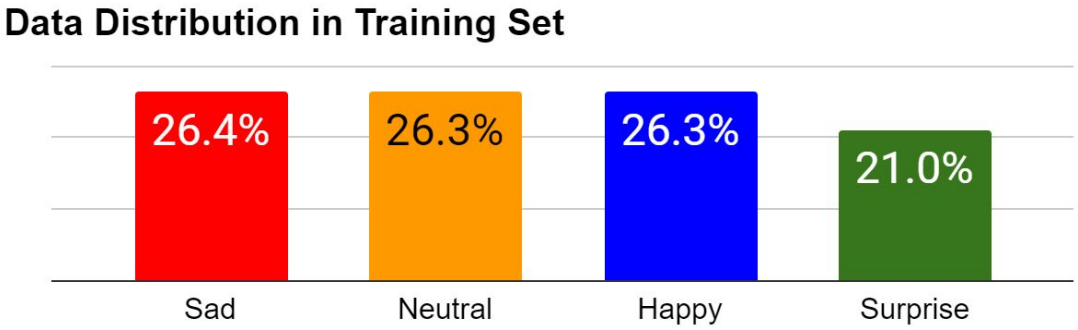
Problem Definition

GOAL

Create a computer vision model that can accurately detect facial emotions. The model should be able to perform multi-class classification on images of facial expressions, to classify the expressions according to the associated emotion.

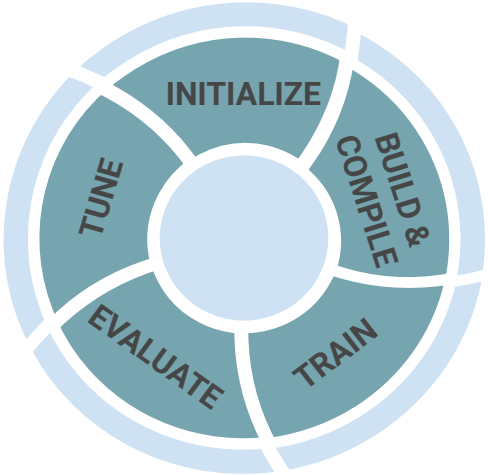
DATASET

Images of size 48X48X3 divided in 3 folders: **Train (15,109 images)**, **Validation (4977 images), and Test (128 images)**.
Each folder has four subfolders: **Happy**, **Sad**, **Surprise**, **Neutral**



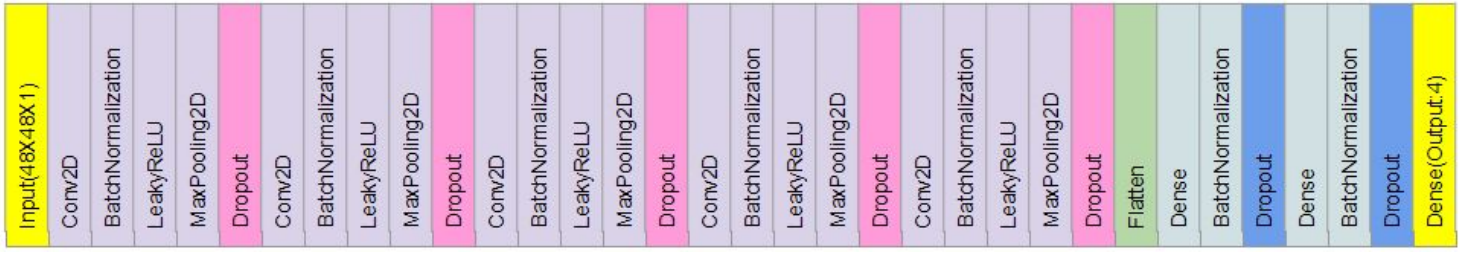
Solution Approach

1. *Problem Definition. EDA.*
2. *Build a Basic CNN Model.* — *12 models*
3. *Experiment Transfer Model* — *6 models*
4. *Build a Complex CNN Model* — *3 models*
5. *Compare, Evaluate & Conclude*



Model Solution

After training 21 models and different approaches, the selected model is a complex convolutional architecture model of 33 layers (batch size 32 and grayscale images set).

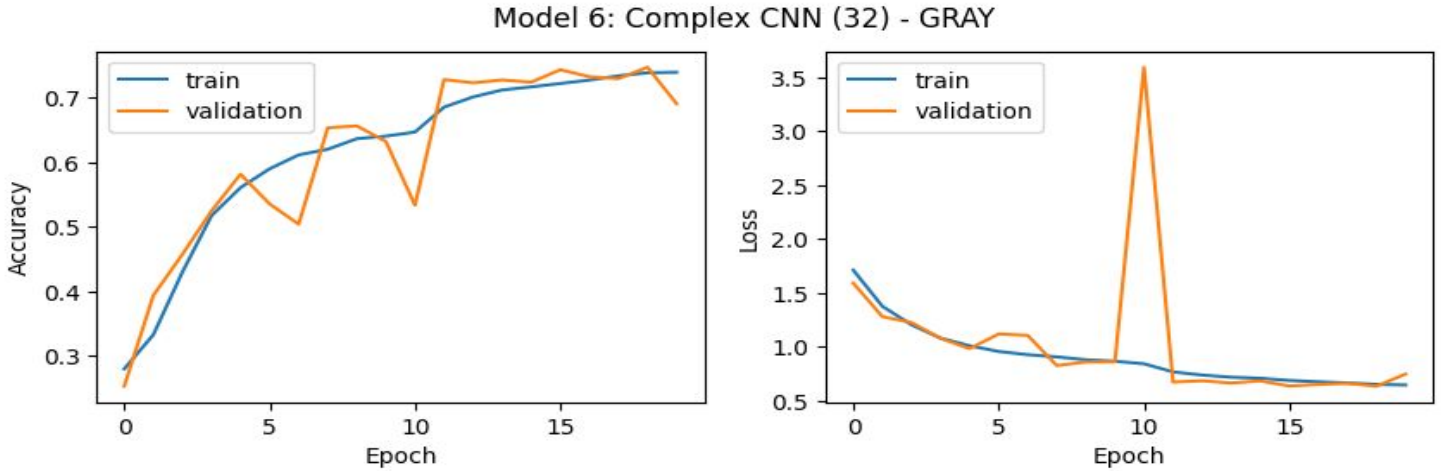


Feature extraction layers
5 blocks
Conv2D - BatchNormalization - LeakyReLu - MaxPooling2D - Dropout

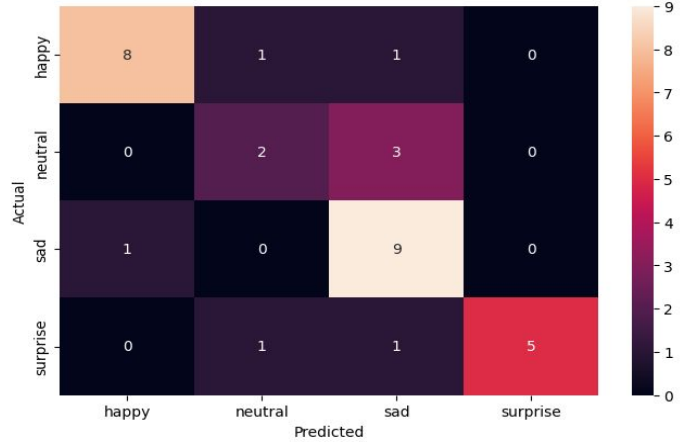
Custom Fully Connected Layers

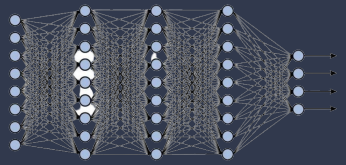
Model Evaluation

The selected model, with 33 layers, achieves 75% of accuracy in the testing set.



	precision	recall	f1-score	support
happy	0.89	0.80	0.84	10
neutral	0.50	0.40	0.44	5
sad	0.64	0.90	0.75	10
surprise	1.00	0.71	0.83	7
accuracy			0.75	32
macro avg	0.76	0.70	0.72	32
weighted avg	0.78	0.75	0.75	32



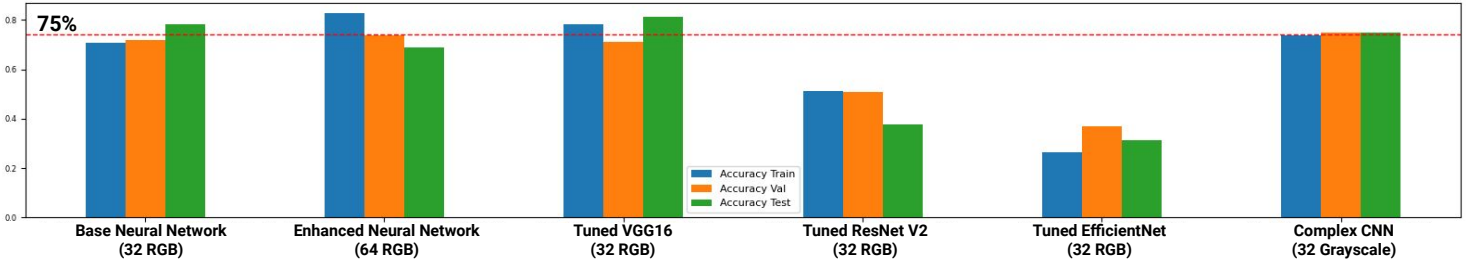


Executive Summary

⏩ 21 Models trained:

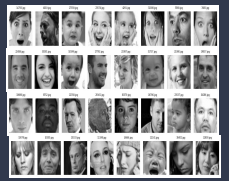
- 6 models related to a **basic CNN model of 13 layers**. Batch sizes 32, 64, and 128, and the images set in RGB or Grayscale.
- 6 models related to an **enhanced CNN model of 20 layers**. Batch sizes 32, 64, and 128, and the images set in RGB or Grayscale.
- 2 models using the **transfer model approach with VGG16 (18 layers)**.
- 2 models using the **transfer model approach with ResNet V2 (546 layers)**.
- 2 models using the **transfer model approach with EfficientNet (561 layers)**.
- 3 models related to a **Complex CNN model of 33 layers**. Batch sizes 32, 64, and 128, and just images in Grayscale.

Best Performed Models



⏩ The model that achieved the greatest accuracy was the model "**A complex model focused on a grayscale dataset and batch size 32**". In only 20 epochs this model achieved approximately 75% of accuracy, looking promising for the future.

Solution Benefits

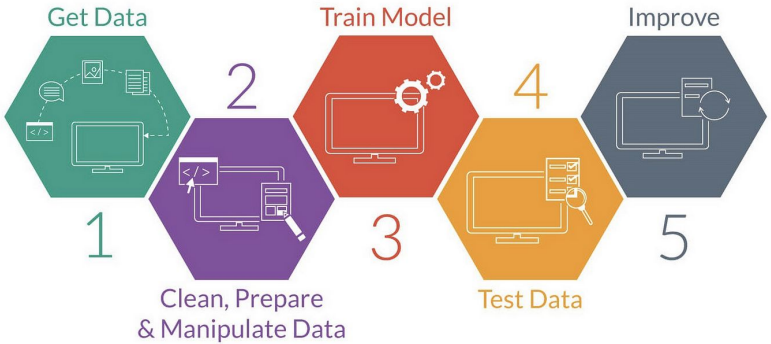


- The recommended model ("A complex model focused on a grayscale dataset and batch size 32") can be applied to the business to start classifying the images to address each image according to its class, this will improve the business operation by reducing time and cost.
- The company will not require hiring people for classification activities, allowing them to focus on their main objectives that take place after addressing each image in the right direction.

	Manual Process	Deep Learning Support Process
Task	Classify 30,000 images	Classify 30,000 images
Time	**21 working days.** Notes: 20 sec. per images, for 30,000 images, one person will need 21 working days to complete the task.	**7 minutes.** Notes: Approx 0.5 sec per batch of 32 images using Google Colab Pro Service. To classify 30,000 images, the model will need approx 7 minutes.
Costs	**$ 1,218.00 per month. No more tasks can be assigned.** Notes: Considering the minimum hourly salary of $ 7.25 USD, the cost of having one person working on the classification of 30,000 images will cost $ 1,218.	**$ 9.99 per month. More task can be assigned.** Notes: The Google Colab Pro Service costs $ 9.99 per month. This provides 100 compute units per month, Faster GPUs and access to higher memory machines.

Next Steps

Risks and Challenges



⏩ Next Steps

- Review the quality of the image dataset, paying attention to classes: "Sad" and "Neutral".
- Explore the results of the recommended model with the RGB dataset.
- The learning curve shows that the recommended model still learns at the end of the 20 epochs, it is necessary to explore if having more than 20 epochs gives a better result.
- Work on the overfitting detected in VGG16 model, and verify if the performance improves. This can be an alternative model to implement.

⏩ Risks and Challenges

- The image dataset needs to be reviewed and ensure acceptable quality.
- The company needs also to explore the possibility to get more images, paying special attention to Neutral and Sad classes.

Thanks

Capstone Project

“Emotions are essential parts of human intelligence. Without emotional intelligence, Artificial Intelligence will remain incomplete.”
— Amit Ray, Compassionate Artificial Intelligence.

References

- He, H., & Ma, Y. (Eds.). (2013). Imbalanced learning: foundations, algorithms, and applications.
- Save, serialize, and export models | TensorFlow Core. (n.d.). TensorFlow. https://www.tensorflow.org/guide/keras/serialization_and_saving
- Team, K. (n.d.). Keras documentation: Keras API reference. Keras API Reference. https://keras.io/api/
- Module: tf | TensorFlow v2.13.0. (n.d.). TensorFlow. https://www.tensorflow.org/api_docs/python/tf
- Colab Subscription Pricing. (n.d.). Colab Subscription Pricing. https://colab.research.google.com/signup
- Minimum wage | USAGov. (2023, June 16). Minimum Wage | USAGov. https://www.usa.gov/minimum-wage