Facial emotion detection

Presented by: Farzana R. Zaki

Overview of the problem

Ref: 1. https://nordicapis.com
2. https://appinventiv.com/blog/top-telehealth-trends/
3. https://www.cleanpng.com/png-sleep-car-controlled-access-highway-driver-driving-216837/
4. https://www.xminstitute.com/launchpads/understanding-human-behavior/
5. shutterstock.com
6. istock.com

Overview of the problem (contd.)

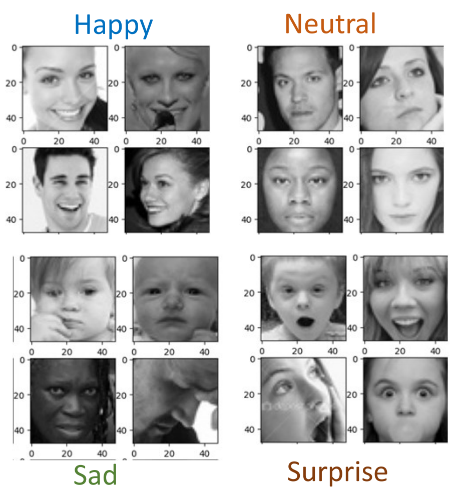
- However, extracting features from face-to-face is a challenging, complicated, and sensitive task for the computer vision techniques, such as deep learning, image processing, and machine learning to perform automated facial emotion recognition and classification.
- Some key challenges in extracting features from the facial image dataset include a variation of head pose, resolution and size of the image, background, and presence of other objects (hand, ornaments, eyeglasses, etc).
- In recent years, deep learning has become an efficient approach with the implementation of various architectures which allow the automatic extraction of features and classification using convolutional neural network (CNN), transfer learning, and recurrent neural network (RNN).

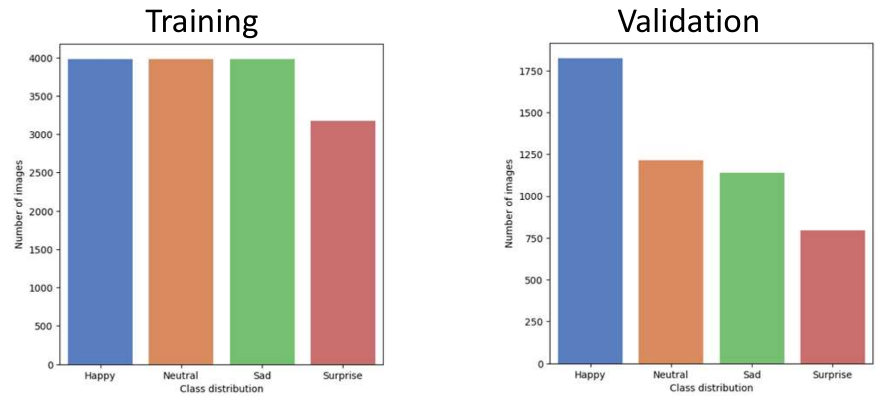
Solution approach

- Objective: To build a CNN model for facial emotion detection accurately.

- The proposed CNN model performs multi-class classification on images of facial emotions to classify the expressions according to the associated emotion.

Key findings and insights





Distribution of dataset

- Class distributions of four classes are slightly imbalanced for the training and validation datasets.
- Moreover, some low-resolution images are present in the training dataset.
- Images of neutral and sad faces are pretty much confusing. Therefore, CNN algorithms have faced the difficulty to correctly detect them properly. Therefore, for all the classifiers, F1 scores of neutral and sad emotions are not satisfactory.

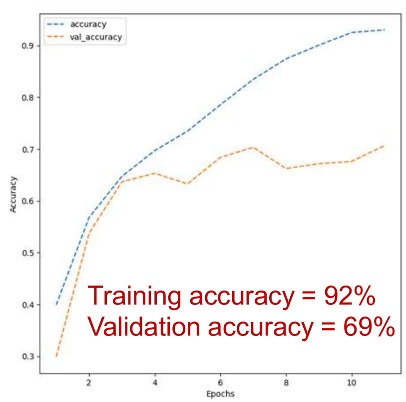
Missing images from training data

Missing images of validation data

Proposed model solution

- In this facial emotion detection project, various CNN models (simple, transfer learning and complex) are employed for training, validation and testing.
- In total, 17 different configuration CNN models are applied and evaluated.
- Simple and transfer learning CNN models are overfitting and have low F1 score.
- The best CNN model with few layers (with 3 convolutional layers and 2 dense layers) has achieved the training accuracy of 92.45% and the validation accuracy of 68.68%.

Training and validation accuracies



Confusion matrix for test data

Performance metrics for test data

	precision	recall	f1-score	support
0	0.82	0.72	0.77	32
1	0.58	0.69	0.63	32
2	0.57	0.62	0.60	32
3	0.93	0.78	0.85	32
accuracy			0.70	128
macro avg	0.72	0.70	0.71	128
weighted avg	0.72	0.70	0.71	128

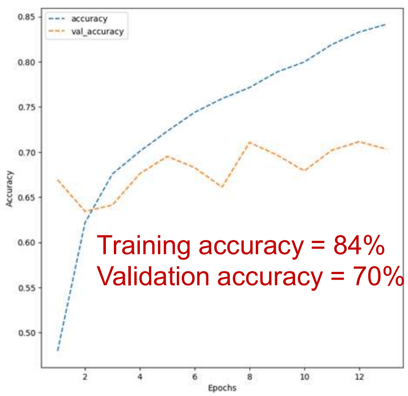
0: Happy, 1: Neutral, 2: Sad, 3: Surprise

CNN model performance (with 3 convolutional layers and 2 dense layers)

Proposed model solution (contd.)

Transfer learning model: VGG 16 performance (Output from block3_pool)

Training and validation accuracies



Confusion matrix for test data

Performance metrics for test data

```
              precision    recall  f1-score   support

           0       0.86      0.78      0.82        32
           1       0.55      0.88      0.67        32
           2       0.74      0.44      0.55        32
           3       0.93      0.84      0.89        32

    accuracy                           0.73       128
   macro avg       0.77      0.73      0.73       128
weighted avg       0.77      0.73      0.73       128
```

0: Happy, 1: Neutral, 2: Sad, 3: Surprise

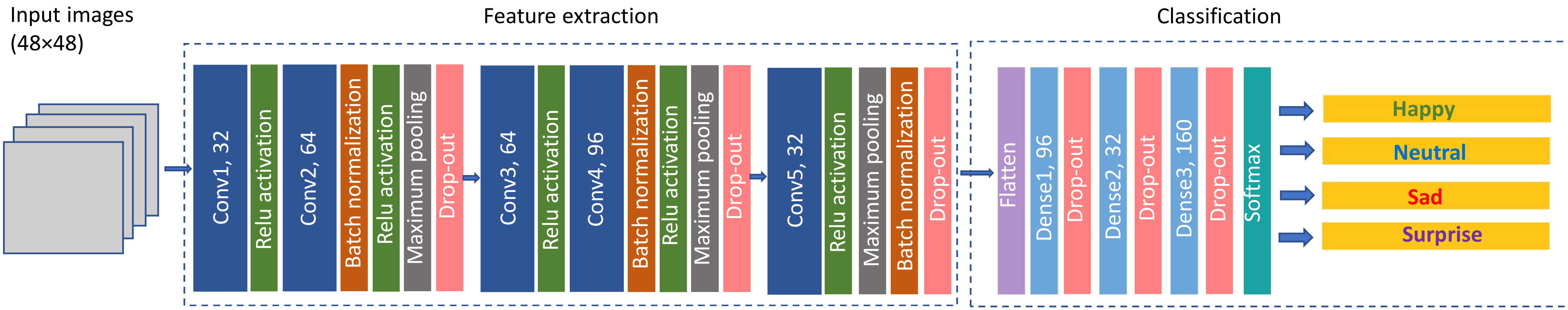
Obs.	Model	Training accuracy	Validation accuracy	Test accuracy	Status	Changes made to architecture
1.	VGG16	59.49	55.54	-	Poor performance	Output from block5_pool
2.	VGG16	77.38	68.05	-	Overfitting	Output from block4_pool
3.	**VGG16**	**84.16**	**70.32**	**73**	**Overfitting**	**Output from block3_pool**

Proposed model solution (contd.)

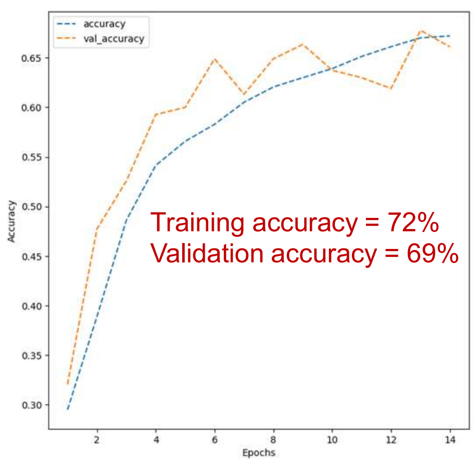
- For building the proposed CNN model, the hyperparameter tuning using random search from the keras tuner was applied to select the building blocks of the complex CNN models.
- As an optimizer, adam with three various learning rates: 0.1,0.01 and 0.001 are used.
- A layer with five convolutional blocks for feature selection and three dense layers for the classification are used for the complex CNN models with batch size of 16,32 and 64.
- Out of three complex CNN models, CNN model with five convolutional blocks for feature selection and three dense layers for the classification, with batch size of 32, learning rate of 0.001 and adam optimizer shows the best performance. This model is selected as the final proposed CNN model for the face emotion detection.

Obs.	Model	Training accuracy	Validation accuracy	Test accuracy	Status	Changes made to architecture
1.	**Complex CNN**	**72.23**	**69.10**	**74**	**Generalized**	**Batch size =32**
2.	Complex CNN	75.92	71.39	73	Generalized	Batch size =64
3.	Complex CNN	67.20	66.10	74	Generalized	Batch size =16

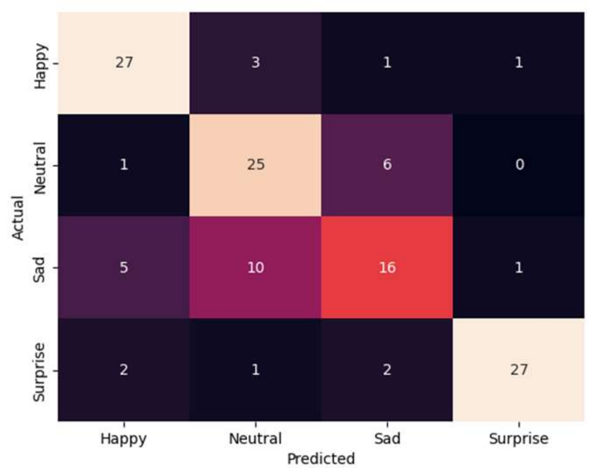
Final model solution



Training and validation accuracies



Confusion matrix for test data



Performance metrics for test data

	Precision	Recall	F1-score	Support
Happy	0.77	0.84	0.81	32
Neutral	0.64	0.78	0.70	32
Sad	0.64	0.50	0.56	32
Surprise	0.93	0.84	0.89	32
Accuracy			0.74	128
Macro avg.	0.75	0.74	0.74	128
Weighted avg.	0.75	0.74	0.74	128

Predicted class labels

Recommendations for implementation

1. We could employ oversampling technique for the 'surprise' dataset to make the balanced dataset and then again train the model and compare the performances.
2. The dataset is pretty small. Data augmentation can be applied to generate a large volume of training dataset by using the transformations of the face images, such as flip, shift, scaling, and rotation.
3. Convolutional with deeper layers (such as 6 or more layers) to verify the effectiveness of the augmented dataset, and the performance of these approached CNN models in comparison with some of the frequently used face recognition methods.
4. The proposed CNN model was implemented using GPU. Using more deeper convolutional layers with millions to trillions of training dataset may increase the implementation cost.
5. Several other transfer learning models can be applied to improve the performance of the facial recognition.

- It is recommended that stakeholders consider these variables in building improved long-term facial emotion detection models, as well as include the full range of environmental implications of various data sources, usage of color images instead of grayscale images, usage of more training data in developing future facial emotion detection.

Thank you

Appendix : Performance evaluation of 17 CNN architectures

Obs.	Model	Training accuracy	Validation accuracy	Test accuracy	Status	Changes made to architecture
1.	ANN	54.20	53.29	-	Poor performance	-
2.	Simple CNN (2-3 conv layers)	95.05	66.81	70	Overfitting	Conv layers added, leakyRelu
3.	Simple CNN (2-3 conv layers)	84.58	68.70	70	Overfitting	Added batch normalization and dropout layers, leakyRelu
4.	Simple CNN (2-3 conv layers)	97.43	65.62	67	Overfitting	Conv layers added, Relu
5.	Simple CNN (2-3 conv layers)	92.45	68.68	70	Overfitting	Added batch normalization and dropout layers, Relu
6.	VGG16	59.49	55.54	-	Poor performance	Output from block5_pool
7.	VGG16	77.38	68.05	-	Overfitting	Output from block4_pool
8.	VGG16	84.16	70.32	73	Overfitting	Output from block3_pool
9.	Resnet v2	26.21	36.67	-	Poor performance	Output from conv5_block3_add
10.	Resnet v2	31.12	22.89	-	Poor performance	Output from conv4_block23_add
11.	Resnet v2	26.08	22.89	25	Poor performance	Output from conv3_block4_add
12.	EffectiveNet	25.85	24.43	-	Poor performance	Output from block6e_expand_activation
13.	EffectiveNet	26.20	36.67	25	Poor performance	Output from block6e_expand_activation
14.	EffectiveNet	25.98	22.89	-	Poor performance	Output from block6e_expand_activation
15.	Complex CNN	72.23	69.10	74	Generalized	Batch size =32
16.	Complex CNN	75.92	71.39	73	Generalized	Batch size =64
17	Complex CNN	67.20	66.10	74	Generalized	Batch size =16