USING CONVOLUTIONAL NEURAL NETWORKS TO CLASSIFY FACIAL EXPRESSIONS

Brijesh Hota

Jenna Cogswell

PROBLEM STATEMENT

The problem that this project will address is the inability of visually impaired persons to read facial expressions of others.

Without an accessibility tool to aid in recognizing emotions, confusion and frustrations from all people involved in the conversation or interaction, may arise.

The models used in this project could theoretically be implemented as a software tool that can take in images and relay the emotions back to the user.

DATASET

The dataset chosen was found on Kaggle, it is the Fer-2013 dataset, made for a classification problem similar to this one.

The data consists of 23 298, 48x48 greyscale images. These images were arranged into folders classified by emotion.

The original dataset was split into train and test sets, for the purposes of this project we saved the images and kept the emotion folders but got rid of the split folders as we planned on splitting the data our own way.

The images are already centered on the faces. Ideally, we would have used images of higher quality and not already centered, but since we needed images already classified by emotion, this was the dataset we went with.

PROPOSED METHODOLOGY

Data Preprocessing

Split data into train, validation and test sets. Ensure number of images in classes are in a good range. Create directory iterators with image generators.

Build model

Add convolutional layers with relu activation function. Add maxpooling. Flatten the layers, then add dense layers with a softmax output layer.

Evaluate model

Compile and fit the model on the training and validation sets. Create graphs to display the results of the accuracy and loss from the model. Get the accuracy and loss from testing the model. Print confusion matrix and classification report.

Make changes to model and save as new one x2

Review evaluation of the last model and make changes accordingly and save as a new model to compare results. Changes include adding convolutional layers and dropout layers.

Test Loss: 1.5281312465667725
Test Accuracy: 0.620939314365387

	precision	recall	f1-score	support
0	0.13	0.10	0.12	743
1	0.00	0.00	0.00	83
2	0.15	0.14	0.14	769
3	0.24	0.22	0.23	1349
4	0.18	0.26	0.21	930
5	0.17	0.15	0.16	912
6	0.09	0.10	0.10	601
accuracy			0.17	5387
macro avg	0.14	0.14	0.14	5387
weighted avg	0.17	0.17	0.17	5387

EVALUATION METRICS

For this project, the evaluation metrics accuracy, precision, recall, and F1-score were researched, and Accuracy was considered.

The accuracy is the main metric we have focused on as the goal is to predict as many correct emotions as possible.

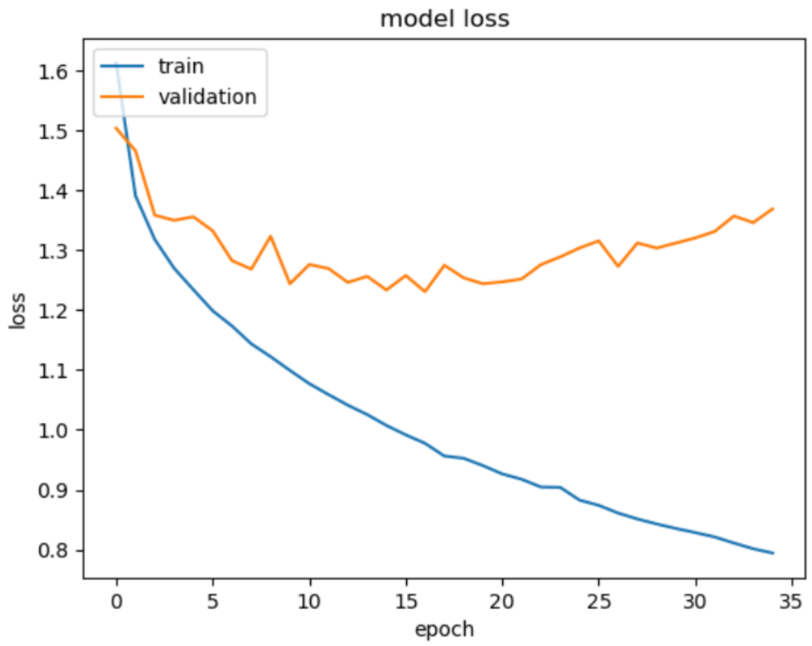
The other metrics are shown in our classification reports.

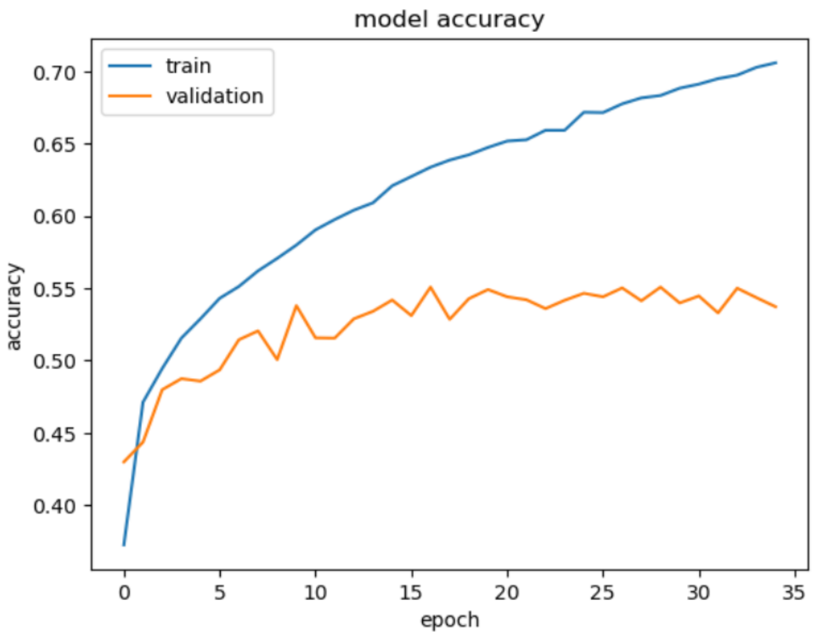
6

RESULTS – MODEL 1

loss: 1.3576 - accuracy: 0.5419

	precision	recall	f1-score	support
0	0.13	0.15	0.14	743
1	0.00	0.00	0.00	83
2	0.16	0.14	0.15	769
3	0.26	0.23	0.25	1349
4	0.16	0.20	0.18	930
5	0.17	0.15	0.16	912
6	0.09	0.10	0.09	601
accuracy			0.17	5387
macro avg	0.14	0.14	0.14	5387
weighted avg	0.17	0.17	0.17	5387

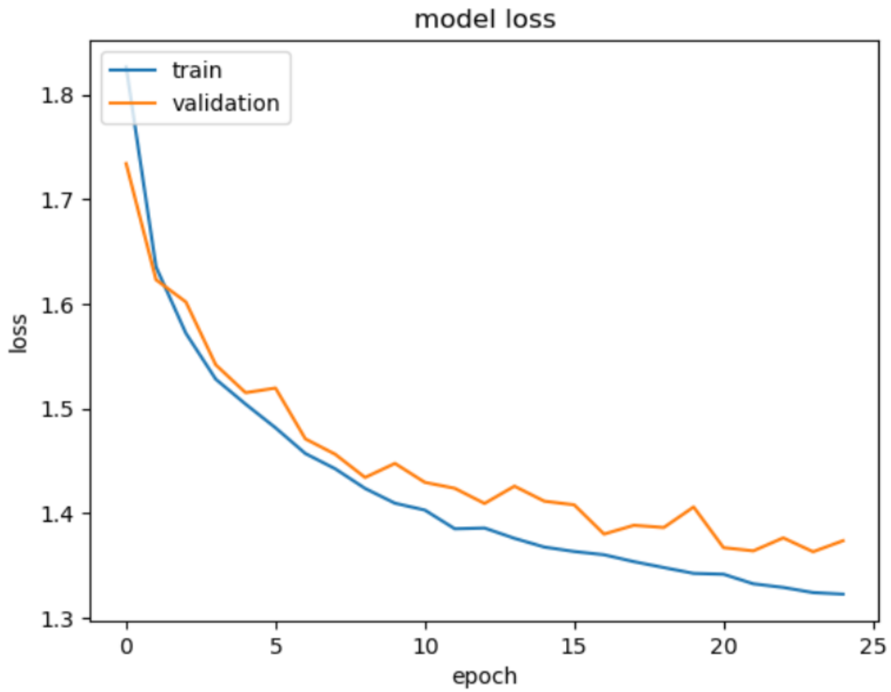


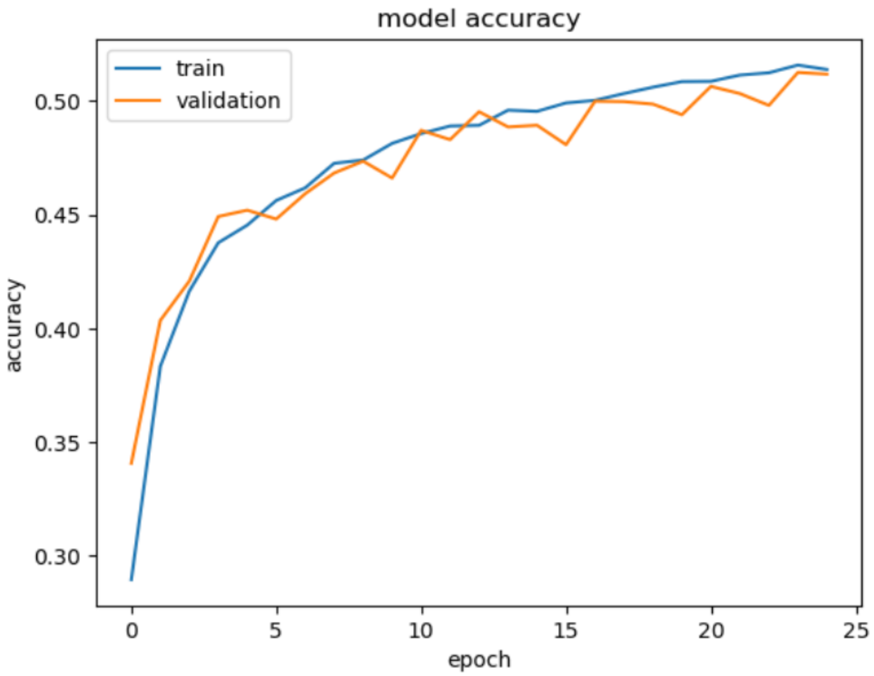


RESULTS – MODEL 2

```
loss: 1.3629 - accuracy: 0.5025
```

	precision	recall	f1-score	support
0	0.14	0.14	0.14	743
1	0.00	0.00	0.00	83
2	0.15	0.05	0.08	769
3	0.24	0.24	0.24	1349
4	0.17	0.20	0.18	930
5	0.18	0.24	0.20	912
6	0.12	0.12	0.12	601
accuracy			0.18	5387
macro avg	0.14	0.14	0.14	5387
weighted avg	0.17	0.18	0.17	5387

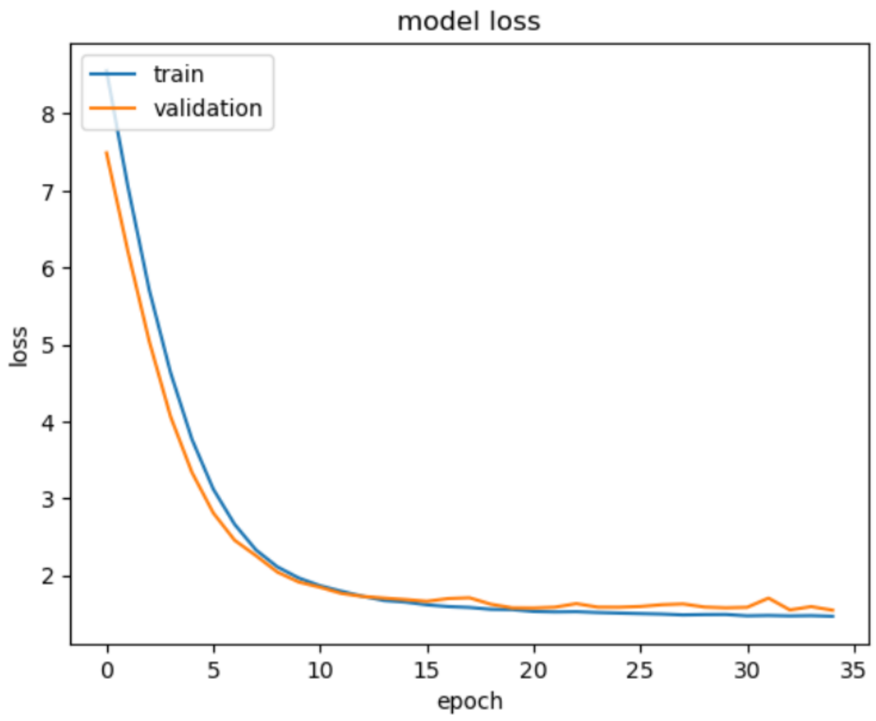


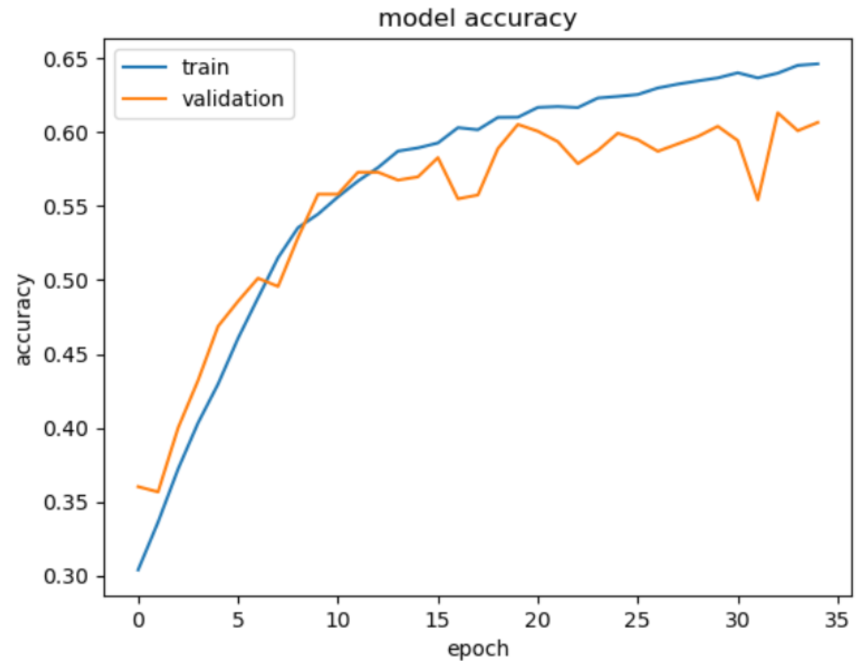


RESULTS – MODEL 3

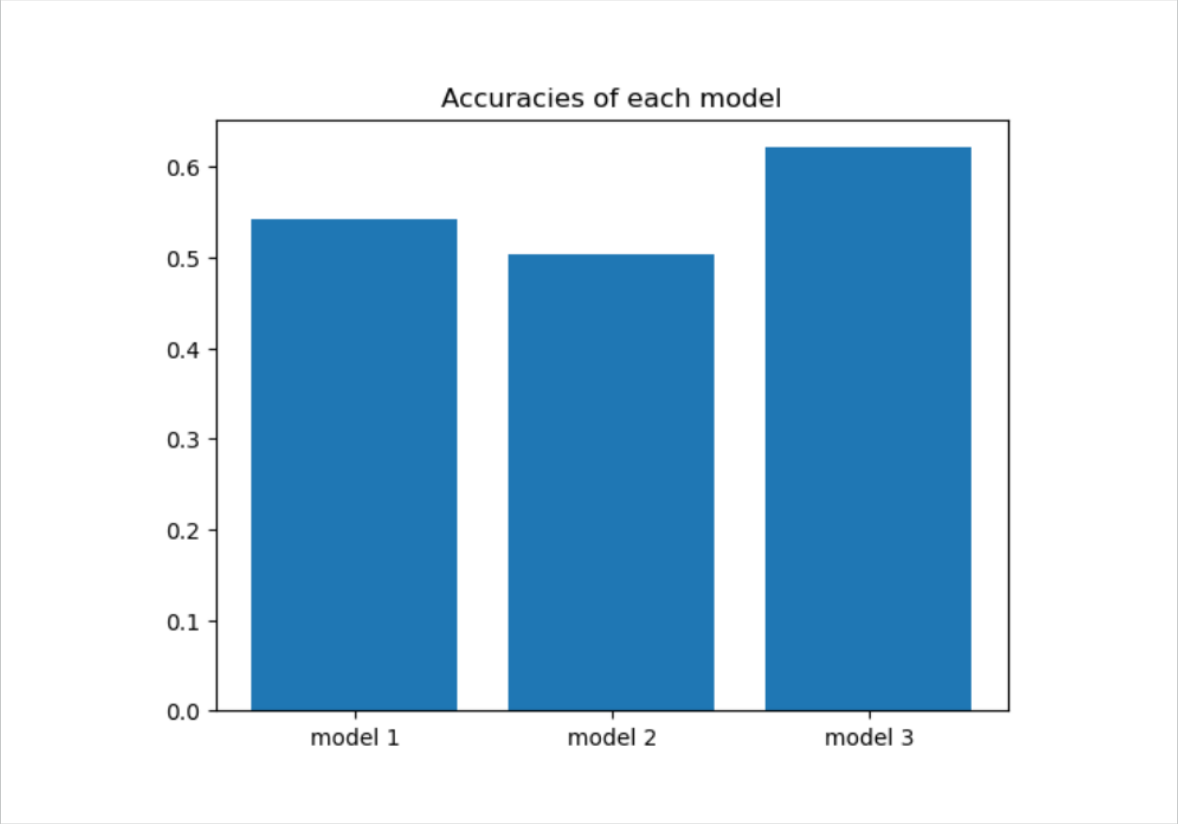
```
loss: 1.5281 - accuracy: 0.6209
```

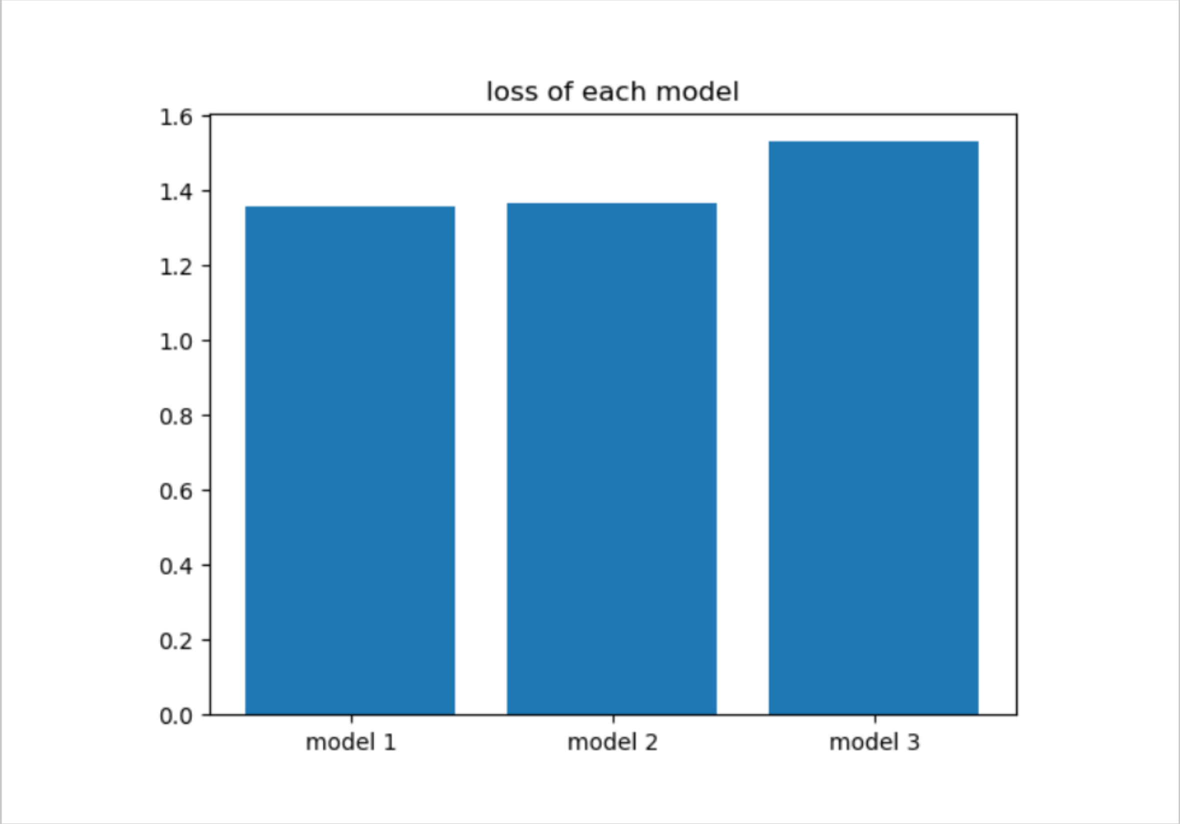
	precision	recall	f1-score	support
0	0.13	0.10	0.12	743
1	0.00	0.00	0.00	83
2	0.15	0.14	0.14	769
3	0.24	0.22	0.23	1349
4	0.18	0.26	0.21	930
5	0.17	0.15	0.16	912
6	0.09	0.10	0.10	601
accuracy			0.17	5387
macro avg	0.14	0.14	0.14	5387
weighted avg	0.17	0.17	0.17	5387





OVERALL RESULTS





CONCLUSION

- From our experiments we understood that more layers without proper parameter tuning is not helpful to increase the accuracy of a model.
- As observed in the case of 2nd and 1st models, where we increased the number of different layers, but the training and test accuracy decreased
- We also observed that greater number of convolutional layers resulted in better accuracy of the training set (and testing set) and greater training time (as each step took around 190 seconds to be trained).

FUTURE SCOPE

- We intend to utilize this project with different visualization software's that can be further integrated with assistive technologies for visually impaired
- The CNN model can be tuned even further into developing a real-time system for automated detection and recognition of human emotion
- Expanding the current model with RGB images and research more about the models that work with different types of datasets like RGB images, audio, speech, text and video.