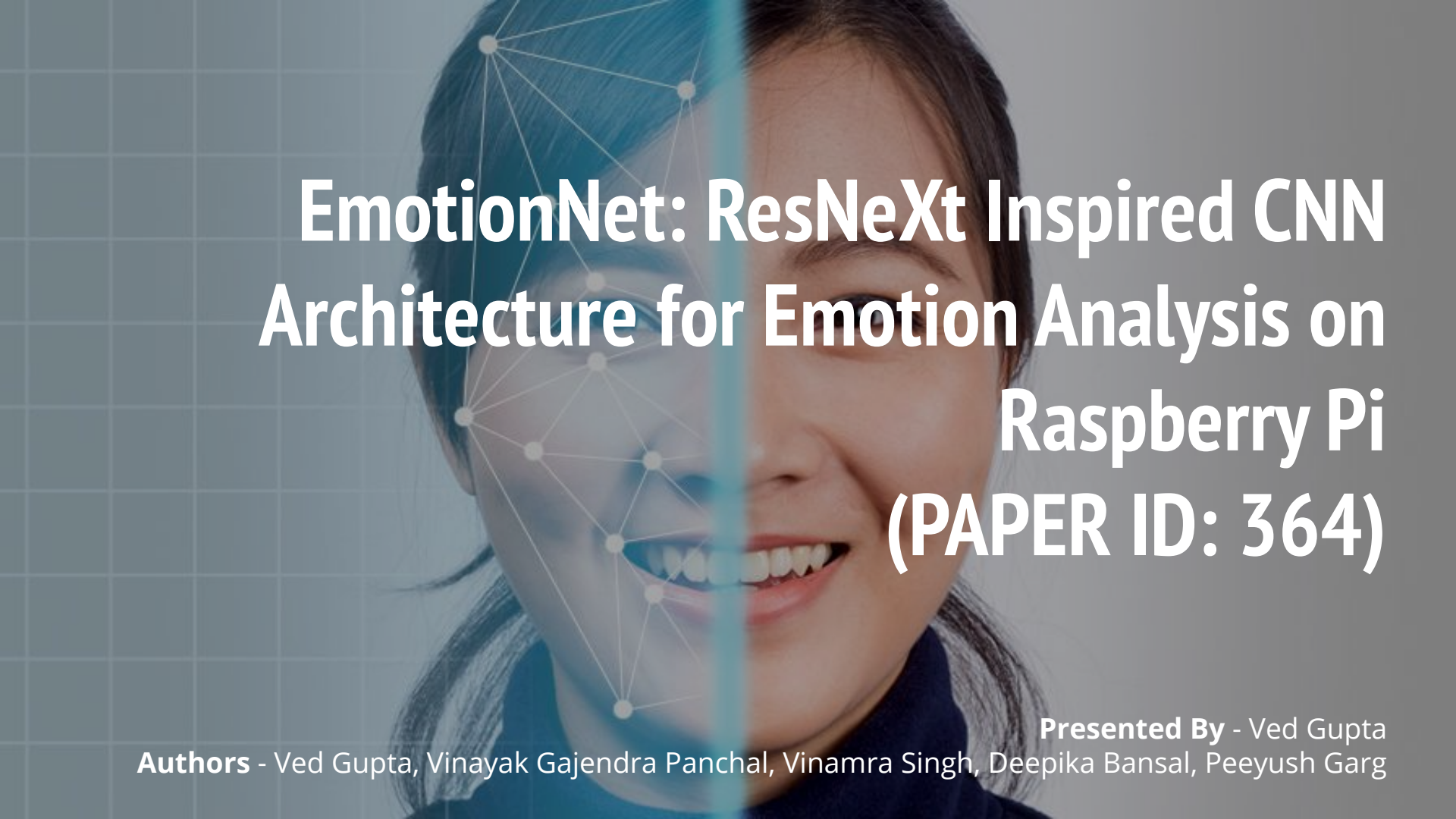


EmotionNet: ResNeXt Inspired CNN Architecture for Emotion Analysis on Raspberry Pi (PAPER ID: 364)

Presented By - Ved Gupta
Authors - Ved Gupta, Vinayak Gajendra Panchal, Vinamra Singh, Deepika Bansal, Peeyush Garg

MOTIVATION

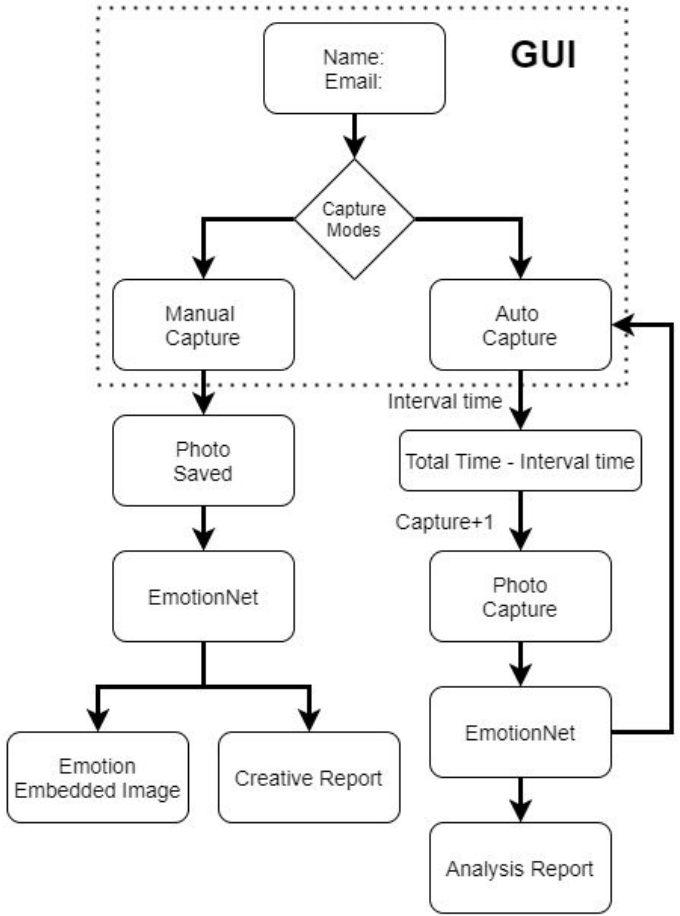
- Gallup 2021 Survey:
  - Adults have poor mental health, high levels of tension.
  - Increased levels of sadness.
  - Higher anger issues.
- World Happiness Report 2021:
  - Steady decline in India’s happiness levels over the past 8 years
  - Down to 3.82 from 4.77.
- Such trends, alongwith recent advancements in Artificial Intelligence, have popularised the field of Facial Emotion Recognition (FER).
- Most algorithms are resource intensive and computationally expensive.

OUR FER SYSTEM

Can be divided into three parts:

- ➔ Graphical User Interface
- ➔ Emotion Classification
- ➔ Report Auto-generation

Classification into five classes of emotions: Neutral, Happy, Sad, Angry, and Surprised.



FACE EXPRESSION DETECTION

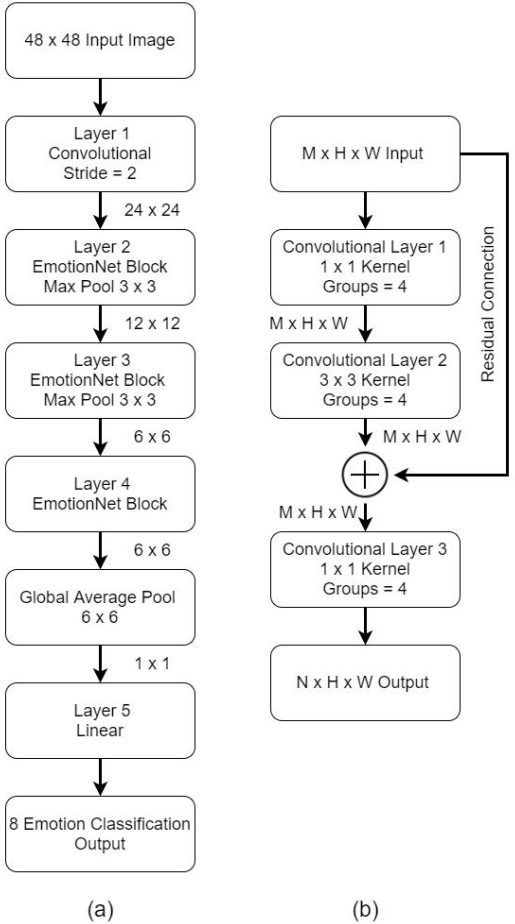
Haar Cascade Classifier: Uses haar filters introduced by Viola and Jones. ROI includes whole face.

EmotionNet: CNN-based model for classification into five emotions namely, neutral, happy, sad, anger and surprised.

EMOTIONNET ARCHITECTURE

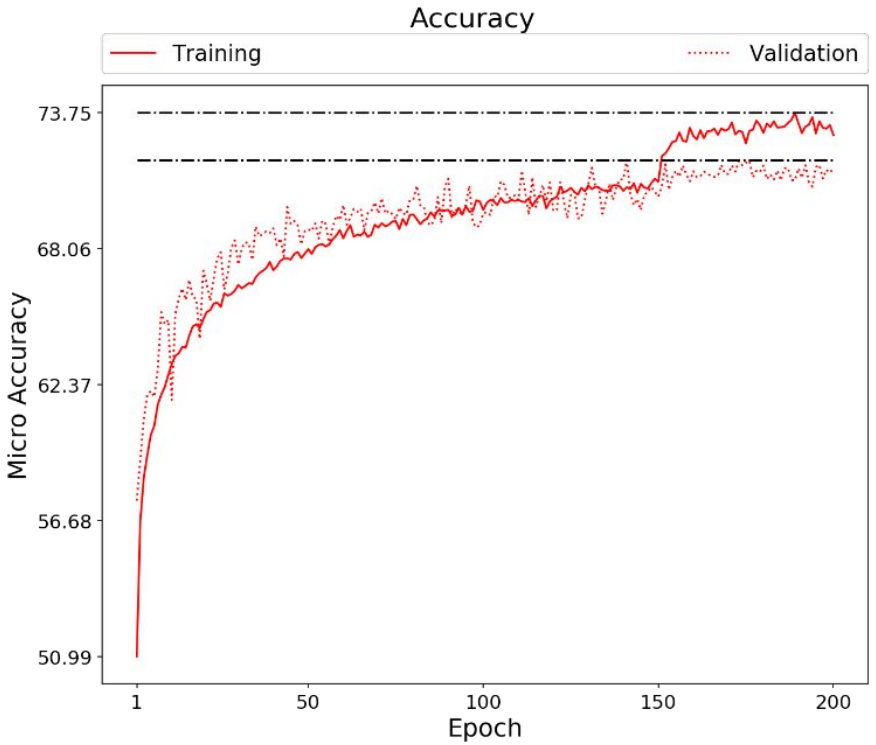
We introduce a novel architecture called EmotionNet.

- ➔ Inspired by ResNeXt architecture.
- ➔ Uses residual (shortcut) connections.
- ➔ Uses grouped convolutions (group = 4).
- ➔ 6 x 6 Global Average Pooling.
- ➔ Other features like Batch Normalization and Dropout.
- ➔ Extremely lightweight model, with only 29.16K parameters.
- ➔ CPU runtime of 4ms.



EMOTIONNET: TRAINING THE NETWORK

- Dataset used is FERPlus.
- Total 27,374 training examples.
- Image size 48 x 48, 5 classes.
- Trained using Adam Optimizer.
- Cost Function: Cross Entropy Loss.
- 200 epochs.



EMOTIONNET: RESULTS

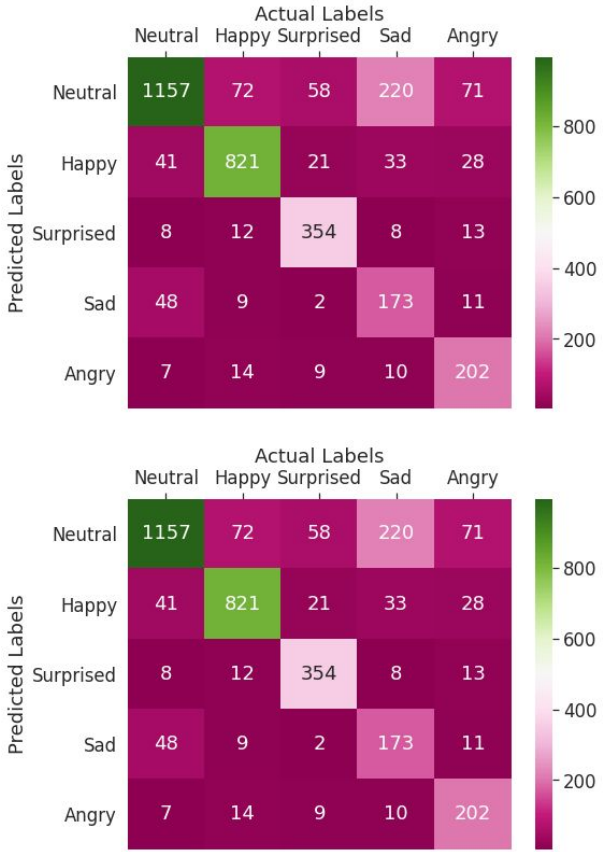
All metrics are micro-averaged.

Validation Set results:

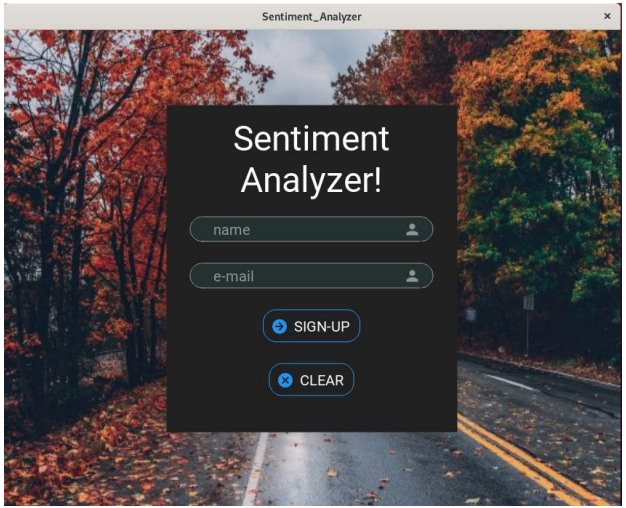
- Accuracy = 71.41%
- Precision = 81.87%
- Recall = 73.08%

Test Set results:

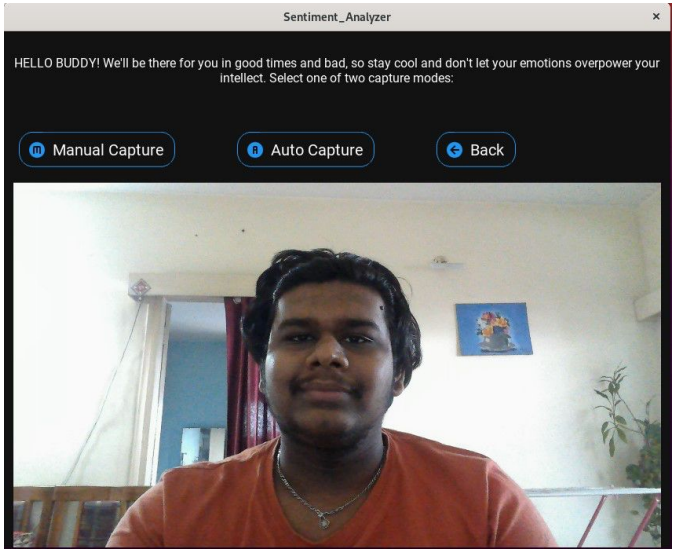
- Accuracy = 70.22%
- Precision = 80.91%
- Recall = 72.21%



GRAPHICAL USER INTERFACE



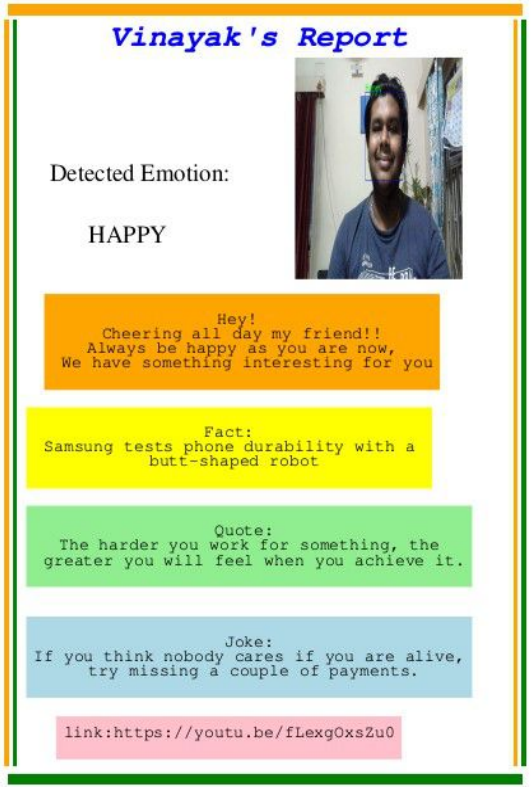
- User-friendly graphical interface, abstract, easy to operate.
- **Login Page:** Input name and email.
- **Capture Mode Page:** Two capture modes: “Manual Capture” and “Auto Capture”, live camera feedback embedded in the page.

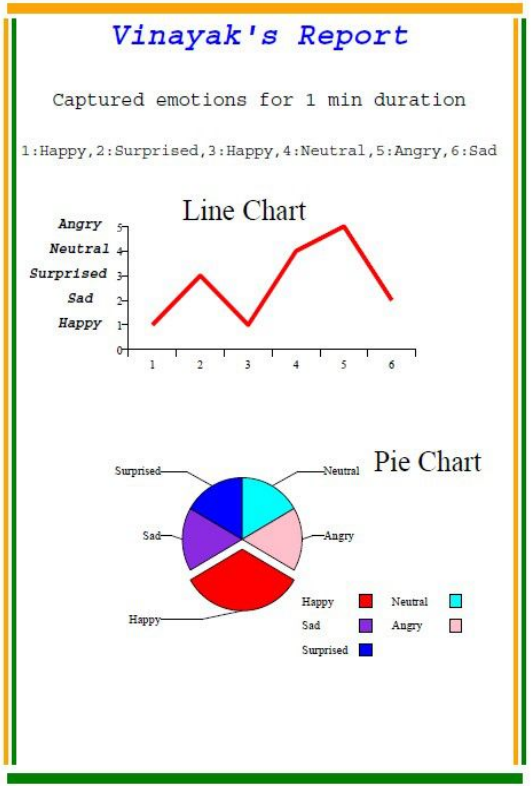


- **Manual Capture:** User controlled.
- **Auto Capture:** Snaps at regular intervals.
- Designed using Python library **KivyMD** and implemented on Raspberry pi.

REPORT AUTO-GENERATION

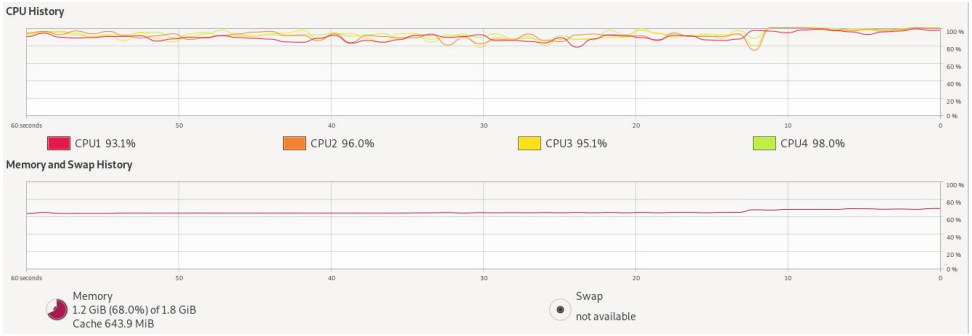
- Created using ReportLab.
- **Manual Capture Report:**
  - Bounded, tagged face image.
  - Messages, facts, quotes, jokes, and links to relevant videos and articles.
- **Auto Capture Report:**
  - Sequential list of detected emotions.
  - Timeline chart.
  - Pie chart.





SYSTEM TESTING

- Performance analysis on Raspberry Pi running Auto-Capture mode: 95.55% average CPU usage, and 68.0% RAM usage of 1.8 GB.
- Rate of 1.33 frames per second. This means that this system can run as a subpart of a project running at 1 fps while still leaving room for other processes.



CONCLUSION

- Compactness of the Emotion Recognition Model, with features like residual connections and grouped convolutions.
- Modularity achieved by integrating the model effectively with Raspberry Pi.
- On the FERPlus dataset, validation accuracy of 71.41% and a test set accuracy of 70.22% was achieved.
- Extremely lightweight model, with only 29.16K parameters, which is orders of magnitude lighter than similar networks.
- CPU time of 4ms, results in the FER system running very smoothly on the Raspberry Pi.
- The GUI is very easily operable due to its simplicity and the report provides all substantial information about the user in one place.
- Utilization in various fields, from small retail shops to large-scale companies.