Activation Layers Implication of CNN Sequential Models for Facial Expression Recognition

M. Shyamala Devi¹, Ankita Sagar², Karan Thapa³, Mreenmoy Hazarika⁴, P. Swathi⁵

¹Professor, Computer Science & Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai.
², ³, ⁴, ⁵Final Year B.Tech Student, Computer Science & Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai.
shyamalapmr@gmail.com

Abstract. Facial Expression Recognition is the critical part in the human emotional detection in the field of image processing. The application tends to soft or hard real time application based on the power of expression detection of idle image or videos. The Social communication with any object can be done by verbal or non-verbal format. Expression and Emotion detection completely rely on the non-verbal communication and facial expression. Machine Learning play an important role in the recognition of facial expression. An attempt is made in this paper to analyse the performance of Convolutional neural network models with diverse activation layer for the performance evaluation. Firstly, the facial expression dataset is extracted from the website http://www.consortium.ri.cmu.edu/ckagree/, http://app.visgraf.impa.br/database/faces/ is subjected with the data processing. Secondly, the data analysis is done for the distribution of expression image in the training and testing dataset. Thirdly, the facial expression images are detected with HAAR cascade and then the images are cropped with (350, 350). Fourth, the facial expression images is applied with normalized and the bottleneck features are created for training and testing data. Fifth, the training dataset is fitted with convolutional sequential neural network models with various activation layers like Sigmoid, Elu, Relu, Selu, Tanh, Softsign and Softplus. Sixth, the performance analysis is done with loss and accuracy for all the epoch of all CNN models for all the activation layers. Experimental results show that CNN sequential model with Relu activation layer is found to have the accuracy of 100%.

1. Introduction
Machine learning technology is used in all the areas of image processing in the applications like feature extraction of parts from the body, angular motion detection of body parts, identification of features from the image, counting of objects from the photo image and video, analysis of type of objects in the image, identification of outcome in gaming application and emotion identification of the object. Facial Recognition is the challenging task in the field of virtual reality where the people are working in various diverse applications like emotion identification, feature extraction, person identification, biometric identification, and fingerprint identification and classification and expression detection.

2. Literature Review
This paper deal with the comparative study of all the available dataset for facial expression images that are captured through idle photo image and video image. It also explores all the research insight that are
carried in the field of facial expression recognition and evaluation [1]. Human computer interaction is the major trending in the field of virtual reality and multimedia. This paper deals with emotional identification of the people interacting the system by the extraction of feature algorithm [2]. Machine learning models identify and predict potential high-cost patients and explore the key variables of the forecasting model, by comparing differences in the predictive performance of different variable sets. Interpretative regression method based on Dempster-Shafer theory, using the Evidence Regression model to predict the cost for insurance [3]. This paper explores the application of facial expression recognition in the field of driving. The hazardous damage occurs due to the unsafe driving by the drivers of the vehicle. It occurs due to the bad habits of the drivers by the alcohol intake. This may also be due to the frustration and depression of the drivers due to problems in their family. Such unaware feelings of the drivers can somehow identify by their facial expression reaction while they drive. This paper provides a solution for identifying the facial expression of the vehicle drivers to avoid the accident [4].

This work is dedicated in finding the action points in the face features to arrive the solution in finding the facial expression recognition. The expression of the face completely depends on the motion of the facial parts [5]. This paper develops the software for the automatic identification of the person face and their expression at real time [6]. It explores the complete facial expression recognition methods and techniques in the real time applications. For example in the crime cases, the police may get the video of the event happened. In such case, we need a face recognition system to find the target person [7]. This paper deals with normalization and the identification of human features in the face recognition objects by using the discrete cosine transform in logarithm domain [8]. This paper analyzes the use of evolutionary algorithm for the features that are extracted from the face. The inner differences between each facial feature are synthesized by using the genetic algorithm [9]. The facial expression recognition can be done and the efficiency can be improved by using machine learning classification algorithms. There is a possibility of combining the facial features for the facial expression recognition during real time and it is carried out by the support vector machine classification algorithm [10].

3. Our Contributions
The main analysis of this paper is to analyze the performance of Convolutional neural network models with diverse activation layer for the performance evaluation. The overall architecture of this paper is shown in Figure.1.

(i) Firstly, the facial expression dataset is extracted from the http://www.consortium.ri.cmu.edu/ckagree/, http://app.visgraf.impa.br/database/faces/ is subjected with the data processing.

(ii) Secondly, the data analysis is done for the distribution of expression image in the training and testing dataset.

(iii) Thirdly, the facial expression images are detected with HAAR cascade and then the images are cropped with (350, 350).

(iv) Fourth, the facial expression images is applied with normalized and the bottleneck features are created for training and testing data.

(v) Fifth, the training dataset is fitted with convolutional sequential neural network models with various activation layers like Sigmoid, Elu, Relu, Selu, Tanh, Softsign and Softplus.

(vi) Sixth, the performance analysis is done with loss and accuracy for all the epoch of all CNN models for all the activation layers.
4. Implementation Setup

4.1. Dataset Exploratory Analysis
The facial expression dataset is extracted from the http://www.consortium.ri.cmu.edu/ckagree/, http://app.visgraf.impa.br/database/faces/ is subjected with the data processing. The python scripting language is coded in Spyder editor with Anaconda navigator for implementation. Data analysis is done for distribution of expression image in training and testing dataset and is shown in Figure. 2 – Figure.4.

---

**Figure 1.** Overall Proposed Workflow

**Figure 2.** Distribution of Emotion of Training Data
5. Results and Discussions

The facial expression images are detected with HAAR cascade and then the images are cropped with the size of (350, 350). The facial expression images is applied with normalized and the bottleneck features are created for training and testing data. The training dataset is fitted with convolutional sequential neural network models with Sigmoid activation layer. Table 1 shows the performance metric for the different Epoch with Sigmoid activation layer. The performance analysis of Epoch VS Accuracy and Epoch VS Loss is shown in Figure. 5-Figure. 6.
The training dataset is fitted with convolutional sequential neural network models with Elu activation layer. Table 2 shows the performance metric for the different Epoch with Elu activation layer. Performance analysis of Epoch VS Accuracy, Epoch VS Loss is shown in Figure 7-Figure 8.
Table 2. Analysis of Performance Metric for different Epoch of ELU

| Epoch | Comb_Train_Loss | Comb_Train_Accuracy | CVHuman_Loss | CVHuman_Accuracy |
|-------|-----------------|---------------------|--------------|-----------------|
| 1     | 2.303932        | 0.185714            | 13.380573    | 0               |
| 2     | 1.719688        | 0.357143            | 23.328525    | 0               |
| 3     | 1.673203        | 0.328571            | 16.39051     | 0               |
| 4     | 1.487079        | 0.414286            | 10.606923    | 0               |
| 5     | 1.375292        | 0.485714            | 6.738069     | 0.028571        |
| 6     | 1.204241        | 0.585714            | 5.71521      | 0.1             |
| 7     | 1.106813        | 0.571429            | 5.330394     | 0.114286        |
| 8     | 0.866124        | 0.657143            | 3.219572     | 0.214286        |
| 9     | 0.839699        | 0.671429            | 2.029013     | 0.357143        |
| 10    | 0.624861        | 0.8                 | 2.621637     | 0.271429        |
| 11    | 0.611234        | 0.785714            | 3.189922     | 0.057143        |
| 12    | 0.443013        | 0.928571            | 2.274216     | 0.214286        |
| 13    | 0.333843        | 0.942857            | 1.771692     | 0.371429        |
| 14    | 0.229817        | 0.971429            | 2.786089     | 0.257143        |
| 15    | 0.242833        | 0.942857            | 2.510545     | 0.414264        |
| 16    | 0.142365        | 0.971429            | 3.314649     | 0.185714        |
| 17    | 0.134086        | 0.985714            | 2.281677     | 0.114286        |
| 18    | 0.099402        | 0.98                 | 1.601833     | 0.3             |
| 19    | 0.058849        | 0.98                 | 1.482906     | 0.428571        |
| 20    | 0.044989        | 0.99                 | 1.698972     | 0.457143        |

The training dataset is fitted with convolutional sequential neural network models with Elu activation layer. Table 3 shows the performance metric for the different Epoch with Tanh activation layer. The performance analysis of Epoch VS Accuracy and Epoch VS Loss is shown in Figure 9-Figure 10.

Table 3. Analysis of Performance Metric for different Epoch of TANH

| Epoch | Comb_Train_Loss | Comb_Train_Accuracy | CVHuman_Loss | CVHuman_Accuracy |
|-------|-----------------|---------------------|--------------|-----------------|
| 1     | 2.162163        | 0.171429            | 2.389536     | 0.2             |
| 2     | 2.008764        | 0.185714            | 2.345485     | 0.157143        |
| 3     | 2.31146         | 0.114286            | 2.060338     | 0.185714        |
| 4     | 1.903016        | 0.257143            | 2.070557     | 0.2             |
| 5     | 2.027521        | 0.185714            | 2.123116     | 0.271429        |
| 6     | 2.192716        | 0.128571            | 2.336197     | 0               |
| 7     | 2.044704        | 0.171429            | 2.610682     | 0               |
| 8     | 2.081102        | 0.114286            | 2.711213     | 0               |
| 9     | 2.021846        | 0.214286            | 2.415396     | 0               |
| 10    | 2.02971         | 0.348571            | 2.312553     | 0               |
| 11    | 1.997615        | 0.392857            | 2.248497     | 0               |
| 12    | 1.975284        | 0.457143            | 2.115869     | 0.042857        |
| 13    | 1.891716        | 0.49                | 2.026604     | 0               |
| 14    | 1.959064        | 0.537143            | 2.248666     | 0.042857        |
| 15    | 1.95536         | 0.6                 | 2.343453     | 0.2             |
| 16    | 1.972222        | 0.652857            | 2.317549     | 0.157143        |
| 17    | 2.005251        | 0.7                 | 2.326868     | 0.114286        |
| 18    | 4.950364        | 0.757143            | 2.418729     | 0               |
| 19    | 5.970067        | 0.794286            | 2.418652     | 0               |
| 20    | 6.916914        | 0.817143            | 2.845998     | 0               |

Figure 9. EPOCH VS LOSS for TANH

Figure 10. EPOCH VS ACCURACY for TANH
The training dataset is fitted with convolutional sequential neural network models with Elu activation layer. Table 4 shows the performance metric for the different Epoch with Selu activation layer. The performance analysis of Epoch VS Accuracy and Epoch VS Loss is shown in Figure. 11-Figure. 12.

| Epoch | Comb Train Loss | Comb Train Accuracy | CVHuman Loss | CVHuman Accuracy |
|-------|-----------------|---------------------|--------------|-----------------|
| 1     | 2.559849        | 0.185714            | 4.90646      | 0.185714        |
| 2     | 1.983575        | 0.242857            | 9.61076      | 0.3             |
| 3     | 1.475739        | 0.428571            | 13.452377    | 0.3             |
| 4     | 1.262215        | 0.514286            | 8.95803      | 0.3             |
| 5     | 1.322643        | 0.514286            | 4.371711     | 0.285714        |
| 6     | 1.102371        | 0.571429            | 3.892981     | 0.185714        |
| 7     | 0.957996        | 0.657143            | 3.54792      | 0.2             |
| 8     | 0.796078        | 0.771429            | 3.509347     | 0.2             |
| 9     | 0.631038        | 0.771429            | 2.292966     | 0.5             |
| 10    | 0.689604        | 0.742857            | 2.271856     | 0.457143        |
| 11    | 0.544641        | 0.814286            | 2.006077     | 0.385714        |
| 12    | 0.397772        | 0.871429            | 1.38054      | 0.442857        |
| 13    | 0.264411        | 0.942857            | 1.206447     | 0.457143        |
| 14    | 0.286198        | 0.928571            | 1.489429     | 0.4             |
| 15    | 0.218547        | 0.914286            | 1.690097     | 0.412486        |
| 16    | 0.163724        | 0.971429            | 2.512727     | 0.128571        |
| 17    | 0.143616        | 0.985714            | 3.048205     | 0.185714        |
| 18    | 0.103006        | 0.985714            | 2.525985     | 0.314286        |
| 19    | 0.111789        | 0.957143            | 2.65915      | 0.314286        |
| 20    | 0.080792        | 0.96                | 4.199882     | 0.112486        |

Figure 11. EPOCH VS LOSS for SELU
Figure 12. EPOCH VS ACCURACY for SELU

The training dataset is fitted with convolutional sequential neural network models with Elu activation layer. Table 5 shows the performance metric for the different Epoch with softsign activation layer. The performance analysis of Epoch VS Accuracy and Epoch VS Loss is shown in Figure. 13-Figure. 14.

| Epoch | Comb Train Loss | Comb Train Accuracy | CVHuman Loss | CVHuman Accuracy |
|-------|-----------------|---------------------|--------------|-----------------|
| 1     | 1.935812        | 0.271429            | 1.949712     | 0.2             |
| 2     | 1.892569        | 0.228571            | 1.897579     | 0.3             |
| 3     | 1.783687        | 0.257143            | 1.909734     | 0.3             |
| 4     | 1.716713        | 0.4                 | 2.066091     | 0.3             |
| 5     | 1.700327        | 0.342857            | 1.959024     | 0.3             |
| 6     | 1.758143        | 0.271429            | 1.945406     | 0.3             |
| 7     | 1.637998        | 0.428571            | 2.053937     | 0.271429        |
| 8     | 1.644241        | 0.49                | 2.230492     | 0.142857        |
| 9     | 1.843184        | 0.542857            | 2.055639     | 0.3             |
| 10    | 1.857275        | 0.657143            | 1.935686     | 0.3             |
| 11    | 1.639422        | 0.657143            | 2.150152     | 0.285714        |
| 12    | 1.686175        | 0.7                 | 2.481322     | 0.028571        |
| 13    | 1.565544        | 0.728571            | 2.778958     | 0.2             |
| 14    | 1.579589        | 0.742857            | 2.918343     | 0.22            |
| 15    | 1.576361        | 0.742857            | 2.956654     | 0.24            |
| 16    | 1.640967        | 0.8                 | 2.716831     | 0.24            |
| 17    | 1.539839        | 0.81                | 2.12686     | 0.23            |
| 18    | 1.389249        | 0.824286            | 2.097143     | 0.25            |
The training dataset is fitted with convolutional sequential neural network models with Elu activation layer. Table 6 shows the performance metric for the different Epoch with Softplus activation layer. Performance analysis of Epoch VS Accuracy and Epoch VS Loss is shown in Figure. 15-Figure. 16.

Table 6. Analysis of Performance Metric for different Epoch of SOFTPLUS

| Epoch | Comb_Train_Loss | Comb_Train_Accuracy | CVHuman_Loss | CVHuman_Accuracy |
|-------|----------------|---------------------|--------------|-----------------|
| 1     | 2.521988       | 0.257143            | 10.769857    | 0.142857        |
| 2     | 1.858173       | 0.328571            | 14.013214    | 0.071429        |
| 3     | 1.586269       | 0.485714            | 8.450514     | 0               |
| 4     | 1.371293       | 0.5                 | 5.912697     | 0               |
| 5     | 1.22164        | 0.528571            | 4.612834     | 0.128571        |
| 6     | 0.994899       | 0.628571            | 1.815635     | 0.571429        |
| 7     | 0.759172       | 0.814286            | 2.047843     | 0.371429        |
| 8     | 0.679914       | 0.785714            | 2.116582     | 0.414286        |
| 9     | 0.456746       | 0.9                 | 1.590425     | 0.457143        |
| 10    | 0.361784       | 0.942857            | 2.683431     | 0.428571        |
| 11    | 0.291873       | 0.957143            | 3.042206     | 0.3             |
| 12    | 0.229183       | 0.971429            | 2.075521     | 0.314286        |
| 13    | 0.141551       | 1                    | 2.790561     | 0.242857        |
| 14    | 0.114004       | 1                    | 2.889609     | 0.271429        |
| 15    | 0.072381       | 1                    | 2.089473     | 0.314286        |
| 16    | 0.084483       | 0.985714            | 1.952255     | 0.371429        |
| 17    | 0.057751       | 1                    | 1.957056     | 0.3             |
| 18    | 0.054487       | 1                    | 2.250003     | 0.342857        |
| 19    | 0.059941       | 1                    | 2.20055      | 0.342857        |
| 20    | 0.089052       | 0.985714            | 2.690978     | 0.3             |
The training dataset is fitted with convolutional sequential neural network models with Elu activation layer. Table 7 shows the performance metric for the different Epoch with Relu activation layer. The performance analysis of Epoch VS Accuracy and Epoch VS Loss is shown in Figure. 17-Figure. 18.

| Epoch | Comb_Train_Loss | Comb_Train_Accuracy | CVHuman_Loss | CVHuman_Accuracy |
|-------|-----------------|---------------------|--------------|------------------|
| 1     | 2.303362        | 0.157143            | 10.111909    | 0.271429         |
| 2     | 1.999738        | 0.257143            | 9.292259     | 0.3              |
| 3     | 1.843891        | 0.357143            | 5.913297     | 0.3              |
| 4     | 1.657243        | 0.314286            | 7.447931     | 0.3              |
| 5     | 1.639881        | 0.342857            | 5.799656     | 0.257143         |
| 6     | 1.515211        | 0.314286            | 5.677063     | 0.2              |
| 7     | 1.355517        | 0.471429            | 3.825484     | 0.142857         |
| 8     | 1.279170        | 0.471429            | 2.265329     | 0.085714         |
| 9     | 1.358319        | 0.457143            | 6.458704     | 0              |
| 10    | 1.229157        | 0.542857            | 7.298427     | 0                |
| 11    | 0.982870        | 0.600000            | 4.836370     | 0.128571         |
| 12    | 0.911965        | 0.657143            | 3.400242     | 0.257143         |
| 13    | 0.771945        | 0.742857            | 2.544070     | 0.114286         |
| 14    | 0.663263        | 0.828571            | 1.747203     | 0.242857         |
| 15    | 0.562241        | 0.857143            | 1.785336     | 0.2              |
| 16    | 0.441646        | 0.928571            | 2.436385     | 0.2              |
| 17    | 0.347169        | 0.971429            | 2.715715     | 0.2              |
| 18    | 0.285406        | 0.971429            | 2.387037     | 0.314286         |
| 19    | 0.262842        | 0.957143            | 2.294834     | 0.328571         |
| 20    | 0.197343        | 1.000000            | 3.707844     | 0.2              |

Table 7 shows the Accuracy performance metric for the different Epoch with all activation layers of CNN modes and the analysis is shown in Figure. 19.
Figure 19. Accuracy Analysis of Activation Layers

### Table 8. Accuracy Analysis of Activation Layers of CNN Models

| Epoch | sigmoid | elu   | tanh  | selu  | Softsign | softplus | relu     |
|-------|---------|-------|-------|-------|----------|----------|----------|
| 1     | 0.142857| 0.185714| 0.171429| 0.185714| 0.271429| 0.257143| 0.157143|
| 2     | 0.3     | 0.357143| 0.185714| 0.242857| 0.228571| 0.328571| 0.257143|
| 3     | 0.342857| 0.328571| 0.114286| 0.428571| 0.514286| 0.657143| 0.471429|
| 4     | 0.4     | 0.414286| 0.257143| 0.242857| 0.328571| 0.528571| 0.471429|
| 5     | 0.428571| 0.485714| 0.185714| 0.185714| 0.271429| 0.257143| 0.342857|
| 6     | 0.457143| 0.585714| 0.128571| 0.271429| 0.21429| 0.628571| 0.314286|
| 7     | 0.442857| 0.514286| 0.114286| 0.514286| 0.257143| 0.528571| 0.457143|
| 8     | 0.414286| 0.657143| 0.257143| 0.428571| 0.328571| 0.528571| 0.49    |
| 9     | 0.428571| 0.671429| 0.214286| 0.771429| 0.542857| 0.528571| 0.342857|
| 10    | 0.428571| 0.3     | 0.348571| 0.742857| 0.657143| 0.928571| 0.542857|
| 11    | 0.385714| 0.785714| 0.392857| 0.814286| 0.657143| 0.957143| 0.6     |
| 12    | 0.3     | 0.928571| 0.457143| 0.871429| 0.971429| 0.657143| 0.657143|
| 13    | 0.428571| 0.942857| 0.49     | 0.942857| 0.728571| 1       | 0.742857|
| 14    | 0.357143| 0.971429| 0.537143| 0.928571| 0.742857| 1       | 0.857143|
| 15    | 0.3     | 0.942857| 0.6     | 0.914286| 0.742857| 1       | 0.857143|
| 16    | 0.542857| 0.971429| 0.652857| 0.971429| 0.985714| 1       | 0.957143|
| 17    | 0.631429| 0.985714| 0.6     | 0.985714| 0.985714| 1       | 0.971429|
| 18    | 0.697143| 0.98     | 0.757143| 0.985714| 0.824286| 1       | 0.971429|
| 19    | 0.704286| 0.98     | 0.794286| 0.957143| 0.82     | 1       | 0.957143|
| 20    | 0.791429| 0.99     | 0.817143| 0.96     | 0.844286| 0.985714| 1       |

6. Conclusion

An attempt is made in this paper to analyze the performance of Convolutional neural network models with diverse activation layer for the performance evaluation. The facial expression images are applied with normalized and the bottleneck features are created for training and testing data. The training dataset is fitted with convolutional sequential neural network models with various activation layers like Sigmoid, Elu, Relu, Selu, Tanh, Softsign and Softplus. Performance analysis is done with loss and accuracy for all the epoch of all CNN models for all the activation layers. Experimental results shows that CNN sequential model with Relu activation layer is found to have the accuracy of 100%.

References

[1] Huang Y Chen F Lv S and Wang X 2019 Facial Expression Recognition: A Survey. Symmetry vol 11 1189.

[2] Cowie R Douglas Cowie E Tsapatsoulis N Votsis G Kollias S Fellenz W and Taylor J G 2001 Emotion recognition in human-computer interaction IEEE Signal Process Mag vol 18 p 32–80

[3] Jabon M Bailenson J Pontikakis E Takayama L and Nass C 2011 Facial expression analysis for predicting unsafe driving behavior IEEE Transactions on Pervasive Computing vol 10 p 84–95

[4] Tian Y I Kanade T and Cohn J F 2001 Recognizing action units for facial expression analysis
IEEE Trans. Pattern Anal. Mach. Intell. 2001, 23, 97–115.

[5] Samal A and Iyengar P A 1992 Automatic recognition and analysis of human faces and facial expressions: A survey Pattern Recognition journal vol 25 p 65–77

[6] Deshmukh S Patwardhan M and Mahajan 2016 A Survey on real-time facial expression recognition techniques. IET Biometric Journal vol 5 p 155–163

[7] Chen W Er M J and Wu S 2006 Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain IEEE Transaction System. Management Cybernetics vol 36 p 458–466

[8] Yu J and Bhanu B 2006 Evolutionary feature synthesis for facial expression recognition Pattern Recognition Letter vol 27 p 1289–1298

[9] Tsai H H and Chang Y C 2018 Facial expression recognition using a combination of multiple facial features and support vector machine Soft Computing vol 22 p 4389–4405

[10] Donato G Bartlett M Hager J Ekman P and Sejnowski T 1999 Classifying facial actions. IEEE Transaction on Pattern Analysis and Machine Intelligence vol 21 no 10 p 974–989

[11] Pantic M and Rothkrantz L 2000 Expert system for automatic analysis of facial expression Image Vis. Comput vol 18 no 11 p 881–905