SUPPORT VECTOR MACHINE WITH RADIAL BASIS FUNCTION FOR FACIAL EMOTION VALENCE RECOGNITION

F. Ludyma Fernando[1], Dr. S. John Peter [2]

[1][2] Department of Computer Science, St.Xavier's College(Autonomous),
Affliated to Manonmaniam Sundaranar University, Tirunelveli, 627012, India

Abstract

The affective quality called Valence refers to the intrinsic goodness (positive valence) or badness (negative valence) of an event, object, or situation. For this purpose, a model for classification and characterization of emotions have been developed which is discussed in this paper. In this model, the images are smoothened using an Average Filter and are first identified through a Convolutional Neural Network which uses the ReLU activation function. Then, the valence is classified using a Support Vector Machine (SVC) classifier, which uses a Radial Basis Function (RBF) kernel. For this reason, the emotions are labeled according to their nature. The positive emotions are labeled 1 (inclusive of the neutral emotion) and the negative emotions are labeled as 0. The images from the FER 2013 dataset is used for Valence Recognition and is given via a RBF Kernel in a SVM, which classifies whether the emotion recognized is positive or negative. The haarcascade algorithm is implemented to detect the face. In this paper, the 7 human emotions (happiness, surprise, fear, anger, fear, disgust, sadness and neutral) have been identified and their valence recognized.

Keywords: CNN, Emotion Recognition, Support Vector Machine, Radial Basis Function

Introduction

A human face plays a significant role in expressing emotions of different kinds[1]. A facial expression can identify the state of mind of an individual. So, in order to identify an emotion, the face has to be detected. Face detection is the important step in emotion detection[2]. It removes the part of the image that is not relevant. Then the emotions are recognized and characterized. The images from the FER2013 dataset are used for the training and the testing purposes.

Before using the data, it is important to go through a series of steps called preprocessing. This makes the data easier to handle. The dataset (FER 2013) is stored in csv (Comma Separated Values) format. Each row in a csv file denotes an instance. Every instance has two column attributes. The pixels of the image and its corresponding target label. There are a total of 35K images, distributed equally across the emotions. The images are then cropped into 48 x 48 grayscale images. The target labels are integers encoded as in **Table 1**.

Value	Emotion
0	Anger
1	Disgust
2	Fear
3	Happiness
4	Sad
5	Surprise
6	Neutral

Table 1 : The Encoded values for each emotion.

The preprocessing also includes the splitting of the data and standardizing it. The dataset is split into the training and the testing sets and is standardized by rescaling it to have a mean of 0 and a standard deviation of 1. This transformation makes the data ready for the removal of noise by the use of an Average Filter. The preprocessed images are given as input to a Convolutional Neural Network, which uses ReLU[3] activation to classify the seven emotions, using the training dataset. As the scope of this paper is about the Support Vector Machine that is used for the actual classification, the rest of the paper deals with the same.

Support Vector Machines with the RBF Kernel

The SVM model discussed in this paper is for a non linearly separable data, because in the recognition of an emotion, the trained data is a set of non linearly separable data.

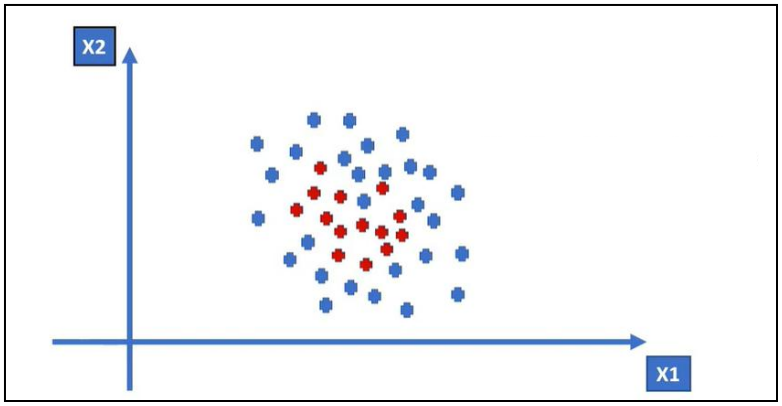


Fig.1. Non Linearly Separable Data

The SVM[4] is a supervised learning model with associated learning algorithms that analyze data for classification. It is one of the robust prediction methods. Classifying such data as ours, there is need for a kernel. The one used in this model is the Gaussian RBF Kernel. The Radial Basis Function(RBF)[5], as it is commonly called, helps the SVM in classifying non linearly separable data. The RBF kernel has been used with other prediction models such as Principal Component Analysis[6] and has been found to work well. Hence in this model, the RBF kernel is used with a SVM. The main function of the kernel is to convert the 2D data into a higher dimension, so as to improve the classification process. The hyper plane is the line which runs through the data. In the case of a linearly separable data, this hyper plane concludes the classification as it passes through the distinguishing part of the data clusters. In the case of non-linearly separable data, the hyper plane has to pass through the data, but by itself it cannot classify the data on the 2D plane. So, there is a need for the RBF kernel (∅), which translates the whole set of data to be classified, into a higher dimension (probably a 3D data).

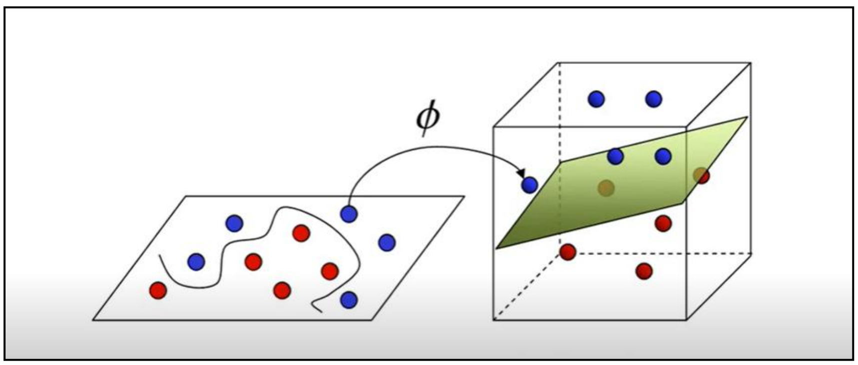


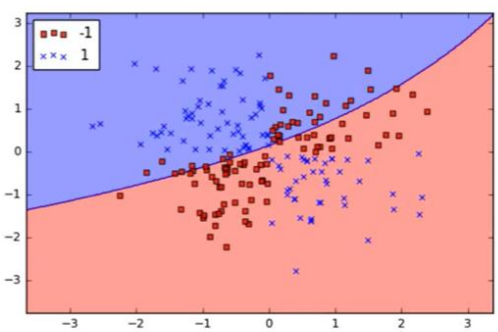
Fig.2.: Data in the Input Space (2D) Vs the Data in the Feature Space (3D)

As shown, in the **Fig.2**. the kernel tries to separate the two classes in the input space without showing any detail of how the classifying is done at the nD space. It actually remains in the 2D space and does the classification as it were in a higher dimensional space. This is what that is called the Kernel Trick. This reduces the processing time and makes the computation simple. The code required to model a SVM with an RBF is given in **Fig. 3**.

```
# Create a SVC classifier using an RBF kernel
svm = SVC(kernel='rbf', random_state=0, gamma=.01, C=1)
# Train the classifier
svm.fit(X_xor, y_xor)
```

Fig. 3. SVM with RBF (code)

‘gamma’ is a parameter of the RBF kernel and can be thought of as the ‘spread’ of the kernel and therefore the decision region. When gamma is low, the ‘curve’ of the decision boundary is very low and thus the decision region is very broad. When gamma is high, the ‘curve’ of the decision boundary is high, which creates islands of decision-boundaries around data points.



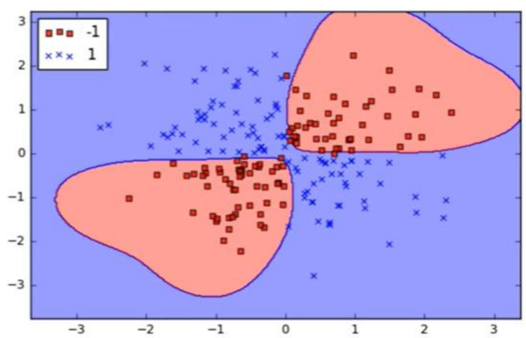
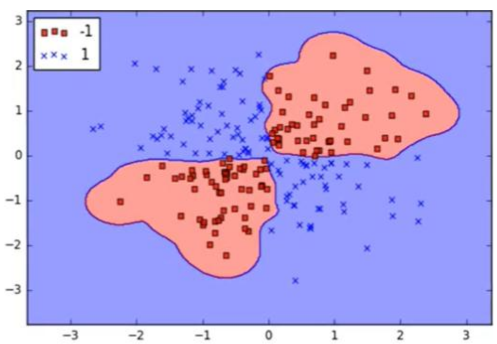


Fig.4. gamma = 0.01 **Fig.5. gamma = 1.0**



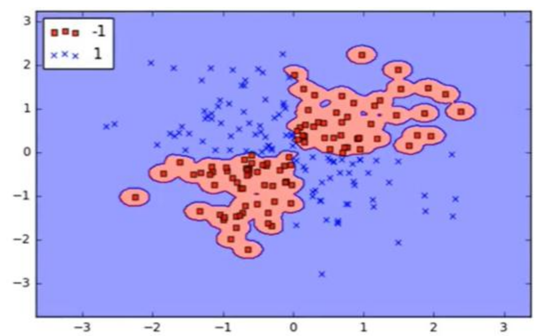
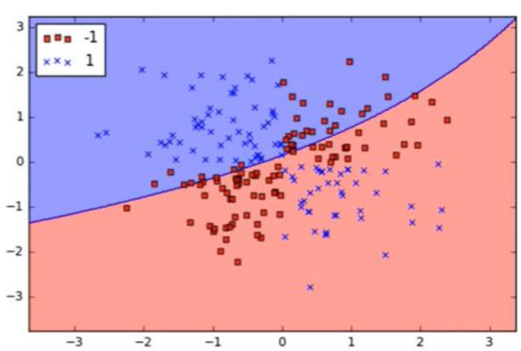


Fig.6. gamma = 10 **Fig.7. gamma = 100**

The states of the data during classification depending on the gamma values and their corresponding curves are depicted in the **Figs 4-7**. C is a parameter of the SVC learner and is the penalty for misclassifying a data point. When C is small, the classifier is okay with misclassified data points (high bias, low variance). When C is large, the classifier is heavily penalized for misclassified data and therefore bends over backwards avoid any misclassified data points (low bias, high variance). The data classification with the ‘c’ parameter is shown in **Figs.8-12.**



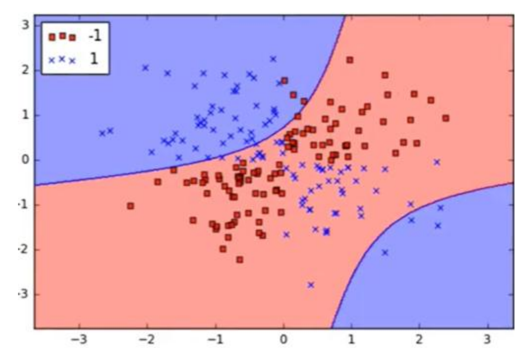
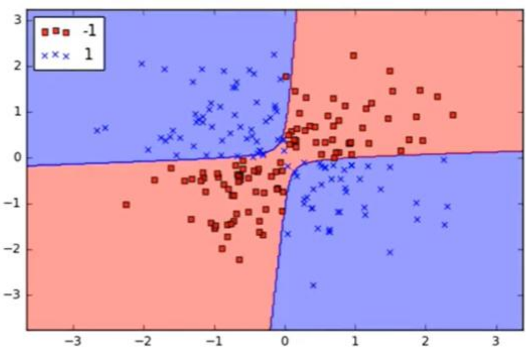
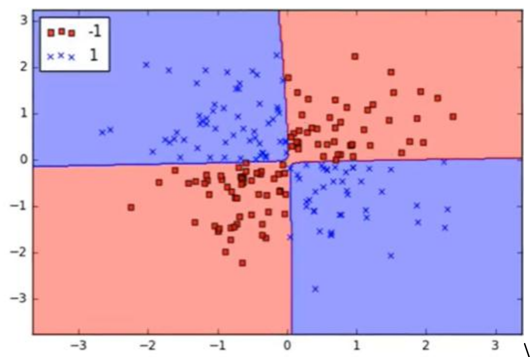


Fig.8. c = 1

Fig.9.c = 10





\

Fig.10. c = 1000

Fig.11. c = 10000

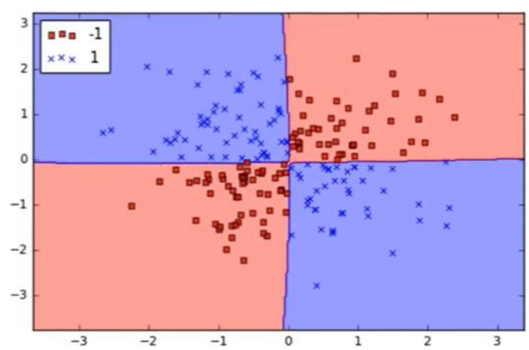


Fig.12. c = 100000

The Proposed Model:

As mentioned in the beginning, the emotions are recognized first and then their valence is calculated. To recognize emotions a Convolutional Neural Network is trained to recognize the emotions and the Support Vector Machine is implemented in the end to calculate the valence of the emotion. **Fig.13** shows the process flow of the proposed model.

ISSN 2515-8260 Volume 09, Issue 04, 2022

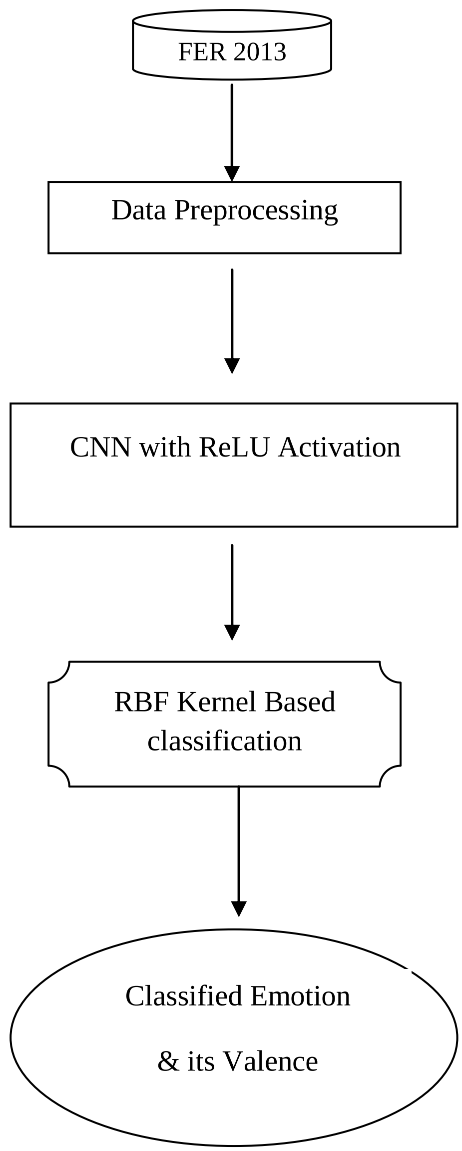


Fig . 13. The Proposed Model

The data is from the FER2013 dataset. It is preprocessed using an Average Filter[9] to remove noise. The images are converted into 48 x 48 sized, grayscale images. The training of the images are done using a CNN which extracts about 512 features of the image and adjusts the weights using the Rectified Linear Unit (ReLU). It has seven output layers which correspond to the seven emotions to be recognized. The Support Vector Machine with the RBF kernel is applied in the end to characterize the emotions. At first, the valency for emotion is defined in a csv file.

Value	Emotion	Valence
0	Anger	0
1	Disgust	0
2	Fear	0
3	Happiness	1
4	Sad	0
5	Surprise	1
6	Neutral	1

Table 2. The Valence of the Emotions

The positive emotions, Happiness, Surprise and Neutral (Neutral considered positive) have their valence as 1, while the negative emotions such as Anger, Fear, Disgust and Sad have the valence as 0. While the CNN recognizes the emotions[8], the SVM with the RBF kernel is trained with this csv file to characterize the emotions as having positive or negative valence.

The face is that which expresses the emotions to be identified. For this reason, the face needs to be detected from the image. Face detection is done using a haarcascade function[7] to display bounding box over the detected faces.

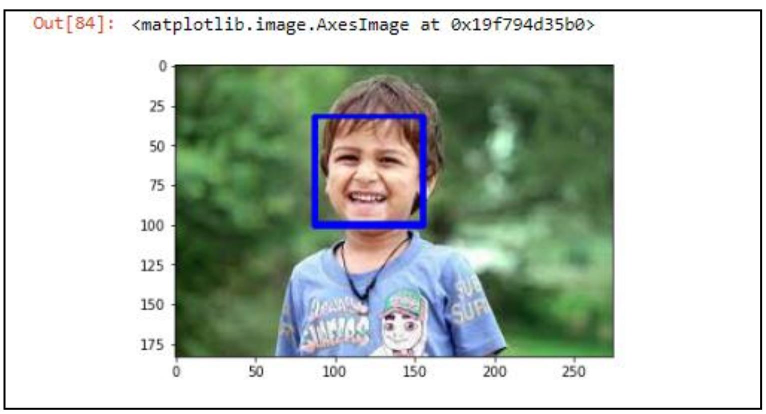


Fig. 14. Face Detection

Now, the area of the image under consideration, i.e. the face, should be cropped so that the features can be extracted.

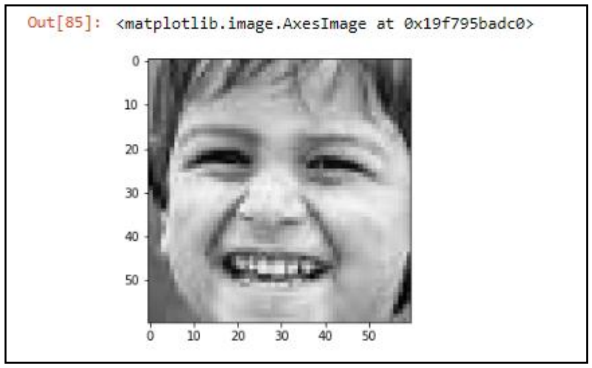


Fig. 15. The area of the image under consideration

Now, the code used to process the image is run and hence the emotion identified and its valence characterized

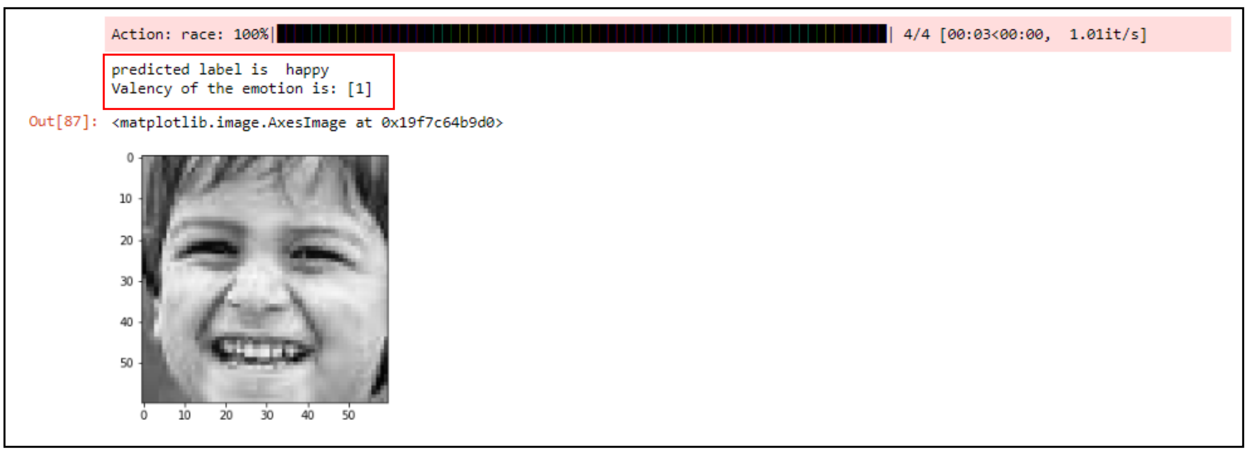


Fig. 16. The Result.

Conclusion:

The proposed model used, has the CNN running on the ReLU activation function and extracts over 512 features. Though the accuracy of characterizing a single emotion is significant, the model could be transformed to identify compound emotions like happily surprised etc. This model works for single defined emotions only. The SVM with RBF has worked well in characterizing the emotions as positive and negative. The scope of this model can be extended to calculating the valence of compound emotions with a higher accuracy.

References

[1] Li, J., Jin, K., Zhou, D., Kubota, N., & Ju, Z. (2020). Attention mechanism-based CNN for facial expression recognition. Neurocomputing, 411, 340-350.

[2]Kusuma, G. P., Jonathan, A. P. L., & Lim, A. P. (2020). Emotion recognition on fer-2013 face images using fine-tuned vgg-16. Advances in Science, Technology and Engineering Systems Journal, 5(6), 315-322.

[3] Weng, L., Zhang, H., Chen, H., Song, Z., Hsieh, C., Daniel, L., Boning, D. & Dhillon, I.. (2018). Towards Fast Computation of Certified Robustness for ReLU Networks.Proceedings of the 35th International Conference on Machine Learning, in Proceedings of Machine Learning Research 80:5276-5285

[4]Zhou, Liqin, et al. "Emerged human-like facial expression representation in a deep convolutional neural network." *Science advances* 8.12 (2022): eabj4383.

[5] Zhang, J., Yin, Z., Chen, P., & Nichele, S. (2020). Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review. Information Fusion, 59, 103-126

[6]S. Thakur, J. K. Sing, D. K. Basu, M. Nasipuri and M. Kundu, "Face Recognition Using Principal Component Analysis and RBF Neural Networks," 2008 First International Conference on Emerging Trends in Engineering and Technology, 2008, pp. 695-700, doi: 10.1109/ICETET.2008.104.

[7] Mehtab, Sidra & Sen, Jaydip. (2020). Face Detection Using OpenCV and Haar Cascades Classifiers. 10.13140/RG.2.2.26708.83840.

[8] Carlos Busso, Zhigang Deng, Serdar Yildirim, Murtaza Bulut, Chul Min Lee, Abe Kazemzadeh, Sungbok Lee, Ulrich Neumann, and Shrikanth Narayanan. 2004. Analysis of emotion recognition using facial expressions, speech and multimodal information. In Proceedings of the 6th international conference on Multimodal interfaces (ICMI '04). Association for Computing Machinery, New York, NY, USA, 205–211. https://doi.org/10.1145/1027933.1027968

[9]Kanchanadevi and P.R. Tamilselvi(2020). Preprocessing using image filtering method and techniques for medical image compression techniquesb,Department of Computer Science, Periyar University, India,Department of Computer Science, Government Arts and Science College, Komarapalayam, India

ISSN 2515-8260 Volume 09, Issue 04, 2022

[10] Keiron O’Shea 1 and Ryan Nash 2 1 Department of Computer Science, Aberystwyth University, Ceredigion, SY23 3DB keo7@aber.ac.uk 2 School of Computing and Communications, Lancaster University, Lancashire, An Introduction to Convolutional Neural Networks, arXiv:1511.08458v2 [cs.NE] 2 Dec 2015