Privacy Preserving Facial Emotion Recognition using Federated Learning

Rekha V S
rekhavsrh@gmail.com
5thYear, MSc Data Science(Integrated),
Department of Computing,
Coimbatore Institute of Technology,
Coimbatore.

Mani Barathi SP S
manibarathids2018@gmail.com
5thYear, MSc Data Science(Integrated),
Department of Computing,
Coimbatore Institute of Technology,
Coimbatore.

Dr.D.Sudha Devi MCA.,M.Phil.,Ph.D.,
sudhadevi.cit@gmail.com
Associate Professor,
Department of Computing,
Coimbatore Institute of Technology,
Coimbatore.

Dr.M.Srividya MCA.,M.Phil.,Ph.D.
srividhya@cit.edu.in
Assistant Professor,
Department of Computing,
Coimbatore Institute of Technology,
Coimbatore.

***Abstract* — With rise in growth of automation, the need for security rises as users data is valuable than an asset. With the widespread use of biometrics, there comes a privacy issues, which would cause a threat to the individual. The face recognition system collects the face data from users. Face data is usually unique and irreplaceable and once leaked, it will cause great damage to the user's privacy. However, while training, data becomes more prone to get leaked. Training a face classification based model requires large pool of private data. In order to protect the data from any vulnerable threats, federated learning based approach is carried out. In federated approach, image embedding is extracted from local servers (offline servers) and model training happens at main server (online). This study proposes to federated learning approach to recognize the face and detect the emotional state of the users.**

Keywords : federated learning, servers, vulnerable threat, face recognition, emotion recognition.

I. INTRODUCTION

One of the main communication of humans happens through face. Facial expression plays an important role to identify the current emotional well-being of the people. In real-life facial expressions are influenced by various characteristics like age, gender, personality traits and more. In order to handle wide plethora of cases, Deep learning model is trained to classify the facial expressions. These models are trained on large datasets but they suffer from bias in the dataset and fails to generalize most of the cases. This leads to the degradation of the performance. Hence there is a need of real scenario data. Due to privacy and sensitivity of the facial expression real case data it has become a bottle neck for research in this area. Data protective steps were incorporated to eliminate security threat of users face

image dataset. Wide variety of open sourced data were removed from the internet due to security threat. To overcome all these threats and to train an effective model federated learning approach is carried out. Federated Learning (FL) is an effective way of training models on private data as well as in private machine. This paper first proposes a face privacy protection scheme based on secure multi-party computing which establishes a face recognition privacy protection model.

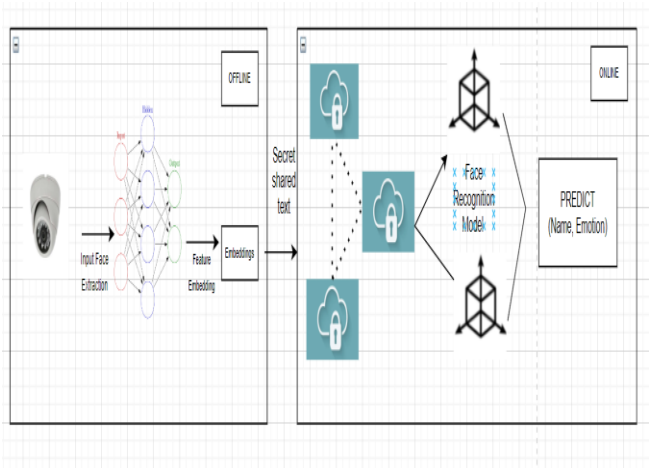

Figure 1: offline and online client

Face recognition can be performed by calculating the sensitive data from multiple resources. On the same hole, the privacy of face data can be guaranteed.

This paper is divided into two phases (online and offline) as in figure 1. In the first phase the face embedding's where extracted and on the next phase privacy preserving face recognition is carried out. The face feature embedding's are extracted by offline client. First, preprocessing of the face images where done, then a deep learning model based on the Siamese neural network as in figure 2 is used. This model process the face features to extract low-dimensional face representations (face embeddings). Prediction and training of face recognition

model is done on online server. There are 2 servers, in the private server model is trained in cloud through joint multifaceted face embedding data. The parameters of the face recognition privacy protection model after training are stored in the form of secret shared secret text. Later, the trained model is used to recognize the face by combining the face embedding data to be recognized. The recognition result will be in the form of secret ciphertext which is then reconstructed and returned to the client. The correctness and security of the scheme is analyzed, and

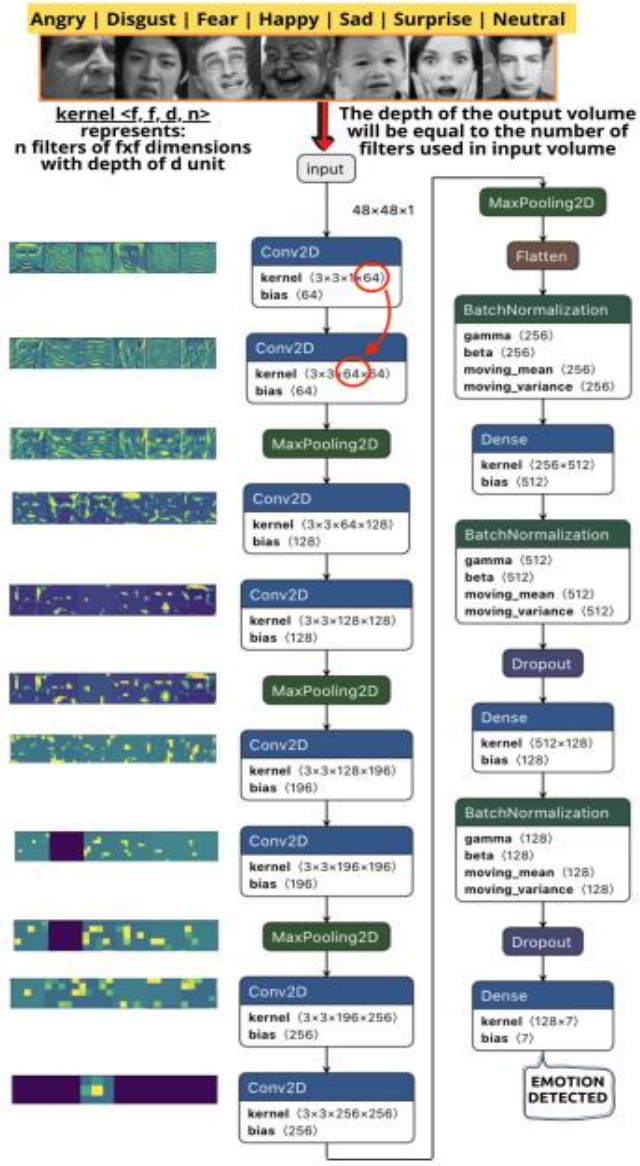

Figure 2: Face Emotion recognition and detection

implemented by conducting various experiments on the scheme. The results show that the face recognition privacy protection not only reliable and secure and also has the advantages of light weight, high accuracy, and computationally efficiency.

II. Related works

Hard et al. introduced the first application of federated learning on the next word prediction. This project relates the

sentences and predicts the next upcoming word. CIFB bases model is trained and best accuracy is obtained. This accuracy outperforms all other previously developed model based on word prediction.

Hartmann et al. used federated learning based approach to improve word suggestions with ranking based approach. Search bar based ranking approach is incorporated in this paper. They built a system and proposed SVM loss to support federated learning and to optimize the ranking of suggestions while preserving the users' privacy.

The above examples performed training of the language models for prediction and recommendation purposes. But this paper proposes the federated learning strategies for a face recognition model.

III. MODEL BUILDING

In this section, how the emotions of the person are recognized with federated learning is dealt. Simple CNN model is trained to classify and recognize the emotions of the face. The face recognition model, as the name suggests detects the emotion of a person based on facial expressions captured. The speech recognition model classifies the mood of the person. In part1 of the model phase, images gets extracted and different alignment and augmentation is done to increase the data quality and quantity. Images are resized into 48 * 48. Then data is convoluted with max pooling and Relu activation function is used. Finally the convoluted layer is flattend and depending on decision tree results the final emotion of the image is obtained. The approach suggested, using preprocessing techniques for image enhancement, then face detection, and finally emotion detection. With the help of the proposed model, the authors were able to reach a test accuracy of 60.12% and validation accuracy of 89.78% on the FER2013 data set. It is a traditional preprocessing step in various facial recognition tasks. The initial step is to identify the face from the image and eliminate the background.

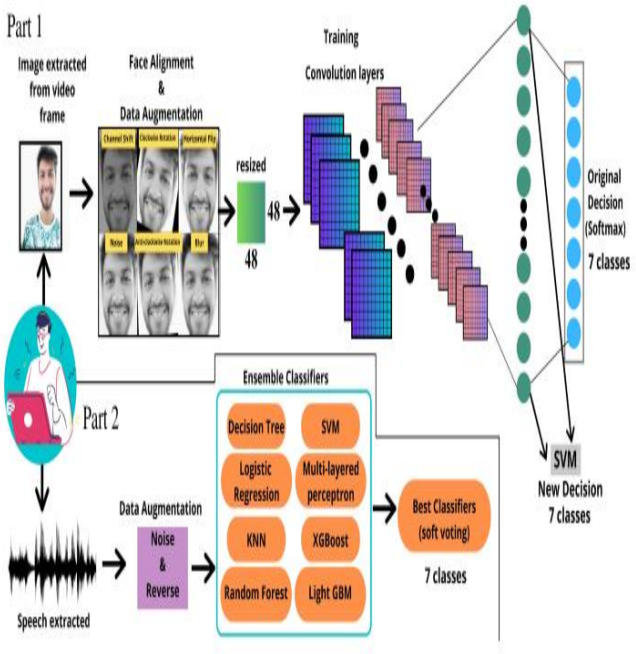

Figure 3 : Model Architecture – Federated learning

IV. EXPERIMENTATIONS

A. DATASETS

In this paper two datasets were used from FER2013 and AffectNet. FER2013 was presented during the ICML 2013 Challenges in Representation Learning. This dataset is gathered from Google images by using Google search API. All images are in grayscale and have been resized to 48×48 pixels after rejecting wrongfully labeled frames and adjusting the cropped region. The data set consists of 28709 training images and 3589 validation images. AffectNet includes both the categorical and dimensional models as in figure 3and it is one of the largest datasets for facial affect in still images. By using 1250 emotion-related tags in six different languages, that are English, German, Spanish, Portuguese, Arabic, and Farsi, the dataset is collected. There are more than one million images with faces in the dataset and also it contains the extracted facial landmark points.

B. HYPERPARAMETER TUNING

The hyper parameters for the classifiers are tuned on a cross-validation set. The cross-validation set is created by using a k-fold cross validation module from the scikit-learn library. The value of "k" is chosen as 5 for all used classifiers in the proposed work. So, the training set is divided into five equal partitions, and training is done of four partitions, taking the left partition for the validation. This activity is performed five times, such that every portion becomes cross-validation once. The hyper parameter tuning was performed using a randomized search and grid search. The optimal hyper parameters with highest validation accuracy is selected.

V. CONCLUSION

Computing capabilities, and the advancements in ML and DL opened up many possibilities for various applications. Traditional cloud-based ML/DL methods require the data to be centralized in a cloud server or a data center. The most powerful, natural, and universal signs for human beings to convey their emotional states is through facial emotions. Also, speech is the most significant mode of communication among human beings and a potential method for emotion detection by using sensors. In this paper, we solved the problem of automatic emotion recognition for an individual. The proposed work shows the feature extraction techniques to extract features from an image. The proposed scheme detects human emotions using extracted facial and speech features. The emotion of an individual is categorized into seven different classes based on the output of both classifiers. The proposed facial and speech emotion recognition classifier gives an accuracy of 71.64% and 85.04% respectively. It can detect the state of depression of an individual at an earlier stage, and based on that the classifier will recommend the individual to meet a counselor. For the individual's data privacy, the concept of FL is used, which allows seamless data sharing and data access while still being secure.

VI. REFERNCES

[1] Ekman, P. & Keltner, D. (1997). Universal facial expressions of emotion: An old controversy and new findings. In Segerstråle, U. C. & Molnár, P. (Eds.), Nonverbal communication: Where nature meets culture (pp. 27-46). Mahwah, NJ: Lawrence Erlbaum Associates.

[2] Matsumoto, D. & Kupperbusch, C. Idiocentric and allocentric differences in emotional expression, experience, and the coherence between expression and experience. Asian Journal of Social Psychology

(3), pp. 113-131 (2001).I.S. Jacobs and C.P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G.T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271-350.

[4] Lucey, P., Cohn, J.F., Kanade, T., Saragih, J. Ambadar, Z. The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression. IEEE Computer Society Conference CVPRW (2010) [4] Zhang, Z. Feature-based facial expression recognition: Sensitivity analysis and experiments with a multilayer perceptron. International Journal of Patten Recognition and Artificial Intelligence 13 (6):893-911 (1999).

[5] Michael J. Lyons, Shigeru Akemastu, Miyuki Kamachi, Jiro Gyoba. Coding Facial Expressions with Gabor Wavelets, 3rd IEEE International Conference on Automatic Face and Gesture Recognition, pp. 200-205 (1998)

[6] Shan C, Gong S, McOwan PW. Facial expression recognition based on local binary patterns: a comprehensive study. Image Vis Comput. 27(6):803–816 (2009)

[7] Carcagnì P, Del Coco M, Leo M, Distante C. Facial expression recognition and histograms of oriented gradients: a comprehensive study. SpringerPlus. 4:645. (2015)