A R DIGITECH
International Journal Of Engineering, Education And Technology (ARDIJEET)
www.ardigitech.in ISSN 2320-883X, VOLUME 13 ISSUE 01, 01/01/2025 pub: 26/02/2025

Machine Learning Techniques For Emotion Recognition From Facial Expressions

[1]Mohini Narkhede, [2]Prof. Pallavi P. Rane, [3]Prof. Nilesh N. Shingne,
[1]mohininarkhede1989@gmail.com, [2]koltepallavi200@gmail.com, [3]shingne.nilesh236@gmail.com
[2]Assistant professor, Rajarshi Shahu College of Engineering, Buldhana, Maharashtra
[3]Assistant professor, Sanmati Engineering College, Washim, Maharashtra

Abstract: Social media's rise has led users to express themselves through images and text, making multimodal content the fastest-growing type. Since user posts often contain sentiment, multimodal sentiment analysis has become a key research area. Existing methods typically extract and combine text and image features separately, often overlooking their interaction. This project introduces a new model that addresses this. It first reduces noise in text and enhances image feature extraction. Then, using an attention-based fusion mechanism, text and image features are symmetrically learned from each other. Finally, these fused features are used for sentiment classification. Experiments on sentiment datasets validate the model's performance. [1]

Keywords: Multimodal, Analysis, Extract, Symmetrical, Validate

1. Introduction

Image classification aims to categorize images, identifying whether they belong to a specific class (e.g., containing certain objects like cars, or representing scenes like cities). A binary classifier can be trained using labelled images, but this process is time-consuming, especially for diverse categories. While large labelled datasets are used in research, this approach is often impractical for applications like personal photo organization. Consequently, self-supervised or naturally-supervised learning has gained traction. These methods leverage non-visual signals linked to images for feature learning. The abundance of websites with images and associated annotations offers a natural source of supervision. Unlike previous image-text embedding methods, the goal here is to learn general, discriminative features in a self-supervised manner, without relying on annotated datasets. Recent research has explored joint image and text embeddings. Combining different data types allows for simultaneous learning, attracting both general and applied research. One approach, the Deep Visual-Semantic Embedding Model, learns to predict Word2Vec representations of image labels instead of ImageNet classes. By using the distributional semantics of a text corpus associated with each image, the model can infer unseen concepts. Even with errors, the model's semantically relevant predictions are valuable, as these errors often generalize to classes outside the labelled training set.

2. Project Objectives:

- To analysed the techniques used for sentiment data and also demonstrate that how individual model works. Extracting the sentiments from different input modes is achieved by different classifying techniques.
- To find out different input modes to generate the model for analysis.
- To analysed challenges of our proposed system and integration of different modes and its effect on emotion reorganization system

3. Literature Review

Alejandra Sarahi Sanchez-Moreno et al state that facial recognition is fundamental for a wide variety of security systems operating in real-time applications. Recently, several deep neural networks algorithms have been developed to achieve state-of-the-art performance on this task. The present work was conceived due to the need for an efficient and low-cost processing system, so a real-time facial recognition system was proposed using a combination of deep learning algorithms like FaceNet and some traditional classifiers like SVM, KNN, and RF using moderate hardware to operate in an unconstrained environment. Generally, a facial recognition system involves two main tasks: face detection and recognition. [1]

HANG DU, et all state that Face recognition (FR) is an extensively studied topic in computer vision. Among the existing technologies of human biometrics, face recognition is the most widely used one in real-world applications. With the great advance of deep convolutional neural networks (DCNNs), the deep learning based methods have achieved significant improvements on various computer vision tasks, including face recognition. In this survey, we focus on 2D image based end-to-end deep face recognition which takes the general images or video frames as input, and extracts the deep feature of each face as output. We provide a comprehensive review of the recent advances of the elements of end-to-end deep face recognition. Specifically, an end-to-end deep face recognition system is composed of three key elements: face detection, face alignment, and face representation. [2]

Madan Lal et al state that with the rapid growth in multimedia contents, among such content face recognition has got much attention especially in past few years. Face as an object consists of distinct features for detection; therefore, it remains most challenging research area for scholars in the field of computer vision and image processing. [3]

4. Proposed Methodology:

Proposed Methodology for face emotion recognition using machine learning techniques

A proposed methodology for face emotion recognition using machine learning techniques typically involves several key stages. Here's a breakdown of a common approach:

4.1. Data Acquisition and Preprocessing:

- **Dataset Collection:** Gather a diverse dataset of facial images labeled with corresponding emotions (e.g., happy, sad, angry, surprised, fearful, disgusted, neutral). Publicly available datasets like FER2013, CK+, RAF-DB, and AffectNet are often used. Consider factors like variations in pose, lighting, age, gender, and ethnicity to ensure robustness.
- **Face Detection:** Employ a face detection algorithm (e.g., Haar cascades, MTCNN, YOLO) to locate and crop the facial region from each image. This step isolates the area of interest for emotion recognition.
- **Preprocessing:** Perform preprocessing steps to standardize the data and improve model performance. This may include:
    - **Resizing:** Resize all face images to a consistent size.
    - **Normalization:** Normalize pixel values to a specific range (e.g., 0-1) to reduce the impact of lighting variations.

- **Data Augmentation:** Apply transformations like rotation, scaling, flipping, and small color adjustments to artificially increase the size of the dataset and improve model generalization. This helps prevent overfitting.

4.2. Feature Extraction:

- **Manual Feature Extraction (Traditional Methods):** Historically, methods like Local Binary Patterns (LBP), Histogram of Oriented Gradients (HOG), or facial landmarks were used to extract relevant features from the face images. These features represent facial expressions.
- **Automatic Feature Extraction (Deep Learning):** Convolutional Neural Networks (CNNs) have become the dominant approach for feature extraction in facial emotion recognition. Pre-trained CNNs (e.g., ResNet, VGG, EfficientNet) can be used as feature extractors, or custom CNN architectures can be trained. CNNs automatically learn hierarchical features from the raw pixel data, eliminating the need for manual feature engineering.

4.3. Emotion Classification:

- **Traditional Machine Learning Classifiers:** If manual features are extracted, classifiers like Support Vector Machines (SVM), Random Forest, or k-Nearest Neighbors (k-NN) can be used to map the extracted features to specific emotions.
- **Deep Learning Classifiers:** When using CNNs for feature extraction, the final layers of the CNN are typically used for classification. A fully connected layer followed by a softmax activation function is common for multi-class emotion classification. Fine-tuning the pre-trained CNN on the emotion dataset is a common practice.

4.4. Model Evaluation and Refinement:

- **Evaluation Metrics:** Evaluate the performance of the trained model using metrics like accuracy, precision, recall, F1-score, and confusion matrices. Consider using cross-validation techniques to ensure robust evaluation.
- **Hyperparameter Tuning:** Optimize the model's hyperparameters (e.g., learning rate, batch size, network architecture) using techniques like grid search or random search to improve performance.
- **Model Selection:** Compare the performance of different models and choose the best one based on the evaluation metrics.

4.5. Deployment (Optional):

- Integrate the trained model into a real-world application, such as a user interface for emotion-based interaction or a system for analyzing customer feedback.

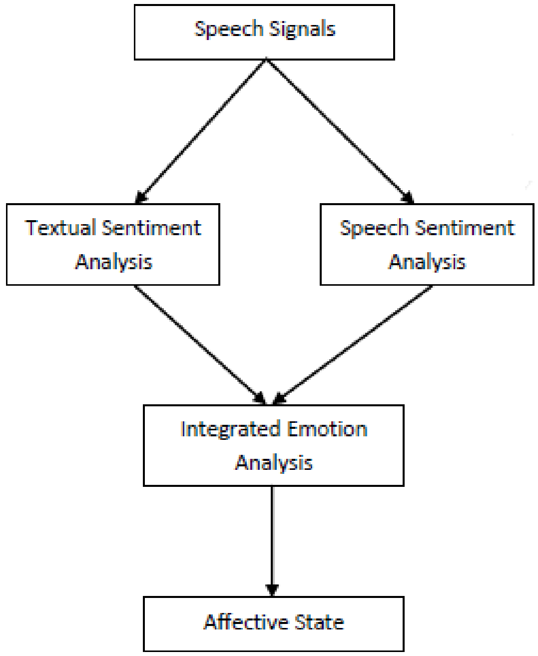

Fig. 1 Flow of the proposed system

6. Algorithm used

Support Vector Machines (SVMs) are a popular supervised learning algorithm used for both classification and regression, though primarily for classification. SVMs aim to find the optimal decision boundary (hyperplane) that separates data points into classes, enabling accurate classification of new data. The algorithm identifies extreme points (support vectors) that define the hyperplane. SVMs can be linear (for linearly separable data) or non-linear (for non-linearly separable data). In essence, SVMs map data points to a high-dimensional space where an optimal separating hyperplane is found.

Haar Cascades are a machine learning-based object detection approach that uses a large number of positive (images containing the object) and negative (images not containing the object) images for training. The algorithm operates in four stages:

- Haar Feature Calculation: Haar features are calculations performed on adjacent rectangular regions within a detection window. These features capture variations in pixel intensities.
- Integral Image Creation: Integral images accelerate Haar feature calculation by creating sub-rectangle references, reducing redundant computations.
- Adaboost Training: Adaboost selects the best Haar features and trains classifiers (weak learners) using them. It combines these weak learners into a strong classifier.
- Cascade Classifier Implementation: A cascade of stages, each containing weak learners, is used. Stages are designed to quickly reject negative samples, focusing on regions of interest. The cascade structure prioritizes minimizing false negatives.

Face recognition technology has advanced significantly. From its early forms to widespread use in smartphones and other applications, it has become a commonplace technology. It's used in various applications, from mobile phone security to crime prevention and travel facilitation.

6. Application

6.1 Prevent Retail Crime

Face recognition is currently being used to instantly identify when known shoplifters, organized retail criminals or people with a history of fraud enter retail establishments. Photographs of individuals can be matched against large databases of criminals so that loss prevention and retail security professionals can be instantly notified when a shopper enters a store that prevents a threat. Face recognition systems are already radically reducing retail crime. According to our data, face recognition reduces external shrink by 34% and, more importantly, reduces violent incidents in retail stores by up to 91%.

6.2 Unlock Phones

A variety of phones including the latest iPhone are now using face recognition to unlock phones. This technology is a powerful way to protect personal data and ensure that, if a phone is stolen, sensitive data remains inaccessible by the perpetrator.

6.3 Smarter Advertising

Face recognition has the ability to make advertising more targeted by making educated guesses at people's age and gender. Companies like Tesco are already planning on installing screens at gas stations with face recognition built in. It's only a matter of time before face-recognition becomes an omni-present advertising technology.

6.4 Find Missing Persons

Face recognition can be used to find missing children and victims of human trafficking. As long as missing individuals are added to a database, law enforcement can become alerted as soon as they are recognized by face recognition—be it an airport, retail store or other public space. In fact, 3000 missing children were Discovered In Just Four Days Using Face Recognition In India!

6.5 Help the Blind

Listerine has developed a ground breaking facial recognition app that helps the blind using face recognition. The app recognizes when people are smiling and alerts the blind person with a vibration. This can help them better understand social situations.

6.6 Protect Law Enforcement

Mobile face recognition apps, like the one offered by FaceFirst, are already helping police officers by helping them instantly identify individuals in the field from a safe distance. This can help by giving them contextual data that tells them who they are dealing with and whether they need to proceed with caution. As an example, if a police officer pulls over a wanted murderer at a routine traffic stop, the officer would instantly know that the suspect may be armed and dangerous, and could call for reinforcement.

8. Conclusion:

Face emotion recognition and sentiment analysis are active research areas in machine learning. Despite existing work, sentiment analysis remains challenging due to cultural influences, linguistic variations, and contextual differences, largely due to the unstructured nature of language. A key challenge lies in incorporating multiple modalities, such as audio and video, alongside text to improve accuracy. Textual emotion classification relies on polarity and lexicon intensity, while audio classification uses prosodic features, and video classification uses postures and gestures. Integrating these modalities can enhance accuracy. Future research should focus on these multimodal challenges, reflecting the shift from unimodal to multimodal approaches.

References:

[1] Efficient Face Recognition System for Operating in Unconstrained Environments Alejandra Sarahi Sanchez-Moreno 1, Jesus Olivares-Mercado 1 , Aldo Hernandez-Suarez 1 , Karina Toscano-Medina 1, Gabriel Sanchez-Perez 1 and Gibran Benitez-Garcia 2,* J. Imaging 2021, 7, 161. https://doi.org/10.3390/jimaging7090161

[2] The Elements of End-to-end Deep Face Recognition: A Survey of Recent Advances Hang Du, Hailin Shi, Dan Zeng, Xiao-Ping Zhang, Tao Mei Accepted for publication in ACM Computing Surveys Computer Vision and Pattern Recognition (cs.CV)

[3] The Elements of End-to-end Deep Face Recognition: A Survey of Recent Advances HANG DU, HAILIN SHI, DAN ZENG, XIAO-PING ZHANG, TAO MEI, JD AI arXiv:2009.13290v4 [cs.CV] 27 Dec 2021

[4] (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 9, No. 6, 2018 42 | P a g e www.ijacsa.thesai.org Study of Face Recognition Techniques: A Survey Madan Lal, Kamlesh Kumar Department of Computer Science Sindh Madressatul Islam University, Karachi, Sindh, Pakistan Rafaqat Hussain Arain, Abdullah Maitlo, Sadaquat Ali Ruk, Hidayatullah Shaikh

[5] Research on Face Recognition Classification Based on Improved GoogleNet Zhigang Yu, Yunyun Dong, Jihong Cheng, Miaomiao Sun , and Feng Su Hindawi Security and Communication Networks Volume 2022, Article ID 7192306, 6 pages https://doi.org/10.1155/2022/7192306

[6] Boiy, E. Hens, P., Deschacht, K. &Moens, M.F., "Automatic Sentiment Analysis in Online Text", In Proceedings of the Conference on Electronic Publishing(ELPUB-2007).

[7] J. Wiebe, T. Wilson, and C. Cardie. "Annotating expressions of opinions and emotions in language", Language Resources and Evaluation, 2005.

[8] Scott Brave and Clifford Nass, Emotion in Human- Computer Interaction. Retrieved from http://lrcm.com.umontreal.ca/dufresne /COM7162/EmotionHumanInteraction.pdf

[9] S. V. Bo Pang, Lillian Lee, "Thumbs up? Sentiment classifcation using machine learning techniques", Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), ACL, pp. 79– 86,July 2002.

[10] Pravesh Kumar Singh and Mohd Shahid Husain "Methodological study of opinion mining and sentiment analysis techniques" International Journal on Soft Computing (IJSC) Vol. 5, No. 1, February 2014