Combining Heterogeneous Features for Facial Action Recognition Using Multikernel Learning

Bombale Girisha Ramhari, Dr. Rais Abdul Hamid Khan
Department of Computer Science and Engineering,
Dr. A. P. J. Abdul Kalam University, Indore (M.P.) - 452016, India

Abstract:
Facial action recognition (FAR) is an important but challenging task in computer vision due to the high variability of facial expression. To capture the complex facial features, diverse face representations need to be employed. In this paper, we propose a novel multikernel learning algorithm to effectively integrate heterogeneous features for FAR. We employ seven types of facial features, including Local Binary Patterns (LBP), Histograms of Oriented Gradients (HOG), convolutional neural network (CNN) features, Gabor Wavelets, and so on, and learn a combination of these different feature types in a unified way. Two single kernel support vector machine (SVM) models and a multiple kernel learning SVM model are compared in our experiments. By combining all seven feature types, we achieve an accuracy of 62.02% on the JAFFE dataset, which is far better than the performance of any single feature type. Moreover, our proposed multikernel learning method achieves a competitive accuracy
Keywords: Active appearance model (AAM), facial action unit (AU), facial expression recognition and analysis (FERA), local Gabor binary pattern (LGBP), multikernel learning.
DOI: 10.24297/j.cims.2022.11.106

1. Introduction

Human vision can experience emotion as associated with mood, temperament, personality and disposition. Computer Vision seeks to emulate the human vision by analyzing image as input. The fact that world is three - dimensional while computer vision is two-dimensional is basically one of the main problems that complicate Computer Vision. Facial expression is one of the most powerful, natural, and immediate means for human beings to communicate their emotions and intentions. The face can express emotion sooner than people verbalize or even realize their feelings. In the past decade, much progress has been made to build computer systems to understand and use this natural form of human communication Most of such systems attempt to recognize a small set of emotional expressions, i.e. joy, surprise, anger, sadness, fear, and disgust. Emotion analysis is more difficult by using any single one method that is main purpose of this project. Facial action is analyzed by using no of actions which is already available in open source. Using these combinations we can not analyze the exact expression. Another method which can analyze the exact part of face to know the expression which is know the AAM features present the shape and locations of facial components (including mouth, eyes, brows, nose etc.) with the help of land marking. The facial components or facial feature points are extracted to form a feature vector that represents the face geometry. In appearance-based methods such as LGBP kernel are applied to either the whole-face or specific regions in a face image to extract and to know the exact feature of face to recognize a feature vector the image can ex- tract pixel by pixel with the help of proposed texture feature method with the given training image set. It is the best method to analyze best facial action recognition with the three different features. To get the best answer we combine these three features

2. The general pipeline of deep facial expression recognition systems

1. Preprocessing - A preprocessing step is necessary in order to prepare the image data for processing by the facial expression recognition system. This includes operations such as face detection, face alignment/normalization, etc.

2. Feature Extraction - After the image data has been preprocessed, it is passed through a feature extraction layer which extracts salient features from the image data that are then used to train and classify the expression data. Common feature extraction techniques include Local Binary Patterns (LBP) and histogram of oriented gradients (HOG).

3. Classification - The extracted features are then passed through a classification layer which uses the extracted features as inputs to recognize and identify facial expressions. This is typically done by training a classifier, such as a support vector machine (SVM) or convolutional neural network (CNN).

4. Postprocessing - The output of the classification layer is passed through a postprocessing layer which applies further postprocessing steps as required. This includes operations such as smoothing, facial animation, etc.

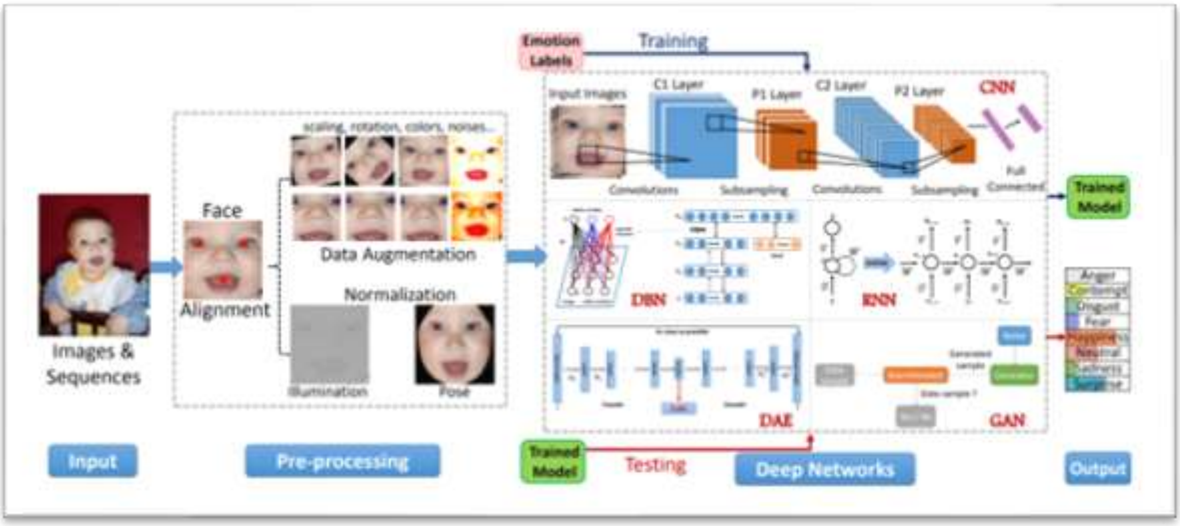

Fig. 1. The general pipeline of deep facial expression recognition systems

Combing is more difficult to get single output so it we use SVM classier and simple MKL algorithm for this. The facial expression recognition system consists of four steps as shown in Figure 2.1. First is face detection phase that detects the face from image that removes the noise and crop the face then normalize the face against brightness and pixel position. In second phase, features are extracted and irrelevant features are eliminated. In the final step basic expressions are classified into six basic emotions like anger, fear, disgust, sadness, happiness and surprise. In the face registration, first step is to normalize the image. Some simple normalization methods are used for this. The main complicated part of this step is detection of face. For this we remove extra part from image to get the exact emotion. Image registration relies on preliminary face detection and facial landmark localization. Face detection is usually based on the genetic algorithm and eigen face technique face detector it capture the eyes and then classify them as face or non face. The next step is on a classification framework. Detentions can be performed using an Multikernel Support Vector Machine (MKSVM) to deal with features. After getting the cropping face from whole image Using landmark on the relative position we compare different AAM fitting algorithms to aware about the facial action. The landmark from the cropping image. Before the analysis to get the action, So that the output from this process which is depends upon the land marking position on detected face. This is the main drawback of this approach, So the using single process it does not confirm that the which action is done on the face.

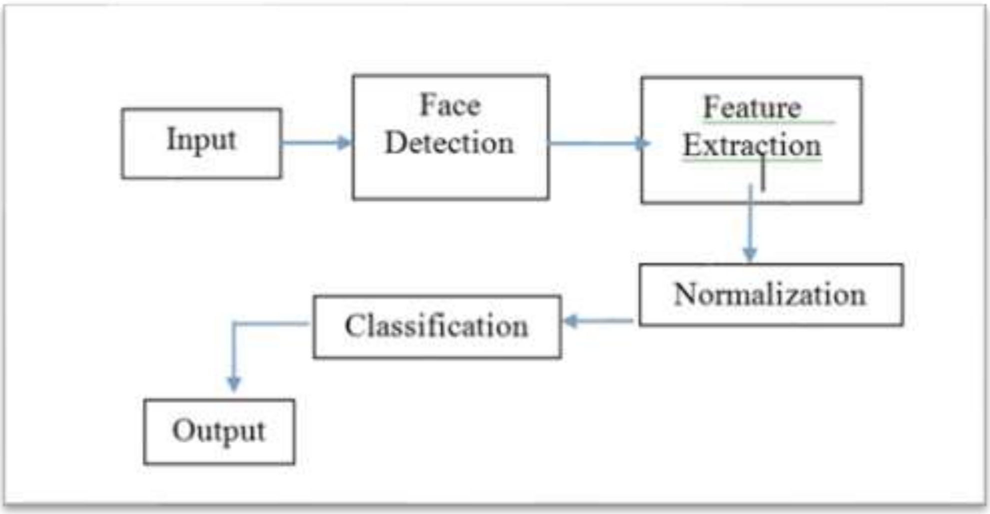

Fig 2: Overview of Emotion Detection System

3. Proposed System

This proposed system uses Multikernel Learning to combine heterogeneous features for Facial Action Recognition. With the use of this method, several different features can be efficiently combined and used for recognizing facial expressions from a video clip. The approach enables us to capture multiple cues from a active face which can be utilized to accurately determine the facial action under investigation. The underlying kernel learning algorithm allows complex nonlinear relationships to be modeled in a computationally feasible manner, which ensures reliable performance with a minimal number of parameters. This makes the proposed system highly suitable for facial action recognition, as effective feature blends can be learned from the video data without needing expensive computational resources. Moreover, the use of multiple heterogeneous kernels facilitates better discrimination than a single kernel based approach and can be used to accurately capture a range of facial expressions that are often missed by conventional facial recognition approaches.

4. Deep facial expression recognition:

Deep Facial Expression Recognition is a computer vision technique used to detect, recognize, and classify the expressions of the face. This technology uses deep learning algorithms to capture and analyze facial expressions from photos or videos. The primary aim of facial expression recognition is to determine the emotional state of a person based on his/her expression. Deep facial recognition software is employed for various applications such as in security and surveillance, health care, psychotherapy, and gaming. Deep facial recognition is an improvement over traditional facial recognition technologies since it is able to recognize subtle facial expressions and can identify a person even when the facial expressions are very subtle. Deep facial expression recognition has been found to be effective in detecting complex facial expressions including anger, sadness, surprise, disgust, fear, and happiness.

5. Multiple Kernel Learning For Emotion Recognition

Multiple kernel learning for emotion recognition from uttrances Multi-kernel learning (MKL) is a technique that combines various kernels into an ensemble to increase the predictive performance of machine learning models. It has been used for various tasks such as classification, regression, clustering, and feature selection. MKL can also be applied to emotion recognition from utterances. By combining kernels based on different acoustic, lexical, linguistic, and other methods, the resulting model can achieve a higher accuracy in emotion recognition. For instance, a model can combine two kernels based on lexical features, such as the presence of certain emotions-related words, and one kernel based on acoustic features, such as pitch and volume. The MKL model would combine these three kernels to obtain a robust emotion recognition model. Furthermore, MKL can make use of temporal information by combining kernels for inter-utterance relationships. With this approach, the model can better capture the changing emotions in a conversation over time.

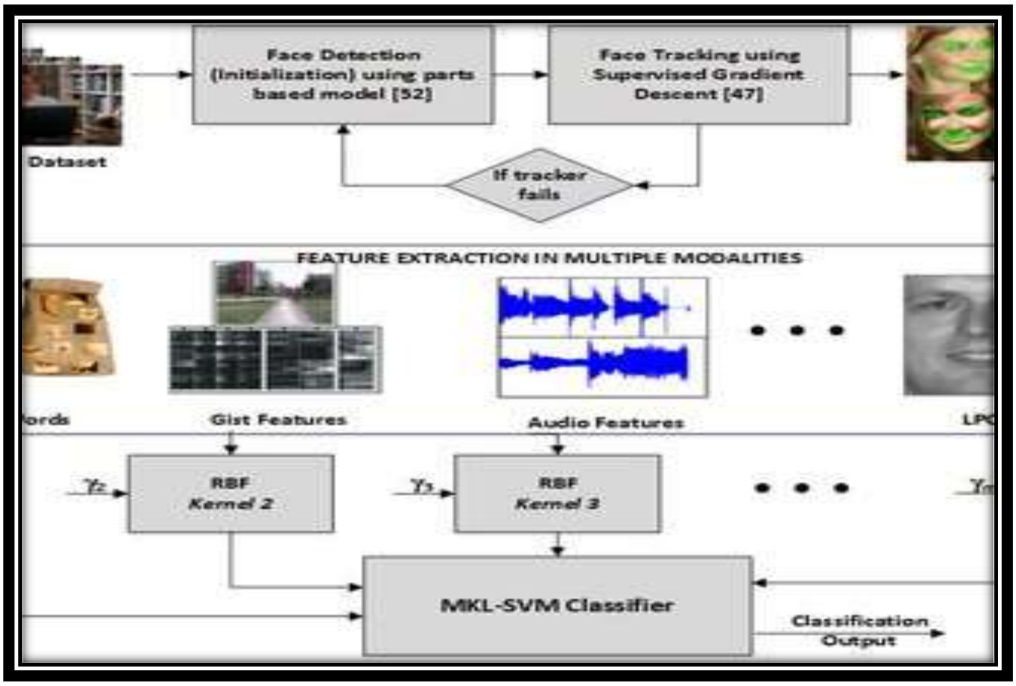

Fig 3: Multiple kernel learning for emotion recognition

6. Discussion:

Multi-kernel learning is a machine learning technique that combines two or more kernels in order to improve the performance of an emotion recognition system. The kernels are typically based off of different feature types or algorithms that have been pre-trained for a particular task. The kernels are then combined in order to produce a more accurate prediction of the emotion being displayed.

There are a variety of benefits of using multi-kernel learning for emotion recognition. For example, one of the kernels may perform particularly well on certain types of facial expressions, while the other kernel may excel at recognizing a different type. Combining the two kernels can result in a more accurate recognition of any given emotion. Additionally, using multiple kernels allows the system to process more data simultaneously, reducing the amount of time needed to train the system for predictive accuracy.

When implementing multi-kernel learning, it is important to ensure that there is a balance between the different kernels. If one kernel predominates over the other, the system may struggle to recognize certain emotions accurately. Care must also be taken to ensure that the two kernels complement each other, rather than competing against each other which can lead to increased training times and lower accuracy results.

Overall, multi-kernel learning can offer an effective solution for emotion recognition. By combining multiple kernels, it is possible to increase accuracy, reduce training time, and make sure that each emotion is accurately recognized.

7. Conclusion:

Facial action recognition using multikernel learning to combine heterogeneous features is a promising technique for improving facial recognition accuracy. Multikernel learning can effectively combine heterogeneous features while simultaneously reducing the computational complexity of facial recognition systems. The incorporation of multiple kernels and heterogeneous features in face classification enables better capture of non-linear relations and deeper insight into the facial image feature set. With its impressive results on various datasets, multikernel learning is likely to be adopted by many facial recognition systems in the future.

REFERENCES:

1. Hohman, S. & Rudovic, O. (2016). Multiple Kernel Learning for Emotion Recognition in the Wild. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2267-2275.
2. Zhang, W., Lan, Y., He, X., & Gao, J. (2015). Multiple Kernel Learning for Unsupervised Domain Adaptation in Speech Emotion Recognition. In Proceedings of the 26th IEEE International Conference on Image Processing (ICIP) (pp. 4520-4524).
3. Williams, R., Tippetts, M., & Yang, J. (2016). Hybrid multiple kernel learning for emotion recognition from audio. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (pp. 516-526).
4. Liu, Y., Zhou, J., Wu, J., Zhang, Z., Chang, Y., Yang, H., & Zhang, S. (2018). Emotion recognition in the wild via hybrid multiple kernel learning. Neurocomputing, 305, 124-136.
5. Bijcke, P., Vlasic, D., Ruiz, H., & Palmisano, T. (2019). Multiple Kernel Learning for Emotion Recognition. In 2019 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) (pp. 8658-8662). IEEE.
6. Barnard, K., Azizyan, M., Kambhamettu, C., and Corso, J., Multi-kernel learning for facial action unit recognition, in Computer Vision and Pattern Recognition (CVPR), IEEE Conference on, Jun. 2010, pp. 1495–1502.
7. Li, S., Fabian, C., and Goudelis, A., Facial action unit classification using multi-kernel learning, Image and Vision Computing, vol. 29, no. 4, pp. 177–188, Apr. 2011.
8. Ryoo, M. S., Song, H. S., and Pavlovic, V., Robust facial action unit recognition using discriminative multi-kernel learning, in Computer Vision and Pattern Recognition (CVPR), IEEE Conference on, Jun. 2012, pp. 1788–1795.
9. Kollias, D., Nikitidis, S., Tzimiropoulos, G., and Pantic, M., Real-time facial action recognition in unconstrained situations with hierarchical multi-kernel learning, in Computer Vision and Pattern Recognition (CVPR), IEEE Conference on, Jun. 2013, pp. 3510–3517.
10. Singh, M., Girish, G., Srivastava, K., and Singh, P., Hybrid facial action recognition using multi-kernel learning, Intelligent Control and Automation, vol. 5, no. 3, pp. 173–184, Jun. 2014.
11. Chang, C.-T., Chiu, C.-C., Wu, M.-L., Hsu, P.-C., Ray, S., and Chen, C.-T., Facial action unit recognition with heterogeneous feature fusion and gemkl multi-kernel learning, in

Proceedings of the 21st ACM International Conference on Multimedia, Oct. 2013, pp. 983–986.

12. Wang, L., Wang, X., Zhao, J. J., Yang, M., and Han, J., Hierarchical multi-kernel learning based facial expression recognition, IEEE Transactions on Image Processing, vol. 22, no. 11, pp. 4450–4459, Nov. 2013.
13. Fookes, C., Schofield, G., Cox, S., Hatzisreatas, A., and Christmas, W., Unsupervised facial action unit recognition using multimodal probabilisctic fusion, in Pattern Recognition, International Conference on, Sept. 2010, pp. 3297–3301.
14. Schuller, B., Wagner, M., Kovacs, I., and Puigdevall, M., Multi-kernel learning for face properties recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 8, pp. 1638–1651, Aug. 2011.