Robust Emotion Recognition on Hand Crafted Features in Static Action Sequences

R.Santhoshkumar, B. Rajalingam, P. Santosh Kumar Patra, M. Narayanan[a)], T. Poongothai, N. Satheesh, G. Jawaherlal Nehru, G. Govinda Rajulu

Dept. of CSE, St. Martin's Engineering College (Autonomous), Secunderabad, Telangana, India.

[a)]Corresponding author: drmnaranyanan-cse@yahoo.com

ABSTRACT: The acknowledgement of human emotions plays a key role in daily life and is necessary for successful social interaction. In many applications of human computer interaction nonverbal communication methods such as human body movements, face expressions, eye movement and gestures are used, among them the recognition of emotions from human body movements, because they convey the emotions and feelings of the person. In this paper Advanced Block Based Intensity Value (ABBIV) feature is proposed for emotion recognition from human body movements and compared with Histogram of Gradient (HoG) feature. The GEMEP corpus videos for five basic emotions are converted into gray frames. Then the HOG feature and Block Based Intensity Value (BBIV) features were extracted from the body movements of the human present in the consecutive frames. Among the two features Advanced Block Based Intensity Value (ABBIV) perform better accuracy than HoG feature. The extracted features are fed to the SVM and KNN and Random Forest classifiers to identify the emotions of the human. The performance measure can be calculated using F-Score value. The five archetypical emotions (angry, fear, joy, sad, pride) from GEMEP corpus dataset are used for this experiment. **Keywords:** Body Movements, Non-verbal communication, Emotion recognition, Histogram of Gradients (HoG), Advanced Block based intensity value (ABBIV), SVM, KNN and Random Forest.

INTRODUCTION

Recent research on experimental psychology demonstrated that emotions are important in decision making and rational thinking. Over the years research in emotion recognition mainly concentrated on facial expression, voice analysis, full-body movements and gestures. The possibility of many scientists in the psyche, psychiatry, neurosciences and behavioural sciences to measure and recognise emotions is also in the interests of many people. Computer systems with exact measurements can significantly improve the quality and acceleration of current research, where many data are manually processed. Emotional conditions are a foundational phase of human interaction and should thus also be used in interaction between humans and computers. Affective countries are motivating and enriching our social interactions. If computing disregards these aspects, a great deal of information received by the user will also be lost in the interplay. The affective computing paradigm suggests that user interfaces should answer not only user orders, but also emotions. Besides thousands of articles and books, emotions were always at the centre of human knowledge and the raw material, but they have now been for technology and science too. Evolving research shows that people can efficiently decode emotional signals in non-verbal communication from others and deduce other people's emotional states. Somebody actions are referred to as gestures. Mostly the head, hands and arm can perform the action. These interactions contain information and the content of the interactions of emotional states. In the past decade, plenty of effort has been taken to recognize emotions automatically through their combinations. The research was conducted with audio, facial expression and followed by bodily expressions. Recognizing emotions from human body movements has numerous applications with the support of psychological studies. Some areas in which automated emotional recognition by body signals are applied are suspicious action recognition to alarm safety personnel, computer interaction, and care and to help autism patients.Many people are also interested to be able to measure and to recognize emotions for many scientists in psycho-psychiatry, neuroscience and behavioral sciences. The quality and acceleration of current research, with many data handled, can be improved significantly by computer systems with exact measures. Emotional conditions are a basic phase of human interaction and should therefore be used in human-computer interactions as well. Affective countries motivate our social interactions and enrich them. If computing does not take into account these aspects, the interaction will also lose a great many information the user receives. Some actions of the body are called gestures. The action can mainly be carried out with the head, hands and arm. These conversations collect details and

Recent Trends in Science and Engineering
AIP Conf. Proc. 2393, 020215-1–020215-12; https://doi.org/10.1063/5.0079770
Published by AIP Publishing. 978-0-7354-4198-9/$30.00

the content of emotional states' interactions. In the past decade, a lot has been done to automatically recognize emotions by their combinations. Audio and facial expressions were conducted and physical expressions followed.The objective for this paper: This article seeks to recognize the emotions of human movements by using handcrafted features with different actions. The motivation of this paper: When the camera is too long for a person, the face is not clear in the surveillance environment. This kind of problem can be corrected by capturing movements of the body to recognize human emotions (head, legs, hands). Contribution of this paper: The human emotion is easily recognized by the use of SVM, KNN, and Random Forest Classifier with the aid of HoG and ABBIV.The rest of this paper organized as follows. The outline of this work is described in Section II. Section III shows the three types of extraction features and related debates. Section IV provides the SVM, KNN and Random Forest classification experimental results and performance assessment. The paper ends in Section V.

RELATED WORKS

R. Santhosh Kumar(2019) proposed emotion recognition from human body movement using DCNN model. R. Santhosh Kumar(2020) developed a video based automatic human emotion recognition using gesture dynamic's features and the HOG – KLT features are evaluated by SVM, and Random Forest classifier. NelleDeal, et al. (2012) developed a body posture and body action coding system from body movement on an anatomical level is different articulations of body parts, direction and orientation of the movement. Ioanna-Ourania Stathopoulos, et al. (2011) conducts a survey on recognizing human emotion from hand/arms, gestures and body movements. M. Melissa Gross, et al. (2011) develops a robust technique for assessing human body expression based on movement characteristics with positive and negative emotions. BokkenCimon, et al. (2013) describes the study to analyse the spatial and temporal information structure of the motion capture data and extract features that are related to affective state descriptors. HarrisZacharias's, et al. (2014) performs the survey on recent advances in developing robust techniques and modalities for automatic human emotion recognition system from body movements. Here the importance of body movement segmentation are discussed and advanced application areas are described. Ginevra Castellano, et al. (2007) proposes an analysis of emotional behaviour system based on classification of time series and dynamics of expressive motion cues. A novel AANN miscounting Rate (AMR) algorithm is used to detect the shot transitions. Joti Joshi, et al. (2013) developed the automatic depression analysis system from human gestures and upper body expressions. The bag of words and space-time interest points is developed for the analysis of facial and upper body movements. Mohamed BachaKandice, et al. (2010) proposed a robust gesture recognition system using learning local motion signatures (LMSs). NaveedDall, et al. (2005) study human detection and robust visual object recognition using adopting linear SVM. They showed the performance of human detection using feature sets of Histograms of Oriented Gradient (HOG) descriptors. Laptev. (2005) proposes a quantized trajectory snippets method for tracked features. This method is a simple feature tracking method and computationally efficient for motion detection.Chang, et al. (2011) proposed real-time vision systems for video analysis to describes SIFT feature extraction algorithms and novel implementations of the KLT feature tracking. R. Santhosh Kumar, et al. (2017) proposed a feedforward deep convolutional network for recognizing the human emotion from static action sequences for the five basic emotions (angry, fear, joy, sad and pride). R. Santhosh Kumar, et al. (2017) developed the body movement feature and is modeled by Support Vector Machine (SVM) and Random Forest (RF) algorithm for recognizing the human emotion from video sequences.

METHODOLOGY

In this paper, the pre-processing step is common for all HoG and ABBIV features. The five basic emotions from GEMEP datasets videos are converts into RGB frames. Then the RGB frames are converted into gray frames and extract human in each frame using the bounding box method. It is generated based on height, width, x and y points of each frame. Figure 1 illustrates videos into RGB frames then RGB into gray frames and detect human using bounding box for all frames.

FIGURE 1. Common Pre-Processing Steps

After pre-processing step, the HoG features are extracted from each sequence of the frame. The figure 1 shows the overview of this work. The pre-processed frame is divided into two blocks (B1 & B2). To select the motion block, the maximum gray value block is chosen by calculating the maximum gray value for B1 and B2. The HoG feature is extracted from the maximum gray value block B1. Figure 2 describes the screen short feature extraction process.The HoG is defined as the appearance and shape of a local object can often be characterized by the distribution of the gradient intensity of the corresponds gradient. (Santhosh Kumar et al., 2018).

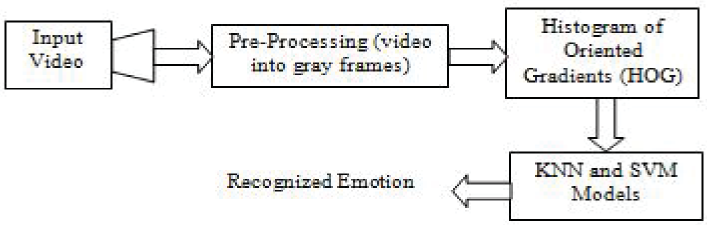

FIGURE 2.Overview of HoG Feature Extraction

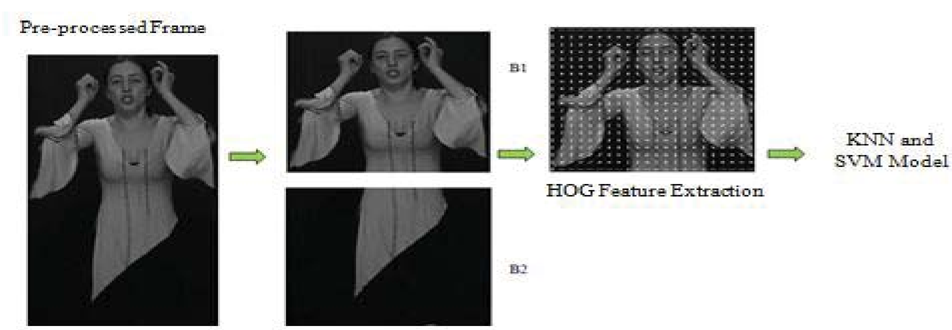

FIGURE 3.Block Diagram for HoG Feature Extraction

The last line indicates that the summary of the HOG technique used in its superior form in Scale Invariant Transformation Features (SIFT) has been largely demotivated in human detection (Sudipta, et al. 2006). The following 1D histogram is built, whose enumeration provides the HOG descriptor with the pixel values of the cell pixels. The intensity of the image L is to be analyzed. Further the image is divided into cells of 3 × 3 pixels size with 9 bins. Therefore 81 dimensional feature vector for each frame. The gradient magnitude g and the gradient orientation θ used to calculate all image gradients pixels on the block.

$$g(a,b) = \sqrt{g_x(a,b)^2 + g_y(a,b)^2} \quad (1)$$

$$\Theta\,(a,b) = \arctan\frac{g_y(a,b)}{g_x\,(a,b)} \quad (2)$$

Compute a feature vector v_{ig}or each cell c_{ij} in the block. Equation 3 defines the weighted gradient magnitude by quantizing the unsigned orientation into K orientation bins.

$$v_{ij} = [v_{ij}(\beta)]^T{}_{\beta\in[1....K]} \quad (3)$$

The equation 4 defines the $v_{ij}(\beta)$

$$v_{ij}(\beta) = \sum_{(a,b)\in c_{ij}} g(a,b)\delta[\mathrm{bin}(a,b) - \beta] \, (4)$$

The index of the orientation bin with the pixel (a, b) returns the function bin (a, b) and the function δ is the Kroenke delta. The coefficient ρ normalize the feature vector in all cells from 2D descriptor of block.

$$\rho = \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{\beta=1}^{K} v_{ij}(\beta) \qquad (5)$$

The videos input are standardized and an image of the difference from an image sequences in condo B is discussed. The image of difference shows motions and the region is seen as a region of interest (ROI). The ROI is extracted in two blocks of 250 x 360 pixel size, respectively. The pixel size block of 50x90 is subsequently divided into 5x4 pixel block. The 20-dimensional feature values are extracted from the block 5x4. The algorithm is discussed here for the recognition of emotions. Figure 4 shows the proposed approach.

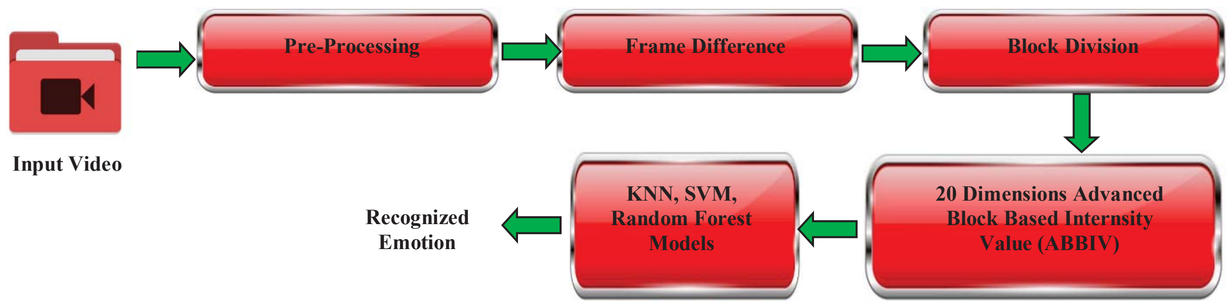

FIGURE 4. Overview of BBIV feature extraction

Algorithm

Step 1: Difference image is obtained from series frames as in Eq. 6.

Step 2: Difference image is calculated with Eq.7, to extract motion information.

Step 3: 20-dimensional feature vector is extracted using advanced block-based intensity value (ABBIV).

Step 4: The 20-dimensional feature vector is fed to SVM and KNN.

Frame Differencing

Frame differencing is the lingering image through subtracting the following frames for video sequence change detection. For creating the difference image, the two time frames t and t+1 are used. The high-amplitude region is seen as differentiating emotional regions.

$$\mathrm{Diff}_k = |\mathrm{Int}_k(i,j) - \mathrm{Int}_{k+1}(i,j)| \qquad (6)$$

$$1 \le i \le w, 1 \le j \le h$$

The information extracted from the motion is deemed the region of interest (ROI). The two following frames for the GEMEP data set are shown in Figure 5(a), 5(b).The resulting difference image is shown in figure 5(c). Diff_k (I, j)is the difference image, Int_k (I, j)is the intensity of the pixel (I, j)in the k^{th} frame, the width is w and height is h of the respective image. Motion information MI_kor difference image is calculated using

$$\mathrm{MI}_k(i,j) = \begin{cases} 1, & if\,\mathrm{Diff}_k(i,j) > t \\ 0, & otherwise \end{cases} \qquad (7)$$

Where, threshold is t

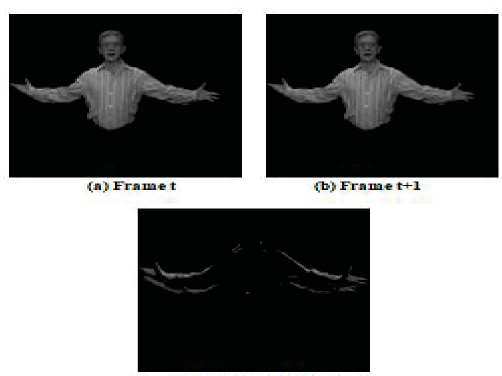

FIGURE 5.(a) Frame t (b) Frame t+1 (c) Difference image

20-Dimensional Feature Vector

The aim of the work is to recognize emotions from symmetric sequences of action. It can be seen that human heads and arms are used to emotion.

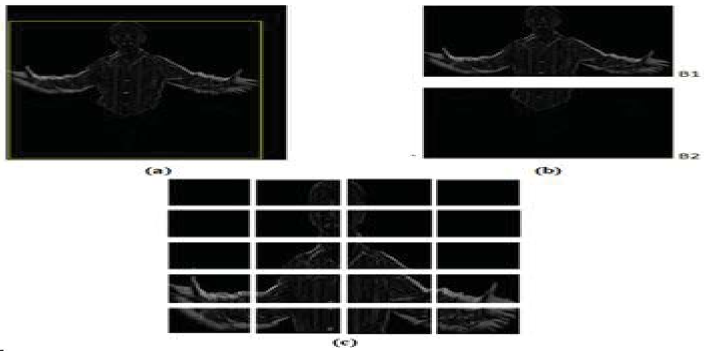

FIGURE 6.(a) Motion information. (b)Extracted ROI from a. (c) 5x4 Block division of B1

The temporal feature is extracted from the 250 x 360 size difference picture as indicated in fig. 6. (a). The ROI is divided into two B1 and B2 blocks consisting of the head, torso, brace and elbow regions of 250 x 360 size each as shown in Fig. 6. (b). The highest intensity region of B1 is identified as the heavy moving region for the extraction of the 20-dimensional vector and the selected block is divided into 5 x 4 sub blocks for further encoding, as indicated in Fig. 6. (c). the maximum intensity of each sub-block is a 20-dimensional vector feature. The Support Vector Machine (SVM) for visual model recognition is an important and effective technique. The kernel learning algorithm uses SVM most extensively. Two classes have been divided into large border hyperplanes by elegant theory. N mutually exclusive classes cannot be extended easily. The popular 'one vs other' approach is used to address the problem of multiple classes where one class is separated from N classes. Typically, the classification task includes training and testing data. The training data are separated by $(s_{1,}\ t_1)$, (s_2, t_2),...... (So, t_n) into two classes, where b_e Є {+1,-1} are the class labels and s_{oju}Є T_N includes n-dimensional vector feature. The objective of Support Vector Machine is to develop a model that forecasts the test value. Wks. + b = 0 is the hyperplane of binary classification, where we R^N, The two classes are separated by b Є R (Laptev, 2005). M = 2/||w|| is the large margin as show in Fig. 7.The Lagrange multipliers α_i (i=1, ...m) to solve the problem of minimization, where v and y are the optimal values from Eq. 8. $h(s) = \mathrm{sgn}\left(\sum_{j=1}^{n} x_j b_j L(s_j, s) + y\right)$ (8)

Maximize the extent and reduce the exercise error by non-negative slack variablesε_j. The Eq. 9 and Eq. 10 obtain the soft margin Classifier.

$$\min_{v,y,\varepsilon} \frac{1}{2} v^R v + D \sum_{j=1}^{k} \varepsilon_j \qquad (9)$$

$$b_j(v^R \emptyset(s_j) + y \geq 1 - \varepsilon_j, \varepsilon_j \geq 0 \qquad (10)$$

When the training sample is not linearly separable, the input space mapped into high dimensional space using kernel function $L(s_j s_k)$= Ø (s_j). Ø (s_k) (Dalal, et al., 2013).

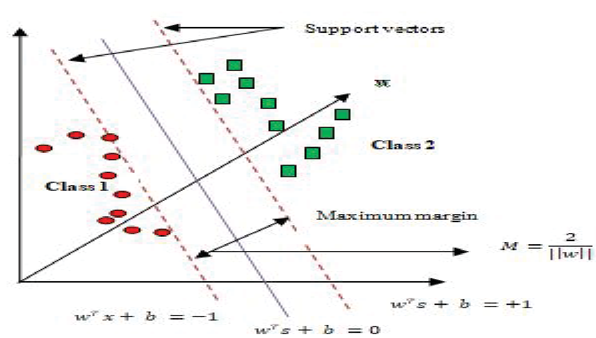

FIGURE 7.Illustration of hyperplane in linear SVM

Linear:$L(s_j s_k) = s^R j^s k$ (11)

Polynomial:$L(s_j, s_k) = (\alpha s^R j^s k + \alpha)^c, \alpha > 0$ (12)

Radial Basis Function (RBF):$L(s_j, s_k) = \exp(-\alpha||s_j - s_k||^2), \alpha > 0$(13)

Sigmoid:$L(s_j, s_k) = \tanh(\alpha s^R j - s_k + t)$ (14)

Where, α, t, and c are parameters of kernel.

K nearest neighbor is a simple and popular technique for pattern recognition, machine learning and data mining field. It is a type of supervised learning method. It is said to be a lazy learning where the function is only approximated locally. K nearest neighbor is a non-parameter algorithm where samples are classified depends on the category of their nearest neighbor. According to the k learning samples, the classification algorithm finds the test sample's categories which are the nearest neighbor to the test sample. However, the classification algorithm needs to compute all distance between the training sample and testing sample. The process of K nearest neighbor algorithm to categorize sample S. Assume training samples A1, A2, Aq. After feature reduction, N is the addition of the training samples and get the n-dimension feature vector. All training samples (S1, S2,and Sn) have the same feature vector of sample S and evaluate the similarities among them. For example taking the PTH sample b_p (b_p1, b_p2, bpn) and the similarity SIM(S, b_p) is:

$$\mathrm{SIM}(\mathrm{S}, \mathrm{b_p}) = \frac{\sum_{q=1}^{n} S_q . b_{pq}}{\sqrt{\left(\sum_{q=1}^{n} S_q\right)^2} . \sqrt{\left(\sum_{q=1}^{n} b_{pq}\right)^2}} \qquad (15)$$

The Larger N similarities, SIM(S, bp), (p=1, 2, N), of k samples are chosen and consider them as a K nearest neighbor collection of S. Then, the probability of S can be calculated using this formula.

$$\mathrm{R}(\mathrm{S}, \mathrm{A_q}) = \sum_b (\mathrm{S}, \mathrm{b_p}). \mathrm{y}(\mathrm{b_p}, \mathrm{A_q}) (16)$$

Where y (b_p, A_q) is a category attribute function.

The detailed random forest technique has shown more precision in classification and regression than the independent decision trees. A random classification of forests consists of a number of trees each cultivated by randomization. The random forest uses a variety of functions to split each node into each tree.

Algorithm:

- For original data to draw t_{ree}bootstrap samples.
- Grow an unspurred grading tree or regression tree, with a modification: select the randomly sample of the predictors and choose the best division between the variables at every node, rather than the best division.
- Predict new data by adding trees trees predictions

EXPERIMENTAL RESULTS

FIGURE 8.Sample frames for basic emotions (Angry, Joy, Fear, Sad, and Pride)

The next five emotional feelings are anger, joy, fear, sadness and pride in GEMEP corpus data set. The tests are done on a computer with Intel Core i7 Processor of 3.40GHz with 8GB of RAM on Windows 10 Operating System using Python and openCV. The HoG, STIP and ABBIV obtained are provided in classification systems for the emotion from GEMEP data set to the KNN, SVM and Random Forest. A collection of audio and video recordings are provided by The Geneva Multimodal Emotion Portrayals (GEMEP). 10 actors can be involved in the 18 affective emotional expressions. They were active in different kinds of speech and verbal contents. For this work were chosen from those five fundamental emotions (Angry, Joy, Fear, Sad and Pride). In each emotional video, there are 10 actors (5 male and 5 female). The recorded videos have a resolution of 720 to 576, and 25 frames a second for each video (fps). The achievements of all three types of features were assessed using supervised methods of learning. The extracted features are supplied by train, rail label, test data and test label one at a time to the KNN and SVM classifiers. The measurement evaluation for this execution is accuracy, accuracy, F-score, specificity and precision. Precision is a precise measure. Remember how a special emotion is accurately recognized. The symphonic mean of accuracy and reminder is called f-pointing. Specificity shows the extent to which a strategy accurately recognizes negative emotions. Finally, Precision shows that movement recognition is generally accurate. The accuracy, recall, F-score, specificity and precision assessments shall be provided as follows.

$$\text{Accuracy} = \frac{\text{tp+tn}}{\text{tn+fp+tp+fn}} \quad (17)$$

$$\text{Recall} = \frac{\text{tp}}{\text{tp+fn}} \quad (18)$$

$$\text{F}-\text{Score} = 2\frac{\text{PrecisionXRecall}}{\text{Precision+Recall}} \quad (19)$$

$$\text{Specificity} = \frac{\text{tn}}{\text{tn+fp}} \quad (20)$$

$$\text{Precision} = \frac{\text{tp}}{\text{tp+fp}} \quad (21)$$

Where tp and TN are the quantity that the classes and fp and FN are true, the amount of false positive and false negative expectations is true. This section shows the experimental results on GEMEP dataset for HoG feature and BBIV features with SVM and KNN classifiers. Experiment Results on KNN, SVM and Random Forest Classifiers for HoG Feature. The normal recognition accuracy of KNN is 89.6% and SVM is 90.2% on the GEMEP dataset and the confusion matrix appeared in Table 1 and 2. The corner to corner of the confusion matrix illustrates the percentage of instance that was classified accurately. Each emotion class occurrence is spoken to by the lines and the emotion class anticipated by the classifier is spoken to by the sections. The emotions like sad, fear and pride are grouped well with precision more noteworthy than 90%. From this, angry and joy emotions are confused as a curve, where these two emotions instinctively appear to be difficult to separate and it needs to promote consideration. The execution assessment comes about are computed for the evaluated HoG feature has better accuracy, recall, F-score and Specificity for KNN, SVM and Random Forest classifiers on GEMEP dataset.

TABLE 1. Individual emotions accuracy percentage for KNN for HoG feature

	ANGRY	JOY	FEAR	SAD	PRIDE
ANGRY	77.6	22.4	0	0	0
JOY	23.9	75.1	0	0	1
FEAR	1	0	97.9	0	1.1
SAD	0	0	0	98.6	1.4
PRIDE	1	1	0.9	1	96.1

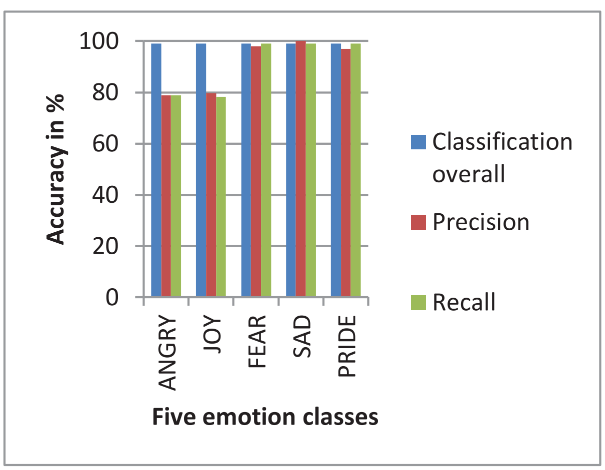

FIGURE 9. Bar chart of Performance Measure for KNN using HoG.

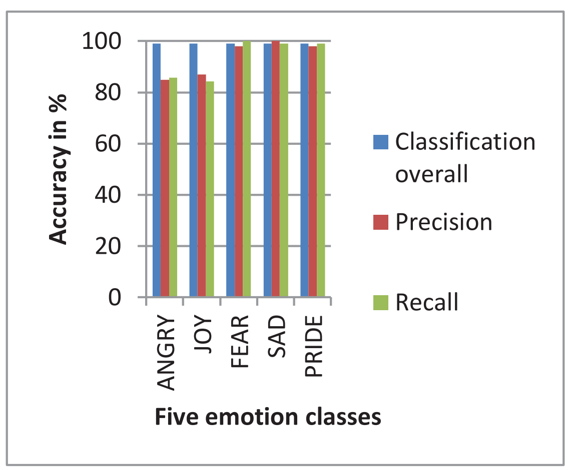

FIGURE 10. Bar chart of Performance Measure for SVM using HoG.

TABLE 2. Performance Measure for KNN using HoG in (%)

	Classification overall	**Precision**	**Recall**
ANGRY	99	77.77	75.49
JOY	99	75.00	76.53
FEAR	99	97.98	100
SAD	99	98.99	98.99
PRIDE	96	96.87	96.87

TABLE 3. Individual emotions accuracy percentage for SVM for HoG feature

	ANGRY	**JOY**	**FEAR**	**SAD**	**PRIDE**
ANGRY	78.3	21.7	0	0	0
JOY	20.5	79.5	0	0	0
FEAR	1	0	97.8	0	1.2
SAD	0	0	0	99.6	0.4
PRIDE	0.1	1	1	1	96.9

TABLE 4. Performance Measure for SVM using HoG in (%)

	Classification overall	**Precision**	**Recall**
ANGRY	99	78.78	78.78
JOY	99	79.79	78.21
FEAR	99	97.98	98.98
SAD	99	100	99
PRIDE	99	96.97	98.96

TABLE 5. Individual emotions accuracy percentage for Random Forest for HoG feature

	ANGRY	**JOY**	**FEAR**	**SAD**	**PRIDE**
ANGRY	84.5	15.5	0	0	0
JOY	13.9	86.1	0	0	0
FEAR	1.6	0	97.1	0	1.3
SAD	0	0	0	99.4	0.6
PRIDE	0.1	1	0	1	97.9

TABLE 6. Performance Measure for Random Forest using HoG in (%)

	Classification overall	**Precision**	**Recall**
ANGRY	99	84.84	85.71
JOY	99	86.86	84.31
FEAR	99	97.98	100
SAD	99	100	99.00
PRIDE	99	97.98	98.98

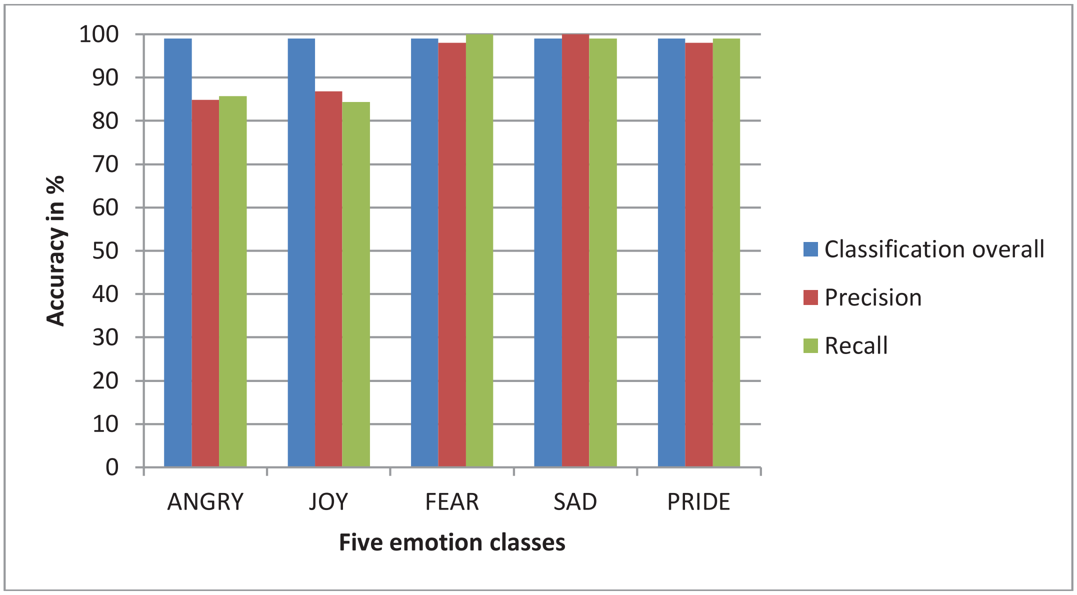

FIGURE 11.Bar Chart of Performance Measure for Random Forest using HoG in (%)

The figure 9, 10 and 11 shows the bar chart description of performance measure using KNN, SVM and Random Forest respectively for HoG feature. And also the table 2, 4 and 6 shows the performance measures like overall accuracy, precision and recall for KNN, SVM and Random Forest classifiers respectively for HoG feature. Among the three classifiers, the random forest shows better overall accuracy when compared with KNN and SVM. The individual emotion recognition accuracy also gives better in random forest classifiers for HoG feature. Experiment results on KNN, SVM and Random Forest classifiers for ABBIV feature The normal recognition accuracy of KNN is 90.3% and SVM is 93.4% on the GEMEP dataset and the confusion matrix appeared in Table 3 and 4. The corner to corner of the confusion matrix illustrates the percentage of instance that was classified accurately. The each emotion class occurrence is spoken to by the lines and the emotion class anticipated by the classifier is spoken to by the sections. The emotions like Angry, happy, sad, fear and pride are grouped well with precision more noteworthy than 90%. The execution assessment comes about are computed for the evaluated block based intensity value (ABBIV) feature has a better accuracy, recall, F-score and Specificity than HoG feature for KNN and SVM with RBF kernel on GEMEP dataset.

TABLE 7. Individual Emotions Accuracy Percentage for KNN for ABBIV Feature

	ANGRY	**JOY**	**FEAR**	**SAD**	**PRIDE**
ANGRY	90.7	5.3	2	1	1
JOY	6.2	90.8	1	0.4	0.6
FEAR	0.5	2.5	93.4	0.6	3
SAD	2.3	1.7	0.3	93.7	2
PRIDE	2.1	1.3	2.9	2.4	91.1

TABLE 8. Performance Measure for KNN using ABBIV in (%)

	Classification overall	**Precision**	**Recall**
ANGRY	99	90.90	90.00
JOY	87	92.78	90.90
FEAR	98	94.89	94.89
SAD	98	94.89	96.87
PRIDE	98	92.85	93.81

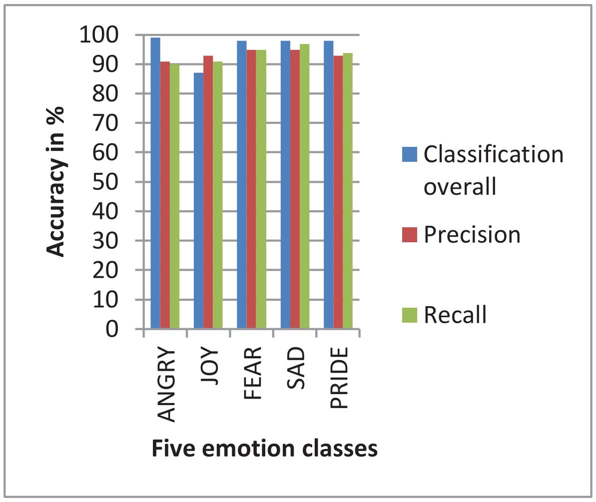

FIGURE 12. Bar chart of Performance Measure for KNN using ABBIV.

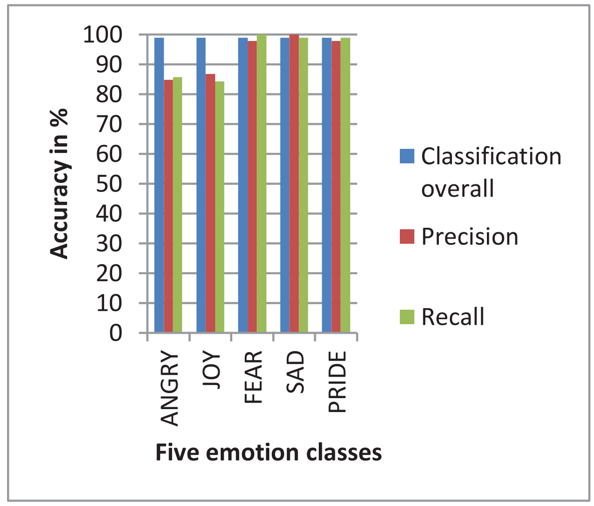

FIGURE 13. Bar chart of performance Measure for SVM using ABBIV

TABLE 9. Individual Emotions Accuracy Percentage for SVM for ABBIV Feature

	ANGRY	**JOY**	**FEAR**	**SAD**	**PRIDE**
ANGRY	91.3	6.7	0	1	1
JOY	5.2	91.8	1	0	1
FEAR	0.3	1.5	94.4	1.8	2
SAD	2.3	1.7	0.3	94.7	1
PRIDE	0.1	1.3	0.9	2.4	95.1

TABLE 10. Performance Measure for SVM using ABBIV in (%)

	Classification overall	**Precision**	**Recall**
ANGRY	99	91.91	92.85
JOY	98	92.85	91.00
FEAR	98	95.91	98.94
SAD	98	95.91	95.91
PRIDE	98	96.93	95.00

TABLE 11. Individual Emotions Accuracy Percentage for Random Forest for ABBIV Feature

	ANGRY	**JOY**	**FEAR**	**SAD**	**PRIDE**
ANGRY	93.0	4.3	1.4	1.3	0
JOY	6.3	92.7	0	1	1
FEAR	0.4	1.4	94.4	1.2	2.6
SAD	1.8	1.9	1	95.1	0.2
PRIDE	0	1.1	1	2.2	95.6

TABLE 12. Performance Measure for Random Forest using ABBIV in (%)

	Classification overall	Precision	Recall
ANGRY	99	93.93	96.00
JOY	100	92.00	92.92
FEAR	98	95.91	96.90
SAD	98	96.93	95.00
PRIDE	99	95.96	96.93

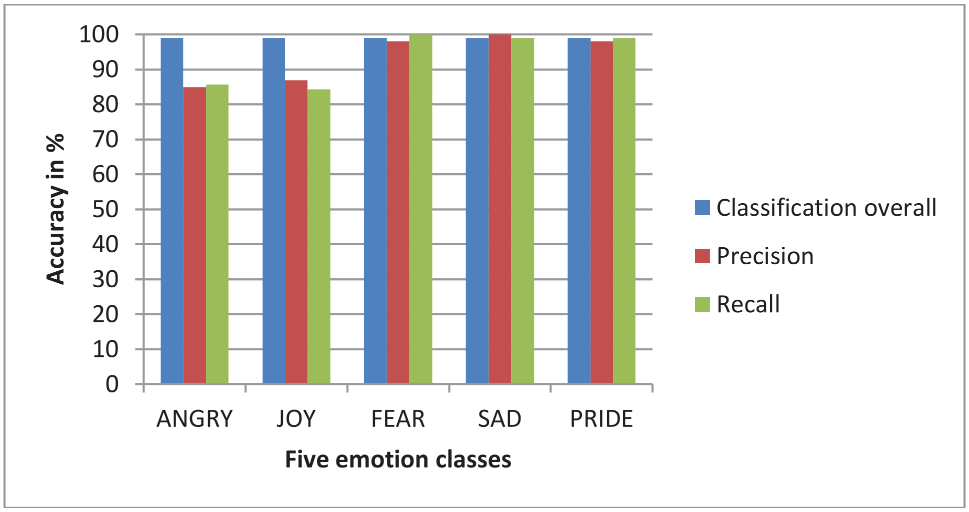

FIGURE 14. Bar chart of Performance Measure for Random Forest using ABBIV

The figure 12, 13 and 14 shows the bar chart description of performance measure using KNN, SVM and Random Forest respectively for ABBIV feature. And also the table 8, 10 and 12 shows the performance measures like overall accuracy, precision and recall for KNN, SVM and Random Forest classifiers respectively for ABBIV feature. Among the three classifiers, the random forest shows better overall accuracy when compared with KNN and SVM. The individual emotion recognition accuracy also gives better in random forest classifiers for ABBIV feature.

TABLE 13. Comparison of HoG, BBIV Features with KNN, SVM and Random Forest Classifiers

Feature	Dataset	KNN classifier	SVM classifier	Random Forest classfier
HoG feature	GEMEP corpus	89.6 %	90.2 %	91.7%
Proposed ABBIV feature	GEMEP corpus	**90.3 %**	**93.4 %**	**95.3%**

Among the three features, 20 dimensional Advanced Block-Based Intensity Value (ABBIV) feature predicts the emotions accurately than HoG features. The table 5 shows the accuracy20-dimensional ABBIV features gives better results than HOG feature with KNN and SVM classifiers. Table 13 shows the proposed work recognition rates with existing works.

TABLE 14. Comparison with Existing Work

Research	Model	Recognition Rates (in %)
Wang et al.	KNN	86.4
Ginevra Castellanoet al.	KNN	73.0
Arunnehru et al.	SVM	90.7
This Work	Random Forest	95.3

CONCLUSION AND FUTURE WORK

In this paper, the comparison of HoG and ABBIV feature for predicting human emotions from body movements on the sequence of frames are discussed. The Advanced Block-Based Intensity Value (ABBIV) feature performs better accuracy than the other two features. The experiments are evaluated on challenging benchmarks GEMEP corpus dataset. The emotions of the human can be identified accurately from the human body movements using ABBIV feature. The identification of human emotions from facial expression is not clear and accurate when the human is not in front of the camera. In any angle of the camera, the identification of human emotion from body movements show better accuracy using ABBIV feature is discussed in this paper. Future work will investigate probabilistic pose estimation and emotion recognition methods in video surveillance data. The research applications of probabilistic deep convolutional neural networks are to recognize the emotions of autism spectrum disorder (ASD) people in the video. In autism children emotion recognition system sends all the recognized emotions to the registered mobile number which makes the parents or caretaker feel disturbed. To avoid this situation, the system can be made for sending only the sad or fear emotions to the registered mobile number.

REFERENCES

1. R. Santhoshkumar, M. Kalaiselvi Geetha, Journal of Mechanics of Continua and Mathematical Sciences (JMCMS) **14**(3), 182-195 (2019).
2. R. Santhoshkumar, M. Kalaiselvi Geetha, Lecture Notes in Networks and Systems, Springer**121**, 261-272.
3. Nele Dael, Marcello Mortillaro, Klaus R. Scherer. Springer Science Business Media (2012).
4. Ioanna-Ourania Stathopoulou, George A. Tsihrintzis. Springer-Verlag Berlin Heidelberg (2011).
5. Melissa Gross, M., Elizabeth A. Crane, Barbara L. Fredrickson. Springer Science + Business Media (2010).
6. Gokcen Cimen, Hacer Ilhan, Tolga Capin and Hasmet Gurcay. Anim. Virtual Worlds **24**, 355–363 (2013).
7. Haris Zacharatos, Christos Gatzoulis and Yiorgos Chrysanthou. Computer Graphics and Applications, IEEE (2014).
8. Ginevra Castellano, Santiago D. Villalba, and Antonio CamurriSpringer-Verlag Berlin Heidelberg (2007).
9. Jyoti Joshi, Roland Goecke, Gordon Parker and Michael Breakspear. International Conference and Workshops on Automatic Face and Gesture Recognition, IEEE (2013).
10. Mohamed Bêcha Kaâniche, François Brémond. IEEE Conference on Computer Vision and Pattern Recognition (2010).
11. Navneet Dalal and Bill TriggsIEEE Computer Society Conference on Computer Vision and Pattern Recognition **1**, 886-893 (2013).
12. Laptev. International Journal of Computer Vision**64**(4), 107–123 (2005).
13. N.Dalal and X.He. IEEE computer society press **1**, 225-232(2005).
14. Santhoshkumar, R., Kalaiselvi Geetha, M.International Journal of Scientific Research in Computer Science Applications and Management Studies. **7**(3), 1-11(2018).
15. Santhoshkumar, R., Kalaiselvi Geetha, M., Arunnehru, J. International Journal of Pure and Applied Mathematics **117**(15), 621-634 (2017).
16. Santhoshkumar, R., Kalaiselvi Geetha, M., J.Arunnehru, J. International Journal of Pure and Applied Mathematics**117**(15), 1185-1194 (2017).