



Facial Expression Analysis Using Multi-label Classification

Trasha Gupta¹, Vidhan Garg², Sarika Kaushal² and Sumedha Seniaray^{1*}¹Department of Applied Mathematics, Delhi Technological University, Delhi, India; ²Delhi Public School, R.K. Puram, Delhi, India

E-mail/Orcid Id:

TG,  trashagupta@dtu.ac.in,  <https://orcid.org/0000-0003-0121-4644>; VG,  vidhangarg12@gmail.com,  <https://orcid.org/0009-0004-8317-4854>;
SK,  sarikakaushal@gmail.com,  <https://orcid.org/0009-0001-3911-3024>; SS,  sumedhaseniaray@dtu.ac.in,  <https://orcid.org/0000-0001-9574-2928>

Article History:

Received: 09th Jun., 2024Accepted: 24th Jul., 2024Published: 30th Jul., 2024

Keywords:

Compound emotions, emotion recognition, facial expressions, multilabel classification

How to cite this Article:

Trasha Gupta, Vidhan Garg, Sarika Kaushal, Sumedha Seniaray (2024). Facial Expression Analysis Using Multi-label Classification. *International Journal of Experimental Research and Review*, 41(spl.), 217-226.

DOI:

<https://doi.org/10.52756/ijerr.2024.v41spl.018>

Abstract: Emotions form the reflection of an individual's mental state. Automatic expression analysis and its recognition based on their facial landmarks are of critical use, particularly to understand the emotional states of physically disabled individuals like autistic children, deaf, dumb, and bedridden individuals. This approach helps in interpreting human emotions, which is crucial for designing computational models and interfaces that cater to their needs. While traditional research has focused on six basic emotion categories (happiness, surprise, sadness, anger, fear, disgust), your work expands this by exploring compound emotion categories. Compound emotions combine these basic categories to create nuanced emotional states. The research utilizes the CFEE_Database_230 dataset of facial expressions captured for analysis and training purposes. The proposed methodology has the following three steps: Analyze the dataset and extract the region of interest (ROI), Extract various statistical and potentially discriminating features and apply a multi-label classification approach to categorize sets of emotions. This involves comparing the feature values across different emotion classes and assigning appropriate labels, particularly focusing on compound emotion categories. Also, the research employs different classifiers and metrics to evaluate the effectiveness of the model. After applying the classification methods, the results are analyzed using various metrics to assess the accuracy and effectiveness of emotion recognition based on facial landmarks. Based on the findings, the Binary Relevance Method yielded the best performance with Mean Average Precision and hamming Loss of 0.7227 ± 0.0344 and 0.1906 ± 0.0227 , respectively. Overall, the work contributes to advancing automatic emotion recognition by considering a broader range of emotional categories beyond the traditional basics. This is particularly beneficial for populations such as the physically disabled and autistic children, where traditional communication methods might be limited or challenging.

Introduction

FACES are keys to individual emotion; they play a significant role in communication. A person's facial expressions are one of the essential ways to communicate. With 43 different muscles, our face can make over 10,000 speeches, many returning to our ancient roots. Although each facet has its unique way of expressing emotions, several selected expressions are always present, regardless of a person's age, race, language, or religion, which is the emotional category of words.

Mehrabian and Ferris, 1967 are a renowned psychologist, found in his research that the emotions people classify as basic are being categorized. We have found that only 7% of the total sensory data are passed through language, and our language resource transports 38%, the Co-Editor who coordinates the review of this manuscript and authorizes it to be published by Rosalia Maglietta. which varies from culture to culture, such as rhythm, tone, tone, etc. To date, the highest percentage of facial data is 55%.

There is a significant amount of research on facial



expressions in computer vision. Perhaps the most fundamental problem in this area is categorizing facial expressions to extract information about the underlying emotional states. In this paper, we have considered the problem of classifying the different compound emotions based on the relative change in people's facial expressions. Most research is limited to these sets of the 6 basic emotions: 'Neutral', 'Happy', 'Sad', 'Angry', 'Surprised' and 'fearful.'

In the present work, we demonstrate that the compound emotion class's production and visual perception constitute the combination of the given basic emotions. The problem of multilabel arises as the repertoire of facial expressions typically used by humans is better described using a rich set of basic and compound categories rather than a small set of basic elements. So, the problem of multi-label emotion classification is of great relevance as humans express their feelings and emotions not just through basic emotions (Mehrabian and Ferris, 1967; Sharma et al., 2023) in the real world. Human emotions can be expressed through several combinational emotions, too.

The proposed methodology can be defined in the following three steps: The first step analyzes the dataset and extracts the region of interest (ROI). The ROI forms the active marks on the face when an individual alters their emotions and facial expressions. In the next step, we extract various statistical features to be able to differentiate the facial landmarks in different individual's emotional states. Finally, in the third step, we apply multi-label classification algorithms to compare the difference in the descriptor's value-set for differentiating emotion classes and then label them as a compound emotion.

This paper uses a machine learning algorithm to train the model to differentiate the emotional category. One of the most widely used techniques for reading machine-reading techniques is content classification, which is used in many cases, such as saying that a given movie review is good or bad or that a cat or dog is in the picture. This work may be divided into three domains: two subdivisions, multiple subdivisions, and multiple label divisions. As our problem seeks to split emotions into an integrated format, it is a problem for many labels that businesses in the database can label multiple. For example, if we talk about the movie genre, then a movie can come under numerous genres as a movie can constitute drama, romance and comedy together.

Similarly, we are supposed to classify the facial expressions as combinations of basic ones. So, we can label an emotion that may be a combination of 2 or maybe 3 basic emotions. The dataset utilized is Compound Facial Expressions of Emotion

(CFEE_database_230), a facial expression database with around 6000 great diverse facial images that constitute mixed emotions images. The class of compound emotions considered is Happily Surprised, Anger Surprise, Fearfully Surprised, SadFear, SadAnger, SadSurprise, FearAnger.

These classes of emotions are more realistic, and we as humans express these emotions very frequently in our day-to-day lives. For instance, Angrily surprised is a compound emotion class. We experience such emotion when a surprising event or action makes us angry, for example, unexpectedly and without provocation, a friend insults us.



Figure 1. Basic Emotions.

Basic Emotion

Many different types of emotions influence the way we live and interact with others. Psychologists have also tried to identify the various kinds of emotions that people experience. Although there are differences of opinion among experts, most psychiatrists agree that there are at least six mainstream emotions listed here. The basic emotion category means different facial expressions: Neutrality, happiness, sadness, anger, fear, and surprise.

Table 1. Six Basic Emotions with Facial Curves.

Basic emotions	Facial movements
Neutral	The face is relaxed and neutral. No frowns or tension on the face
Happy	The eye muscles tightened. "Raven's feet" are wrinkles around the eyes. The cheeks are raised, and the corners of the lips are raised diagonally.
Sad	The inner corners of the eyebrows are raised eyelids. The corners of the lips are pulled down.
Anger	Eyebrows pulled down, upper eyelids pulled up, and lower eyelids pulled up—the edges of the folded lips. The lips can be tightened.
Fear	Eyebrows are pulled up and together, upper eyelids are pulled up, and the mouth is stretched.
Surprise	The entire eyebrow is pulled up, the eyelids pulled up, and the mouth hangs open pupils dilated

Compound Emotion

Integrated emotions can be created by combining the components of the basic components to create new ones. For example, pleasant surprise and annoyance are two different combined categories of emotions.

Joint emotions are expressed more often than just, for example, being threatened by someone emotionally close to you, or you are really disappointed when something scary makes you feel more than that emotion you may feel will be sad and scared about the event. Such situations are prevalent in our lives, so it is a big job for machines to be able to sense human emotions using facial movements.



Figure 2. Showing 12 Compound Emotions.

Table 2. Compound Emotions.

Compound emotions	Description
Happily Surprised	An emotion we feel when we receive excellent or joyful or unexpected news/outcomes
Anger Surprise	An emotion we feel when a surprising event or action makes you angry
Fearfully Surprised	An emotion we feel when a surprising event makes you fearful.
Sad Fear	An emotion we feel when a fearful event or action makes us sad.
Sad Anger	An emotion we feel when an event or action makes us angry and leaves us disappointed or sad.
Sad Surprise	An emotion we feel when a surprising event makes you sad
Fear Anger	An emotion we feel when a fearful event or action makes us angry.

Multi-Label Classification

Multilabel refers to the problem of classification, in which each sample may belong to a few previously defined categories at once.

- In classifying multiple labels, a training set is made up of conditions associated with a set of labels, and the task is to predict event label sets that can be identified by analyzing training conditions with known label sets.
- For example, the division into several categories makes it possible to assume that each sample is assigned to only one label: the fruit can be orange or pear but not both simultaneously. An example of the division of many labels would be that a movie could be of any kind, such as romance, comedy, drama, or action at the same time.
- The difference between multiple categories and multiple labels is that in the problems of many classes, the classes are different. In the problems of many labels, each label represents a different classification function, but the functions are related in some way.

The problem considered in this paper is Emotional analysis using multi-label classification. In the present work, we demonstrate that the production and visual perception of compound emotion class constitutes the combination of the given basic emotions. The problem of multilabel arises as the repertoire of facial expressions typically used by humans is better described using a rich set of basic and compound categories rather than a small set of basic elements. The state-of-the-art research works for facial emotion recognition and considers facial emotion recognition as a single class classification with compound emotions (comprising multiple basic emotions) as a single label for a given instance. Rather, we break the compound emotions into basic emotions and assign multiple basic emotions to the instance. So, we have multiple labels associated with a given sample. We have used intricate facial markers to extract the statistical features of the problem. Also, we have not limited the labels to only two emotions for a sample, as in the case of compound emotions. A sample may have more than 2 labels, for example {disgusted, angry and sad}.

Applications

Human emotions are not limited to the six basic categories of emotions but rather an important category of mixed emotions that clearly and informally express the attitudes and personalities that we constantly express ourselves as human beings. I find that this category of emotions is broad. An important application for the medical profession. Examples include classifying people with physical disabilities (deaf, mute, and bedridden) as

well as children's emotional states of Autism based on global facial features. So, reading someone else's emotions can have several applications for psychologists as they can better understand their patient's attitudes and feelings, especially in the case of patients who cannot express what they are feeling directly in words or actions. Along with the given problem of detecting a combined category of emotions, it increases the spectral of limited sensory acquisition.

Problem Statement

The scenario is about classifying different compound emotions, which are combinations of the basic emotions, into various classes. While previous research works identified a facial condition associated with one label that is not enough to express the complete feeling of a person, in this paper, we have studied an essential class of facial expressions and observed broken compound emotions.

For instance, the facial expression of a happily surprised when we feel delighted at an amazing birthday party. So, here we have identified 7 classes of compound emotions that are consistently produced across cultures, suggesting that the number of facial expressions is higher than previously believed. The hypothesis is that human emotions are not restricted to the class of the basic six emotions. Instead, an essential class of compound emotions exists that is more expressive and informative about a person's mental state and mood, which we humans express almost regularly. It is of significant use to detect such a class of emotions. Most studies have focused on the six most common emotions experienced in many lands - happiness, surprise, sadness, anger, fear, and neutrality- while this paper explores the category of combined emotions as a multilabel classification problem.

Related work

Emotion recognition from facial expressions is a challenging problem with many applications (Dalvi et al., 2021; Sharma et al., 2024). The related work of the problem typically focuses on categorizing a set of predetermined facial motions, such as in FACS (facial action coding system), which are all visually distinguishable facial movements.

The authors developed the Face Coding System (FACS) to explain subtle changes in facial expressions (Ekman and Friesen, 1978). FACS consists of 49 action units, including those of the head and eye. Thirty of these

are related to body composition and the shortening of a particular set of facial muscles. Although there are only a small number of atomic action units, more than 7,000 active combinations have been detected. FACS provides the information needed to define facial expressions.

Another proposed method to solve the problem was to divide the face into lower, middle, and upper levels (Black and Yacoob, 1995). At lower levels, we take the regions corresponding to the face, mouth, eyebrows, and eyes and model the strongest and non-strongest of these regions using a set of parameter flow models and unstable facial movements using a set of local models with optical flow parameters.

These approaches are a bit complex as action units and are a difficult problem because there are no plural definitions and they come from complex combinations in this paper (Tian et al., 2000) rather than using the prior method of action units to recognize any respective emotion, we have approached the problem by a different procedure, here we firstly detected the facial landmarks which further differentiated the face into 8 facial regions which are the most sensitive areas when facial emotion changes thus are of greater relevance. The following 8 facial regions are mouth, inner mouth, right eye, left eye, right eyebrow, left eyebrow, nose and Jaw. Then, the respective features from each of the face regions are considered tools to classify facial emotions.

The research work performed facial Micro-expression analysis using Multi-label Classification (Zhang et al., 2021). Jiang and Deng (2023) used a class activation map to analyze the multi-label compounded Facial expression recognition. They improved the interpretability of the model. Kollias (2023) also proposed a multi-task learning method for facial compound expression recognition. They also contributed a video-based dataset C-EXPR-DB, consisting of 400 videos of 200K frames. The study used Multi-Label Graph Convolutional Networks to identify the Dynamic Facial Expressions for basic emotions (Wang et al., 2022).

Materials and methods

This section explains the step-by-step methodology used in this work. It consists of three phases: (1) Acquire the dataset and identify the critical facial landmarks that give key facial features or regions. (2) Extracting facial features. (3) Classification of extracted features. The Figure 3 illustrates the flow of the methodology.

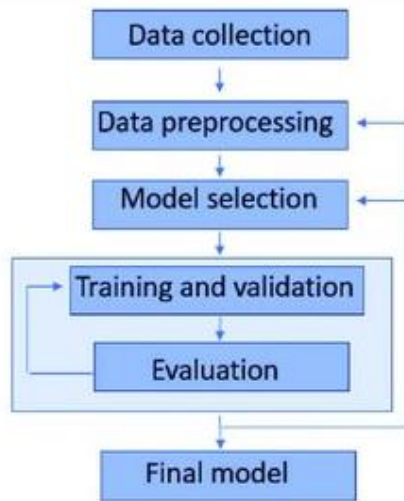


Figure 3. Flow of the methodology.

Dataset

Compound Facial Expressions of Emotion (CFEE_database_230) is a large face website with 5972 different faces, freely available online. The face images of individuals have variations in age, education, gender and race, shape, light conditions, closure (e.g., glasses, facial hair or closure), post-processing activities (e.g., various filters and special effects), etc. CFEE_DB has a wide variety, large numbers, and rich annotations. It comprises two different subset sets: one subset is annotated with one basic emotion; another sub-set is annotated with 12 categories of integrated emotions like HappySurprise, AngerSurprise, FearSurprise, SadFear, SadAnger, SadSurprise, FearAnger. We have broken these labels into separate labels for instance the label “HappySurprise” is read as labels [Happy, Surprise] to convert the problem into multilabel classification.

Facial Landmark Detection

Face detection is a computer-aided operation where we want to detect and track key points on a person's face. Face detection is a subset of the shape prediction problem. Shape forecast attempts to identify critical points of interest in a given image (and usually an ROI describing an object of interest) (Saha et al., 2015). Our goal in the context of global facial features is to use predictive methods like posture predictors (Beumer et al.) to identify the face and critical facial features. Face landmarking is the recognition and localization of specific vital points on the face, perhaps the most crucial stage in the face-processing process (Figures 4 and 5). The issue is locating several landmarks on a face based on a photograph of the face. The eye corners, nose tip, nostril corners, mouth corners, terminal points of the eyebrow arcs, jaw and other landmarks are commonly employed.

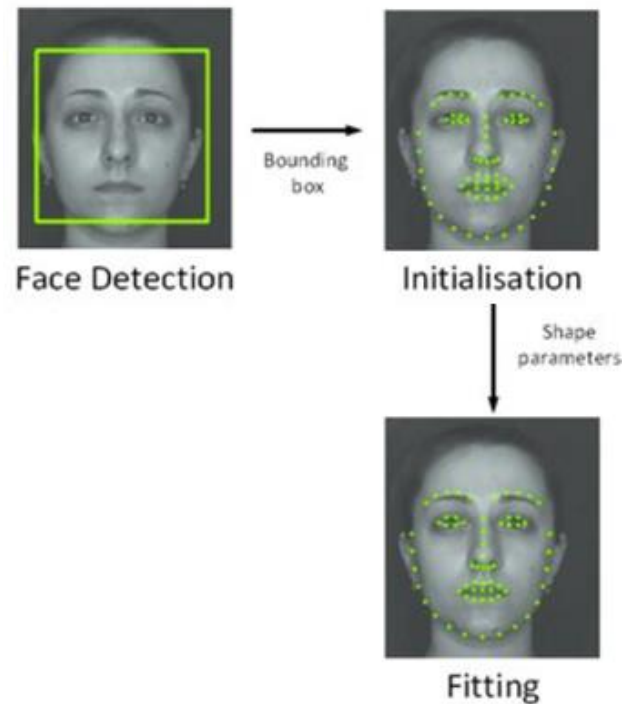


Figure 4. Flowchart For Facial Landmark Detection (Valstar, 1015).

The *Dlib* library's pre-trained facial landmark detector estimates the location of 68 (x, y)- coordinates corresponding to facial structures. Detecting the face's Region of Interest entails segmenting the face into the parts we're interested in, which are the most active on the face when a person alters their facial expressions.

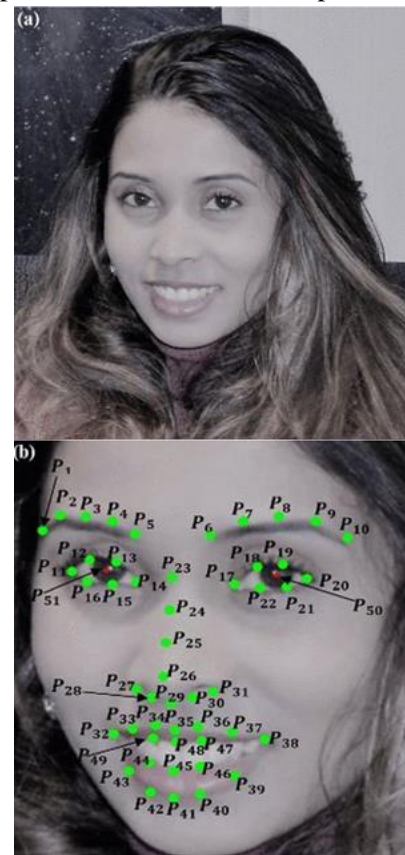


Figure 5. Detecting facial landmarks (Ugail and Al-Dahoud, 2018).

Feature Extraction

Feature extraction is a technique for extracting the visual content of photographs so that they may be indexed and retrieved. Feature extraction is a term that refers to a piece of data that is useful in completing a computational task for a particular application. The purpose is to turn the image into quantitative data that can be further processed for emotion labelling, classification and recognition.

First-order Statistics (FOS)

First-order feature extraction is a retrieval approach based on image histogram features. The histogram depicts the likelihood of a given value of grayscale pixel degree occurring in a picture. Following first-order parameters are derived from the numbers provided by the histogram: Minimal Gray Level, Maximal Gray Level, Mean, Variance, Kurtosis, Energy, Entropy, Coefficient of Variation, 10th Percentile, 25th Percentile, 75th Percentile, 90th Percentile, Median, Mode, Skewness and Histogram Width.

Gray level Co-occurrence Matrix (GLCM)

The GLCM functions calculate the frequency of pairs of given pixels and the specified local relationships from an image (Mohanaiah et al., 2013). First, we constructed the GLCM and then subtracted the statistical measurements from this matrix to reflect the image's texture (Mohanaiah et al.). After creating GLCMs with *graycomatrix*, we used *graycoprops* to get more second-order statistics from them. We calculated the following features from GLCM: Contrast, Correlation, Inverse Difference Moment, Difference Variance, Difference Entropy, Information1, Sum Average, Sum Variance, Angular Second Moment, Sum Entropy, Sum of Squares Variance, Entropy, Information 2 and Maximal Correlation Coefficient.

Local Binary Pattern (LBP)

Local Binary Pattern (LBP) is a primary but effective texture operator that labels pixels in an image by shortening the area of each pixel. This method is commonly used to investigate their local qualities and to discover the characteristics of specific image sections. We calculated the following features from LBP: Energy, Entropy and Mean features (Murugappan and Mutawa, 2021).

Multi-Label Classifier

It is used when there are two or more categories, and the data we want to classify may be one class or all at once, e.g., to distinguish what road signs are contained in the image or to classify the multiple emotions on a human face.

Multilabel K-Nearest Neighbor (MLKNN)

MLKNN is performed in two phases. In the first

phase, the prior probabilities and conditional probabilities are calculated as follows.

- 1) Compute the distance between training instances.
- 2) Then compute the prior probability $P(C_i)$ and $P(\sim C_i)$ for the i th class C_i .
- 3) Then Compute the Conditional probability for the i th class C_i , $P(k|C_i)$ and $P(k|\sim C_i)$, ($0 < k \leq \text{Num}$), i.e., k nearest neighbours of an instance in C_i and not in C_i , respectively, that will belong to C_i are computed.
- 4) The number of the ' Num ' nearest neighbours of the i th instance, which belong to the j th instance, is computed.

In the testing phase, the distance between the training and testing instances is calculated first, as in training. Then, Num neighbours are calculated for each testing instance. Finally, the number of the Num nearest neighbours of the i th instance, which belongs to the j th instance, is computed. Then, the posterior probability is calculated for the instances in C_i and not in C_i belonging to C_i .

Back Propagation Multilabel Learning (BPMLL)

Let the training set be composed of m training samples $\langle x_i, Y_i \rangle$ where x_i is of d -dimension and Y_i are the labels associated with x_i . A single-layer neural network has d input nodes corresponding to the d -dimensional training set, a hidden layer with M nodes and Q output units corresponding to Q classes. The input layer is fully connected to the hidden layer with weights $V = [v_{ih}]$ where $i \in [1, d]$ and $h \in [1, M]$. The hidden layer is full connected to the output layer with weights $W = [w_{ho}]$ where $h \in [1, M]$ and $o \in [1, Q]$. The bias parameters of hidden units b_h and output units p_o are weights from extra units of fixed value 1. For training example x_i with associated labels as Y_i , output of the o th output unit, $c_o = \tanh(g_h w_{ho} + p_o)$ and output of the h th hidden unit, $g_h = \tanh(a_i v_{ih} + b_h)$. To the trained BPMLL network, when an unseen sample is fed, a threshold function decides which labels should be associated with the sample using the actual outputs. The threshold function $t(x)$ is modelled as a linear function, $w_T.c(x) + b$.

Random K Labelsets (RakEL)

RAKEL proposes an approach that iteratively constructs an ensemble of k Label Powerset (LP) classifiers. Each LP classifier is trained using a different small random subset of the set of labels. It takes label correlations into account.

- 1) Randomly break a large set of labels into a number (n) of subsets of small size (k), called k -labelsets.
- 2) For each of them, train a multi-label classifier using the LP method

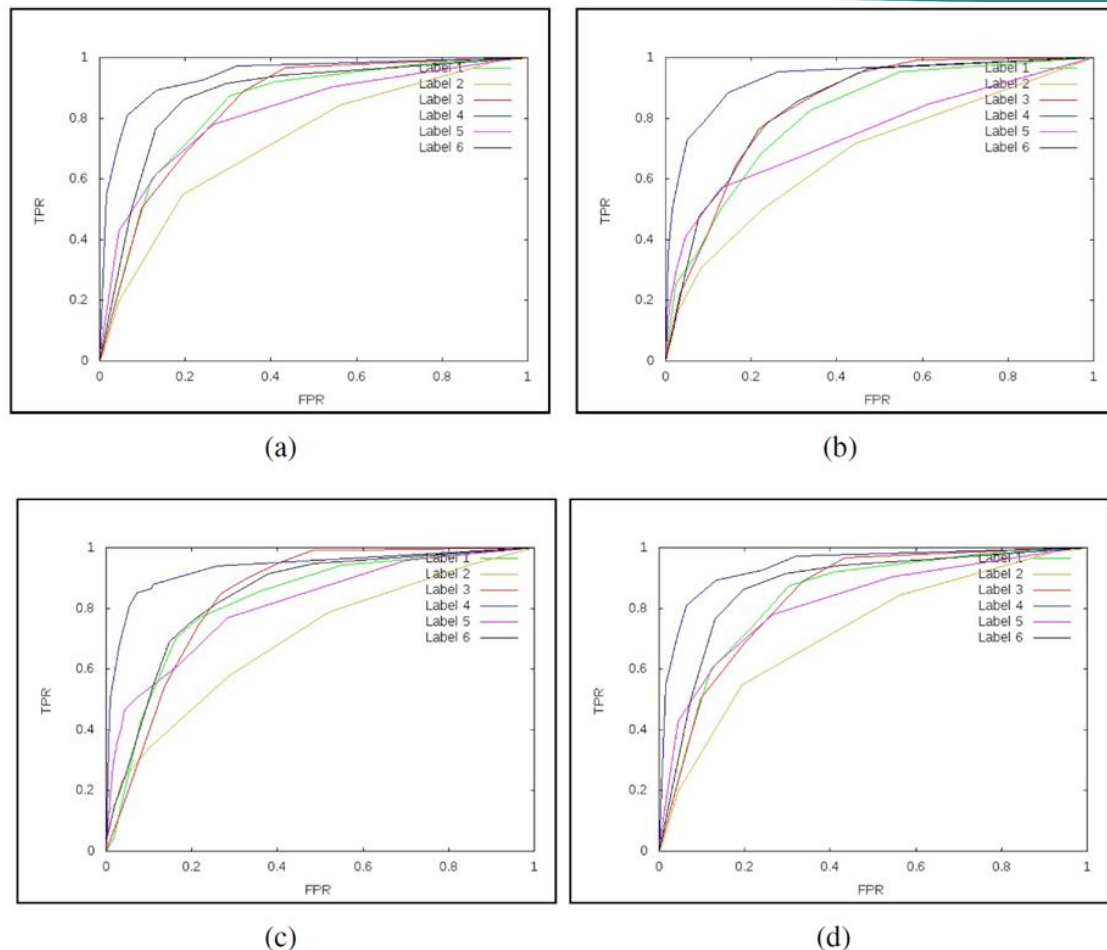


Figure 6. ROC curves for various Multilabel Classifiers (a) MLkNN, (b) BPMLL, (c) RakEL, and (d) BR.

- 3) Given a new instance, query models and average their decisions per label
- 4) Thresholding to obtain the final model

Computationally it breaks into simpler subproblems. In the Ensemble combination phase, the mean number of votes per label is nk/q . The larger it is, the higher the effectiveness. It characterizes RAKEL as an ensemble method. k should be small enough to deal with LP's problems, and n should be large enough to obtain more votes.

Binary Relevance Method (BR Method)

BR is a decay method based on the learning assumption that labels are independent. Therefore, each label is considered appropriate or unimportant according to the two categories studied on that label independently of other labels. Converts multi-labeled L-label problems into L-label separation problems with the same label using the same base separator provided by the manufacturer. Output predictive union for all dividers per label.

Multiple method Analysis (MULAN)

Mulan (Tsoumakas et al.) is a Java library for open-source learning for databases with multiple labels.

Multiple label data sets include training examples of targeted activity with a wide range of targeted binary options. This means that each database item with various labels can be a member of several categories or defined by multiple labels (classes). The library provides the following features:

- Feature selection. Simple, basic methods are currently supported.
- Testing. Classes calculate various test scores by checking for delays and counter-validation.
- Mulan is a library as it is, providing only a structured API for library users.

Results and Discussion

In this section, we performed experiments using four multi-label classifiers to analyze their performance for emotion analysis. We utilized the free source machine learning framework "MEKA" for this. It helps create, test, and assess multi-label and multi-target classifiers. In MLKNN, $k = 5, 10, 30, 50$ neighbors (smoothing factor of 1) are used. In RAKEL, the size of label sets (i.e. 'k') used for experimentation is 2, 3 and 4. BPMLL uses a single hidden layer with 13 hidden neurons, 36 hidden neurons and 50 hidden neurons (100 training epochs and

0.05 learning rate). For sampling 10, 10-fold cross-

We found that the MLKNN performs best for $k = 10$,

Table 4. Algorithms' Best Performances With A subset Of Attributes Selected.

Metrics	BR	MLKNN=66	RAKEL=67
HL	0.1861±0.0189	0.1889±0.0191	0.2187±0.0238
SA	0.2882±0.0397	0.3168±0.0558	0.2647±0.0515
mAP	0.7385±0.0434	0.7279±0.0470	0.6561±0.0496
mAR	0.6243±0.0484	0.6321±0.0269	0.6282±0.0404
mAR	0.6756±0.0376	0.6759±0.0283	0.6412±0.0408
MAP	0.7207±0.0550	0.7417±0.0551	0.6548±0.0499
MAR	0.6049±0.0442	0.6099±0.0258	0.6186±0.0394
MAF	NaN±NaN	0.6383±0.0231	0.6271±0.0407
AP	0.8169±0.0340	0.8053±0.0350	0.7811±0.0350
MAvP	0.7179±0.0373	0.7268±0.0497	0.7013±0.0459
mAAUC	0.8551±0.0239	0.8588±0.0227	0.8313±0.0239
MAAUC	0.8389±0.0272	0.8366±0.0310	0.8165±0.0269

validation was performed for all the algorithms. We analyzed the performance using the following metrics:

Subset Accuracy (SA), Micro-averaged AUC (mAAUC), Macro-averaged Precision (MAP), Micro-averaged Precision(mAP), Micro-averaged Recall(mAR), Micro-averaged F-Measure (mAF), Macro-averaged Recall(MAR), Macro-averaged F-Measure(MAF), Macro-averaged AUC (MAAUC), Mean Average Precision (MAvP), Hamming Loss (HL) and Average Precision (AP).

A data instance can be connected with numerous labels in the multi-label problem. This contrasts the typical task of single-label classification (i.e., multi-class or binary), which assigns a single class label to each instance. The multi-label context is gaining popularity, and it may be used in a wide range of domains, including text, music, photos, and video, as well as bioinformatics. The result of our experiment with different classifiers is as follows :

RAKEL for $k = 4$, BPMLL for the single hidden layer with 36 neurons and BR classifier. Among all four algorithms, BR performs better in the majority of measures (Table I). The ROC plot for all four algorithms is shown in Figure 6 for all the multiple labels.

The results presented in this study reveal that, in the case of visual acuity, its features contain essential information about emotions that cannot be extracted from visual information. This unwanted information is necessary to improve the functioning of the emotional alert system. Most importantly, the caption-based features have made reading small details from facial areas possible whenever emotions are expressed through the facial emotions category.

Conclusion and Future Scope

To examine the class of compound emotions, this study looked at several parts of the face that make up the

Table 3. Algorithms' Best Performances.

Metrics	RakEL (k =4)	MLKNN (k=10)	BPMLL (Single hidden layer- 36 neurons)	BR
HL	0.2210±0.0333	0.2083±0.0207	0.1993±0.0123	0.1906±0.0227
SA	0.2676±0.0838	0.2510±0.0539	0.3016±0.0369	0.2730±0.0444
mAP	0.6550±0.0633	0.7087±0.0344	0.6638±0.0215	0.7335±0.0474
mAR	0.6197±0.0584	0.5627±0.0425	0.7312±0.0245	0.6110±0.0481
mAR	0.6357±0.0543	0.6268±0.0366	0.6955±0.0163	0.6658±0.0413
MAP	0.6527±0.0624	NaN±NaN	0.6625±0.0223	0.7060±0.0492
MAR	0.6150±0.0508	0.5394±0.0354	0.7157±0.0238	0.5933±0.0435
MAF	0.6245±0.0510	NaN±NaN	0.6797±0.0140	NaN±NaN
AP	0.7812±0.0481	0.7883±0.0278	0.8053±0.0359	0.8211±0.0324
MAvP	0.6991±0.0634	0.6940±0.0381	0.7135±0.0280	0.7227±0.0344
mAAUC	0.8296±0.0365	0.8368±0.0186	0.8587±0.0193	0.8553±0.0210
MAAUC	0.8147±0.0415	0.8144±0.0262	0.8450±0.0214	0.8400±0.0244

prominent features. Detecting facial emotion is an intrinsically challenging problem due to individual variability in facial features, as well as ethnic and cultural differences. Detecting compound emotions is even more difficult because the distinction between dominant and complimentary emotions is often thin. The class of compound emotions is frequently misclassified, including by humans. However, the findings of this work suggest that a careful investigation of a face's region of interest (ROI) can clear up most of these misunderstandings.

This paper analyzes the winner's approach—segmenting the face into the 8 crucial sub-parts that are most effective regions when a person's facial emotion changes and detecting the image texture-based features that can differentiate the facial details in a particular emotional state. The proposed method includes three significant contributions:

- Spot the facial landmarks and derive the suitable ROI
- Detecting the features for every sub-region of the face
- Analyzing the difference in the features and classifying and labeling the emotion class.

The results presented in this study suggest that it is possible to recognize a person's complex emotional state rather than just the basic emotion class, which is ineffective in expressing the person's true feelings and catches limited information visually.

Conflict of interest

We hereby declare that all the authors contributed equally to this work.

References

- Beumer, G. M., Tao, Q., Bazen, A. M., & Veldhuis, R. N. (2006). A landmark paper in face recognition. IEEE, In *7th international conference on automatic face and gesture recognition (FGRO6)*, 2006, 6-78. <https://doi.org/10.1109/FGR.2006.10>
- Banerjee, M., Goyal, R., Gupta, P., & Tripathi, A. (2023). Real-Time Face Recognition System with Enhanced Security Features using Deep Learning. *Int. J. Exp. Res. Rev.*, 32, 131-144. <https://doi.org/10.52756/ijerr.2023.v32.011>
- Black, M. J., & Yacoob, Y. (1995). Tracking and recognizing rigid and non-rigid facial motions using local parametric models of image motion. IEEE, In *Proceedings of IEEE International Conference on Computer Vision*, pp. 374-381.
- Dalvi, C., Rathod, M., Patil, S., Gite, S., & Kotecha, K. (2021). A survey of ai-based facial emotion recognition: Features, ml & dl techniques, age-wise datasets and future directions. *IEEE Access*, 9, 165806-165840. <https://doi.org/10.1109/ACCESS.2021.3131733>
- Ekman, P., & Friesen, W. V. (1978). *Facial action coding systems*. Consulting Psychologists Press. <https://doi.org/10.1037/t27734-000>
- Jiang, J., & Deng, W. (2023). Improving Multi-Label Facial Expression Recognition with Consistent and Distinct Attentions. *IEEE Transactions on Affective Computing*, pp. 1-11. <https://doi.org/10.1109/TAFFC.2023.3333874>
- Kollias, D. (2023). Multi-label compound expression recognition: C-expr database & network. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 5589-5598. <https://doi.org/10.1109/CVPR52729.2023.00541>
- Mehrabian, A., & Ferris, S. R. (1967). Inference of attitudes from nonverbal communication in two channels. *Journal of Consulting Psychology*, 31(3), 248. <https://doi.org/10.1037/h0024648>
- Mohanaiah, P., Sathyanarayana, P., & GuruKumar, L. (2013). Image texture feature extraction using GLCM approach. *International Journal of Scientific and Research Publications*, 3(5), 1-5.
- Murugappan, M., & Mutawa, A. (2021). Facial geometric feature extraction based emotional expression classification using machine learning algorithms. *Plos One*, 16(2), e0247131. <https://doi.org/10.1371/journal.pone.0247131>
- Saha, P., Bhattacharjee, D., De, B. K., & Nasipuri, M. (2015). An approach to detect the region of interest of expressive face images. *Procedia Computer Science*, 46, 1739-1746. <https://doi.org/10.1016/j.procs.2015.02.123>
- Sharma, A., Yaduvanshi, E., Sharma, A., & Saha, P. (2024). Mitigating SAD States and Maladaptive Coping in Law Enforcement: Enhance Emotional Competence. *International Journal of Experimental Research and Review*, 40(Spl Volume), 132-141. <https://doi.org/10.52756/ijerr.2024.v40spl.010>
- Sharma, R., Mohan, M., & Gupta, S. (2023). Emotions in retail setting: a systematic literature review based on current research. *Int. J. Exp. Res. Rev.*, 30, 416-432. <https://doi.org/10.52756/ijerr.2023.v30.039>
- Tian, Y. L., Kanade, T., & Cohn, J. F. (2000). Recognizing lower face action units for facial expression analysis. IEEE, In *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580)*, pp. 484-490.

- Tsoumakas, G., Spyromitros-Xioufis, E., Vilcek, J., & Vlahavas, I. (2011). Mulan: A Java library for multi-label learning. *The Journal of Machine Learning Research*, 12, 2411-2414.
- Ugail, H., & Al-Dahoud, A. (2018). Is gender encoded in the smile? A computational framework for the analysis of the smile driven dynamic face for gender recognition. *The Visual Computer*, 34(9), 1243-1254. <https://doi.org/10.1007/s00371-018-1494-x>
- Valstar, M. (2015). Automatic facial expression analysis. *Understanding Facial Expressions in Communication: Cross-Cultural and Multidisciplinary Perspectives*, pp. 143-172. https://doi.org/10.1007/978-81-322-1934-7_8
- Wang, K., Lian, Z., Sun, L., Liu, B., Tao, J., & Fan, Y. (2022). Emotional reaction analysis based on multi-label graph convolutional networks and dynamic facial expression recognition transformer. In *Proceedings of the 3rd International on Multimodal Sentiment Analysis Workshop and Challenge*, pp. 75-80. <https://doi.org/10.1145/3551876.3554810>
- Zhang, L., Arandjelovic, O., & Hong, X. (2021). Facial action unit detection with local key facial sub-region based multi-label classification for micro-expression analysis. In *Proceedings of the 1st Workshop on Facial Micro-Expression: Advanced Techniques for Facial Expressions Generation and Spotting*, pp. 11-18. <https://doi.org/10.1145/3476100>.

How to cite this Article:

Trasha Gupta, Vidhan Garg, Sarika Kaushal, Sumedha Seniaray (2024). Facial Expression Analysis Using Multi-label Classification. *International Journal of Experimental Research and Review*, 41(spl.), 217-226.

DOI : <https://doi.org/10.52756/ijerr.2024.v41spl.018>



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.