American Journal of Engineering Research (AJER)
e-ISSN: 2320-0847 p-ISSN : 2320-0936
Volume-7, Issue-2, pp-236-249
www.ajer.org
Research Paper Open Access

A Review of Facial Expression Recognition in Human Using Particle Swarm Optimization and Support Vector Machine

[1]agboola,F.F; [2]dr.Oye, N. D.
[1]*Department of Computer science, MAUTECH, Yola- Nigeria+2348053269499*
[2]*Department of Computer Science, MAUTECH- Yola Nigeria, +2347037460352*
Corresponding author: agboola,F.F

ABSTRACT *Facial expression is a natural nonverbal communication language. A person can express his or her sentiments/ state of mind through facial expressions but sometimes these expressions are not good enough for recognition systems, they have to be more refined to get right results. This issue still needs an attention, but many algorithms have been proposed so far to handle these vague expressions. Consequently, developing a robust facial expression recognition system which can recognize facial expression in humans and can serve as an important component of natural human-machine interfaces is highly required. Support Vector Machine (SVM) among other algorithms has a very good generalization capability and dynamic classification scheme which makes it suitable for facial expression recognition. Support vector machines have previously been successfully employed in a variety of classification applications including identity and text recognition as well as DNA microarray data analysis.This paper presents the review of particle swarm Optimization (PSO) and support Vector machine. The review takes into consideration, the contribution of many authors on this issue. Facial expression analysis is an interesting and challenging problem and impacts important applications in many areas such as human interaction and data-driven animation. Deriving an effective facial representation from original face images is a vital step for successful facial expressing recognition.*

KEYWORD: Biometrics, Facial expression recognition, Particle swarm optimization, Support vector machine
----
Date of Submission: 12-02-2018 Date of acceptance: 27-02-2018

I. BACKGROUND OF STUDY

Facial expression recognition is a kind of cognitive nonverbal task used in various applications to understand the human internal feelings (Hegde & Seetha, 2017). Emotions are a natural and powerful way of communication among living beings. Humans express their emotions by voice, face, body gestures, and behavioral changes. A reliable emotion perception scheme is required in order to translate human expression and behavioral changes into useful commands to control systems. Emotion recognition is a challenging task because humans do not always express themselves by words and gestures. Automatic human emotion recognition is a multidisciplinary area including psychology, speech analysis, computer vision, and machine learning (Qayyum, Majid, Anwar & Khan, 2017). Currently, with regard to technological advancements, facial expressions recognition has a lot of uses such as communication of humans with computers. With regard to extensive application in this field, automatic facial recognition has received significant attentions in recent years. Although there have been a lot of advancements in this field, facial expression recognition with high accuracy can be hardly achieved because of the existence of complexities and changeability of facial expressions (Gholami, 2017).

Facial expression is considered as a powerful mean of one to one communication after speech signals and plays a pivotal role in human computer interaction (HCI). Human emotional state is provoked by external stimuli resulting in changes in facial dimensions. Development of an affective facial recognition system still remains a challenging task. Facial images and videos are affected by illumination conditions, human age, and variations in how the emotion is expressed (Qayyum *et. al.*, 2017). Several techniques have been adopted to recognize facial expression and most of these work have achieved promising results. This research tends to apply Particle swarm optimization (PSO) along with Support Vector Machines (SVM) for an improved facial expression recognition.

Particle swarm optimization (PSO) is a nature-inspired algorithm that draws on the conduct of flocking birds, social interactions among humans, and the schooling of fish. Specifically, PSO is a metaheuristic algorithm that was inspired by the collaborative or swarming behavior of biological populations (Bouallegue, Haggege & Benrejeb, 2011). PSO is becoming one of the most important swarm intelligent paradigms for solving global optimization problems (Fuzhang, 2016). This algorithm has unfathomable intelligence background and is appropriate for scientific research and engineering application. Therefore, PSO algorithm has triggered the widespread attention of researchers in the field of evolutionary computation, and has attained a lot of research results over the years (Liu, 2016). Similarly, PSO is easy to implement and has been effectively functioned to resolve a varied collection of optimization problems. Thus, because of its easiness and effectiveness in directing huge search spaces for optimal solutions and its dominance with respect to other Evolutionary algorithm techniques; PSO algorithm is engaged in this research to achieve an optimal parameter for SVM classification.

In this research, an evaluation of the performance of the application of Particle Swarm Optimization on Support Vector Machine algorithm for facial expression recognition will be carried out. PSO has an obvious advantages of very high convergence speed. Face database with variation in facial expressions and facial details will be used. Principal Component Analysis (PCA) will also be used to perform dimensionality reduction by projecting into Eigen faces to serve as input image for the PSO model. The PSO will be used to select optimum facial feature subset for Support Vector Machine (SVM) which will perform feature classification. The efficiency of the application of the PSO technique on SVM in facial expressions recognition will be discovered through the evaluation of its recognition rates.

II. STATEMENT OF PROBLEM

Facial expression is a natural nonverbal communication language. A person can express his or her sentiments/ state of mind through facial expressions but sometimes these expressions are not good enough for recognition systems, they have to be more refined to get right results. This issue still needs an attention, but many algorithms have been proposed so far to handle these vague expressions (Hsieh *et al*., 2010). Facial images and videos are affected by illumination conditions, human age, and variations in how the emotion is expressed (Qayyum *et. al.*, 2017).

Most existing facial expression recognition system are computationally expensive in terms of training and recognition time. The inherent property of high dimension in facial images; as a result of highly correlated pixels, leads to redundant information which causes computational burden in terms of processing speed and memory utilization (Babatunde, Olabiyisi, Omidiora & Ganiyu, 2015). Support Vector Machine (SVM) among other algorithms has a very good generalization capability and dynamic classification scheme which makes it suitable for facial expression recognition. Various studies have shown that the performance of SVM drops with increasing data samples (Abdulameer *et. al.*, 2014). Also, SVM consumes large amount memory and time we applied to problems with large complexities. Nevertheless, adequate parameter selection for SVM will not only reduce its computational burden but also increase its generalization capability and accuracy. An improved feature selection technique and swarm-optimized facial expression classification system are major sub-problems that will be addressed.

III. AIM AND OBJECTIVES

The aim of this work is to develop a technique for facial expression recognition in human using Particle Swarm Optimization and Support Vector Machine.

The specific objectives are to:

i. develop a Particle Swarm Optimization technique to optimize the parameter of Support Vector Machine for facial expression recognition in human,

ii. implement the developed technique in Matrix Laboratory (MATLAB) software,

iii. evaluate the performance of the technique using metrics such as false positive rate, sensitivity, specificity, precision, recognition accuracy, training time and average recognition time.

Significance of the Study

Facial expressions play a key role in communication and understanding human behavior. Automatic human emotion recognition is a multidisciplinary area including psychology, speech analysis, computer vision, and machine learning (Qayyum *et. al.*, 2017). Facial expression is considered as a powerful mean of one to one communication after speech signals and plays a pivotal role in human computer interaction (HCI). Facial emotion recognition has a myriad of applications in video security, surveillance, advertising, and robotics.

Especially for human-robot interaction, facial expressions play a key role in communication and understanding human behavior.

Similarly, Facial emotion recognition has opened up a new era for human–computer interaction, and has provided benefits to a wide range of computer vision applications, such as healthcare, surveillance, event detection, personalized learning, and robotics (Zhang et. al., 2015). Robust emotion classification relies heavily on effective facial representation. However, it is still a challenging task for identifying significant discriminative facial features that could represent the characteristics of each emotion because of the subtlety and variability of facial expressions. Fast and accurate facial expression recognition is hence crucial to these applications, particularly with respect to the challenges associated with facial images.

Consequently, developing a robust facial expression recognition system which can recognize facial expression in humans and can serve as an important component of natural human-machine interfaces is highly required. Theswarm-optimized based technique for recognizing facial expression in human will give an improved performance with respect to the evaluation of the performance metrics used.

IV. LITERATURE REVIEW

Facial expression analysis is a rapidly growing field of research, due to the constantly increasing interest in, and feasibility of applying automatic human behaviour analysis to all kinds of multimedia recordings involving people. Applications include classical psychology studies, market research, interactions with virtual humans, multimedia retrieval, and the study of medical conditions that alter expressive behaviour (Valstar, 2014). Given the increasing prominence and utility of expression recognition systems, it is is increasingly important that such systems can be evaluated fairly and compared systematically. Most Facial Expression Recognition and Analysis systems proposed in the literature focus on analysis of expressions from frontal faces. While it can be argued that in many scenarios people’s faces will indeed be largely frontal most of the time, there are also many conditions in which either the camera angle is such that obtaining frontal views is unrealistic, or where the head pose with respect to the camera varies widely over time (Valstar *et. al.*, 2017).

The research in the field of automatic facial expression analysis is focused more to six prototypic emotional facial expressions - fear, sadness, disgust, anger, surprise, and happiness. However, these basic expressions represent only a small set of human facial expressions. In fact, human emotion is composed of hundreds of expressions and thousands of emotional words, although most of them differ in subtle changes of a few facial features.

V. BIOMETRICS

Biometrics refers to the identification of humans by their characteristics or traits. It is used in Computer Science as a research area that deals with identification, access control and surveillance (Jain, Rose & Flynn, 2008). Biometric identifiers are the distinctive measurable characters used to label and describe individuals (Jain, Hong and Pankanti, 2000). They are often categorized as physiological versus behavioral characteristics. Physiological characteristics are related to the shape of the body and include finger print, face, DNA, palm print, hand geometry, iris, retina, and odour /scent. Behavioral characteristics are related to the pattern of behavior of a person, including but not limited to typing rhythm, gait and voice (Jain et al2000). Some researchers have coined the term “behavrometrics” to describe the latter class of biometrics (Jain et al, 2012).

VI. FACE RECOGNITION

Omidiora et al (2008) describe face recognition as an active area of research which has aroused interest of researchers from security, psychology, neuroscience and image processing to computer vision. Furthermore, it is also regarded as one of the biometric techniques that identify people by “who they are” and not by “what they have” or “what they know”.Face recognition identify a specific individual in digital images, surveillances and access restricted cameras. Facial recognition systems are commonly used for security purpose but are increasingly being used in a variety of other applications such as in telecommunication system and in social networks. Currently a lot of facial recognition development is focused on smart phones application. Smart phone facial recognition capacities include image tagging and other social networking integration.

Some facial recognition algorithms identify facial features by extracting landmark or features from an image of the subject’s face, for example, an algorithm may analyze the relative position, size, and or shape of the eye, nose, cheekbone and jaw. These features are then used to search for other images with matching features (Bonsor, 2008). Other algorithms normalize a gallery of face images and then compress the face data, only saving data in the images that is useful for face recognition. Face Recognition algorithms can be divided into two main approaches: Geometric approach and photometric approach. Geometric approach looks at distinguishing features while photometric approach is the statistical approaches that distills an image into valves and compares the value with templates to eliminate variance (Brunelli, 2009). Some approaches define face recognition system as a three step process: face detection, feature extraction and the face recognition phases. The first two could run simultaneously

Face detection is defined as the process of extracting faces from scenes. So, the system positively identifies a certain image region as a face. This procedure has many applications like face tracking, pose estimation or compression. The next step -feature extraction- involves obtaining relevant facial features from the data. These features could be certain face regions, variations, angles or measures, which can be human relevant (e.g. eyes spacing) or not. This phase has other applications like facial feature tracking or emotion recognition. Finally, the system does recognize the face (Zhao et al, 2003).

VII. FACE EMOTION RECOGNITION

Emotions can be described as a discrete and consistent response to internal and external events which have a particular significance on human; emotions are being expressed using the face, the speech or body language. Face emotions is defined as a positive or negative experience that is associated with a particular pattern of physiological activity. However, they are brief in duration and consist of coordinated set of response which may include verbal, physiological, behavioral and neural mechanisms; it's often a driving force behind motivation either positive or negative.

These are the basic types of face emotion according to Paul et al (2002); Anger, Disgust, Fear, Happiness, Sadness, and Surprise. Although these lists of emotion were later expanded by adding a range of positive and negative emotions which includes shame, sensory pleasure, satisfaction, relief, pride in achievement, guilt, excitement, embarrassment, contentment, contempt, and amusement. Some researchers categorize Neutral as part of emotion in fact it's consider as one of basic emotion.

Facial action code system (FACS) is a system to taxonomies human facial movement by their appearance on the face based on a system originally developed by a Swedish anatomist named Carl – Herman Hjortsjo (Paul *et. al*, 2002). Movements of individual facial muscles are encoded by FACS from slight different instant changes in facial appearance (Hamm et al 2011). It is a common standard to systematically categorize the physical expression of emotions, and it has proven useful to psychologists and to animators. Due to subjectivity and time consumption issues, FACS has been established as a computed automated system that detects faces in videos, extracts the geometrical features of the faces and then produces temporal profile of each facial movement (Hamm *et.al*, 2011).

Human coders manually code nearly any anatomically possible facial expression, deconstructing it into the specific action units (AU) and their temporal segments that produced the expression (Freitais, 2012). Action Units are independent of any interpretation; they are used for any higher order decision making process including recognition of basic emotions or pre- programmed commands for ambient intelligent environments. FACS has been proposed for use in the analysis of depression (Reed et al, 2007) and the measurement of pain in patients unable to express themselves verbally (Lints *et. al.*, 2007).

Concept of Facial Expression Recognition System

The main approaches embedded in the components of an automatic expression recognition system are reviewed below.

i) Image Acquisition: Images used for facial expression recognition are static images or image sequences. An image sequence contains potentially more information than a still image, because the former also depicts the temporal characteristics of an expression. With respect to the spatial, chromatic, and temporal dimensionality of input images, 2-D monochrome (grey-scale) facial image sequences are the most popular type of pictures used for automatic expression recognition. However, colour images could become prevalent in future, owing to the increasing availability of low-cost colour image acquisition equipment, and the ability of colour images to convey emotional cues such as blushing.

ii) Pre-processing: Image pre-processing often takes the form of signal conditioning (such as noise removal, and normalisation against the variation of pixel position or brightness), together with segmentation, location, or tracking of the face or its parts. Expression representation can be sensitive to translation, scaling, and rotation of the head in an image. To combat the effect of these unwanted transformations, the facial image may be geometrically standardised prior to classification. This normalisation is usually based on references provided by the eyes or nostrils. Segmentation is concerned with the demarcation of image portions conveying relevant facial information. Face segmentation is often anchored on the shape, motion, colour, texture, and spatial configuration of the face or its components (Yang *et al*, 2002). The face location process yields the position and spatial extent of faces in an image; it is typically based on segmentation results. A variety of face detection techniques have been developed (Yang *et al*, 2002). However, robust detection of faces or their constituents is difficult to attain in many real-world settings. Tracking is often implemented as location, of the face or its parts, within an image sequence, whereby previously determined location is typically used for estimating location in subsequent image frames.

iii) Feature Extraction: Feature extraction converts pixel data into a higher-level representation of shape, motion, colour, texture, and spatial configuration of the face or its components. The extracted representation is used for subsequent expression categorisation. Feature extraction generally reduces the dimensionality of the input space. The reduction procedure should (ideally) retain essential information possessing high discrimination power and high stability. Such dimensionality reduction may mitigate the 'curse of dimensionality' (Jain et al., 2000). Geometric, kinetic, and statistical or spectral-transform-based features are often used as alternative representation of the facial expression prior to classification (Pantic & Rothkrantz, 2000).

iv) Classification: Expression categorisation is performed by a classifier, which often consists of models of pattern distribution, coupled to a decision procedure. A wide range of classifiers, covering parametric as well as non-parametric techniques, has been applied to the automatic expression recognition problem (Pantic & Rothkrantz, 2000). Ekman (1982) defines two main types of classes used in facial expression recognition to be action units (AUs) and the prototypic facial expressions. The 6 prototypic expressions relate to the emotional states of happiness, sadness, surprise, anger, fear, and disgust (Ekman, 1982). However, it has been noted that the variation in complexity and meaning of expressions covers far more than these six expression categories (Lien *et al.*, 1998). Moreover, although many experimental expression recognition systems use prototypic expressions as output categories, such expressions occur infrequently, and fine changes in one or a few discrete face parts communicate emotions and intentions (Tian *et al.*, 2001).

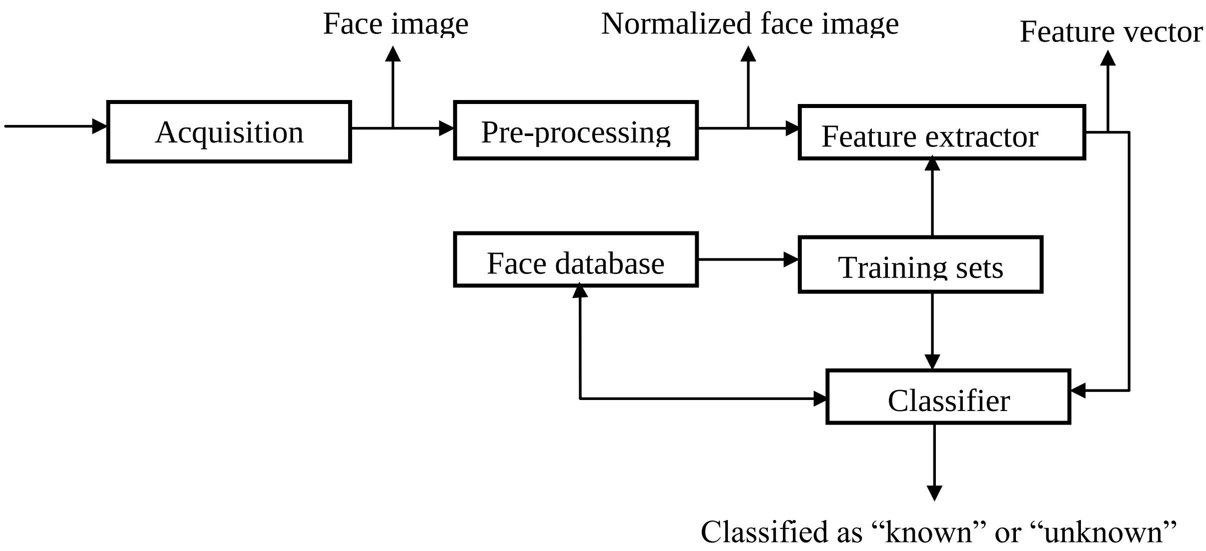

Figure 1: Phases of Face expression recognition system architecture

(Source: Samiksha *et. al.*, 2014)

An AU is one of 46 atomic elements of visible facial movement or its associated deformation; an expression typically results from the agglomeration of several AUs (Donato *et.al.* 1999). AUs are described in the Facial Action Coding System (FACS). Sometimes, AU and prototypic expression classes are both used in a hierarchical recognition system for example, categorisation into AUs can be used as a low-level of expression classification, followed by a high-level classification of AU combinations into basic expression prototypes (Pantic & Rothkrantz, 1999).

v) Post-processing: Post-processing aims to improve recognition accuracy, by exploiting domain knowledge to correct classification errors, or by coupling together several levels of a classification hierarchy, for example. The approaches to facial expression recognition can be divided into two classes in many different ways. In one way, they can be classified into static-image-based approaches and image sequence-based approaches (Yeasin *et al.*, 2006). While the static-image-based approach classifies expressions based on a single image, the image sequence-based approach utilizes the motion information in an image sequence. In another way, they can be classified into geometrical feature-based approaches and appearance-based approaches (Tian et al., 2002). The geometrical feature-based approach relies on the geometric facial features such as the locations and contours of eyebrows, eyes, nose, mouth, etc. As for the appearance-based approach, the whole-face or specific regions in a face image are used for the feature extraction via some kinds of filters or transformations (Mu-Chun *et al.*, 2013). Some approaches can fully automatically recognize expressions but some approaches still need manual initializations before the recognition procedure.

VIII. FACE IMAGE PRE-PROCESSING

Anitha and Radha (2010) described image pre-processing as a necessity and requirement for improving the performance of image processing methods like image transform, segmentation, feature extraction and fault detection. Image preprocessing is the techniques of enhancing data images prior to computational processing. The aim of preprocessing is an improvement of the image data that suppresses unwanted distortion or enhances some image features important for further processing.There are four categories of image pre-processing methods according to the sizes of the pixels' neighborhood that is used for the calculation of new pixel brightness: Pixel Brightness Transformations modify pixels' brightness that is the transformation depends on the classes of brightness corrections: position dependent and gray scale transformations.Position dependent brightness correction is the sensitivity of image acquisition and digitization devices should not depend on position in the image, but this assumption is not valid in many practical cases. Sources of degradation: uneven sensitivity of light sensors and uneven object illumination. Systematic degradation can be suppressed by brightness correction.

Gray Scale Transformation do not depend on the position of the pixel in the image, it can be performed using look up tables. Also it mostly used if the result is viewed by a human. Histogram Equalization: it equally distributes brightness levels over the whole brightness scale.Image pre-processing methods use the considerable redundancy in images. Neighboring pixels corresponding to one object in real images have essentially the same or similar brightness value. Thus distorted pixels can often be restored as an average value of neighboring pixels, one example of this is filtering impulse noise.If pre-processing aims to correct some degradation in the image, the nature of a priori information is important: knowledge about the degradation; knowledge about the image acquisition devices, and condition under which the image was obtained. The nature of noise (usually its spectral characteristic) is sometimes known. Knowledge about objects that are searched for in the image, if knowledge about object is not available in advance it can be estimated during the processing. Miljkovic (2009), applied image pre-processing on images at the lowest level of abstraction and its aim is to reduce undesired distortions and enhance the image data which is useful and important for further processing.

IX. FACIAL FEATURES EXTRACTION

Feature extraction process can be defined as the procedure of extracting relevant information from a face image. This information must be valuable to the later step of identifying the subject with an acceptable error rate. The feature extraction process must be efficient in terms of computing time and memory usage. The output should also be optimized for the classificationstep. Feature extraction involves several steps - dimensionality reduction, feature extraction and feature selection. These steps may overlap, and dimensionality reduction could be seen as a consequence of the feature extraction and selection algorithms. Both algorithms could also be defined as cases of dimensionality reduction (Ion, 2010).Dimensionality reduction is an essential task in any pattern recognition system. The performance of a classifier depends on the amount of sample images, number of features and classifier complexity. One could think that the false positive ratio of a classifier does not increase as the number of features increases. However, added features may degrade the performance of a classification algorithm. This may happen when the number of training samples is small relative to the number the features. This problem is called "curse of dimensionality" or "peaking phenomenon".

Principal Component Analysis (PCA)

One of the features extraction algorithm is principal component analysis PCA, the objectives of principal component Analysis is to discover or reduce the dimensionality of the data set and to identify new meaningful underlying variables. PCA is a classified statistical method which has been widely used in data analysis and compression. It's the simplest of the true eigenvector based multivariate analysis. Often its operation can be thought of as revealing the internal structure of the data in a way that the dataset is visualized as a set of co –ordinates in a higher dimensional data space. PCA can supply the user with a lower dimensional picture, a projection or shadow of this object when view from its most informative view point. This is done by using only the first few principal components so that the dimensionality of the transformed data is reduced.

Kirby and Sirovich (1990) were among the first to apply principal component analysis (PCA) to face images, and showed that PCA is an optimal compression scheme that minimizes the mean squared error between the original images and their reconstructions for any given level of compression. Turk and Pentland (1991) popularized the use of PCA for face recognition. They used PCA to compute a set of subspace basis vectors (which they called "eigenfaces") for a database of face images, and projected the images in the database into the compressed subspace. New test images were then matched to images in the database by projecting them onto the basis vectors and finding the nearest compressed image in the subspace (eigenspace).

Boualleg et al (2006) defined PCA as a useful statistical technique that has found application in fields such as face recognition and image compression, and is a common technique for finding patterns in data of high dimension. It is a way of identifying patterns in data, and expressing the data in such a way as to highlight their similarities and differences. It was discovered that having found these patterns in the data, and being compressed i.e. then PCA reduced the number of dimensions, without much loss of information.

PCA is a mathematical procedure that uses an orthogonal transformation to convert a set of M face images into a set of K uncorrelated variables called eigenfaces. The number of eigenfaces is always less than or equal to the number of original images (i.e K<M). This transformation is defined in such a way that the first eigenface shows the most dominant "direction"/" features", under the constraint that it is uncorrelated to the preceding eigenface. To reduce the calculations needed for finding these eigenfaces, the dimensionality of the original training set is reduced before eigenfaces are calculated.

These eigenfaces (eigenvectors) are in fact the principal components of the training set of face images generated after reducing the dimensionality of the training set. Once they are selected, each training set image is represented in terms of these eigenfaces. Then, an unknown face, used for recognition purposes is represented in terms of the eigenfaces. The eigenface representation of this unknown face is compared with each training set face image. The "distance" between them is then calculated. If the distance is above a specified threshold value, the unknown face is recognized as that person and vice versa.

(i) PCA eigenfaces method considers each pixel in an image as a separate dimension. i.e. N x N image has N^2 pixels, therefore has N^2 dimensions.
(ii) For recognition, a training set (dataset) of face images was introduced.
(iii) All face images have the same size exactly (dimensions)
(iv) PCA eigenfaces method does not work on images directly but rather converts them to matrix (vector) form.

X. SUPPORT VECTOR MACHINE

Support Vector Machines (SVM) are classification and regression methods which have been derived from statistical learning theory (Vapnik, 1995). The concept is based on optimal linear separating hyperplane that is fitted to the training patterns of two classes within a multi-dimensional feature space. The optimization problem that has to be solved relies on structural risk minimization and is aiming at a maximization of the margins between the hyperplane and closest training samples. Support vector machine method classifies both linear as well as non-linear data. It transforms the data into higher dimension. The SVM finds hyperplane using support vectors and margin define by support vector. The data transform into dimension equals to the number of attribute in data. Hyperplane with maximum margin is classifying the data with high accuracy. There is high classification accuracy of support vector machine (Cervantes *et. al.*, 2007).

Given a training set of instance-label pairs $(x_i, y_i), i = 1, 2, . . ., l,$ where $x_i \in R^n$ and $\in y \in \{1, -1\}^l$, the support vector machines (SVM) require the solution of the following optimization problem (Chih-Wei *et. al.* 2003):

$$\min_{w,b,\xi} \frac{1}{2}\boldsymbol{w^T w} + C\sum_{i=1}^{l}\xi_i$$

$$Subject\ to \quad y_i(\boldsymbol{w^T}\phi(x_i) + b) \geq 1 - \xi_i$$

$\xi_i \geq 0.$ (2.1)

From above the training vectors x_i are mapped into a higher dimensional space by the function ϕ. SVM finds a linear separating hyperplane with the maximal margin in this higher dimensional space. $C > 0$ is the penalty parameter of the error term. Moreover, $K(x_i, x_j) \equiv \phi(x_i)^T\phi(x_j)$ is called the kernel function. Despite the fact that new kernels are being proposed by new researchers, SVM uses for basic kernels (Chih-Wei *et. al.* 2003):

- ✓ Linear : $K(x_i, x_j) = x_i^T x_j$ (2.2)
- ✓ Polynomial: $K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0.$ (2.3)
- ✓ Radial Basis Function (RBF) : $K(x_i, x_j) = \exp\left(-\gamma\|x_i - x_j\|^2\right), \gamma > 0.$ (2.4)
- ✓ Sigmoid: $K(x_i, x_j) = tanh(\gamma x_i^T x_j + r)$ (2.5)

Where $\gamma, r,$ and d are kernel parameters.

Kernel functions are used to efficiently map input data that may not be linearly separable to a high dimensional feature space where linear methods can then be applied (Cristianini & Shawe-Taylor, 2000). Previous research work from (Keerthi & Lin, 2003) (Joachims, 1998) showed optimal performance with the polynomial kernel, which lead to this experiment adjusting the degree of the polynomial to try to increase accuracy. For the binary classification case, the optimal hyperplane was a line, independent of the polynomial degree.

The biggest difficulties in setting up the SVM model are choosing the kernel function and its parameter values. If the parameter values are not set properly, then the classification outcomes will be less than optimal (Morik, Brockhausen & Joachims, 1999). In complex classification domains, some features may contain false

correlations, which impede data processing. Moreover, some features may be redundant, since the information that they add is contained in other features. Redundant features can lengthen the computational time, influencing the classification accuracy. Hence, the classification process must be fast and accurate using the minimum number of features, which is a goal attainable through the use of feature selection. Feature selection has been applied to enhance classification performance, and to reduce data noise (Lee & Estivill-Castro, 2007).
If the SVM is adopted without feature selection, then the dimension of the input space is large and non-clean, lowering the performance of the SVM. Thus, the SVM requires an efficient and robust feature selection method that discards noisy, irrelevant and redundant data, while still retaining the discriminating power of the data. Features extracted from the original data are adopted as inputs to the classifiers in the SVM (Shih-Wei *et. al.*, 2008).Support vector machines have previously been successfully employed in a variety of classification applications including identity and text recognition as well as DNA microarray data analysis. Support Vector Machine (SVM), developed from statistical learning theory, is a widely used classifier for facial expression classification (Ying & Zhang, 2009). It has several advantages in solving small sample size, non-linear, or high dimensional classification problems. Bartlett *et al*. (2005) developed a technique using AdaBoost to choose a subset of features extracted from Gabor filters. The SVMs are then employed to perform classification. Their system achieves 93% recognition accuracy in recognizing seven facial expressions on their own database, which is higher than other classification methods tested on the same database.

Evolutionary Algorithms for Feature Selection

Problem solving approaches that use computational models which are based on the principles of natural evolution are called evolutionary algorithms (EA) (Saleem, 2001). In evolutionary algorithms an individual would have to go through the process of selection, recombination and mutation. Then the individuals who have the high fitness values would most likely be selected to form the next generation. Some of the popular evolutionary algorithms are Ant Colony Optimization (ACO), Genetic Algorithms (GA), Particle Swarm Optimization (PSO) and Cultural Algorithms (CA).

Ant Colony

Ant Colony Optimization (ACO) takes inspiration from the foraging behaviour of ants (Dorigo *et al.*, 2006). Ants deposit a chemical substance called pheromone on the path when walking to and from a food source. Other ants sense the presence of pheromone and tend to follow a trail that is rich in pheromone. Thus they are able to find the shortest path from a food source to the nest. Also ants are capable of adapting to changes in the environment. For example, being able to find a new shortest path when the old path is no longer feasible due to obstacles.
In the Ant Colony Optimization Algorithms proposed by Dorigo *et al.*, (2006) the ants are defined as computational agents, which iteratively build solutions to an optimization problem. ACO has been applied to optimization problems in areas such as asymmetric and symmetric travelling salesman problem, scheduling, routing and partitioning problems.
In (Peng, 2005) each ant will move from a state **ι** to another one ψ corresponding to a more complete solution. Each agent at each step will compute a set of feasible expansions to its current state by using the following probability distribution;

$$\rho^{k}{}_{\iota}\Psi = \begin{cases} \dfrac{\alpha.\tau\,\iota\,\psi + (1-\alpha).\eta\,\iota\,\psi}{\sum(\tau\,\iota\,v) + (1-\alpha).\eta\,\iota\,v} & if\ \iota\,\psi\ \epsilon\ tabuk\ \notin tabuk \\ 0, & otherwise \end{cases} \qquad (2.6)$$

Will move to the new position only if it is better than the previous step. **tabuk** represents a set of feasible moves for an agent **k**. Parameter α defines the relative importance of the trail. Then after each iteration **t** of the algorithm, the trails are updated by the following formula,

$$\tau\,\iota\,\psi(t) = \rho\,\tau\,\iota\,\psi(t-1) + \Delta\,\tau\,\iota\,\psi \qquad (2.7)$$

ρ is a user defined co-efficient and Δ τ ι ψ represents the sum of the contributions of all agents that used the move ι ψ to find the next position. The pseudo code from (Maniezzo, 2000) describes how the basic Ant Colony optimization works.

1. (Initialization)
Initialize τ ι ψ, " ι, ψ
2. (Construction)
For each ant k **do**
Repeat
Compute η ι ψ, " ι, ψ
Choose in probability the state to move into append the chosen move to the kth ant 's set tabuk

Until ant k has completed its solution
[Apply a local optimization procedure]
End do
3. (Trail update)
For each ant move (ι, ψ) **do**
Compute Δ τ ι ψ and **update** the trail values
4. (Terminating condition)
If not (end condition) and **go to** step 2.

Particle swarm optimization (PSO)

Particle swarm optimization (PSO) is a nature-inspired algorithm and a population-based optimization technique inspired by the behaviour of schools of fish, herds of animals or flocks of birds (Eberhart & Kennedy 1995). In fish schooling, bird flocking, and human social interactions, the population is called a swarm and candidate solutions, corresponding to the individuals or members in the swarm, are called particles. Birds and fishes normally travel in a group without collision. Consequently, using the group information for finding the shelter and food, each particle adjusts its corresponding position and velocity, representing a candidate solution. The position of a particle is influenced by neighbors and the best found solution by any particle (Ying-Yi, Angelo & Arnold, 2016).

Particle Swarm Optimization (PSO) has been applied to solve a wide range of optimization problems, such as constrained and unconstrained problems, multi-objective problems, problems with multiple solutions, and optimization in dynamic environments. Particle swarm optimization through it application in various domain has proven to be computationally efficient in terms of effective convergence, parameter selection, simplicity, flexibility, robustness, ease of implementations (Poli, Kennedy & Blackwell, 2007), ability to hybridize with other algorithms and many others (Ying-Yi, Angelo & Arnold, 2016).In PSO, individuals are referred to as particles. Each particle has a position represented by a position-vector $\bar{x}_n$ and a velocity represented by a velocity-vector $\bar{v}_n$. At each iteration, each particle changes its searching course based on: its preceding velocity $\bar{v}_n(t)$ its best position usually called *local orpersonal best(* P_b*).* it has met so far and the best position G_b attained so far by all particles in the swarm (called *global best*). That is, each particle updates its velocity and position according to equation (2.7) and (2.8) respectively (Attiya & Zhang, 2017):

$$\bar{v}_{mn}(t+1) = \omega\bar{v}_{mn}(t) + c_1 r_1\left(P_b - \bar{x}_{mn}(t)\right) + c_2 r_2\left(G_b - \bar{x}_{mn}(t)\right) \quad (2.7)$$

$$\bar{x}_{mn}(t+1) = \bar{x}_{mn}(t) + \bar{v}_{mn}(t+1) \quad (2.8)$$

Where $\bar{v}_{mn}(t+1)$ velocities of particle mat iterations n,$\bar{x}_{mn}(t+1)$, positions of particle m^{th} at iterations n^{th}. $\omega$is inertia weight to be employed to control the impact of the previous history of velocities. t denotes the iteration number, c_1 is the cognition learning factor, c_2 is the social learning factor, r_1 and r_2 are random numbers uniformly distributed in [0, 1].

The Stepwise procedure of the PSO algorithm are presented below (Alam *et. al*., 2015):

1. Set parameter ω_{min} , ω_{max} , c_1 and c_2 of PSO
2. Initialize population of particles having positions x_n and velocities v_n
3. Set iteration n = 1
4. Calculate fitness of particles $F_{mn}(t) = f(\bar{x}_{mn}(t))$ and find the index of the best particle b
5. Select $Pbest_{mn}(t) = \bar{x}_{mn}(t)$ and $Gbest_n(t) = x_{bn}(t)$
6. $\omega = \omega_{max} - n \times (\omega_{max} - \omega_{min})/Maxnum$
7. Update velocity and position of particles

$$\bar{v}_{mn}(t+1) = \omega\bar{v}_{mn}(t) + c_1 r_1\left(P_b - \bar{x}_{mn}(t)\right) + c_2 r_2\left(G_b - \bar{x}_{mn}(t)\right)$$

$$\bar{x}_{mn}(t+1) = \bar{x}_{mn}(t) + \bar{v}_{mn}(t+1)$$

8. Evaluate fitness $F_{mn}(t+1) = f(\bar{x}_{mn}(t+1))$ and find the index of the best particle b_1
9. Update $Pbest$ of population
   If $F_{mn}(t+1) < F_{mn}(t)$ then $Pbest_{mn}(t+1) = \bar{x}_{mn}(t+1)$ else
   $$Pbest_{mn}(t+1) = Pbest_{mn}(t)$$
10. Update $Gbest$ of population
   If $F_{bn}(t+1) < F_{bn}(t)$ then $Gbest_n(t+1) = Pbest_{bn}(t+1)$ and set $b = b_1$ else
   $$Gbest_{bn}(t+1) = Gbest_n(t)$$
11. If $n < Maxnum$ then $n = n + 1$ and goto step 6 else goto step 12
12. Output optimum solution as $Gbest_n(t)$.

Review of Related Works

Xiang *et al*, (2008) employed Fourier transform to extract features to represent an expression. For further expression processing the Fuzzy C Means (FCM) was used to generate a spatio-temporal model for each

expression type. In this research the Cohen-kanade database is used for training and testing the model and the best test results was 88.8%. Smile and sadness have the best and worst recognition rates respectively.
Poon Bruce *et.al* (2011) have evaluated PCA based FER methods for distorted images. They are worked with the different database, like CMU (Carnegie Mellon University) and ORL database which is in same illuminating condition and calculate Eigen faces. In this paper, the comparison of recognition rate, which outcome with CMU database is 100% and ORL database is 90%.
Thai *et. al.* (2011) presented a novel approach for canny edge detection, PCA and ANN (Artificial Neural Network). They detect the local region of the face such as an eyebrow, eye and mouth then reduce the dimension using PCA. ANN applied for classification on JAFFE database which have 85.7% recognition rate.
Samad and Hideyuki (2011) have presented use of edge based feature extraction with Gabor Features. Image was filter by Gabor and convolute with multiple edge detector. They were use multiple edge detector because each edge detection problem of manual selection of threshold value. The features are reduced in dimension by PCA and classify by SVM, which resulted 91.7% recognition rate for FEEDTUM (Facial Expressions and Emotions from Technical University of Munich) database among all subject dependent recognition.
Rahulamathavan *et. al* (2013) developed FER with encrypted domain using LFDA (Local Fisher Discriminant Analysis). This method was applied to JAFFE and MUG (Multimedia Understanding Group) database which have a recognition rate respectively 94.37% and 95.24%.
Mliki et.al. (2013) developed a data mining-based facial expressions recognition system. They segment facial feature contours using Vector Field Convolution (VFC) technique, extract facial feature points which go with facial-expression deformations and modelled a set of distances among the detected points to define prediction rules through data mining technique. They recorded 81.42% of classification success rate with five classes and 75.32% with seven classes and also found that the class JOY correspond to SMILE class.
Abdulrahman *et. al*. (2014) propose a method which is implemented using Gabor wavelet transform with PCA and LBP. This hybrid approach gives 90% average recognition rate for JAFFE database. Meher *et. al.* (2014) proposes a PCA for Face Recognition and FER. For PCA classification is matters for performance. In this study, result of recognition rate was 81.36% for CSU (Charles Sturt University) dataset and 85.5% for ATT dataset.
Sobia *et. al.* (2014) investigated FER using PCA based interface for wheel chair. In this study, for pre-processing colour space transform, skin region detection, noise removal and morphological operation apply on facial images. Then based on PCA, Eigen vectors and Eigen faces are calculated and classify expression using Euclidean Distance. This approach resulted 96.667% recognition rate for the 60 Eigen faces for JAFFE database.
Qayyum *et. al.* (2017) proposes a facial expression recognition using Stationary Wavelet Transform Features. In this work, stationary wavelet transform is used to extract features for facial expression recognition due to its good localization characteristics, in both spectral and spatial domains. Feature dimensionality is further reduced by applying discrete cosine transform on these subbands. The selected features are then passed into feed forward neural network that is trained through back propagation algorithm. An average recognition rate of 98.83% and 96.61% is achieved for JAFFE and CK+ dataset, respectively. An accuracy of 94.28% is achieved for MS-Kinect dataset that is locally recorded.
Gholami (2017) proposes a facial expressions recognition based on a combination of the basic Facial Expression Using weighted the Local Gabor Binary Pattern (LGBP). Local Gabor Binary Pattern (LGBP) algorithm were used for facial expression recognition of emotion (happiness, sadness, anger, disgust, surprise and fear). The input images were partitioned into 9 equal areas and extract the LGBP features. KNN classifier was used in this research as a weighting fuzzy and without using fuzzy weighting. The result achieved reveals that the fuzzy mode is 1.66 percent more accurate than other. However, in some of the reviewed work the volume of the dataset has significant effect on the recognition accuracy and other performance metrics. The proposed technique which involves the application of PSO to improve the recognition performance and other metrics such as computation time.

Facial Feature Selection and Classification

This research intends to apply PSO along with SVM in facial expression recognition. The intrinsic property of the fast convergence ability of PSO will be evaluated with SVM. PSO algorithm will be applied with the aim of achieving a faster convergence speed and better precision, specificity and accuracy when optimizing the parameters of SVM. PSO automatically will chose parameter for SVM to reduce the number of feature subset in entire features that will be extracted from the images to reduce noise and redundant data. The Block diagram showing the process flow of the developed technique is shown in Figure 6 below.

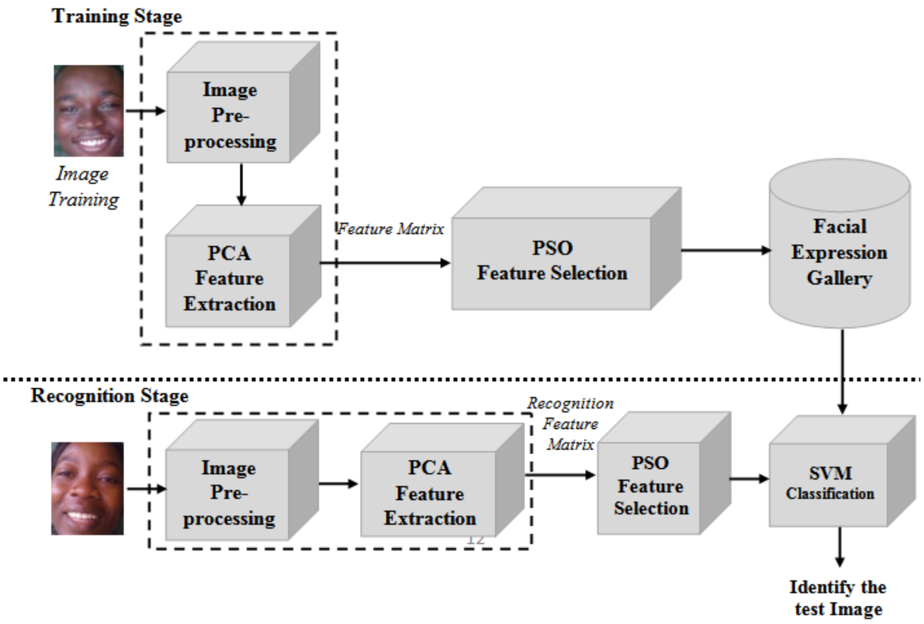

Figure 2: The Block diagram representing the process flow of the Proposed techniques

Implementation for training and testing Facial Expression

In order to determine the performance of the developed technique on the facial expression recognition; this work will be implemented in two phase.

In the learning phase (PSO-SVM) the train dataset will undergo preprocessing stage (i.e. conversion into gray scale and face vector, normalization of the face vector) after which PCA will be used for dimension reduction and extraction of facial features. PSO will be used to perform the selection of the optimum feature subset from the entire facial feature extracted by PCA. The trained dataset will be stored in the gallery.

Duringthetesting phase, the input test image (test dataset) will undergo the pre-processing stage after which PCA will be used for dimension reduction and extraction of facial features. PSO will be used to perform the selection of the optimum feature from the entire facial feature extracted by PCA and is fed to SVM for classification. The result of classification was compared with trained feature stored in the library for recognition. The final output of the system based on the recognition of the test images will be displayed. The stepwise procedure of the first phase is shown in Figure 2. The test dataset will be used to test the performance of the proposed system using SVM classification schemes for both phases. The classification schemes are binary and multiclass classification. The binary classification scheme involves one against one i.e. the datasets were classified in pairs while the multiclass classification involves one against all i.e. multiple binary classification.

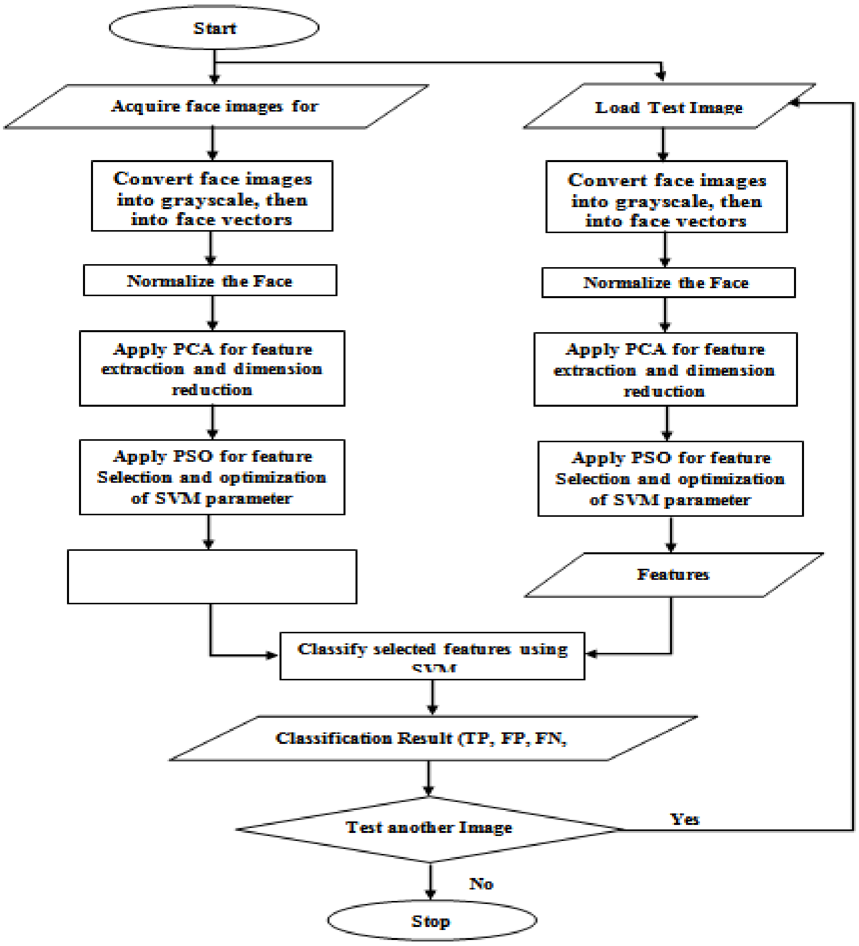

Figure 3: Flowchart showing trained and tested faces with CPSO-SVM

REFERENCES

[1]. Abdulameer M. H., Abdullah H. S. & Othman A. Z. (2014). Support Vector Machine Based on Adaptive Acceleration Particle Swarm Optimization.*Hindawi Publishing Corporation, the Scientific World Journal.* 1-8.

[2]. Abdulrahman, M., Gwadabe, T. R., Abdu, F. J., & Eleyan, A. (2014). Gabor wavelet transform based facial expression recognition using PCA and LBP. *Signal Processing and Communications Applications Conference (SIU)*, pp. 2265-2268.

[3]. Anitha, S. and Radha, V., (2010). Comparison of Image Preprocessing Techniques for Textile Texture Images. International Journal of Engineering Science and Technology, 2(12), pp. 7619-7625.

[4]. Attiya I. & Zhang X. (2017). A Simplified Particle Swarm Optimization for Job Scheduling in Cloud Computing. *International Journal of Computer Applications (0975 – 8887) 163(9), 2017, 34-41.*

[5]. Babatunde R. S., Olabiyisi S. O., Omidiora E. O. and Ganiyu R. A. (2015). Local Binary Pattern and Ant Colony Optimization Based Feature Dimensionality Reduction Technique for Face Recognition Systems. *British Journal of Mathematics & Computer Science*, 11(5), pp. 1-11.

[6]. Bartlett M. S., Littlewort G., Frank M., Lainscsek C., Fasel I., and Movellan J. (2005). Recognizing facial expression: machine learning and application to spontaneous behaviour. *IEEE Computer Society Conference on Computer Visionand Pattern Recognition*, 2(1), pp. 568–573.

[7]. Bonsor. K, (2008). How facial recognition system work. Retrieved 2008-06-02

[8]. Boualleg, A.H., Bencheriet, Ch. and Tebbikh, H., (2006). Automatic Face Recognition Using Neural Network- PCA. *In: Proc. 2nd Information and Communication Technologies ICTTA '06*, Syria: Damascus, pp. 1920 – 1925.

[9]. Bouallegue S., Haggege J. & Benrejeb M. (2011). Particle swarm optimization-based fixed structure H∞ control design. *InternationalJournal of Control, Automation and Systems, 9(2), 258–266.*

[10]. Brunelli.R, (2009). Template matching techniques in computer vision: Theory and Practice, *willey ISBN 978-0-470-51706.*

[11]. Cervantes J., XiaoOu L. & Wen Y. (2007). SVM Classification for Large Data Sets by Considering Models of Classes Distribution, *Artificial Intelligence - Special Session, Sixth Mexican International Conference; MICAI 2007.*

[12]. Chibelushi, C. C., & Bourel, F. (2003). Facial expression recognition: A brief tutorial overview. *CVonline: On-Line Compendium of Computer Vision, 9.*

[13]. Chih-Wei H., Chih-Chung C., & Chih-Jen L. (2003). A Practical Guide to Support Vector Classification. *National Taiwan University, Taipei 106, Taiwan* http://www.csie.ntu.edu.tw/~cjlin. (Last updated: May 19, 2016)

[14]. Cristianini N. & Shawe-Taylor J. (2000). *An Introduction to Support Vector Machines* (Cambridge: Cambridge University Press, 2000).

[15]. Donato G, Bartlett M. S., Hager J. C., Ekman P., Sejnowski T. J. (1999). Classifying Facial Actions. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 21(10), pp. 974-989.

[16]. Dorigo, M.; Birattari, M.; Stutzle, T. (2006). Ant colony optimization. *IEEE Computational Intelligence Magazine,* 1(4), pp. 28 – 39.
[17]. Eberhart R. C, Kennedy J. (1995). "A New Optimizer Using Particle Swarm Theory. *In Proceedings of the 6th International Symposium on Micro-machine Human Science*.: 39–43.
[18]. Ekman P. (1982). "Emotion in the Human Face. *Cambridge University Press*.
[19]. Freitais. M.A, (2012). Micro expression and macro expression. *In US Raamachandran (Ed). Encyclopedia of Human behaviour*. l2:173-183.
[20]. Fuzhang Z. (2016). Optimized Algorithm for Particle Swarm Optimization. *World Academy of Science, Engineering and Technology International Journal of Physical and Mathematical Sciences,* 3(3): 1-6.
[21]. Gholami F. (2017). Facial Expressions Recognition Based on a Combination of the Basic Facial Expression Using Weighted the Local Gabor Binary Pattern (LGBP). *International Journal of Computer Science and Network Security (IJCSNS)*, 17(4): 352-360.
[22]. Hamm J., Kohler C. G., Gur R.C. & Verma. R, (2011). Automated facial action coding system for dynamic analysis of facial expression in neuropsychiatric disorders. *Journal of Neuroscience methods*, 200(2): 237-256.
[23]. Hegde G. P. & Seetha M. (2017). Subspace based Expression Recognition Using Combinational Gabor based Feature Fusion. *I.J. Image, Graphics and Signal Processing*, 1, 50-60.
[24]. Holland, J. H. (1975). Adaptation in natural and artificial systems. *Ann Arbor: The University of Michigan Press*.
[25]. Hsieh C. K., Lai S. H., and Chen Y. C. (2010). An optical flow-based approach to robust face recognition under expression variations. *IEEE Transactions on Image Processing,* 19(1): 233-240.
[26]. Ion. M, (2010) "Face recognition algorithms" A handbook of biometrics springer.pg1-22 ISBN. 978-0-387-71040-2
[27]. Iris. B & Thao. N, (2005). Investigation of Combining SVM and Decision Tree for Emotion classification. Proceedings of the Seventh IEEE International Symposium on Multimedia, 267 : 235-237.
[28]. Jain A, Rose. A. & Flynn.R, (2008). Introduction to Biometrics. *A handbook of biometrics springer ISBN. 978-0-387-71040-2*. 1-22.
[29]. Jain A., Hong. L. & Pankanti. S. (2000). Biometrics identification. *communication of the ACM*, 43(2);91-98.
[30]. Jain. A, Rose. A. & Flynn. R. (2012). Biometrics for secure authentication. retrieved 2012-07-29.
[31]. Jain.A, Duin.R.and Moa. J, (2000) "statistical pattern recognition" A Review IEEE Transactions on Pattern Analysis and machine intelligence 22(1):4-37.
[32]. Joachims T. (1998). Text categorization with support vector machines: learning with many relevant features. In Claire N´edellec and C´eline Rouveirol, editors, *Proceedings of ECML-98, 10th European Conference on Machine Learning*, Springer Verlag, Heidelberg, DE, 1998, 137–142.
[33]. Keerthi S. S. & Lin C. J. (2003) Asymptotic behaviors of support vector machines with Gaussian kernel, *Neural Computing. 15, 2003, 1667–1689.*
[34]. Kirby, M. and Sirovich, L., (1990). Application of the Karhunen-Loeve Procedure for the Characterization of Human Faces, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12, pp. 103-107.
[35]. Lai, L. L., Zhang, Q. P., Wu, Q., Ma, Y. N., Sun, W. C., & Yang, L. (2004). Intelligent weather forecast. *Machine Learning and Cybernetics, 2004*. IEEE *Proceedings of 2004 International Conference,* 7, pp. 4216-4221.
[36]. Lee K. & Estivill-Castro V. (2007), Feature extraction and gating techniques for ultrasonic shaft signal classification, Appl. Soft Comput. 7, 2007, 156–165.
[37]. Lien J. J., Kanade T., Cohn J. F., Li C. C. (1998). Automated Facial Expression Recognition Based on FACS Action Units", *Proc. Third IEEE Int. Conf. Automatic Face and Gesture Recognition*, pp. 390-395.
[38]. Lints .M, Hadjistavropoulos. T, Barber. B, & Gibson. S. J. (2007). A Psychophysical investigation on the facial Action coding system as an index of pain variability among older adults with and without Alzheimer's disease. *Pain medicine, IEEE Conference on Computer Vision and Pattern.*
[39]. Liu Yi. (2016). Study on an Improved PSO Algorithm and its Application for Solving Function Problem. *International Journal of Smart Home* 10 (3), 51-62.
[40]. Maniezzo, V. (2000). Ant Colony Optimization: An overview". http://www3.csr.unibo.it/~maniezzo/didattica/Vienna/ACOintro.pdf Retrieved on March 12, 2005.
[41]. Meher, Sukanya Sagarika, and Pallavi Maben (2014). Face recognition and facial expression identification using PCA. *Advance ComputingConference, 2014 IEEE International*, pp. 1093-1098.
[42]. Miljkovic, O., (2009). Image Pre-processing Tool. Kragujevac J. Math. 32, pp. 97-107.
[43]. Mliki H., Fourati N., Hammami M. and Benabdallah H. (2013). "Data Mining-Based Facial Expressions Recognition System" In: *Twelfth Scandinavian Conference on Artificial Intelligence;doi:10.3233/978-1-61499-330-8-185*.
[44]. Morik K., Brockhausen P. & Joachims T. (1999). Combining statistical learning with a knowledge-based approach - a case study in intensive care monitoring. In *Proc. 16th International Conf. on Machine Learning*, Morgan Kaufmann, San Francisco, CA, 1999, 268–277.
[45]. Mu-Chun Su, Chun-Kai Yang, Shih-Chieh Lin, De-Yuan Huang, Yi-Zeng Hsieh, and Pa Chun Wang (2013). An SOM-based Automatic Facial Expression Recognition System. *International Journal on Soft Computing, Artificial Intelligence and Applications (IJSCAI)*, 2(4).
[46]. Omidiora E. O., Fakolujo A. O., Ayeni R. O. and Adeyanju I. A. (2008). Optimised Fisher Discriminant Analysis for Recognition of Faces Having Black Features. *Journal of Engineering and Applied Sciences*, 3 (7); 524-531.
[47]. Padgett C., Cottrell G.W (1996). Representing Face Images for Emotion Classification. *Proceedings of Conference on Advances in Neural Information Processing Systems*, pp 894-900.
[48]. Pantic M. & Rothkrantz L. J. M. (2000). Automatic Analysis of Facial Expressions: the State of the Art. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(12), pp. 1424-1445.
[49]. Pantic M. and Rothkrantz L. J. M. (2004). Facial action recognition for facial expression analysis from static face images. *IEEE Transactions on Systems,Man, and Cybernetics, Part B: Cybernetics*, 34(3), pp. 1449–1461.
[50]. Pantic M., Rothkrantz L. J. M. (1999). An Expert System for Multiple Emotional Classification of Facial Expressions. *Proc. 11th IEEE Int. Conf. on Tools with Artificial Intelligence*, pp. 113-120.
[51]. Paul Ekman, Wallace. V. Friesen, & Joseph C. Hager (2002), Facial action coding system: The manual on CD Rom. *A human face Salt Lake City.*
[52]. Peng, B. (2005). Knowledge and Population Swarms in Cultural Algorithms for Dynamic Environments. *P.hD Dissertation, Wayne State University, Detroit, Michigan*.
[53]. Poli R., Kennedy J. & Blackwell T. (2007). Particle swarm optimization: an overview. *Swarm Intelligence*, 1(1), 33–57.

[54]. Poon B., Ashraful A. M., and Hong Y. (2011). Performance evaluation and comparison of PCA based human face recognition methods for distorted images. *International Journal ofMachine Learning and Cybernetics*, 2(4), pp. 245-259.

[55]. Qayyum H., Majid M., Anwar S. M. & Khan B. (2017). Facial Expression Recognition Using Stationary Wavelet Transform Features. *Mathematical Problems in Engineering Hindawi, 1-9.*

[56]. Rahulamathavan, Y., Phan R. W., Jonathon A. C., & David J. P. (2013). Facial expression recognition in the encrypted domain based on local fisher discriminant analysis. *Affective Computing, IEEE Transactions*,4(1), pp. 83-92.

[57]. Rashedi, E., Nezamabadi-Pour, H., Saryazdi, S. (2009). "GSA: A Gravitational Search Algorithm." Information sciences, Vol 179, No. 13, pp. 2232-2248.

[58]. Reed.L. I, Sayette. M. A. & Cohn, J. F. (2007). Impact of Depression on Response to Comedy; A Dynamic Facial Coding Analysis. *Journal of Abnormal Psychology* 116(4): 804-809.

[59]. Saleem, S, M. (2001). Knowledge Based Solution to Dynamic Optimization Problems using Cultural Algorithms.

[60]. Samad, Rosdiyana, and Hideyuki Sawada (2011). Edgebased Facial Feature Extraction Using Gabor Wavelet and Convolution Filters. In *MVA*, pp. 430-433.

[61]. Samiksha A., Pallavi K., & Shashikant G. (2014). Facial Expression Recognition Techniques: A Survey; in: *International Conference on Electrical, Electronics, Computer Science and Mathematics Physical Education & Management, (ICEECMPE).*

[62]. Sebe N, Lew M. S., Cohen I., Garg A., and Huang T. S. (2002). Emotion recognition using a cauchy naive bayes classifier. *Proceedings of 16th InternationalConference on Pattern Recognition*, 1, pp. 17–20.

[63]. Shih-Wei L., Lee Z., Chen S. & Tseng T. (2008). Parameter determination of support vector machine and feature selection using simulated annealing approach. *In: Applied Soft Computing 8, 2008, 1505–1512.*

[64]. Sobia, M. C., Brindha V., and Abudhahir A. (2014). Facial expression recognition using PCA based interface for wheelchair. *Electronics andCommunication Systems, InternationalConference*, pp. 1-6.

[65]. Susskind, J.M., Littlewort G., Bartlett M. S., Movellan J., Anderson A. K. (2007). "Human and computer recognition of facial expressions of emotion", *Neuropsychologia,* 45, pp 152-162.

[66]. Taisir E. & Al Qasim R. (2013). "On The Performance of the Gravitational Search Algorithm." International Journal of Advanced Computer Science and Applications (IJACSA).Vol. 4, No. 8, pp.74-78.

[67]. Thai, Le Hoang, Nguyen Do Thai Nguyen, and Tran Son Hai (2011). A facial expression classification system integrating canny, principal component analysis and artificial neural network. *arXivpreprint arXiv: 1111.4052*.

[68]. Tian Y. I., Kanade T., and Cohn J. F (2002). Evaluation of gabor-wavelet-based facial action unit recognition in image sequences of increasing complexity, *Proceedings of Fifth IEEE International Conference on Automatic Face andGesture Recognition*, pp. 229–234.

[69]. Tian Y. L., Kanade T, Cohn J. F (2001). Recognizing Action Units for Facial Expression Analysis", IEEE *Trans. Pattern Analysis and Machine Intelligence*, 23(2), pp. 97-115.

[70]. Tistarelli, M. and Akarun, L., (2005). Report on face state of art, BioSecure, Biometrics for secure Authentication, pp. 1-76.

[71]. Turk, M. and Pentland, A., (1991). Eigenfaces for Recognition, *Journal of Cognitive Neuroscience*, 3, pp. 71-86.

[72]. Valstar M. F., S´anchez-Lozano E., Cohn J. F., Jeni L. A. , Girard J. M., Zhang Z., Yin L., & Pantic M. (2017). FERA 2017 - Addressing Head Pose in the Third Facial Expression Recognition and Analysis Challenge. arXiv:1702.04174v1 [cs.CV]; 1-2.

[73]. Valstar V. (2014). Automatic behaviour understanding in medicine. *In Proceedings of the 2014 Workshop on Road mapping the Future of Multimodal Interaction Research including Business Opportunities and Challenges,* pages 57–60. ACM, 2014.

[74]. Vapnik V (1995). *The Nature of Statistical Learning Theory*. (Springer-Verlag, New York, 1995)

[75]. Wiskott, L., Fellous, J. M., Krueuger, N and Malsburg, C. V, (1999). Face recognition by elastic bunch graph matching. In: L.C. Jain et al., eds. *Intelligent Biometric Techniques in Fingerprint and Face Recognition*, CRC Press, pp. 355-396.

[76]. Xiang, T., Leung M.K.H., Cho S.Y. (2008). Expression recognition using fuzzy spatio-temporal modelling, *Pattern Recognition*, 41, pp. 204-216.

[77]. Yang Ming-Hsuan, Kriegman D. J., Ahuja N. (2002). Detecting Faces in Images: a Survey, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 24(1), pp. 34-58.

[78]. Yeasin M., Bullot B., and Sharma R. (2006) Recognition of facial expressions and measurement of levels of interest from video, *IEEE Trans. on Multimedia*, 8, pp. 500-508.

[79]. Ying Z. and Zhang G. (2009). Facial expression recognition based on NMF and SVM. *International Forum on Information Technology and Applications*, 3, pp. 612–615.

[80]. Ying-Yi H., Angelo A. B. & Arnold C. P. (2016). A Chaos-Enhanced Particle Swarm Optimization with Adaptive Parameters and Its Application in Maximum Power Point Tracking. *Mathematical Problems in Engineering Hindawi Publishing Corporation, 1-19.*

[81]. Zhang L., Mistry K., Jiang M., S. Neoh S. C., and Hossain M. A. (2015). Adaptive facial point detection and emotion recognition for a humanoid robot. *Comput. Vis. Image Understand.*, vol. 140, pp. 93–114.

[82]. Zhang Y. and Ji Q. (2005). Active and dynamic information fusion for facial expression understanding from image sequences. *IEEE Transactions on PatternAnalysis and Machine Intelligence*, 27(5), pp. 699–714.

[83]. Zhang Z., Lyons M., Schuster M., and Akamatsu S. (1998), Comparison Between Geometry-Based and Gabor-Wavelets-Based Facial Expression Recognition Using Multi-Layer Perceptron. *proc. Third IEEE International Conference on Automatic Face & Gesture Recognition*, pp 454 – 459.

[84]. Zhao W., Chellappa R., Rossenfeld A. & Phlipps. P. (2003). Face recognition literature survey. *ACM Computing surveys*, 399-458.

.

agboola,F.F "A Review of Facial Expression Recognition in Human Using Particle Swarm Optimization and Support Vector Machine" American Journal of Engineering Research (AJER), vol. 7, no. 2, 2018, pp. 236-249.