Research article

Pandemic disease detection through wireless communication using infrared image based on deep learning

Mohammed Alhameed[1], Fathe Jeribi[1], Bushra Mohamed Elamin Elnaim[2], Mohammad Alamgir Hossain[1],* and **Mohammed Eltahir Abdelhag[1]**

[1] College of CS & IT, Jazan University, Jazan, Saudi Arabia
[2] College of Science and Humanities, Al-Sulail, Prince Sattam bin Abdulaziz University, Saudi Arabia

*Correspondence: Email: alamgirhossain.nit@gmail.com.

Abstract: Rapid diagnosis to test diseases, such as COVID-19, is a significant issue. It is a routine virus test in a reverse transcriptase-polymerase chain reaction. However, a test like this takes longer to complete because it follows the serial testing method, and there is a high chance of a false-negative ratio (FNR). Moreover, there arises a deficiency of R.T.–PCR test kits. Therefore, alternative procedures for a quick and accurate diagnosis of patients are urgently needed to deal with these pandemics. The infrared image is self-sufficient for detecting these diseases by measuring the temperature at the initial stage. C.T. scans and other pathological tests are valuable aspects of evaluating a patient with a suspected pandemic infection. However, a patient's radiological findings may not be identified initially. Therefore, we have included an Artificial Intelligence (A.I.) algorithm-based Machine Intelligence (MI) system in this proposal to combine C.T. scan findings with all other tests, symptoms, and history to quickly diagnose a patient with a positive symptom of current and future pandemic diseases. Initially, the system will collect information by an infrared camera of the patient's facial regions to measure temperature, keep it as a record, and complete further actions. We divided the face into eight classes and twelve regions for temperature measurement. A database named patient-info-mask is maintained. While collecting sample data, we incorporate a wireless network using a cloudlets server to make processing more accessible with minimal infrastructure. The system will use deep learning approaches. We propose convolution neural networks (CNN) to cross-verify the collected data. For better results, we incorporated tenfold cross-verification into the synthesis method. As a result, our new way of estimating became more accurate and efficient. We achieved 3.29% greater accuracy by incorporating the "decision tree level synthesis method" and "ten-folded-validation method". It proves the robustness of our proposed method.

1. Introduction

A fast diagnosis is an emergent issue in the present situation like-COVID-19 test [1,2]. It is a routine test of viruses following the R.T.–PCR method. However, this test is carried out sequentially. So there is a chance of a high FNR ratio. A test like this will take a longer time to complete. Besides that, there arises a shortage of R.T.–PCR test kits. So there is a severe need for alternative tactics to diagnose patients more quickly to manage these situations. The infrared image is self-sufficient for identifying these diseases by measuring temperature as a fast finding. C.T. scan and other pathological tests are essential in evaluating a patient with a suspected pandemic infection [3,4]. Moreover, for detection deep precognitive analysis is applied for the presentation of diseases detection during pandemic situations. In such situation biologically-Inspired convolution fuzzy network would be more effective [5]. However, a patient's radiological findings may be expected at first. In this paper, we tried to collect information by an infrared camera of a patient's eye-retina to measure temperature by using a smartphone with an advanced camera enriched with infrared features. Almost every mobile phone with an exemplary configuration has this feature. If not, an infrared-enabled application can be installed from the Google Play Store. Then it will be noted, and the next course of action will be taken. Instead of using a regular infrared camera, we want to incorporate a mobile camera with its features. So this technique ensures both applications (through mobile and infrared cameras).

There are numerous applications of infrared images for detecting humans and their body parts for visual surveillance, human-activity tracking [6], medical applications for eye-diseases detection [7], driver safety, homeland security, etc. An essential possibility is monitoring people by CCTV and surveillance camera-based systems. Infrared images are also related to authenticating human facial parts [8,9]. The infrared camera can capture the images in any lighting conditions, while an RGB camera frequently needs proper lighting to capture high-resolution-based clear images. Meanwhile, infrared imaging schemes use infrared light sources to produce a healthier image without natural or artificial light. The traditional face recognition systems depend on self-reporting of manual measurement schemes [10], but the infrared technique depends on an automated system [11]. The traditional system is weak in checking it continuously and quickly. Researchers looked into contactless sensors, infrared cameras, and infrared cameras to solve these issues. Since temperature is directly connected to the physical parts of a human face like the eyes, forehead, cheek, lips, etc., few researchers applied the physical parts to identify temperature. Lips, mouth, and eyebrows are intimately related to different temperature measurement schemes, so a projected model is addressed. In work [10] proposed a method to identify human faces by visual features like the blinking ratio of closed and open eyes. In additional work, researchers noticed by tracing 3D face images using mobile phones. In their recent study M .Kim, B. H. Kim, S. Jo [12] developed a contactless-based real-time scheme for recognizing a driver operating a car in a sleeping or drowsing mood.

The use of a digital camera is a normal phenomenon. However, due to various good characteristics, the infrared camera and its imaging techniques are used in various applications like pandemic disease identification. It is proposed in recent literature that the computational timing of an infrared imaging technique is less than that of an RGB imaging technique [13] to work by taking advantage of these

changes. Using different imaging types, the researchers have shown their efficiency in quantitative measurement. For example, in work [14] researchers achieved 65% accuracy utilizing images, but researchers [15] achieved more than 91% accuracy using infrared images. It demonstrates that it is possible to achieve a greater accuracy rate in recognizing anything by employing infrared imaging techniques.

The following are the key contributions of our manuscript.

- We proposed an infrared imaging technique to identify the pandemic and make accurate and quicker measurements.
- Few parts of a face are considered for measuring temperature. So a segmentation technique is used in dividing a face into various classes.
- A novel searching technique will be incorporated to search the regions from left to right.

The paper is separated into sub-sections as follows:

Section 2 illustrates related works. The proposed tracking scheme is described in section 3 in detail. The image registration process with its sub-sections (Pre-processing, Segmentation by improved gradient method, Feature extraction, Hybrid adaptive optimized classifier system, Head pose estimation and correction, Feature weights extraction, Regions detection, Vector formation from the segmented region) has been demonstrated in section 4. The feature extraction technique is described in section 4. The findings of the experiment are presented in section 5. In the final step, a decision is reached, and then we discuss the path that our work will take in the future.

2. Related works

Mobile and other electrical gadgets are used for face tracking [16] from a short distance in many applications nowadays. However, it may be unsuitable in many situations. Many scholars took facial signals for security and safety to control some surveillance-related issues. A video-based human identification system might have offered a proficient nonintrusive explanation for our day-to-day use. A video-based imaging approach routinely uses two imaging methods, i.e., infrared and color. In work [17], projected an improved resolution-based infrared facial image database by extensive manual explanations and clarified the flexibility in various applications. They proposed a set of correlated algorithms for detection and approximation within a facial image. They advised that a multi-process description by networking middleware resolve be used. It might be used for real-time face authentication and recognition using infrared pictures. By evaluating the frontal view of facial regions, scholars [18] proposed how infrared video can be exploited with the help of the SIFT flow approach. Furthermore, eye-tracking using video and its effectiveness has been presented along with the efficiencies suggested [19]. Finally, the evaluation is finished using the face-temperature histogram value.

Surveillance systems that use network devices like cameras to monitor and track activities generate massive volumes of data [20]. Both the migration of data over bandwidth constraints and the accumulation of lags in network technologies are problems that need to be solved. It proposes constructing a decentralized facial recognition algorithm applying PCA and LBP [21] for a distributed surveillance system using wireless networks and cloud computing. In the regionalized face tracking approach, face recognition, feature extraction and matching [22] are done in two steps. First, face detection and feature extraction are done on a specified cloudlet adjacent to the security cameras, avoiding spending vast amounts of data on a distant processing center. Face matching, on the other hand, is done using the facial features vector in any private cloudlet environment. According to this

study, the suggested approach works effectively in the Wireless Network-Cloud architecture of its usefulness in finding "lost" people.

In recent years, the essential part of human life has been security. The most significant consideration at this point is cost. This technology is quite advantageous in minimizing the cost of external movement monitoring. We provide a real-time recognition method in this study that would allow us to handle photographs fast [23]. The main objective of this manuscript is to recognize people so that the house and office can be secured. A PIR sensor is utilized to detect movement in a defined area. The Raspberry Pi will next capture the photographs. The face in the captured image will then be detected and recognized [24]. Finally, the photographs and notifications will be uploaded to a smartphone-based Wireless Network via the Telegram program. The proposed systems calculate in real-time, are rapid, and are low-cost. The experiments suggest that the proposed facial recognition system may be employed in real-time.

This work proposes a higher efficiency in face recognition approach based on characteristics that use a newly created process named Floor of Log (FoL). This method has the benefit of conserving space and energy while maintaining precision. The scholars [25] used K-Nearest Neighbours (KNN) and Support Vector Machine (SVM) [26] approach to discover the optimal factor of the FoL technique utilizing cross-validation. The correctness and size far ahead of compression of the suggested approach were assessed. In the Extended YaleB, AR, LFW and CelebA face datasets, the FoL produced better results than a technique with the equivalent classifiers [27] without compressed features, with 86% to 91% relative to the same data size. This study provides a robust and easy feature compression technique for FER applications for various parts recognition [28]. The FoL is a supervised compression approach that may be modified to get better results and is compatible with edge computing schemes.

The wireless network is a concept that integrates technology into our day-to-day activities by applying deep neural networks and convolution neural network (CNN) learning techniques [29,30]. In their FER system, the authors of work [31] consider introducing convolutional neural network (CNN) and lengthy as well as short-term memory (LSTM) techniques. On the other hand, researchers [32] implemented CNN using fewer data and the eight-folding approach, yielding improved results. One of the key categories in which skill assists us is safety and privacy. Smartphones may be used as a safety alert scheme because they are the most extensively used smart gadgets. Artificial intelligence (A.I.)-enabled intelligent wireless network devices have grown in popularity in recent years. This study developed an innovative network security solution for the smart home using the neuro-fuzzy optimized classifier [33]. A security system is built around a Raspberry Pi and a NoIR(no Infrared) Pi camera unit that records and captures images[34,35]. A PIR-MS (Passive infrared motion sensor) is also used to identify motion. We propose combining images and motion sensor records from the NoIR(no Infrared) Pi camera unit to identify a safety threat using our algorithm's facial recognition classification technique[36,37]. In the event of an emergency, the system can notify the user. The proposed system has a 95.5% accuracy and 91% precision in detecting any security threat.

The Visual Internet of Things (VIoTs) and wireless communication have gotten much interest in recent years because of their capacity to extract object position from scene picture information, attach an optical tag to the item, and then return scene object information to the wireless network. Face recognition is one of the ideal visual network methods since a person's face is an intrinsic label [38]. The researchers [39] developed a pose estimation technique to resolve the problem due to long-range pixels leading to poor performance in FER. However, due to a lack of processing resources, existing state-of-the-art facial acknowledgement methods based on huge deep artificial neural networks (ANN) [40]

are challenging to implement in the implanted podium for the visual wireless network. To overcome this problem, we provide a small deep ANN-based facial recognition system for the VIoTs. The proposed technique employs deep neural networks with minimal complexity [41] to function in an embedded set.

Moreover, it can withstand changes in lighting and position. We exhibit comparable correctness and enactment results designed for the LFW authentication benchmark using the mobile facial acknowledgement dataset. Scholars [42] have demonstrated that a Facial Action Coding System (FACS) is used in the detection of the characterization of a human being and that this same system could be applied to the detection of the categorization of varied consumers' goods by employing the affective reactions of those selective consumers. Work [43] proved and implemented a method in real-time on an Android-built platform for panic-face detection [44], fatigue detection [45] by incorporating mutation, genetic algorithm [46,47] and many more in expression recognition. Moreover, scholars [48] have proved that a histogram approach becomes very appropriate in these predictions for quicker identification of depth measurement and various expression recognition. Alzheimer's disease is a progressive degenerative neurological illness. It is now incurable, and those who suffer from it are denied the freedom to leave their houses compared to the general population. This article aims to develop and form an IoT prototype that can identify Alzheimer's victimized people, thereby increasing the high worth of life and relaxing caregivers' jobs. The patient wears a small dorsal belt that contains a Node MCU ESP8266 board, a GPS module, and a small portable WiFi modem/router. The patient's location is tracked via a web application and an Android/iOS mobile application. This research also allows the Kalman Filter to track the patient's movement and estimate his position, especially when the patient wanders outside.

Pneumonia causes a high morbidity and mortality rate in infants. This sickness affects the lungs' tiny air sacs, necessitating rapid diagnosis and treatment. One of the most popular diagnostics for diagnosing pneumonia is a chest X-ray. This paper explains how to identify pneumonia in chest X-ray pictures using a real-time Wireless Network system. Three medical experts reviewed the information, including 6000 images of children's chest X-rays. This study adopted twelve alternative image Net-trained Convolutional Neural Networks (CNNs) [49,50] architectures as resource extractors. The CNN and deep neural network are very emergent if they are used in FER for disguised and distorted expression recognition [51]. Many prototype designs are proposed by scholars [52] based on deep CNN for identifying people's faces, iris, and fingers vein in our daily life. CNNs were then integrated with other learning approaches like KNN, Naive Bayes (N.B.), Random Forest (R.F.), Multilayer Perceptron (MLP) [53], and SVM. The best model for diagnosing pneumonia in these chest radiographs was the VGG19 architecture with the SVM classifier and RBF kernel. The accuracy precision scores for this combination were 96.47%, 96.46%, and 96.46%, respectively. Paralleled to other articles in the collected works, the proposed approach yielded better results for the measures used. These findings imply that using a real-time Wireless Network system to detect pneumonia in children is beneficial and could be utilized as a diagnostic tool in the future. Doctors will obtain faster and more precise findings using this technology, allowing them to provide the best treatment possible.

Moreover, many works use machine learning-based recognition approaches. Researchers employ machine learning models such as SVM, CNN, and ANN (Artificial Neural Network), Genetic Algorithm for facial identification. A CNN is a deep learning approach [54,55] with a high-performance level and can extract features from training data. It achieves features from side to side using several convolutional layers and is often validated by a series of totally related layers [30]. On

the CK+ database, investigators [32] raised CNN accuracy to 96.76%. Improved pre-processing processes such as sample formation and intensity normalization have enriched accuracy. A BDBN with numerous facial expression classifications was employed in another investigation. The accuracy on the CK+ database was 96.7 %, whereas it was 91.8 % on the JAFFE database. On the other hand, the execution period lasts for eight days.

In In this study, the proposed technique saves processing time while improving recognition performance using wireless communication accompanied by a cloudlets server. The proposed method uses numerous pre-processing techniques in conjunction with CNN to achieve excellent accuracy. It is also recommended that A.R.s focus on critical data in categorization to forecast projected expressions. This also helps to cut down on processing time. In addition, parallelism [30] is frequently employed to improve speed and precision.

Finally, we conclude that techniques based on appearance can obviate the need for meticulously created visual pieces to characterize a gaze in the eye-tracking system. However, we will incur a substantial penalty in execution time and storage space if we apply the complete input image to a classifier to forecast the gaze. Moreover, training eye image data, including poses and locations, is required during the development phase.

3.   Proposed tracking scheme

Classifying distinct face portions is the primary problem while tracking a human face and extracting features from an infrared-facial image. Poses have a direct impact on a person's facial expression. The head movement and poses of a gaze vector are inextricably linked. We suggest a two-part flow diagram. We illustrated it by our experiment's results in Figure 2. First, we use all strategies to build Experiment-I's face region feature vectors.

| Original Infrared Image | Image after segmentation into Six classes |

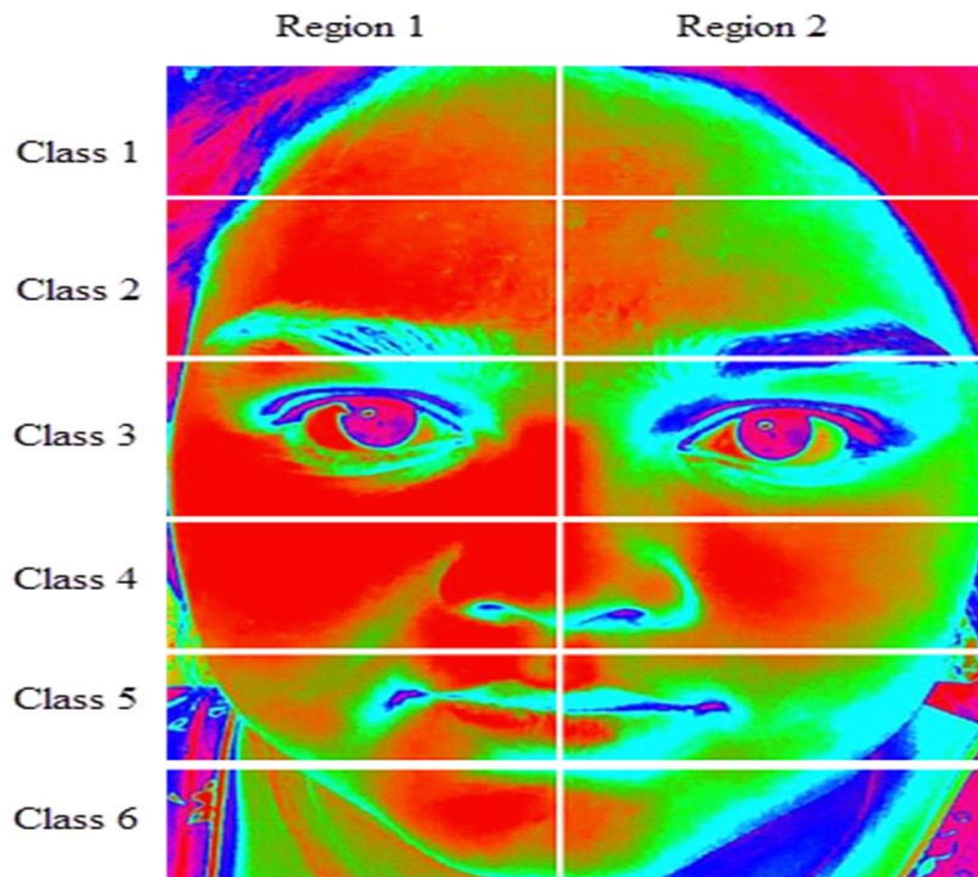**Figure 1**. Infrared image and its divisions into six classes.

Figure 2. Facial image is separated into Six Classes and Twelve Regions.

Here we have followed the processes of registering an image, Sequence Maintenance of its sequence, Central-point and facial regions detection. Then, using the mapping function obtained from getting processes, we examined the correlation among the class calibrations in Experiment II. The system will proceed to Experiment II after completing the calibration process. Otherwise, it returns to the beginning of Experiment I. Finally, from the grouped regions, we proceeded to retrieve attributes.

4. Image registration

Image registration is required to guarantee acceptable infrared image dataset series collected from a facial image. First, the image must be registered in the database if it has not already been done. As a result, the proposed technique can help detect picture duplication during image registration. During the image registration process, video images are converted into frames.

4.1. Pre-processing

Due to noises in the images, traditional recognition methods have proved that obtaining correct portions of any part of a human face in unusual settings (low lighting, dark, rainy period, or natural calamity) is exceedingly challenging. We applied the notion of using histograms to avoid difficulties like this, which assures that any lighting effect does not cause too many complications. Normal and sensitive histograms can be used to implant 3D data. The noises can be removed here in this stage from

the infrared image. It is common practice to use nonlinear optical filtering to remove noise from an image. Pre-handling is done to work on a different type of filtering used in the infrared image. In our application, we applied Adaptive Weiner Filter (AWF) [43] to remove Gaussian Noise and other noises accompanying the infrared images we have used in our application.

Algorithm 1. Pose estimation (PE) algorithm

$P_m = PE(P_{2D}, P_{3D}, f_w, f)$

$Input: P_{2D}, P_{3D}, f_w, f$

$X = Q[P_{2D}]; c = R[X]$

$Y = P_{2D_x/f}; X = P_{2D_y/f}$

$h = [P_{3D}, c]; o = Pinv(h)$

$while(true)$

$\{ \quad j = oxY; k = oxX;$

$L_z = \dfrac{1}{\sqrt{(\dfrac{1}{j} x \dfrac{1}{k})}}$

$p_{m_1} = jxL_z; \ p_{m_2} = kxL_z;$

$R_{m_1} = p_{m_1}(1:6); \ R_{m_2} = p_{m_2}(1:6);$

$R_{m_3} = \dfrac{R_{m_1}}{\|R_{m_1}\|} x \dfrac{R_{m_2}}{\|R_{m_2}\|};$

$p_{m_3} = [R_{m_3}, L_z];$

$c = hx\dfrac{p_{m_3}}{L_z};$

$YY = Y; XX = X;$

$Y = c.fw.P_{2D_x}; \ X = c.fw.P_{2D_y};$

$E_x = Y - Y^2; \ E_y = X - X^2$

$if(\|E\| < E_x)$

$\{ p_{m_6} = [p_{m_1}(1:6)p_{m_2}(1:6), p_{m_3}(1:6), p_{m_4}(1:6), p_{m_5}(1:6), p_{m_6}(1:6), L_z, 1]^{tm};$

$break;$

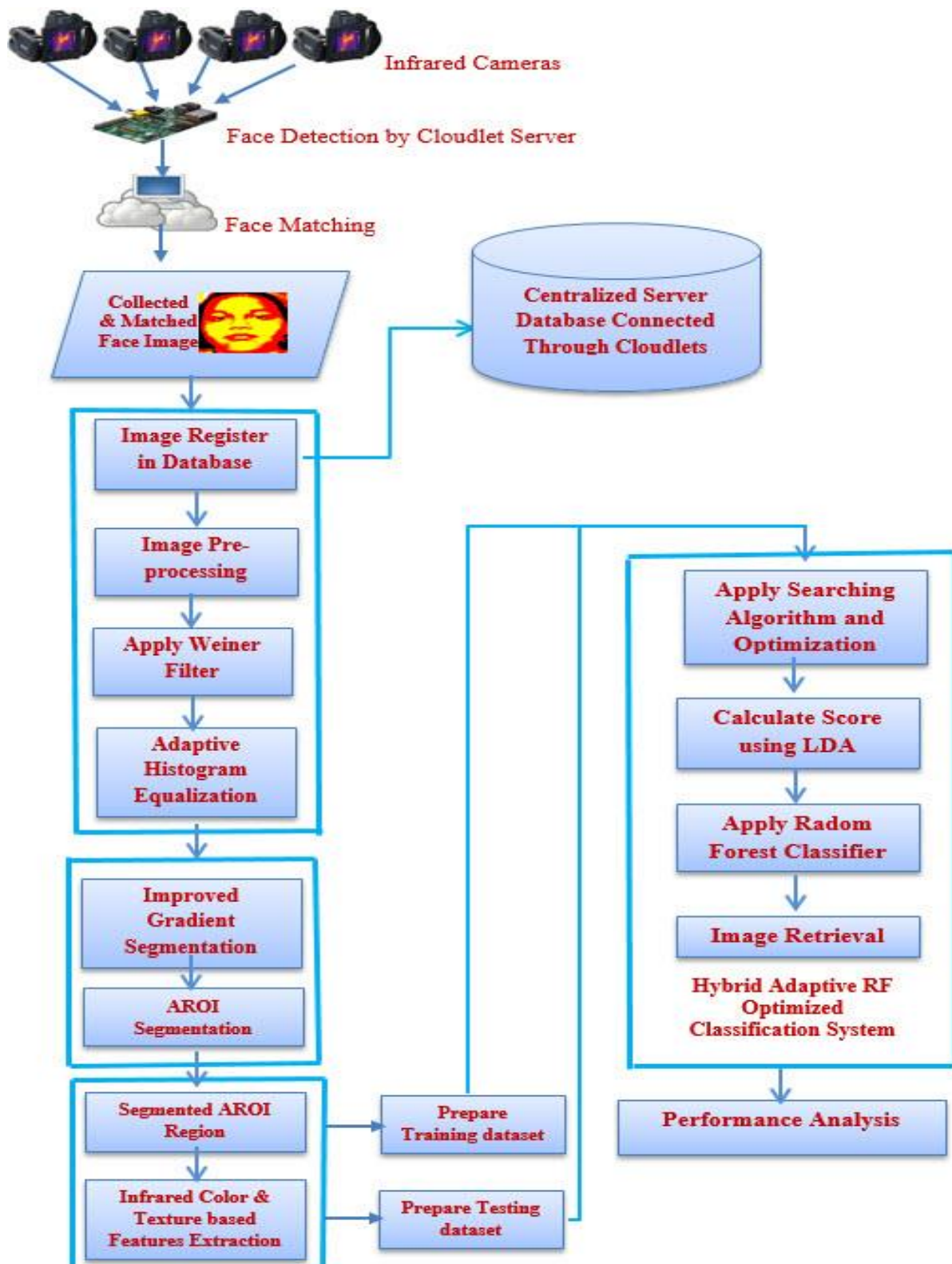$\quad \}$

$Output: P_m$

Figure 3. Proposed flow diagram for implemented infrared image retrieval classification and analysis system.

This is a straightforward way of determining the safest method for re-establishing a magnificent sign. The suggested study employs AWF to handle photo positions efficiently. The AWF is used to simplify the image with the most negligible fluctuation. Histograms are commonly balanced to

improve the picture's uniformity. Histogram correction is a computer-assisted process for enhancing visual contrast. The prime consciousness esteems been fundamentally increased, i.e. the variety of image strength has been broadened. It creates a less close association with the development of district ties. As a result, following histogram correction, the usual picture contrast increases.

$$N^x = \frac{P}{\text{TP}} \tag{1}$$

Where x = 0,1 & -1; P = number of pixel, TP => Total number of pixels

If we want to calculate the histogram equalized image then we may follow the equation below

$$HQ_{m,n} = log^{e^{(1-x)}}(\sum_{x=0}^{b^{m,n}}(N^x)) \tag{2}$$

Where $log^{e^{(1-x)}}$ represent nearest neighbourhood integer value.

The similar representation with respect to the pixel intensity value is as follows:

$$\frac{\partial P}{\partial x}(\int_0^P HQ(x) = \partial P(x)\partial y = \partial P(P^{\frac{1}{x}}.\text{P})^{\delta}/_{\delta P} \tag{3}$$

Finally, the probability distribution function (PDF) can be illustrated uniformity as $\frac{\partial P}{\partial x}$.

However, the outcome shows that the equalization method can smooth and improve histograms.

4.2. Segmentation by improved gradient method

Image line boundary detection [47], also known as edge identification is critical in visual interpretation. Edges store massive amounts of info. As a result, the image size is drastically decreased, and less restorative material is combed through, preserving the image's essential core elements. An edge location is extensively employed in image separation because borders frequently appear at picture object boundaries. We incorporated the AROI by selecting from the six divided classes of the face regions. The restrictions of an aim depicted on an image or volume are calculated here. This method of dividing determines whether or not neighboring pixels of starting seeds should be added. They are combining pixels or sub-regions with a local tool. The simplest of these processes is pixel amalgamation, which starts with a collection of "Images" focuses and develops regions by linking pixels with common characteristics.

$$AROI^S = \sum_{m,n \in Q2}^{S} Vel.(Hp_m, Lp_n)P.logd_m + \delta \int e_m \partial P \tag{4}$$

Where $AROI^S$ is the separated AROI, Vel is the velocity measured by gradient value, Hp_m, Lp_n are the pixels values of low and high, $logd_m$ represents the spatial image size, δ signifies the image frequency coefficient, d_m is the distance between two pixels.

4.3. Feature extraction

We used a grey level co-occurrence matrix to understand texture characteristics. It represents the grey level. The input image's spatial information determines the probability of matches with values obtained. Irrespective of these traits, this method examines 18 texture attributes. We applied the image retrieval method by providing a sample image file that retrieval images from a vast dataset that seems to be adjacent. A range of infrared images is used to evaluate the algorithm's performance upon that texturing dataset. We propose sixteen texture classes. Each texture image is broken into six sub-images for all these examples. Many images are obtained depending on the distance between the queries among the data set. The image features are extracted and used in the investigation. Grey level co-occurrence matrix can choose the pixel frequency inside the individual result. The segmentation's directional value can then be used to erase the image attributes utilized in the segmentation. The following is an example of the grey layer co-occurrence matrix technique:

$$(m, n) = Vc(m, n, u, v) \sum_{m=1}^{H} Vc(m, n, u, v) \qquad (5)$$

Where Vc is the vector, m; n; u; v; are the pixel values with respect to high and low; C is the image characteristics. We used a grey layer co-occurrence matrix to gain the various attributes for feature extraction. Finally, the features are chosen based on texture and color.

4.4. Hybrid adaptive optimized classifier system

We incorporated the Enhanced Cuckoo Search Optimization (ECSO) algorithm [33] to gain the Adaptive Optimized classification (AOC). The Cuckoo Search algorithm (CSA) evaluates the inconsistency characteristics. Moreover, CSA is proposed to minimize the cost of network congestion we may face while collecting input images. In the ECSO algorithm, we have a guideline that we will place a random value on each node during selecting an arbitrary node on the cloudlets. The next cloudlet will move to the most vital node with the most significant number of images. The host server is static, and the cuckoo value possibly will be calculated in coincidence with the possibility of Pr[0; 1] by the host value. To overcome the challenge posed by network congestion and picture optimization variables, ECSO is a required method.

Algorithm 2 ECSO

$Input : Image\_Features(\ Img_{ftr}), Image\_Cordinate(Img_{cor})$

$Output : Classified\ Value(CL_{val})$

$Begin\ to\ compute\ The\ Random\ Value(RV)$

$for\ m = 1: Range(Img_{ftr}, 1)$

$for\ n = 1: Range(Img_{ftr}, 1)$

$Distance(m, n) = \sqrt{(Img_{ftr}(m, 1) + Img_{ftr}(m - 1) + (Img_{ftr}(m, 1) - Img_{ftr}(n, 1))^2}$

End

$Img_{ftr} = \lfloor Img_{ftr} Distance(m, n) \rfloor$

$$RV = (Img_{ftr}{}^{Distance(m,n)})$$

$$Class\ label(CL) = unique(node)$$
$$L = lenth(CL)$$
$$for\ m = 1:L$$
$$\quad T = mean\_all\_cloudlet\_node(Img)$$
$$\quad X(Img,n) = -\frac{1}{2}xTxmean\_all\_cloudlet\_node + \log(m)$$
$$\quad X(Img,2:end) = T$$
$$End$$

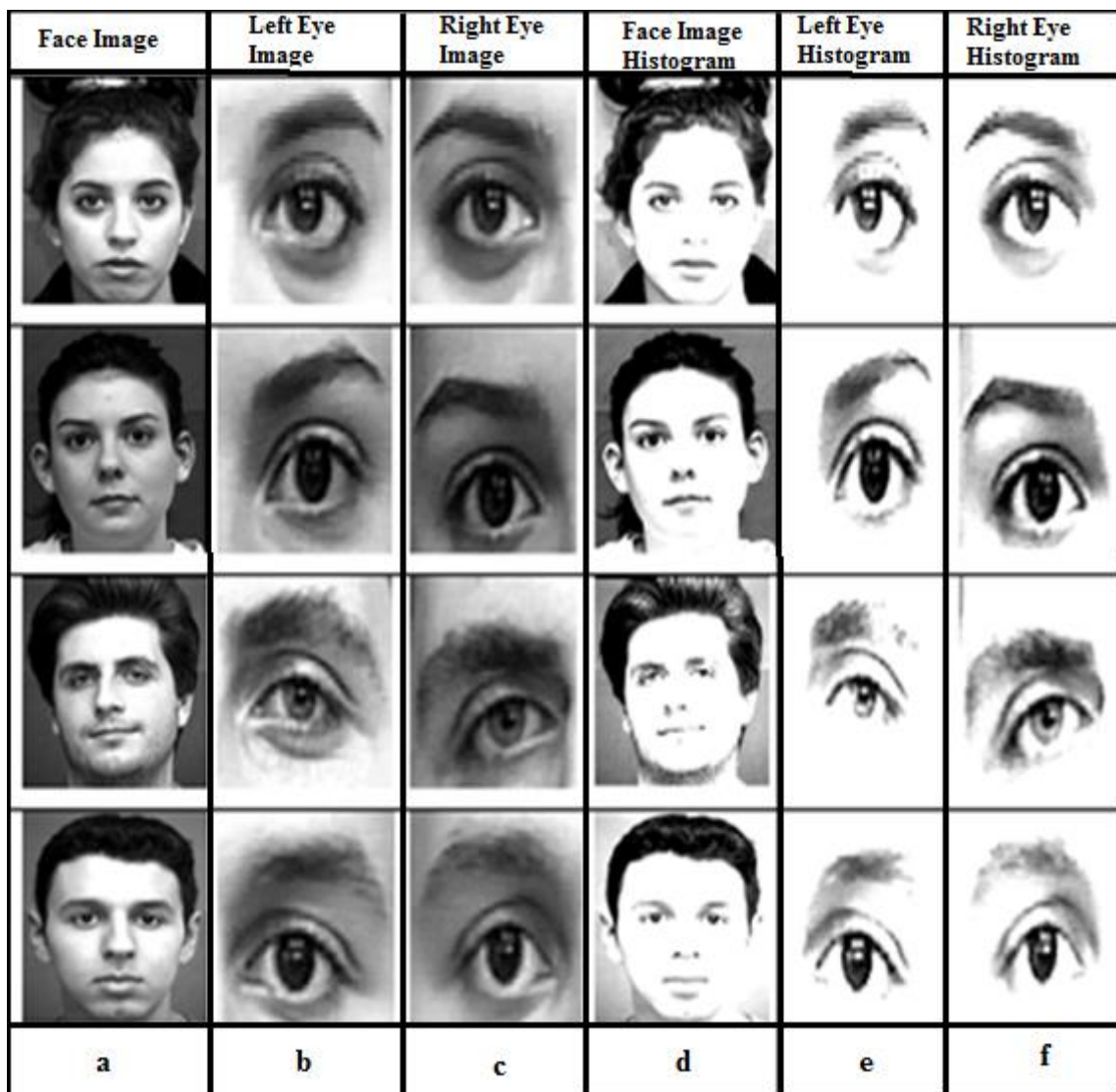| Face Image	Left Eye Image	Right Eye Image	Face Image Histogram	Left Eye Histogram	Right Eye Histogram
a	b	c	d	e	f

Figure 4. Dissimilar illuminations with image histograms equalization.

Using linear discrimination analysis (LDA), the significance of respective image is considered after the infrared images are categorized. The results of the used classifier are given to the LDA. This approach measures the precision of picture statistics. The infrared images classification involves a remarkable role. It is usual practice to utilize LDA, a data inquiry tool, to reduce the dimension of

numerous interconnected variables while maintaining the maximum amount of relevant data. As a type of image categorization, LDA analysis is valid. Matrix creation is demonstrated for use in operators performing image processing. We illustrated the LDA investigation processed by the following equations. The LDA function collects samples in the first phase to prepare properties from testing datasets. This dataset will be built from the six divided classes. The training datasets will be prepared and collected from the twelve regions. The LDA has the K classes (K<=6) and R regions(R<=12). There will be vector representation for each class. It will be represented in multidimensional space. In choosing the features from the regions, we applied the searching algorithm (ECSO)) as stated by Algorithm 2. The sensitive picture histogram is exhibited as shown in Figure 4. There are four images in this set, each of which has been converted utilizing sensitive histograms and various illuminations.

4.5. Features weights extraction

As shown in Figure 5, the features' weights (F.W.) must now be determined, establishing the image's critical pixels and weights. The importance of the features represented by the pixels is reflected in the weights. For example, a set of features represents an A.R. The method of locating POI and important A.R.s such as the nose-tip, RE, L.E., and regions of lips are depicted in Figure 5. Essential qualities are given larger weights to improve the dependability of information, which improves recognition.

Figure 5. Recognition from Active Regions.

The following phase separates the features. The Stochastic Face Shape Model (SFSM) and Optimized Principal Component Analysis (OPCA) were used to compare patterns on original and altered pictures (after the corrections of angles). From there, the features are retrieved and expressed in vector form. When the vectors are formed, phase-I is done. It is worth mentioning that an image-based sequence should be kept. The procedure of extracting facial features and preparing vectors is described below.

This section explains how to track facial features in video sequences and use the head posture estimation technique. Previous head pose estimate (P.E.) methods have relied on a stereo camera to provide correct 3-D information for head pose and make the necessary correction by rotation and normalization [35]. However, a head model's complex illustrations and exact starting value make a real-time clarification perfect. For example, the human Head is frequently revealed by an ellipsoid. Therefore, the cylindrical head model, CHM, calculates head position with 3-D positions on the corresponding sinusoidal surface. The 2-D to 3-D alteration scheme is utilized to gain the head posture statistics when a 2-D facial characteristic is traced in the individual video frame. Pose scaling with iterations is a two-dimensional to three-dimensional adaptation. Because the 2-D face characteristics have a variety of ramifications when recreating the position.

4.6. Different regions detection

Our proposed method will detect the regions based on the feature weights of the corresponding regions. First, we smoothed the ocular region using L0 as per Gradient Minimization Method (GMM) to estimate the radius because it supports to remove noise on an image pixel. Then, we used the canny edge detector on the ocular areas. We get a few invalid edges here, which we can filter out with a filter. Finally, we collect the related information from the identified regions depending on E1 and E2.

$$E_1 = \sum (R_E \times C_R) \qquad (6)$$

$$E_2 = \sqrt{G_X^2 + G_Y^2} \qquad (7)$$

Where R_E and C_R represent the regions, we measure any region's radius based on the values of R_E and C_R. The pixels are measured horizontally and vertically. G_X and G_Y, respectively, represent these two values. To detect any region, we have to lessen the intensity of that region and make the most of the strength or weightage of that region. The parameter τ controls the trade-off. That is

$$U_c, V_c = \min_{(x,y)} \left\{ E_1(U,V) - T. \left(\int_{-x/5}^{x/5} E_1(U,V)\delta s + \int_{4x/5}^{6x/5} E_2(U,V)\delta s \right) \right\} \qquad (8)$$

Where (U_c, V_c) is the coordinate of that region. It establishes the intervals between ($[-\frac{1}{5}\pi, \frac{1}{5}\pi]$ & $[\frac{4}{5}\pi, \frac{6}{5}\pi]$) because the regions do not intersect with each other.

4.7. Vector Formation from segmented region

When G represents the identified region and p represents the vector of that region. To execute a mapping function, it transmits gaze data. The user will give a calibration procedure, and the region vectors will be registered. The mapping function links the vector of that region. Its coordinate's value is measured by (U, V). We applied the SLM, i.e., a simple linear SVM model. We also incorporated the polynomial model (PM) to establish the appropriate mapping functions. In addition, we applied the second-order and third-order polynomial functions in the calibration stage to gain better-synthesized results. As a result, a mapping algorithm can accurately be determined based on a segregated frame.

5. Experimental evaluation

We completed our experiment based on our quantitative and qualitative evaluation algorithm. Therefore, we have organized our entire procedure into five regions, and the execution will be done in parallel in a time-sharing manner. In this testing phase, we are looking into forehead, eyes, nose-tip, and lip feature detection.

5.1. Active regions detection

Recognition and feature extraction of an eye is a thrilling job. We propose the EPDF function to detect the eye area accurately. In this respect, the CK+ and NAVIE datasets are used in our experiment. The CK+ dataset has 1000 grayscale images using ten subjects and different lighting conditions with different scales. Our experiment has a higher recognition accuracy, i.e. 91.73% in RGB images and 92.39% accuracy in detection by twenty-one infrared images.

5.2. Classification based on CNN-Tree-Level Method

Recognition and feature extraction of an eye is a thrilling job. We propose the EPDF function to detect the eye area accurately. In this respect, the CK+ and NAVIE datasets are used in our experiment. The CK+ dataset has 1000 grayscale images using ten subjects and different lighting conditions with different scales. Our experiment has a higher recognition accuracy, i.e., 91.73% in RGB images and 92.39% accuracy in detection by twenty-one infrared images. FER is a new field that requires a high rate of identification for accuracy. This study combines the AROI and CNN methods for identifying and classifying face expressions [35]. Five optimized active regions [30] of interest (OAROI), namely-The forehead, LE, RE, Nose-tip, and Lip region, are taken into account. The mentioned three AROIs are trained for CNN. The ultimate classification outcome is gained utilizing a decision tree-level synthesized method. The figure of the applied method is shown in Figure 6.

The extents of OAR (optimized active regions) differ for each image. Therefore, the following steps are taken to achieve a similar result.

1. OAR of interests (OAROI) is measured by 120*80 pixels.

2. The features and their respective feature-weights of the used images have been learnt based on the CNN density layers and then by the sub-sampling layers simultaneously.

3. Lastly, it will pass through the Fully Connected Layer.

The fully-connected layers have applied the learned feature-weights to gain the classified expressions. Finally, one (1) is used as a hot code to mean that it will be labelled. There are five regions.

The hot code (1) size is five. An individual bit of the hot code resembles one class of regions. The hot index code (1) resulting in one area is showed. Here zero (0) bit implies false, and one (1) implies true. Figure 6 displays an example of the resultant region is the nose tip.
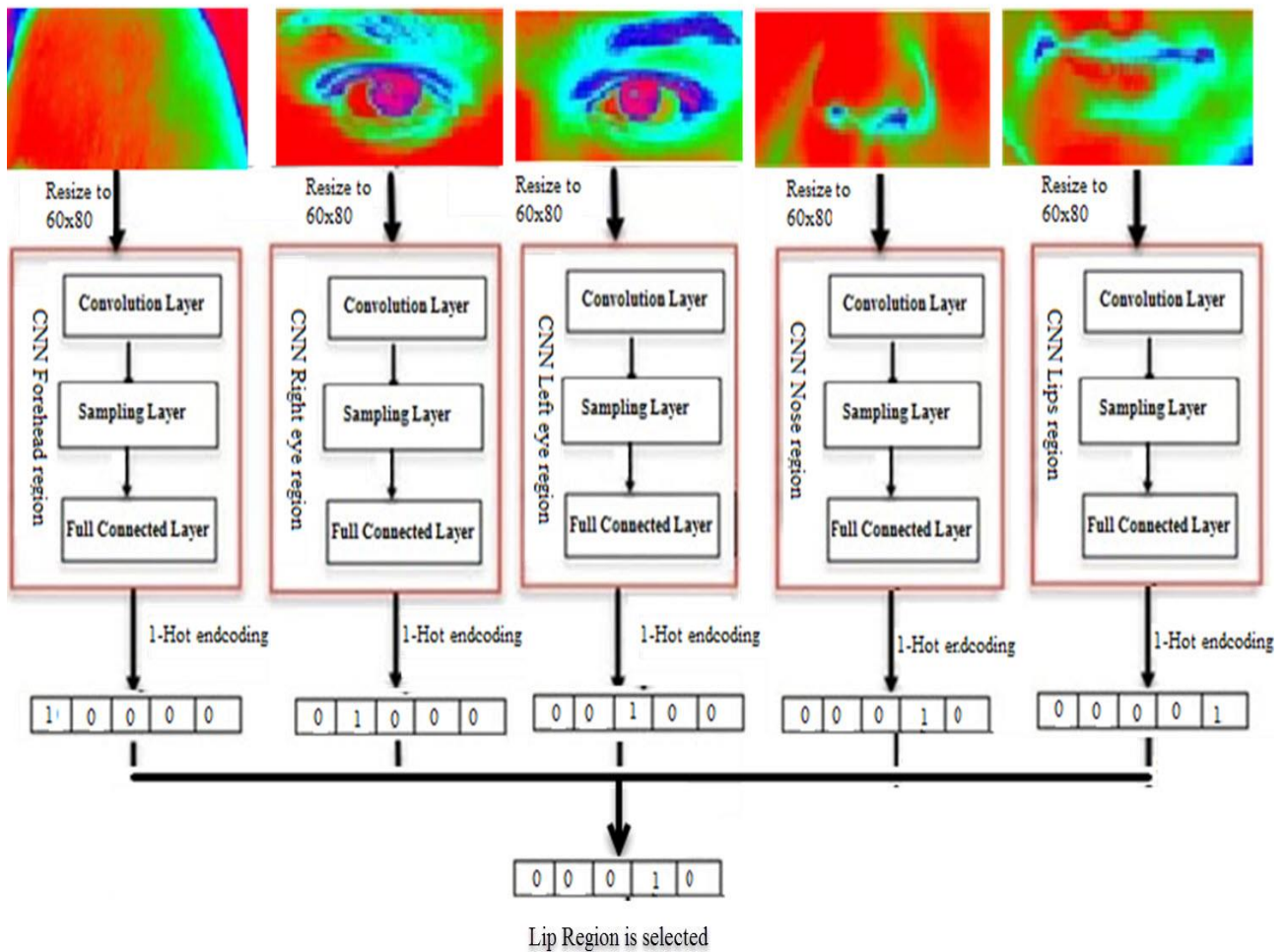

Figure 6. Schematic diagram of learning feature weight, classification of expression based on decision-tree level synthesis scheme.

The final recognition result is attained in the last phase built on decision-tree level synthesized scheme. The decision-level fusion technique is being supported in the following way.

From the above, we can understand that if two or more classifiers need to be used to categorize the expression as the similar class i, i∈I, then the ultimate results after classification will be considered as j. Therefore, if two CNNs classifications are high, the synthesis result will be increased. On the other hand, when the results of five CNNs all will be dissimilar, we have to choose the outcome of the CNN among the five AROI (Forehead, LE, RE, Nose and Lip) based on the maximum accurateness, as it is shown in "Experimental Section".

$$
Final\ Result = \begin{cases} i,\ if\ 2\ or\ more\ CNNs \\ classify\ the\ expression \\ as\ i,\ i \in I \\ The\ result\ of\ CNN\ for\ AROI \\ of\ Forehead,\ Left\ Eye, Right\ Eye,\ Lip \\ ,\ otherwise \end{cases}
$$

5.3. CNN Structure with Ten Fold Validity

The suggested CNN configuration is displayed in Figure 7. The CNN accepts 80*60 input infrared images. We incorporated 3 convolution layers. These three layers are being synthesized into subsampling layers. Sub-sampling layers correlate with the layers. The kernel sizes inside convolution layers, as shown in Figure 7, are divided into 5*5 and 2*2. The stride step point is fixed to 2 for each subsampling layer. As seen in Figure 7, the dimension and quantity of feature maps afterwards the convolution is 80*60*32. After that, an identical style is also applied to the remaining layers. In each of the two fully-connected layers, there are 1024 neurons. Finally, CNN generates the appropriate outputs based on the vote results of five different classed statements. The N (output), as shown in Figure 7, is calculated based on the images stored in the database.
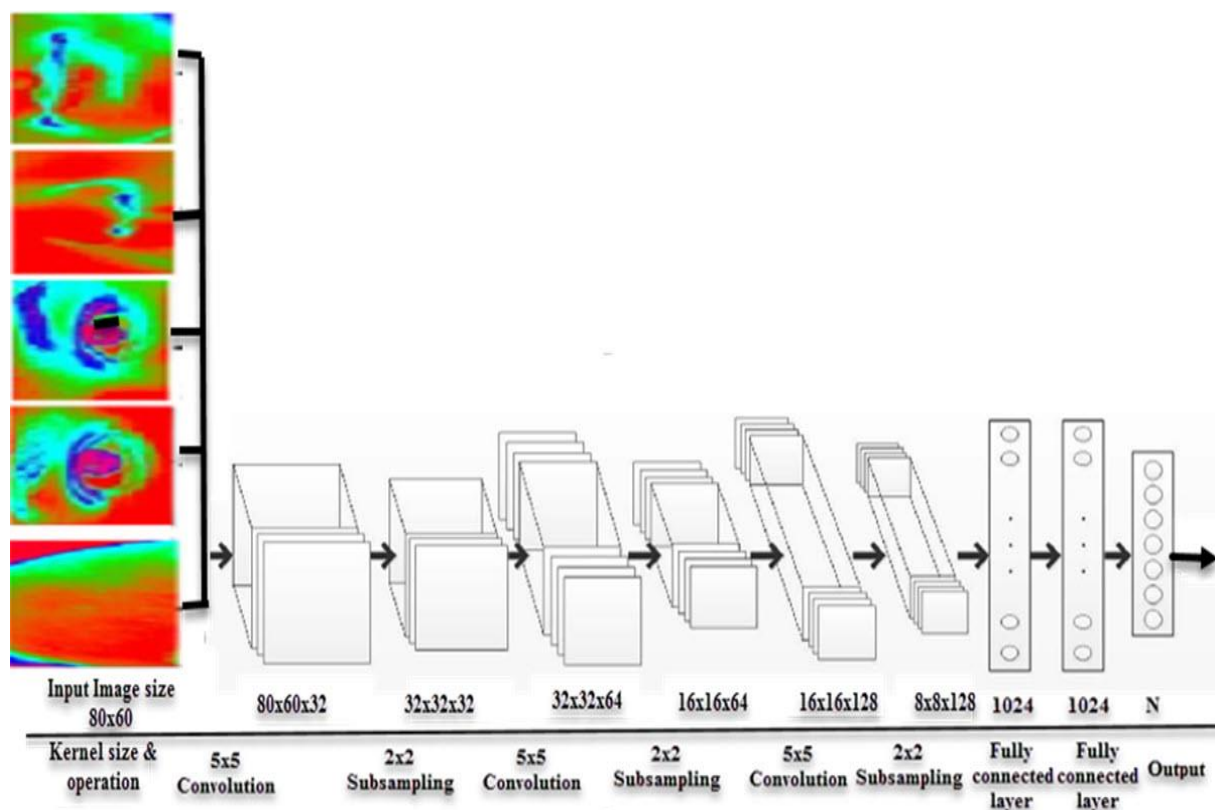

Figure 7. CNN structure with convolution layers, sub-sampling layers and fully connected layers.

6. Performance analysis

The testing is accepted from ground truth and CK+ datasets. The performance in tracking accuracy is shown in Table 1. It is shown that CK+ image-set detection in the nose-tip region is always higher, i.e., 93.98%. At the same time, the accuracy in the forehead, left eye, and right eye religion is about 89.71%, 90.89%, and 92.57%, respectively. The lip region has better accuracy, i.e., 93.02%, than the eyes regions. The result of tracking by using infrared images became higher than RGB images. Using twenty-one infrared image sets, we achieved 94.35% on the nose tip region. In the same way, the accuracy in the forehead, LE, RE, and nose-tip regions are 91.21%, 91.47%, 93.31%, and 94.16%, respectively. So the average detection became 92.96% instead of 92.03%, which is comparatively higher than the RGB image.

Table 1. Recognition accuracy.

Applied images	Forehead region accuracy	Left eye region accuracy	Right eye region accuracy	Nose tip region accuracy	Lip region detection accuracy
RGB images	89.71%	90.89%	92.57%	93.98%	93.02%
Infrared images	91.21%	91.47%	93.31%	94.35%	94.46%

Table 2. Performance result analysis with different investigators.

Investigators	Applied Method	Accuracy found by applied Datasets					
		JAFFE	CK+	MNI	NVIE	SFEW	OWN
S. Happy et al. [17]	SFP	85.06%	89.64%				
Y. Liu et al. [21]	AUDN + 8-fold validity	63.40%	92.40%				
L. Zhong, et al. [23]	CSPL		89.89%	73.53%			
S.H. Lee, et al. [27]	SRC	87.85%	94.70%	93.81%			
A.Mollahosseini et al. [29]	DNN		93.20%	77.90%			
R. Saranya et al. [31]	CNN + LSTM with 10-fold validity			81.60%		56.68%	95.21%
A.T. Lopes et al. [32]	Active appearance method (AMM)				66.24%		
	Integrated AAM				71.72%		
	DAF				79.16%		
	Integrated DAF				91.52%		
J.A. Zhanfu et al. [39]	AAM + infrared thermal + KNN				63.00%		
	Bayesian networks (BNs)				85.51%		
P. Shen, et al. [34]	KNN				73.00%		
Our applied method	CNN method applied on						
	1. Forehead region	89.32%					
	2. Left-Eye region	91.14%					
	3. Right-Eye region	90.75%					
	4. Nose region	91.15%					
	5. Lips region	93.32%					
Our proposed method	Decision tree level synthesized technique and 10-fold validation method applied on						
	1. Forehead region		91.36%				
	2. Left-Eye region		94.14%				
	3. Right-Eye region		93.49%				
	4. Nose region		94.44%				
	5. Lips region		97.26%				
Our Achievement	**We achieved 3.29% increased accuracy by incorporating decision tree level synthesized scheme and 10 folded validity.**						

7. Conclusion

A quicker method is applied in recognizing facial temperature by applying the infrared image because the infrared image is self-sufficient to measure temperature quickly. Moreover, we incorporated an AI-based machine intelligence scheme that accelerated the diagnosis process. In choosing the active regions from the six classes and twelve regions, we applied the Enhanced Cuckoo Search Optimization (ECSO) algorithm by incorporating the wireless network using a cloudlets server so that processing becomes more accessible with minimal infrastructure, as stated by Algorithm 2. The devices (infrared or mobile camera) will help measure temperature first. Then, deep learning CNN will be applied during the record processing and analysis to yield better-synthesized results. Table 1 shows the results regarding the temperature measurement accuracy of five regions. We demonstrated that the recognition rate is always greater when using CK+ image-set identification within the nose-tip region, i.e., 93.98 %. The accuracy of the left and right eye religions is around 90.89 % and 92.57 %. The accuracy of the lip region is higher, at 93.02 %, than those of the eyes. The tracking outcome improved when infrared images were used instead of RGB images. We gained 94.35 % on the nose tip region by applying twenty-one infrared images.

Similarly, accuracy is 91.47 %, 93.31 %, and 94.16 % in the left eye, right eye, and nose-tip regions. As a result, the average recognition rate increased to 92.39 % from 91.73%, more significant than the RGB image set. For example, it is shown in Table 2 that by the CNN method applied to Forehead, LE, RE, Nose, and Lips region, we achieved an accuracy of 89.32%, 91.14%, 90.75%, 91.15%, and 93.32%, respectively. On the other hand, by incorporating our Decision tree-level synthesis method and ten-folded-validation technique applied to the Forehead, LE, RE, Nose, and Lips region, we achieved an accuracy of 91.36%,94.14%,93.49%,94.44%, and 97.26% respectively. However, we achieved 3.29% greater accuracy by incorporating the "decision tree level synthesis scheme" and "ten-folded-validation method."

Acknowledgments

Conflict of interest

The authors declare there is no conflict of interest.

References

1. M. Karnati, A. Seal, G. Sahu, A. Yazidi, O. Krejcar, A novel multi-scale based deep convolutional neural network for detecting COVID-19 from X-rays, *Appl. Soft Comput.*, **125** (2022), 109109. https:// doi.org/10.1016/j.asoc.2022.109109

2. S. Vyas, A. Seal, A comparative study of different feature extraction techniques for identifying COVID-19 patients using chest X-rays images, in *2020 International Conference on Decision Aid Sciences and Application*, (2020), 209–213. https:// doi.org/10.1109/DASA51403.2020.9317299

3. G. N. Ahmad, S. Ullah, A. Algethami, H. Fatima, S. M. H. Akhter, Comparative study of optimum medical diagnosis of human heart disease using machine learning technique with and without

sequential feature selection, *IEEE Access*, **10** (2022), 23808–23828. https://doi.org/10.1109/ACCESS.2022.3153047

4. G. N. Ahmad, H. Fatima, S. Ullah, A. S. Saidi, Imdadullah, Efficient medical diagnosis of human heart diseases using machine learning techniques with and without GridSearchCV, *IEEE Access*, (2022), 1–24. https://doi.org/10.1109/ACCESS.2022.3165792

5. A. Chharia, R. Upadhyay, V. Kumar, C. Cheng, J. Zhang, T. Wang, et al., Deep-precognitive diagnosis: Preventing future pandemics by novel disease detection with biologically-inspired conv-fuzzy network, *IEEE Access*, **10** (2022), 23167–23185. https://doi.org/10.1109/ACCESS.2022.3153059

6. M. A. Hossain, S. A. Turkey, G. Sanyal, A novel stochastic tracking approach on human movement analysis, *Int. J. Comput. Appl.*, **86** (2014), 36–40. https://doi.org/10.5120/15089-3488

7. M. A. Hossain, D. Samanta, G. Sanyal, Eye diseases detection based on covariance, *Int. J. Comput. Sci. Inform. Secur.,* 2 (2012), 376–379.

8. N. M. Moacdieh, N. Sarter, The effects of data density, display organization, and stress on search performance: An eye tracking study of clutter, *IEEE Trans. Human Mach. Syst.*, **47** (2017), 886–895. https://doi.org/10.1109/THMS.2017.2717899

9. M. A. Hossain, B. Assiri, An enhanced eye-tracking approach using pipeline computation, *Arabian J. Sci. Eng.,* **45** (2020), 3191–3204. https://doi.org/10.1007/s13369-019-04322-7

10. K. Kurzhals, M. Hlawatsch, C. Seeger, D. Weiskopf, Visual analytics for mobile eye tracking, *IEEE Trans. Visual. Comput. Graph.*, **23** (2017), 301–310. https://doi.org/10.1109/TVCG.2016.2598695

11. M. A. Hossain, B. Assiri, Facial emotion verification by infrared image, *IEEE*, (2020), 12–14. https://doi.org/10.1109/ESCI48226.2020.9167616

12. M. Kim, B. H. Kim, S. Jo, Quantitative evaluation of a low-cost noninvasive hybrid interface based on EEG and eye movement, *IEEE Trans. Neural Syst. Rehab. Eng.*, **23** (2015), 59–168. https://doi.org/10.1109/TNSRE.2014.2365834

13. M. A. Hossain, H. Zogan, G. Sanyal, Emotion tracking and grading based on sophisticated statistical approach, in *International Conference on Science, Technology, Engineering and Mathematics*, (2018), 21–22.

14. D. Kumar, A. Dutta, A. Das, U. Lahiri, SmartEye: Developing a novel eye tracking system for quantitative assessment of oculomotor abnormalities, *IEEE Trans. Neural Syst. Rehab. Eng.*, **24** (2016), 1051–1059. https://doi.org/10.1109/TNSRE.2016.2518222

15. A. H. Mohammad, A. Basem, Emotion specific human face authentication based on infrared thermal image, in *International Conference on Communication and Information Systems*, (2020), 13–15. https://doi.org/10.1109/ICCIS49240.2020.9257683

16. Z. Kang, S. J. Landry, An eye movement analysis algorithm for a multielement target tracking task: Maximum transition-based agglomerative hierarchical clustering, *IEEE Trans. Human Mach. Syst.*, **45** (2015), 13–24. https://doi.org/10.1109/THMS.2014.2363121

17. W. Zhang, H. Liu, Toward a reliable collection of eye-tracking data for image quality research: Challenges, solutions, and applications, *IEEE Transact. Image Process.*, **26** (2017), 2424–2437. https://doi.org/10.1109/TIP.2017.2681424

18. S. Happy, A. Routray, Automatic facial expression recognition using features of salient facial patches, *IEEE Trans. IEEE Trans. Autom. Control*, **6** (2014), 1–12. https://doi.org/10.1109/TAFFC.2014.2386334

19. X. Zhang, S. M. Yua, An eye tracking analysis for video advertising: Relationship between advertisement elements and effectiveness, *IEEE Access*, **6** (2018), 10699–10707. https://doi.org/10.1109/ACCESS.2018.2802206

20. M. A. Hossain, G. Sanyal, Tracking humans based on interest point over span-space in multifarious situations, *Int. J. Software Eng. Appl.*, **10** (2016), 175–192. https://doi.org/10.1109/TAFFC.2014.2386334

21. Y. Liu, Y. Cao, Y. Li, M. Liu, R. Song, Y. Wang, et al., Facial expression recognition with PCA and LBP features extracting from active facial patches, *IEEE,* (2016), 368–373. https://doi.org/10.1109/RCAR.2016.7784056

22. M. A. Hossain, G. Sanyal, A novel approach to extract region from facial expression based on mutation, *Int. Janit. Clean. Serv. Assoc.*, **2** (2012), 15–18. https://doi.org/10.1109/RCAR.2016.7784056

23. M. A. Hossain, A. M. A Bamhdi, G. S. Sanyal, A new tactic to maintain privacy and safety of imagery information, *Int. J. Comput. Appl.*,**110** (2015), 6–12. https://doi.org/10.5120/19310-0764

24. L. Zhong, Q. Liu, P. Yang, B. Liu, J. Huang, D. Metaxas, Learning active facial patches for expression analysis, in *IEEE Conference on Computer Vision and Pattern Recognition*, (2012), 16–21. https://doi.org/10.1109/CVPR.2012.6247974

25. T. Wu, P. Wang, Y. Lin, C. Zhou, A robust noninvasive eye control approach for disabled people based on Kinect 2.0 sensor, *IEEE Sensors Letters*, **1** (2017), 1–4. https://doi.org/10.1109/LSENS.2017.2720718

26. A. H. Mohammad, G. Sanyal, Object tracking podium on region covariance for recognition and classification, *Int. J. Emerg. Technol. Comput. Appl. Sci.*, **2** (2012), 68–73.

27. S. H Lee, K. N. Plataniotis, Y. M. Ro, Intra-class variation reduction using training expression images for sparse representation based facial expression recognition, *IEEE Trans. Autom. Control*, **5** (2014), 340–531. https://doi.org/10.1109/TAFFC.2014.2346515

28. H. A. Mohammad, S. Samanta, S. Sanyal, Extraction of panic expression depending on lip detection, in *2012 International Conference on Computing Sciences*, (2012), 137–141. https://doi.org/10.1109/ICCS.2012.35

29. A. Mollahosseini, D. Chan, M. H. Mahoor, Going deeper in facial expression recognition using deep neural networks, in *2016 IEEE Winter Conference on Applications of Computer Vision*, (2016), 1–10. https://doi.org/10.1109/WACV.2016.7477450

30. M. A. Hossain, B. Assiri, Facial expression recognition based on active region of interest using deep learning and parallelism, *PeerJ Comput. Sci.*, **8** (2022), e894. https://doi.org/10.7717/peerj-cs.894

31. R. Saranya, C. Poongodi, D. Somasundaram, M. Nirmala, Novel deep learning model for facial expression recognition based on maximum boosted CNN and LSTM, *IET Image Process.*, **14** (2020), 1373–1381. https://doi.org/10.1049/iet-ipr.2019.1188

32. A. T. Lopes, E. Aguiar, A. F. De Souza, T. Oliveira-Santos, Facial expression recognition with convolutional neural networks: Coping with few data and the training sample order, *Pattern Recogn.*, **61** (2017), 610–628. https://doi.org/10.1016/j.patcog.2016.07.026

33. R. Janarthanan, E. A. Refaee, K. Selvakumar, M. A. Hossain, R. Soundrapandiyan, M. Karuppiah, Biomedical image retrieval using adaptive neuro-fuzzy optimized classifier system, *Math. Biosci. Eng.*, **19** (2022), 8132–8151. https://doi.org/10.3934/mbe.2022380

34. P. Shen, S. Wang, X. Liu, Facial expression recognition from infrared thermal videos, in *Intelligent Autonomous Systems 12. Advances in Intelligent Systems and Computing* (eds S. Lee, H. Cho, K. J. Yoon and J. Lee), Springer, (2013), 323–333. https://doi.org/10.1016/j.imavis.2011.07.002

35. S. Pastel, J. Marlok, N. Bandow, K. Witte, Application of eye-tracking systems integrated into immersive virtual reality and possible transfer to the sports sector-A systematic review, *Multimed. Tools Appl.*, (2022). https://doi.org/10.1007/s11042-022-13474-y

36. M. A. Hossain, D. Samanta, G. Sanyal, Statistical approach for extraction of panic expression, in *2012 Fourth International Conference on Computational Intelligence and Communication Networks*, 420–424, https://doi.org/10.1109/CICN.2012.189

37. S. S. Alam, R. Jianu, Analyzing eye-tracking information in visualization and data space: From where on the screen to what on the screen, *IEEE Trans. Visual. Comput. Graph.*, **23** (2017), 1492–1505. https://doi.org/10.1109/TVCG.2016.2535340

38. D. H. Jiang, Y. Z. Hu, D. Lei, P. Jin, Facial expression recognition based on attention mechanism, *Sci. Program.,* **2021** (2021), Article ID 6624251. https://doi.org/10.1155/2021/6624251

39. Z. An, W. Deng, J. Hu, Y. Zhong, Y. Zhao, Adaptive pose alignment for pose-invariant face recognition, *IEEE Access*, **7** (2019), 14653–14670. https://doi.org/10.1109/ACCESS.2019.2894162

40. M. D. H. Alamgir, D. Samanta, G. Sanyal, Automated smiley face extraction based on genetic algorithm, *Comput. Sci. Inform. Technol.*, (2012), 31–37. https://doi.org/10.5121/csit.2012.2304

41. J. Y. Choi, B. Lee, Ensemble of deep convolutional neural networks with gabor face representations for face recognition, *IEEE Transact. Image Process.*, **29** (2020), 3270–3328. https://doi.org/10.1109/TIP.2019.2958404.

42. Z. Lei, L. Ji, Z. Bob, Z. David, Z. Ce, Deep cascade model-based face recognition: When deep-layered learning meets small data, *IEEE Transact. Image Process.*, **29** (2020), 1016–1029. https://doi.org/10.1109/TIP.2019.2938307

43. A. H. Mohd, S. Gautam, A new improved tactic to extract facial expression based on genetic algorithm and WVDF, *Int. J. Adv. Inform. Technol.,* **2** (2012), 37–44. https://doi.org/10.5121/ijait.2012.2504.37

44. A. C. Elizabeth, K. J. Nai, E. D. Susan, A. B. Martha, L. Jacob, L. G. Daniel, et al., The facial action coding system for characterization of human affective response to consumer product-based stimuli: A systematic review, *Front. Psychol.*, **11** (2020), 920. https://doi.org/10.3389/fpsyg.2020.00920

45. S. M. Lajevardi, Z. M. Hussain, Automatic facial expression recognition: Feature extraction and selection, *Signal Image Video Process.*, **6** (2010),159–169. https://doi.org/10.1007/s11760-010-0177-5

46. M. H Alamgir, D. S. S. Goutam, A novel approach for panic-face extraction based on mutation, in *International Conference on Advanced Communication Control & Computing Technology*, (2012), 473–477. https://doi.org/10.1109/ICACCCT.2012.6320825

47. F. Zhang, F. Wang, Exercise fatigue detection algorithm based on video image information extraction, *IEEE Access*, **8** (2020), 199696–199709. https://doi.org/10.1109/ACCESS.2020.3023648

48. M. A. Hossain, G. Sanyal, Extraction of panic expression from human face based on histogram approach, in *International Conference on Image Processing*, (2012), 411–418. https://doi.org/10.1007/978-3-642-31686-9_48

49. C. Vincenzo, G. Antonio, P. Gennaro, V. Mario, Age from faces in the deep learning revolution, *IEEE Trans. Pattern Anal. Mach. Intell.*, **42** (2020), 2113–2132. https://doi.org/10.1109/TPAMI.2019.2910522.

50. B. Jin, C. Leandro, G. Nuno, Deep facial diagnosis: Deep transfer learning from face Recognition to facial diagnosis, *IEEE Access,* (2020). https://doi.org/10.1109/ ACCESS.2020.3005687

51. L. Daqi, B. Nicola, Y. Shigang, Deep spiking neural network for video-based disguise face recognition based on dynamic facial movements, *IEEE Trans. Neural Netw. Learn. Syst.,* **31** (2020), 1843–1855. https://doi.org/10.1109/TNNLS.2019.2927274

52. Y. Said, M. Barr, H. E. Ahmed, Design of a face recognition system based on convolutional neural network (CNN), *Eng. Technol. Appl. Sci. Res.,* **10** (2020), 5608–5612. https://doi.org/10.1109/CAC48633.2019.8996236

53. A. Nada, H. A. B. Heyam, Deep learning approach for multimodal biometric recognition system based on fusion of iris, face, and finger vein traits, *Sensor*, **20** (2020), 5523–5539. https://doi.org/10.3390/s20195523

54. A. Jawad, A. AlBdairi, Z. Xiao, M. Alghaili, Identifying ethnics of people through face recognition: A deep CNN approach, *Sci. Program.,* **2020** (2020), article ID 6385281. https://doi.org/10.1155/2020/6385281

55. W. W. R. Almeida, F. A. Andaló, R. Padilha, G. Bertocco, W. Dias, R. S. Torres, et al., Detecting face presentation attacks in mobile devices with a patch-based CNN and a sensor aware loss function, *PLoS One*, (2020), 1–24. https://doi.org/10.1155/2020/6385