



Race Detection using Mutated Salp Swarm Optimization Algorithm based DBN from Face Shape Features

Ayesha Hojage

RMIT University, Melbourne, Australia
 ayeshahojage@gmail.com

Abstract: Humans can be easily recognized because of their uniqueness in nature, also with their demographic characteristics namely gender, ethnicity such as race, and age correspondingly. In the present decades, numerous numbers of studies have been conducted in the areas such as biological, psychological and cognitive sciences in order to identify how the human brain can memorize, perceives and characterizes human faces. In addition, specific computational developments are performed to attain numerous aspects of this problem. This work attempts to present a novel race detection technique by exploiting the face shape features. Here, the developed model comprises two stages such as feature extraction and detection. In the primary phase, the feature extraction is performed, here the face shape, and face color-based feature, is mined. Particularly, Speeded-up Robust Transform (SURF), and Maximally Stable Extremal Regions (MSER), is extracted in shape features and dense color features are extracted as a color feature. In large dimensions, extracted features are presented; they are changed in the Principle Component Analysis (PCA) technique that is considered the strongest technique to solve the “curse of dimensionality”. Subsequently, dimensional minimized features are fed to Deep Belief Network (DBN), whereas the race is identified. Furthermore, regarding the prediction to make the developed model highly efficient, the DBN weight is fine-tuned with a novel technique called mutated Salp Swarm Optimization Algorithm. The adopted technique is evaluated with existing techniques regarding the accuracy and error performance.

Keywords: DBN, Face Detection, Feature Extraction, PCA, Race.

Nomenclature

| Abbreviations | Descriptions |
|-------------------|--|
| SURF | Speeded-up Robust Transform |
| mutated-SOA | mutated Salp Swarm Optimization Algorithm |
| BS | Base Station |
| PCA | Principle Component Analysis |
| DCF _s | Discriminative Complete Features |
| DBN | Deep Belief Neural Network |
| PAD | Presentation Attack Detection |
| FHEDN | Feature Hierarchy Encoder– Decoder Network |
| MSER | Maximally Stable Extremal Regions |
| DCNN _s | Deep Convolutional Neural Networks |
| PCA | Principle Component Analysis |
| NN | Neural Network |
| DEA | Differential Evolution Algorithm |

1. Introduction

In intelligent communication identity examination, face detection plays an important role. Because of some situations, human face detection tasks attained numerous challenges. Conversely, the human face is attained by other parts in the image namely general masks, scarves, glasses so on. On contrary, because of the camera angle, the face occluded. In order to recognize the human face and to find the face position turns out to be the important work in face recognition applications [1].

Using statistics, an important rule of missing data is attained that possesses applications in numerous fields. It is very necessary to enhance the occluded face detection performance because it possesses numerous problems, which need to be settled immediately. In current days, conventional

researches on occluded face detection are importantly represented in 2 models as a technique on the basis of the occlusion repair and a technique by exploiting the local auxiliary recognition. In other areas, as information repair techniques, the occlusion repair technique initially aspires to attain the database of the face feature using a huge number of face samples learning and tries to retain the features which are missing of the occluded samples [2].

In the face recognition system scenario, spoofing attacks, in general, consider as print attacks. During the last two years, diverse face PAD techniques were presented, which achieves an extraordinary performance on benchmarking datasets. Mainly, the development is performed was aided in part by the EPSRC Programme Grant be obliged to two kinds: the deployment, and model of more efficient indications that can enhanced detain the differences among fake and real biometrics traits, the next one, is performed by exploiting highly powerful two -class classifiers. In much literature, many techniques to face spoofing detection were worked and it formulates the issues as a two class pattern identification issue. These techniques attempt to learn an appropriate classifier to distinguish between spoofing attempts and real access [10].

In the last two decades, numerous techniques for face detection were elaborated deeply. It was done with the advantage of the big data era and computing platforms; deep learning attains extraordinary performance in numerous tasks namely computer vision, natural language processing, and recognition of speech, etc [11]. In the object detection field, the DCNNs based approaches were initiated to apply and attained superior performances. Since the attained object does not cover small-sized faces, the recognized techniques are generally not efficient for recognizing the little-size faces [3].

The major objective of this research is to develop a race detection technique by exploiting face shape features. Here, the developed technique comprises two important stages such as detection and feature extraction. The feature extraction is the most important phase; the image is obtained from the facial image. In this phase, the face shape, and color-based features, are extracted. Here, SURF and MSER features are extracted. Then the PCA is exploited to minimum dimension and they are subjected to DBN. The weight of DBN is tuned using the novel mutated-SOA algorithm.

2. Literature Survey

In 2020, Soroush Fatemifar et al [1] presented a training model for one-class client-precise classifiers (both discriminative and generative) by exploiting the indications attained from pre-trained DCNN. To integrate client-specific information, a different threshold was set for every client based on the subject-specific score distributions that were subsequently exploited to decision making during test time. In 2018, Zexun Zhou et al [2], developed a DCNN called FHEDN. It was enthused by two examinations from contextual semantic information and a multi-scale face detection model. The developed network was scale-variant style model and single-phase that was decoder composed and encoder sub-networks. On the presumption of that contextual semantic information approximately face being auxiliary for faces, a remaining model to combine context prior-based information was developed. In 2020, Qi Jin et al [3], worked on face detection which was considered as an initial step. In many applications, the face recognition model was introduced such as identity verification, intelligent interaction, and mobile social networks. Nevertheless, face occlusion in conventional models was presented for detection techniques, ensuing in the complexity by exploiting these techniques in practical cases. A novel detection technique called FSG-FD with attention method, and aspire to distinguish face with occlusion was presented. In 2018, GuanJun Guo et al [4], presented a rapid face detection model on the basis of the DCFs extracted using a highly modeled CNN, whereas face detection was directly carried out on entire feature maps. DCFs have exhibited the capability of scale invariance that was helpful for face detection with a high speed and shows potential performance. Hence, multi-scale features were extracted on an image pyramid used in existing techniques was not needed in a developed technique that can seriously enhance its effectiveness for detection of face. In 2020, Abdulaziz Zam et al [5], worked on the detection phase at Visual Sensor Network and the recognition phase was attained at the BS. This paper worked on three phases such as energy-aware face detection technique, then extracting the lightweight discriminative vector with minimum transmission cost. At the BS, an effectual face recognition technique was carried out at the received vectors.

3. Proposed Model procedure

The main contribution of the present research is to model a new race detection technique from facial images. Hence, in this paper, two important main stages are considered such as feature extraction well as classification. In the primary phase, the feature extraction is considered; from the input facial

image Im_{face} where the face shapes, and color-based features, are extracted. Specifically, the MSER and SURF are used to extract the shape feature and the color feature is used for the dense color feature. The extracted SURF is F^{SURF} , MSER is F^{MSER} , the dense color feature is F^{DCR} respectively. Here, F denotes the overall feature, and it is represented as the integration of, F^{SURF} , F^{MSER} , and F^{DCR} . Here, the features extracted (F) are high in dimensions; the model might complexity boost up. Therefore, to compress the extracted datasets, PCA is applied which compresses the data sets into lower-dimensional features by preserving appropriate datasets. F_{DIM} represents the dimensional minimized feature. The extracted appropriate features (F_{DIM}) are fed to classification through DBN.

Furthermore, it is tried to optimally fine-tune the DBN weight (Wg) with a novel optimization technique. At last, the results from DBN show exact races of humans. Fig 1 demonstrates an architectural model of the developed race detection.

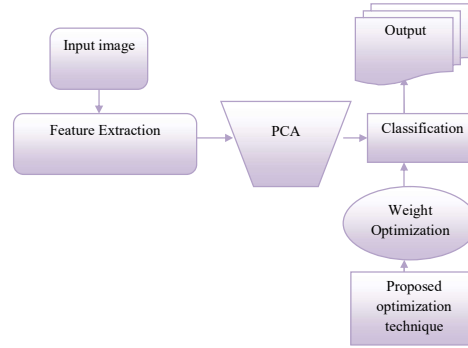


Fig. 1. Architecture model of the developed Facial Race Detection technique

4. DBN Based Classification Technique and Feature Extraction

4.1 Feature Extraction

Im_{face} indicates the input facial image, with that the feature extraction is performed. Moreover, to maximize the classification task accuracy two kinds of features are extracted such as color and shape features.

a) Shape Features:

For the shape feature, the MSER and SURF are considered as the color features that are extracted from input facial images Im_{face} .

b) SURF features [6]:

It is an in-plane rotation invariant feature which covers interest detector and descriptor point. In the facial image, using the detector, the point of interest is localized, and using the descriptor features of these points is shown.

Interest Point Detection: The Hessian approximation matrix's determinants are considered as SURF detector bases. During positions where the determinants are maximum, with a blob-like structure, the interest point is marked. To minimize the computational time, the Hessian approximation matrix exploits integral images. In Im_{face} , at point $S = (a, b)$, hessian matrix $H(S, \sigma)$ subsists. In eq. (1), the Hessian matrix is indicated at scale σ in S . Moreover, the parameters $L_{ab}(S, \sigma)$, $L_{aa}(S, \sigma)$, and $L_{bb}(S, \sigma)$ indicates 2nd order Gaussian partial derivatives at the point S .

$$H(S, \sigma) = \begin{bmatrix} L_{aa}(S, \sigma) & L_{ab}(S, \sigma) \\ L_{ab}(S, \sigma) & L_{bb}(S, \sigma) \end{bmatrix} \quad (1)$$

Furthermore, in lower scales, the blob response matrix is calculated with the purpose of minimizing the computational time. This higher special resolution is indicated as, $D_{ab}(S, \sigma)$, $D_{aa}(S, \sigma)$ and $D_{bb}(S, \sigma)$. To enhance the computation effectuality, the weights used are maintained easier in the rectangular shape. It is stated in eq. (2). Moreover, ω indicates energy conservation amid Gaussian kernels and approximated Gaussian. The formulation for ω is stated in Eq. (3).

$$\det(H_{app}) = D_{aa} \cdot D_{bb} - (\omega D_{ab})^2 \quad (2)$$

$$\omega = \frac{|L_{aa}(\sigma)|_f |D_{bb}(\text{Size of box filter})|_f}{|L_{bb}(\sigma)|_f |D_{ab}(\text{Size of box filter})|_f} \quad (3)$$

Interest Point Description: To show the feature of point of interest Haar wavelet response is used. Moreover, by modeling a square region with the point of interest as its center the descriptor is extracted. The region is divided into lesser sub-regions. From the input image, Im_{face} the SURF feature extracted is indicated as F_{SURF} .

MSER feature [7]: It is a blob recognition technique and it aids in extracting a number of co-variances from the image. In input facial image Im_{face} , consider threshold image be Im_{th} and th indicates the sequence equivalent to the thresholded images. Moreover, at first, the black image is analyzed subsequently the white spots are described.

➤ MSER: Let $\mathfrak{R}_1, \dots, \mathfrak{R}_{i-1}, \mathfrak{R}_i, \dots$ signifies series of nested extreme areas, $\mathfrak{R}_i \subset \mathfrak{R}_{i+1}$. There is utmost stability in the extremal region \mathfrak{R}_{i^*} while $h(i) = |\mathfrak{R}_{i+\Delta} \setminus \mathfrak{R}_{i-\Delta}| / |\mathfrak{R}_i|$ indicates the least local at i^* . In this technique, the parameter is indicated as $\Delta \in J$.

Therefore, extracted MSER feature of the image Im_{face} is indicated as F_{MSER} .

Dense color features: It is the concatenation of the dense SIFT features and a dense color histogram. From the image is indicated as F_{DCF} , dense color feature extracted.

Furthermore, extracted MSER feature (F_{MSER}), surf feature (F_{SURF}), and dense color feature, (F_{DCF}) is together indicated as F . This complete feature F can be indicated as stated in Eq. (4).

$$F = \{F_{MSER}, F_{SURF}, F_{DCF}\} \quad (4)$$

4.2 Dimensionality Reduction using PCA

The extracted features F are almost certainly enormous in dimensions and therefore their requirements as the minimization. PCA is the renowned technique for dimensional reduction as it resolves the issue of the “curse of dimensionality” without loss of information. The mathematical explanation on PCA technique is stated as follows:

Mean is the central tendency measure. Q indicates an arbitrary count, and sample size, are indicated as l in Eq. (5).

$$\text{Mean}(\bar{Q}) = \frac{1}{l} \sum_{i=1}^l Q_i \quad (5)$$

Standard Deviation (SD) exhibits regarding the rapidity of data. The average distance amid mean and point therefore the data is estimated by squaring them and it is stated in Eq. (6).

$$SD = \sqrt{\frac{1}{l} \sum_{i=1}^l (Q_i - \bar{Q})^2} \quad (6)$$

Covariance: It estates regarding the variations quantity in dimension from the mean is stated in Eq. (7).

$$\text{Cov}(Q, R) = \frac{\sum_{i=1}^l (Q_i - \bar{Q})(R_i - \bar{Q})}{l} \quad (7)$$

Formerly, the extracted features are minimized with their dimensions; they are normalized which are fed to classification.

4.3 Classification using DBN

DBN was introduced [8] in 2018. From Boltzmann networks, from stochastic neuron results (σ) are considered as the probabilistic and it possesses a probability function $pf(\tau)$, it is indicated. Furthermore, $pf(\tau)$ is sigmoid-shaped function and stated in Eq. (9). PT indicates the pseudo-temperature and with PT , the stochastic technique is converted to deterministic model and it is stated in Eq. (10). Furthermore, the Boltzmann machine is referred the recurrent network and it is similar to stochastic NN with stochastic neurons. As per Eq. (11), the Boltzmann machine’s energy in the configured state of a neuron ns is described. The weight amid neurons x and y is indicated by $w_{x,y}$ and biases are indicated as ψ_x .

$$pf(\tau) = \frac{1}{1 + e^{\frac{-\tau}{PT}}} \quad (8)$$

$$o = \begin{cases} 0 & \text{with probability of } 1 - pf(\tau) \\ 1 & \text{with probability of } pf(\tau) \end{cases} \quad (9)$$

$$\lim_{PT \rightarrow 0^+} pf(\tau) = \lim_{PT \rightarrow 0^+} \frac{1}{1 + e^{\frac{-\tau}{PT}}} = \begin{cases} 0 & \text{for } \tau < 0 \\ 1 & \text{for } \tau = 0 \\ 2 & \text{for } \tau > 0 \end{cases} \quad (10)$$

$$En(cs) = - \sum_{x < y} cs_x cs_y wt_{x,y} - \sum_x \psi_x cs_x \quad (11)$$

In solitary unit referred to as cs_x the global energy computation stated in the eq. (12). Moreover, the gradient descends technique aids in calculating the mainly least probable energy of the neurons during training.

$$\Delta En(cs_x) = \sum_y cs_x wt_{x,y} + \psi_x \quad (12)$$

On the basis of Eq. (14), (15), and (16), energy descriptions compared with the joint composition of visible and hidden neurons are stated. The binary state of the hidden unit y and visible unit x is indicated as g_x and k_y , correspondingly.

$$En(\vec{g}, \vec{k}) = \sum_{(x,y)} wt_{x,y} g_x k_y - \sum_x u_j a_i - \sum_y v_j b_i \quad (13)$$

$$\Delta En(g_x, \vec{k}) = \sum_y wt_{xy} k_y + d_x \quad (14)$$

$$\Delta En(\vec{g}, k_y) = \sum_y wt_{xy} g_x + m_y \quad (15)$$

Eq. (16) indicates weight obligation performed for DBN to produce maximum probability. In this paper, the weight assignment is indicated by Wg_q , which is optimally tuned by exploiting a novel optimization technique. The RBM probability assignment is indicated to all visible pairs and hidden neurons regarding energy in Eq. (17). The (Z_p) represents the partition function and it is obtained by summation of all neuron probable energy states, Z_p is indicated as Eq. (18).

$$Wg_q = \max_{\vec{g} \in V_t} \Pi pf(\vec{g}) \quad (16)$$

$$pf(\vec{g}, \vec{k}) = \frac{1}{Z_p} e^{-En(\vec{g}, \vec{k})} \quad (17)$$

$$Z_p = \sum_{\vec{g}, \vec{k}} e^{-En(\vec{g}, \vec{k})} \quad (18)$$

The training samples (g) selection is attained and they are braced on to visible conditions. Calculate hidden neuron probability (pf_k) by multiplying visible vector as g by means of weight matrix as Wg as $pf_k = \sigma(v.Wg)$ and it is stated as eq. (19).

$$pf(k_y \rightarrow 1 | \vec{g}) = \sigma \left(m_y + \sum_x g_x wt_{x,y} \right) \quad (19)$$

5. Proposed Optimization Algorithm

5.1 Objective Model

The main objective of this paper is to optimize the DBN classifier by tuning optimal weights, which is taking into consideration the optimization issue, therefore a novel optimization approach is developed. As stated in eq. (20), the objective model of the developed technique is stated in eq. (20), whereas Acc denotes the classification accuracy.

$$Obj = \max(Acc) \quad (20)$$

5.2 Proposed Mutated-SSO Algorithm

In [9], the salp swarm updates the position initially, and after the salps fitness is estimated and optimal salp is selected as optimal salp amid salps.

Additionally, when the fitness $Salp_i(Fit_i(S_i))$ is enhanced then the fitness of optimal $Salp(Fit_i(S_i))$ it is selected as optimal salp for the subsequent generation. The adopted optimization technique is the same as the conventional SSA yet while a salp is not elected as the optimal $Salp(F)$ a mutation is employed on it. There is a higher tackle of mutation model here, salps are mutated by employing mutation of DEA:

$$S_i = X_{r1} + F_m^*(X_{r2} - X_{r3}) \quad (21)$$

whereas, F_m^* signifies constant-coefficient amid [0, 2], and X_{r1}, X_{r2}, X_{r3} states 3 vectors that are arbitrarily chosen.

The Proposed method steps are described as below:
Step 1: Initialize population and parameters: At first, method parameters like DV, XU, NS, XL, MNI, etc are set. Subsequently, population (salps) are produced amid XL and XU as below:

$$S_i^j = XL_j + (XU_j - XL_j) * \text{rand}[0,1] \quad i = 1, 2, \dots, NS, j = 1, 2, \dots, DV \quad (22)$$

Step 2: Updating locations of salps: initially, leader location is updated and after the other salps pursue leader and they create a chain.

$$S_i^j = \begin{cases} XL_j + c_1 \left[(XU_j - XL_j) * r_1 \right] & r_2 \geq 0 \\ XL_j - c_1 \left[(XU_j - XL_j) * r_1 \right] & r_2 < 0 \end{cases} \quad (23)$$

whereas, r_1 and r_2 indicates random count amid [0, 1], S_i^j , F_i^j indicates the leader's location (initial salp) and optimal search agent in j^{th} dimension. c_1 indicates the important parameter to update leader and it is calculated on the basis of eq. (24).

$$c_1 = 2e^{-\left(\frac{4itr}{MNI}\right)} \quad (24)$$

whereas, itr and MNI indicates the current iteration and an utmost count of iteration, correspondingly. The salp position i^{th} S_i^j is updated on the basis of the previous salp as eq. (25):

$$S_i^j = \frac{X_j^i - X_j^{i-1}}{2} \quad (25)$$

Step 3: Estimating objective approach: At first amend salp position i^{th} amid XU and XL. Afterward, calculate it on basis of objective technique.

Step 4: Applying mutation: fitness of salp i^{th} is evaluated to optimal search agent (F). If a salp i^{th} is superior to optimal search agent (F), it is will be validated. Else, a mutation is used on location of salp i^{th} as below:

$$S_{is} = X_{r1} + F_m^*(X_{r2} - X_{r3}) \quad (26)$$

whereas, F_m^* indicates constant-coefficient amid [0, 2], X_{r1}, X_{r2}, X_{r3} indicates "3" vectors that are arbitrarily chosen.

Step 5: Halting circumstance: If $itr < MNI$, steps 2 to 4 are repeated, or the method is halted.

Step 6: return F

6. Result and Discussion

The adopted race detection model was developed in this paper and the results attained were recorded. The graphical representations of the outcomes of the adopted model were shown in Fig. 2, 3, and 4. Here, the performance of the classifier was evaluated with the conventional classifiers regarding the positive and negative measures.

Fig 2, 3, and 4 show the Learning Percentage is LP which possesses a higher influence on the adopted technique performance. Here, the performance of the adopted technique regarding the positive and negative **metrics** is varied with the learning percentage. In some cases, the proposed model tends to minimize with minimizing in LP or vice versa. Here, the LP is varied 50, 60, and 70, the accuracy of the adopted technique is better than traditional techniques. Here, the race detection lies to maximize while increase with the LP.

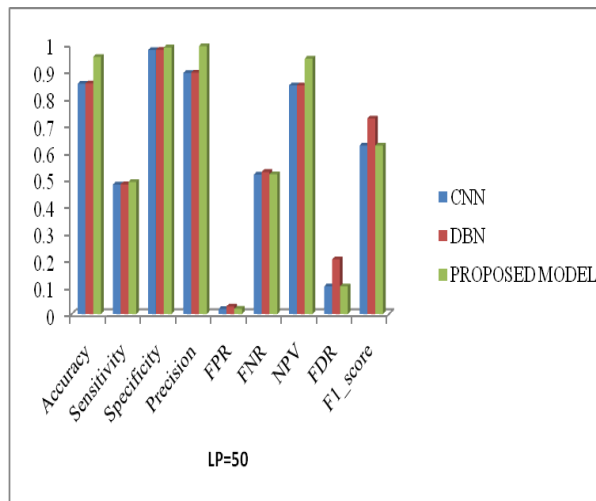


Fig 2. Performance evaluation of adopted and traditional techniques regarding LP=50

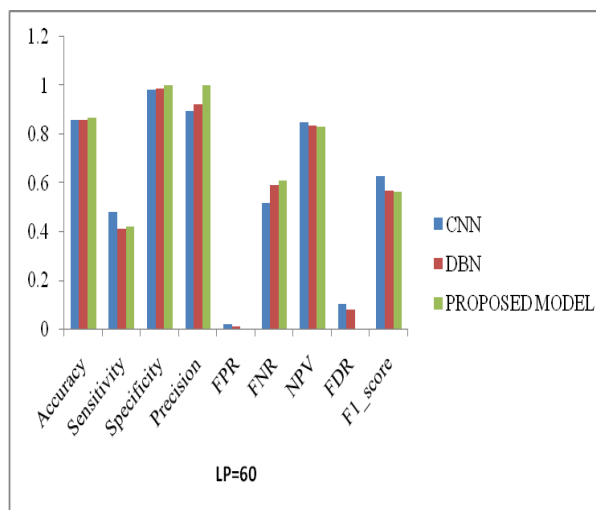


Fig 3. Performance evaluation of adopted and traditional techniques regarding LP=60

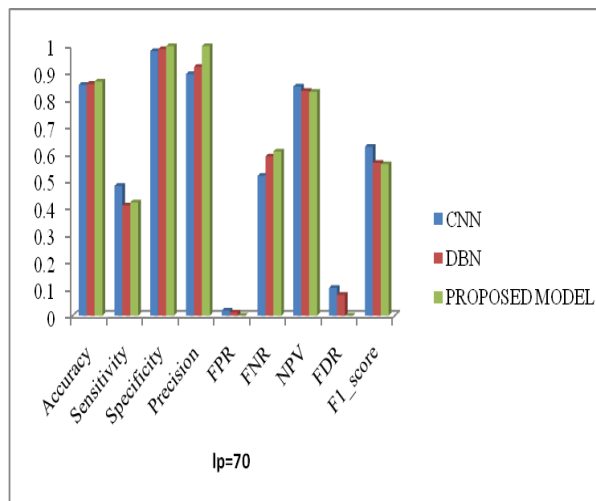


Fig 4. Performance evaluation of adopted and traditional techniques regarding LP=70

7. Conclusion

In this work, a novel race detection model was adopted using two important stages such as detection and feature extraction. In the initial phase, from the facial image, the input image was obtained which was fed to the feature extraction. Here, face shape, and color-based features, were extracted. Specifically, SURF and MSER features were extracted in dense color features, and shape features were extracted as a color feature. As extracted features were large in dimensions; they were reduced in PCA technique. Using DBN, dimensional minimized features were fed to classification. To produce the race recognition in a precise way, the DBN weight was fine-tuned using a novel optimization technique called the mutated-SOA algorithm.

Compliance with Ethical Standards

Conflicts of interest: Authors declared that they have no conflict of interest.

Human participants: The conducted research follows the ethical standards and the authors ensured that they have not conducted any studies with human participants or animals.

References

- [1] Soroush FatemifarShervin Rahimzadeh ArashlooJosef Kittler,"Client-specific anomaly detection for face presentation attack detection",Pattern Recognition, vol. 112, 26 October 2020.
- [2] Zexun ZhouZhongshi HeZiyu Chen,"Context prior-based with residual learning for face detection: A deep convolutional encoder–decoder network",Signal Processing: Image Communication, vol.88, 25 July 2020.
- [3] Qi JinChong MuFankai Ran,"A Region Generation based Model for Occluded Face Detection",Procedia Computer Science, vol. 174, pp.454-462, 27 July 2020.
- [4] GuanJun GuoHanzi WangBo Li,"A fast face detection method via convolutional neural network", Neurocomputing, vol. 395, pp. 128-137, 13 July 2019.
- [5] Abdulaziz ZamMohammad Reza KhayyambashiAli Bohlooli,"Energy-efficient face detection and recognition scheme for wireless visual sensor networks", Applied Soft Computing, vol.89, 15 January 2020.
- [6] Xiangnan Zhang, Xinbo Gao, Chunna Tian,"Text detection in natural scene images based on color prior guided MSER",Neurocomputing,vol.307,pp.61-71,September 2018
- [7] Geng Du, Fei Su, Anni Cai,"Face recognition using SURF features",Pattern Recognition and Computer Vision,2009.
- [8] Md. Zia Uddin, Mohammed Mehedi Hassan, Ahmad Almogren, Mansour Zuair, Jim Torresen,"A facial expression recognition system using robust face features from depth videos and deep learning",Computers & Electrical Engineering,vol.63, pp.114-125,October 2017.
- [9] Mirjalili S, Gandomi AH, Mirjalili SZ, Saremi S, Faris H, Mirjalili SM. Salp Swarm Algorithm: A bio-inspired optimizer for engineering design problems. Adv Eng Softw, pp. 114:163–91, 2017.
- [10] K.Srinivas,"Prediction of E-Learning Efficiency by Deep Learning in EKhool Online Portal Networks",Multimedia Research, vol 3, no 4, October 2020.
- [11] Arvind Madhukar Jagtap,"Developing Deep Neural Network for Learner Performance Prediction in EKhool Online Learning Platform", Multimedia Research, vol 3, no. 4, October 2020.