Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 2, No. 3, February : 2024

IMAGE RECOGNITION OF VOTERS ON POLLING DAYS USING PCA-CNN TECHNIQUES (VAM MODEL)

Mr. Adarsh Lalit Chaubey, Dept. Of Computer Science, Mithibai College, Mumbai University.
Mr. Amol Jogalekar, Dept. Of Computer Science, Mithibai College, Mumbai University.

Abstract
With the rise of focus on AI, there emerge benefits and challenges hand in hand in various fields. This research work focuses on one of such fields i.e. Elections. Election plays a crucial role in every citizen's life and to maintain its stability one has to overcome the challenges of booth capturing and bogus votes. To put light on the view; Paper proposes a model, the Voters Attendance Machine (VAM), which is to be used simultaneously with the Electronic Voting Machine (EVM) during the election days to maintain the democracy of the nation and integrity of election. As the images and voter IDs dataset per booth are large in numbers and it becomes a tedious task to verify authorized voters on election day, hence, the VAM model will act as an attendance machine to recognize voters on polling days. VAM model will be trained with the PCA-CNN technique which will first use Principal Component Analysis (PCA) tools for classifying the images on the features (say Gender, Age group, etc), later it will reduce the features under a domain with similar configurations to low down the computational time of CNN and though it gets a filtered clean data.

Secondly, the VAM model will use the resultant output image of PCA as input data for CNN where it will go through the process of Convolutional Neural Network (CNN) tools by filtering the images using max-pooling and then connecting it to a single-layer network i.e., Fully-connected Layer to recognize voters present face with the image of voter stored in voter id. The motive is to train the model with PCA-CNN techniques to reduce the bogus vote and booth capturing as the VAM model acting as an attendance machine will detect the face of the voter on polling days and will mark them present in his attendance database. Using PCA first is to get more precise results by reducing the dimension of the dataset in starting itself so that the workload of CNN reduces and its efficiency increases. By this, a model can store a large number of image datasets in reduced dimension format which will require less storage space with higher accuracy. Overall, by this combination, the process gets simplified in structure and also time taken to train the models decreases as compared to traditional CNN with an increase in accuracy and speedy results.

Keywords:
Convolutional Neural Network (CNN), Principal Component Analysis (PCA), Max-Pooling Eigenvectors

I. Introduction

By making use of a camera for a facial recognition system as Attendance Machine has applications in various sectors like political (Elections), Economic (Business Markets), Educational (attendance), social (welfare schemes), etc. The design and model of facial attendance systems using cameras are very effective, less expensive, and efficient to further improve the standard of user data and are useful for fulfilling the needs of a diversity of users who use it. A facial attendance system using a camera is shielded with no harm and accurate for recognizing the users because it has a more precise data process so that it can produce a system that is well-founded and robust to identify human faces by its stored structure of images. [3].

Research on Face Recognition for human interests has been carried out, especially for various interests such as security systems, government health I-cards, surveillance, General identity verification, Image database investigations, MGNREGA attendance, Criminal justice systems, "Smart Card" applications, Video indexing, multi-media environments, and Witness face reconstruction [3].

There are two categories in intensity mages -based Face Recognition, feature-based and holistic face recognition. Feature-based face recognition is a face recognition method that processes input images to identify and extract features of faces such as nose, mouth, eyes, etc. then calculates the geometric relationships between these face points thereby reducing the face input images to geometric feature vectors. The initial face recognition research based on feature-based was conducted by [3].

II.What is CNN & PCA

2.1 Convolutional Neural Network (CNN)

Convolutional Neural Networks (CNN) come under the category of Deep Learning algorithms as Neural Network Architectures mostly help in image recognition as a vision to the computer eye. Neural networks are used in various models of datasets like pictures, videos, audio, and text. Different types of Neural Networks are used for different purposes, for example for predicting the text more precisely we use the algorithm Autoencoders, similarly for Image detection, recognition, or classification we use CNN. Commonly there are 3 different types of layers in order such as: Input Layer Hidden Layer - Output Layer.

CNN is a hierarchical neural network, in which the input is a digital image, while the output is a prediction through the distinctive features extracted via convolutional layers alternate with subsampling layers. It has been widely used in handwritten or digital character recognition, face recognition, pedestrian attribute classification, driving posture recognition, human pose search, vehicle model recognition, bank notes serial numbers recognition, chest tube detection in radiographs, and many other applications [2].

Inspired by a cat's visual cortex, it can extract fundamental visual features through the connectivity pattern between its neurons. To do this, each neuron in the intermediate layers receives inputs from the previous layer. These features are then processed and combined by the subsequent intermediate layers to extract higher-level features. To produce the strongest response to a spatially local input pattern, connection weights are shared between the nodes. This weight-sharing set is called the convolutional kernel or is more commonly known as a filter. These filters will then be used to convolve with input, producing feature maps. The pooling layer is another important asset of CNN to reduce the computational complexity. The network is usually followed by a fully-connected layer to generate the desired outcome [2]

A. Convolutional Layer

As the name suggests, the convolutional layer handles a crucial task for how CNN works. Variables of layer focus around the use of kernel. For example, if the input to the network is an image of size 32 × 32 × 3 (a VGR-colored image with a dimensionality of 32 × 32) and we set the field size as 6 × 6, we would have a total of 108 weights on each neuron assigned within the convolutional layer i.e. (6 × 6 × 3 = 108, where 3 is the magnitude of connectivity across the depth of the volume).

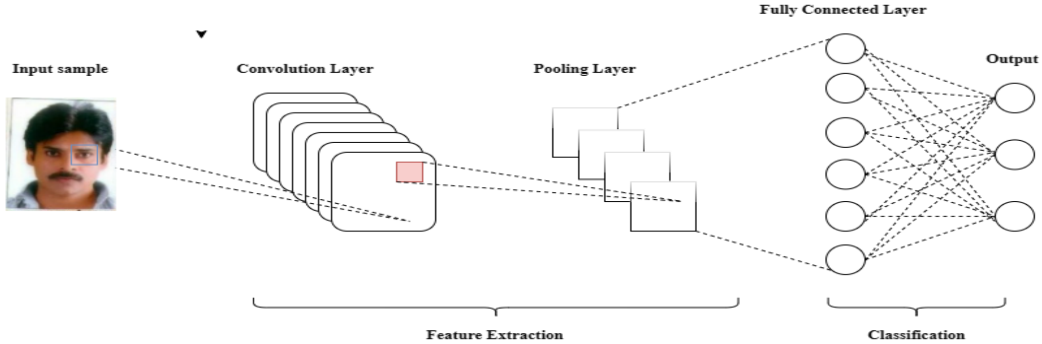

Fig. 1. CNN model

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 2, No. 3, February : 2024

B. Pooling Layer

The main of pooling is down-sampling to reduce the complexity for further layers. In the image processing domain, it can be considered as similar to reducing the resolution. Pooling does not affect the number of filters. Max-pooling is one of the most common types of pooling methods. It partitions the image to sub-region rectangles, and it only returns the maximum value of the inside of that sub-region [6].

• One of the most common sizes used in max-pooling is 2×2. As can be seen in Fig. 2, when pooling is performed in the top-left 2×2 blocks (pink area), it moves 2 and focuses on the top-right part. This means that stride 2 is used in pooling. To avoid down-sampling, stride 1 can be used, which is not common [6].

It should be considered that down-sampling does not preserve the position of the information. Therefore, it should be applied only when the presence of information is important (rather than spatial information). Moreover, pooling can be used with non-equal filters and strides to improve efficiency. For example, a 3x3 max-pooling with stride 2 keeps some overlaps between the areas.

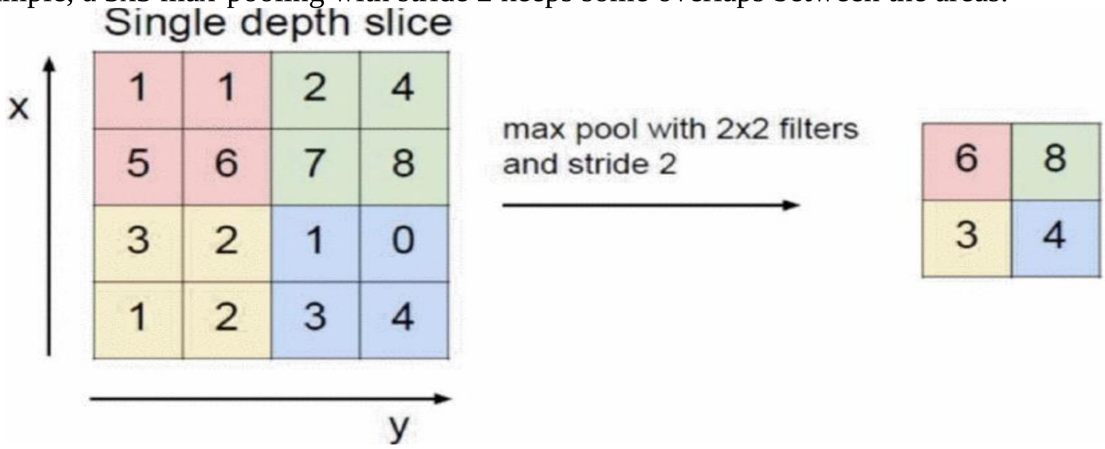

Fig. 2. Extraction of Pooling Layer

C. Fully-Connected Layer

The fully-connected layer is similar to the way that neurons are arranged in a traditional neural network. Therefore, each node in a fully-connected layer is directly connected to every node in both the previous and the next layer as shown in Fig. 1, From this figure we can note that each of the nodes in the last frames in the pooling layer is connected as a vector to the first layer from the fully-connected layer. These are the parameters used with the CNN within these layers and take a long time to train. The loophole of a fully connected layer is that it includes a lot of variables that need high computational in training the model. Therefore, we try to eliminate or merge the number of nodes and connections into a single domain by using PCA [6].

2.2 Principal Component Analysis (PCA)

PCA is one of the methods of Feature Extraction which belongs to the family of Dimensionality Reduction. As the quantity of attributes in a database expands, the quantity of data required to get a notable result also increases. With this, there emerges issues such as overfitting, and an increase in the computation of time, bringing down accuracy for a machine to learn models. All this is studied as a Curse of Dimensionality issue that enlarges with high-dimensional working data. As the dimensions boost up, the number of feasible combinations of attributes also tends to boost up, which tends to raise the difficulty for computers to obtain a representative sample of the data and it becomes tedious to execute tasks like clustering and classification.

To overcome the Curse of Dimension, Features Engineer tools are applied which carries Features Select (Selecting the correlated important features) and Feature Extraction (Merging more than one important feature into a single domain). Feature Extraction is a type of Dimension Reduction technique that aims to transform the number of input features into a single domain using Eigen decomposition and retaining as much of the original information as possible. PCA is a handy statistical tool that carries out Eigen decomposition transformation to bring about a set of new axes called Principal Components which can cover maximum variance.

Through PCA it's possible to change each original picture of the working out set into an equivalent Eigen face. An important feature of PCA is any particular you can reconstruct any unique image from a working out set by combining the Eigenfaces. The initial face picture might be reconstructed from Eigen's faces if most of the eigenfaces (features) are added in the proper proportion. Each Eigen's face represents only certain top features of the face area that might or may probably not be present. In the initial picture, if the function is within the unique picture to an elevated level, the reveal of the equivalent Eigen experience in the "sum" of the eigenfaces should be greater. [5]

If contrary, this feature isn't (or almost not) within the initial image, a corresponding Eigen face should contribute an inferior (or not at all) part to the quantity of eigenfaces. Therefore, to be able to reconstruct the first picture from the eigenfaces, you've got to generate some sort of measured amount of all eigenfaces. [5]

Steps to construct Eigenfaces:

1) *Standardize Dataset:* All the trained images to be stored in a file or dataset say 'A'

$$A = (x1, x2, x3, \ldots, xN) \qquad (1)$$

Equation 1 represents the dataset of stored images and 'N' represents the last stored image

2) *Finding Mean of the Dataset:* The below equation is to find the Mean value of the Dataset to be later used for the covariance matrix

$$\mu = \frac{1}{N}\sum_{i=1}^{N} xi \qquad (2)$$

3) *Covariance Matrix:* Matrix with a relation between its variance i.e., its features or attributes where trace/diagonal elements express the variance of the dataset. Range of Covariance lies between -1 to 1 where positive covariance indicates that the two variables have a positive correlation relationship, whereas a negative covariance suggests an opposite/negative relationship. If it stands at 0, two elements do not vary and don't have any relation, there represents a zero-covariance relationship.

$$Cov(x, y) = \frac{1}{N}\sum_{i=1}^{N}[(x - \bar{x}) * (y - \bar{y})]$$

4) *Eigen Value:* Equation 3, represents a formula for extracting eigenvalues:

$$|A + \lambda I| = 0 \qquad (3)$$

here,
A = Matrix of A i.e., equation (1)
λ = Scalar multiple
I = Identity Matrix

5) *Eigen Vector:* For equal values of 'lambda' we will get the respective eigen-vector i.e. If lambda has 2 values, then we will be getting 2 eigen-vectors for the sample. Equation 4, represents a formula for extracting eigenvectors:

$$(A - \lambda I)\, V = 0 \qquad (4)$$

here,
A = Matrix of A i.e., equation (1)
V = V is the eigenvector of matrix A

6) *Principal Component:* A facial structure contains a certain position of features and these attributes are known as Principal Component (PC). Attributes can be drawn out from the primary image with the assistance of the Principal Component.

III. Literature Review

Concerning image detection there are noteworthy works done by many authors using different algorithms of Deep Learning (CNN, ANN, GAN, LSTMs, etc.), but integrating CNN with PCA to achieve better results is put forward by some of the authors.

The paper presented by Edy Winaro, and others (2019) put light on the feature extraction with more accuracy by combining CNN-PCA methods for the attendance system. They proposed a methodology with pseudo-code for facial recognition stages and resulted in the model which produced accuracy in the range of 90%-98% by marking the attendance of the recognized face. Authors have objected to the comparison table of Accuracy with PCA and CNN-PCA which shows more accuracy when the model is connected with CNN-PCA recognition as compared to the model connected with PCA recognition.

In paper proposed by Hui Ying Khaw, and others (2017), introduced the recognition of different types of noise (impulse, Poisson, Gaussian, Speckle) which should be de-noised by the CNN technique to classify images as some of the images contain different level of noise. The authors also used the technique of PCA to reduce the computational time and to make it a cost-effective model. To check for every type of noise and to de-noise it, authors have used a large set of databases of around 1650 for testing and training purposes so that the model gets habitual with different real-life cases. Models have achieved an accuracy of 99% while recognizing the noise and 99.7% as training accuracy.

The work presented by Xu-Dire Ren, and others (2016) determined the weight-sharing features of CNN and Dimensionality Reduction of PCA; along with-it the paper proposed the use of PCA for extracting eigenvectors and initializing the role of convolutional kernels combined with the training process. The authors experimented on the Cifar-10 and Mnist datasets in which the model reduces the process of iterating and optimizing. With the help of plotting the graphs for easy understanding and visualization, the authors have well-defined the model by achieving an accuracy of 98.92% on Mnist and 65% on Cifar.

The paper proposed by Jaya Prakash Sahoo and Sarat Kumar Patra (2019) used three algorithms to propose a model on Hand Gesture i.e., CNN, Support vector Machine (SVM), and PCA by using a pre-trained AlexNet. In this the author used CNN for extracting deep features of the image; SVM as a classifier for classifying the gesture of hand and PCA for reducing the feature dimensions. A dataset of different hand gestures was used and various tests were performed by authors like the Leave One Subject Out test and Holdout test; visualized the accuracy by a plot histogram with an achievement of a range of more than 85%.

IV.Methodology

In this section, the actual process of a facial recognition model as an Attendance Machine for storing the face image database before the election and using it to recognize voters on the day of the election is shown. This will be done by engaging two algorithms to work i.e., PCA-CNN which will be integrated procedure-wise. Firstly, PCA will extract the feature and will reduce the dimension for the model, later CNN will convert the 2D images into 3D and will undergo a process of filtering & pooling of images which will bring out results as an output from the Fully Connected Layer.

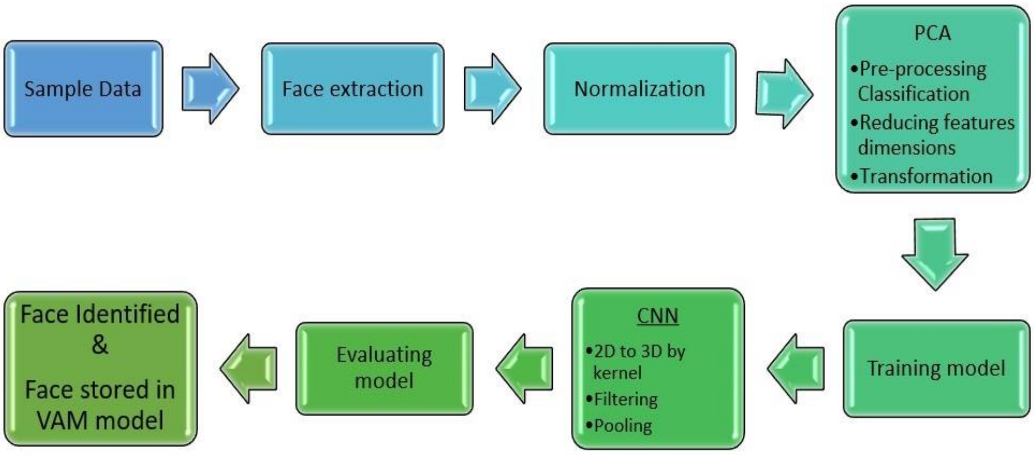

Fig. 3. Pseudocode for PCA-CNN working

A stepwise process is shown in Figure 3:

1) Collecting the sample data (Voter ID)
2) Extracting image from the Voters ID
3) Normalizing/ Standardizing the data according to pooling booths (Area wise storing of data).
4) PCA: As the dataset may be huge with raw data, so here PCA will pre-process the data into features and will reduce the dimensions of features for better and more accurate results.
5) Training the Model to fit the best with the real-life application.
6) Using the model as an input to CNN for further recognition of image by layers of CNN i.e., Convolutional Pooling - Fully connected.
7) Evaluating: Model evaluated at the Fully connected layer of CNN.
8) Now, the output image will be identified and will be stored in the VAM model.

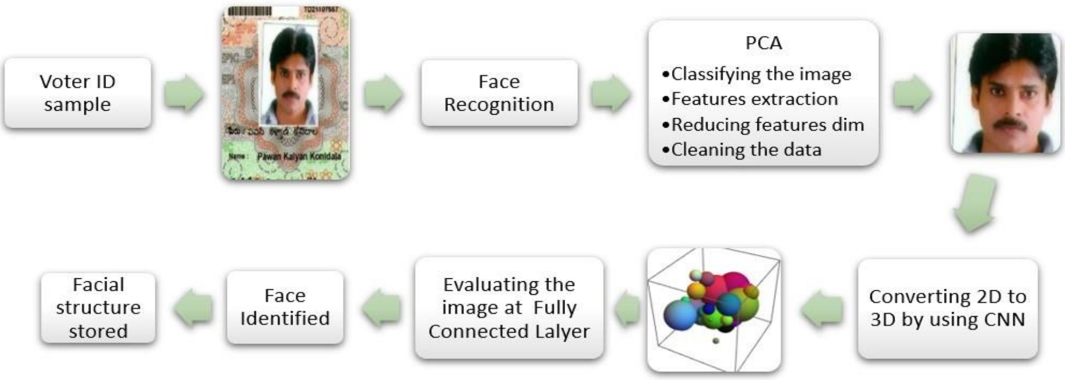

Fig. 4. Sample working model

Figure 4 represents a working model of PCA-CNN together in which a sample voter ID of famous actor Pawan Kalyan is forwarded as input. Firstly, PCA will perform the pre-processing steps by classifying the data, extracting its features, and reducing the features under an important domain. It will segregate the data of images into a domain of Gender (Male, Female, Transgender), later reducing the feature by transforming/merging similar features into a smaller number of domains (e.g., Eyes + nose + ears + lips = Face parts and forehead + jawline = Face Structure). Again, reducing the features (i.e., Face parts + Face structure = Facial attribute) and forwarding the cleaned data to the CNN as its input.

Now, as the image is in 2-dimension, the CNN kernel will convert it into a 3-dimension for better analysis, and after passing through the Convolutional and Pooling Layer, data will be stored in the Fully connected Layer where the evaluation of features will take place. Once the output is predicted it will be stored in the VAM model for recognition of similar faces during the election days.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 2, No. 3, February : 2024

V. Experimental Results

A minimum sample of voter IDs was collected to test our trained model to extract features and recognize images. Fig 5 represents the data of 5 voter IDs from booth A.

Fig. 5. Voter IDs of Booth-A

A model that is already trained with PCA-CNN features will now extract the images from the voter IDs and will go through the procedures of recognition and storing. The output of the model will be stored in a Voters Attendance Machine for the election days. Figure 6 represents the output of the model from the input (voter ID) given to it.

Fig. 6. Images extracted from voter ID

Now; On election day, Images of voters are clicked at the pooling booth A by the camera which is connected to our VAM to recognize the already stored images in VAM and to mark the voter as present on election day. Figure 7 represents a dataset of 30 images of the voters from Booth-A.

Fig. 7. Booth-A images

Voter Attendance Machine collected the images and searched for the images similar to its stored output images of voter IDs. Figure 8 presents the recognized images of VAM from the raw dataset of 30 images.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 2, No. 3, February : 2024

Fig. 8. Recognized images

VI.Conclusion

This paper has focused on Indian democracy that delivers: of the people, by the people, for the people; providing a direction to the model of the Face Recognition Attendance Machine used as Voters Attendance Machine (VAM) which shall be used along with Electronic Voting Machine (EVM) at every pooling booth during the election day, so that any malfunction can be prevented. We can surely get a result for face detection just by only using the CNN algorithm but it's a matter of democracy, choosing our representation who decides our role in governance for the next 5 years; so, for that, we need to be more accurate with our results. Hence, the PCA algorithm has also been used in this paper along with CNN to obtain a more precise result.

The model which is used as experimental results i.e. VAM, accurately detected the face click on the day of elections by recognizing the store image of the voter from the voter ID. This model can achieve 98% accuracy, 2% for technical and human error like 1) Machine unable to store complete features for image, 2) Image on voter ID has not been updated for many years.

Dataset	Accuracy (in %)		
	CNN	PCA	CNN-PCA
10	90.5	90.00	91.55
100	92.67	92.00	94.37
500	95.00	94.67	95.50
1000	97.00	96.37	97.37
1400	97.37	97.00	98.00

The confusion of a model can be shown as a confusion matrix in Figure 9.

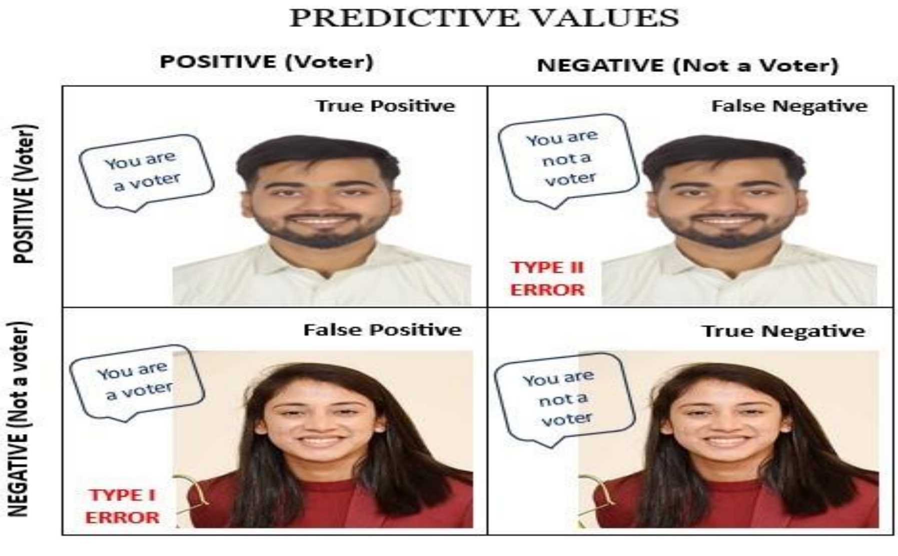

Fig. 9. Confusion Matrix for model

References

[1]Chen, Rongyu, et al." Image retrieval based on deep feature extraction and reduction with improved CNN and PCA." Journal of Information Hiding and Privacy Protection 2.2 (2020): 67.

[2]Khaw, Hui Ying, et al." Image noise types recognition using convolutional neural network with principal components analysis." IET Image Processing 11.12 (2017): 1238-1245.
[3]Winarno, Edy, et al." Attendance system based on face recognition system using cnn-pca method and real-time camera." 2019 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI). IEEE, 2019.
[4]Albawi, Saad, Tareq Abed Mohammed, and Saad Al-Zawi. " Understanding of a convolutional neural network." 2017 international conference on engineering and technology (ICET). Ieee, 2017.
[5]Kaur, Ramandeep, and Er Himanshi." Face recognition using principal component analysis." 2015 IEEE international advance computing conference (IACC). IEEE, 2015.
[6]D. Stutz and L. Beyer, "Understanding Convolutional Neural Networks," 2014.
[7]Sahoo, Jaya Prakash, Samit Ari, and Sarat Kumar Patra." Hand gesture recognition using PCA based deep CNN reduced features and SVM classifier." 2019 IEEE International Symposium on Smart Electronic Systems (iSES) (Formerly iNiS). IEEE, 2019.
[8]Mazloom, Masoud, and Saeed Ayat." Combinational method for face recognition: wavelet, PCA and ANN." 2008 Digital Image Computing: Techniques and Applications. IEEE, 2008
[9]Ren, Xu-Die, et al." Convolutional neural network based on principal component analysis initialization for image classification." 2016 IEEE first international conference on data science in cyberspace (DSC). IEEE, 2016.