Identification and Classification of Face Emotions Based on Age and Linguistic approach Using Convolution Neural Network and Analytic Hierarchy Process

Dasharath.K.Bhadangkar[1]

S.D.M.College of Engg. &Tech, Dharwar – 580 008, INDIA,
dashrathb@gmail.com

Jagadeesh D. Pujari[2]

S.D.M.College of Engg. &Tech, Dharwar – 580 008, INDIA,
jaggudp@yahoo.com

_____

Abstract— Face recognition has been an important topic of study for the last few decades. Human facial expressions show a lot of information instead of telling people what they are saying. It is important for people and machines to be able to recognise each other's faces. Computers still have a hard time figuring out how to recognise facial expressions with a high level of accuracy. In this paper we have used Convolution neural network (CNN) algorithm to recognise and classify face emotions based on age group. Then this paper focuses on Analytic hierarchy process (AHP) to recognise and classify face emotions based on linguistic approach.

Index Terms— Face emotion detection, CNN, Deep learning algorithms, AHP.

_____

I.   INTRODUCTION

One of the most important ways people show their emotions is through facial expressions. A powerful, natural, and immediate way for people to communicate their feelings is through facial expression recognition, which is one of the best ways for people to do this. The recognised emotions can be used in a wide range of applications in the field of human-computer interaction. In order to reach this goal, there are a lot of ways that use a lot of different types of sensors, like cameras and microphones, as well as body sensors. More specifically, visual-based approaches try to figure out how people feel based on how they look, for example, by noticing facial features, gestures, postures, motion, and so on [1], [2]. There are two types of audio-based approaches: those that use speech, and those that use either low- or middle-level features [3] or even natural language understanding methods [4]. Body sensors can measure things like temperature, heart rate, muscle activity, skin conductance response, or even brain activity. These sensors are called "biosensors."

Figure 1: The six cardinal emotions (happiness, sadness, anger, fear, disgust, and surprise) and neutral. The images in the first, second, and the third rows belong to the FER, JAFFE, and FERG datasets, respectively.

Detecting how people feel about their faces is one of the most common tasks in computer vision. First, it is easy to

write down a well-thought-out question and get the data they need because each person has a name tag. If you want to learn more about fine-grained classification, this is a good place to start. The intra-class variation caused by different poses, expressions, and ages can often outweigh the inter-class variation between two different people. Finally, the face recognition problem is very important because it can be used in a lot of different situations. Since face recognition is so important in the field of vision, it's become a lot of fun for people to work on. Most of the time, face recognition is used to do two kinds of things. In this case, given a gallery set and a query set, for a given image in the query set, they wanted to look for the most similar face to it in the gallery set and use the identity of that face as the identity of their query image. This is called face identification. Second, face verification is used to figure out if two images are of the same person.

A lot of different things can be done with facial features. For example, they can be used to recognise faces and figure out how old someone is. There is also the fact that a person's face changes over time as they get older. There are a lot of things that happen as we get older, like wrinkles. This is one of the most difficult things about face applications. Detecting facial expressions is one of the challenges that computer science will face in the future. This is especially true in areas like identification systems, verifying applications, and security. There can be small changes to the way the face looks, like wrinkles, or big changes, like changes to the eyes and nose. This is where age estimation comes in. It makes an automatic model to figure out the exact age or range of ages that someone is. There are some things that need to be explained in order to be able to tell them apart. It is the person's real age, which can be found out by how many years they've been alive. If you look at your face, you can figure out how old you are based on how you look [6]. Everybody ages at some point in their lives; it's a normal thing. Depending on how time affects each person, the process is different. It is not a process that can be controlled. As a process, ageing can be used to help find missing children through law enforcement and forensic investigation tools. This study looked at a lot of different ways to recognise faces because there are already a lot of problems in the field of face recognition. The survey in this paper gives more information about the rules, methods, architecture, advantages, and limitations of these techniques. You can read more about them here. Many technologies that deal with the face are used to help the world, like in forensics, security, and cosmetology. Human forensics use technology to make a person's face look like they did when they were alive [8]. When it comes to security, face recognition is one of the most important parts, especially when it comes to biometric identification and verification. Face techniques have a lot of problems with them. One of these problems is that there isn't enough data to figure out how old someone is in real life. This is more difficult than figuring out your age or spotting your face. [11]. The hard part is that people make a lot of mistakes when they try to figure out how old a person is. It's higher than what computer software can figure out. When people annotate databases of faces with their real ages, it can be hard to trust them.

Motivation

Research on facial age simulation and the Linguistic approach has become important because it is used in many real-world applications, such as public and commercial systems. This is why it is important to study these subjects. As a part of many public systems, facial age simulation is used a lot. If you want to make forensic montages of long-term unsolved cases, look for missing children or separated families, or use face recognition that doesn't get worse with age, this is a good tool for you! Face emotion recognition is based on linguistic descriptors that can be used to solve the problem of face recognition or face retrieval, especially the problem of identifying criminals. Automatic face recognition systems are used in many important applications around the world. Face recognition is a big part of figuring out if ID cards, like driving licences and passports, have been re-issued. This way, people can't get multiple ID cards under different names. Today, there are a lot of commercial uses for automatic face recognition. These include suggestions for "tags" on Facebook, how to organise personal photo collections, and how to unlock mobile phones, among other things.

II.   LITERATURE REVIEW

In [12], authors proposed a new network model deriving from A group of computer programmes called "Where-What networks" (WWNs) can approximate the information processing pathways of a human brain's visual cortex and recognise different types of faces with different locations and sizes in a complex background. To make it easier to recognise, a synapse maintenance mechanism and a mechanism to make new neurons are both added. Synapse maintenance is used to cut down on background noise, and a mechanism called neuron regenesis is used to control the amount of neuron resources available to improve network efficiency. Human face images of five types, 11 sizes,

and 225 locations have been used in experiments. Experiments show that the proposed WWN model can learn about three things (type, location, and size) at the same time. The experiment results also show that the improved WWN-7 model for face recognition is better than many other methods that have been used in the past.

In [13], the author looked through and analysed the literature on skin detectors (SD) to figure out where there are gaps in the research and how to fill them. An extensive search is done to find articles that deal with skin detection, skin segmentation, skin tone detection, and skin recognition. These articles are thoroughly reviewed and a taxonomy of these articles is put together. ScienceDirect, IEEE Xplore, and Web of Science databases are checked for articles about skin detectors. This is how it works: Over the course of seven years, 2803 papers are gathered from 2007 to 2018. The set had 173 articles in it. The majority of the papers (158/173) = 91 percent are in Development and Design, which is trying to figure out how to make a skin classifier that can tell the difference between skin and other things work better. Some (n = 5/173) = 3% of the papers are about Evaluation and Framework, and (n = 10/173) = 6% of the papers are about Comparative Study. This paper talks about the open problems, motivations, and suggestions of the works that are similar to it. Another way to show that this study is different from others is to do a statistical analysis of previous studies, like datasets, colour spaces, features, image types, and classification techniques. This shows that this study is different from other studies.

In [14], the first GAN-based method for ageing faces was proposed. In contrast to previous works that used GANs to change facial attributes, these new ones put a lot of emphasis on preserving the original person's identity in the old version of his or her face. To do this, they came up with a new way to "Identity-Preserving" optimise the latent vectors of GANs. The state-of-the-art face recognition and age estimation tools used to look at the resulting images show that the proposed method has a lot of potential.

In [15], researchers gave an up-to-date look at face recognition and age estimation research. There are a lot of reasons why this survey is being done. First, this paper tried to give an up-to-date look at the literature and methods that are used for facial recognition and age estimation. Second, the study aims to outline research challenges and make suggestions for future research in the field of face detection and age estimation techniques. This is what the study is about. Third, the paper wants to map the research landscape based on the literature that has been read and put together into a critical and coherent taxonomy. For its method, the study looked for every article that talked about 1) face recognition, 2) age estimation, and 3) facial features in major databases like Web of Science, ACM, IEEE, Science Direct, and Springer. The study found that the final set had 72 articles in it. Another thing that was found by the study was that 32 of the 71 articles were reviews and surveys of models used for face recognition. 39/72 of the articles also looked at the models used to figure out how old someone is. The study found that research on face techniques is different when it comes to different datasets. This review study, on the other hand, helps us understand the options and gaps for more studies to join this field of research.

In [16], a group of people looked at how local descriptors used in face recognition methods changed with the ages of the people they were comparing them to. They measured how well local descriptors used in face recognition worked when they were used in age discrimination. The FG-NET database is used to see how well the descriptors work. They show the results for different age groups and for different age differences in the people who were in the training and test images. The values of recognition accuracy are shown with a number of different similarity measures that are used to classify things. Furthermore, the performance of the descriptors when used with Gabor wavelet images is also tested.

In [17], the authors skipped the synthesis step and directly looked at how facial features change with age from a purely matching point of view. Our study is based on the fact that people's faces change in a logical way as they get older. They came up with ways to measure this coherency in feature drifts. Illustrations and tests show that this method is good at matching faces as they get older.

In [18], the authors came up with a way to deal with face matching when there is a lot of difference in age. In this framework, they first design a densely sampled local feature description scheme for each face. SIFT and multi-scale local binary patterns (MLBP) are used as the local descriptors. By sampling a lot of the two kinds of local descriptors in the whole facial image, enough discriminatory information, including the distribution of the edge direction in the face image (which is expected to be age invariant), can be gleaned for further study. Because both SIFT-based local features and MLBP-based local features cover a lot of space, they came up with an algorithm called multi-feature discriminant analysis (MFDA) to process these two local feature spaces in the same way. The MFDA is an improvement and extension of the LDA. It uses more features and two different random sampling methods in both feature and sample space. LDA-based classifiers are built by randomly sampling the training set and the feature

space. Then, they are combined to make a more robust decision by following the fusion rule. A commercial face-recognition engine isn't as good as ours when it comes to two public-domain face-aging data sets: MORPH and FG-NET. There is also a comparison made between the proposed discriminative model and a model that looks at how old people get. Discriminative and generative models are combined in a way that improves face matching accuracy even more when there is ageing in the picture.

III. PROPOSED WORK

The proposed research work is given below:

3.1 Face Emotion Detection based on Age

Deep learning algorithms have been used in facial expression recognition (FER) to solve problems like the ones above, as well as different learning tasks [19]. In deep learning algorithms, the process of finding and extracting unique features is done automatically. Deep learning algorithms have a layered way of representing data. Those at the top of the networks are called high-level feature extractors, and those at the bottom are called low-level feature extractors [20]. A CNN architecture is proposed that has a lot of connectivity across pooling, but doesn't have a fully connected layer to make sure features are shared. Under limited training data, it does well when it comes to representing facial expressions effectively. The DenseNet40 and DenseNet121, which were already trained, show a decrease in performance because they were overfitted. To solve the problem of not having enough data in the field of deep learning, a method is shown that uses new cropping and rotation strategies to make data more plentiful and useful for feature extraction with a simplified CNN. The method of cropping and rotating removes unnecessary parts of the face and keeps important facial information, and the results on CK+ and JAFFE are competitive.

The way we think about how we process information has changed completely thanks to these deep learning approaches. Deep learning is thought to be a better option for problems with vision and classification because it can learn on its own very quickly [21]. Other ways to classify things include pre-trained networks, which speed up the process of long training by using weights that have already been learned [22]. However, learning here requires a lot of tweaking of a lot of network parameters and a lot of labelled data for training. Because FER is important in a lot of different fields, we think FER with deep features can be used to figure out what an instructor's facial expressions mean in a classroom setting.

3.1.1 Data collection

In order to train any convolutional neural network, you need to get a lot of data (CNN). If you want to know how old something is, you need to label the data. This paper is about how to do that. Getting labelled data for some tasks, like real age estimation, is a lot more difficult than getting data for popular classification [23] or detection [24] tasks. This discrepancy is caused by the fact that humans have a lot of trouble figuring out how old people are. One can't rely on human annotators to label faces with their real ages, so this discrepancy is caused by this. Below, we show some statistics about the data we used to train our models.

3.1.2 Face recognition

There are over 4 million images of more than 40, 000 people that we use to train the base model for our facial recognition. The many different images of each person make our deep model more resistant to problems with face recognition.

3.1.3 Age estimation

In the last few years, there have been some efforts to collect data with age labels [25]. Rothe et aldataset .'s in [26] is the largest one, with 523, 051 images. It can be used for research purposes. However, the dataset is not properly annotated and has a lot of mistakes in it. In addition, the data is spread out very unevenly across different ages. Because of this, the authors used only half of the data for training in the first paper. In order to better solve this problem, we came up with a large dataset of about 600,000 images with age labels. Before, our dataset has a more even distribution of different ages. Our dataset has over 120,000 people with ages over 70 or younger than 20. We used a team of human annotators to help us clean up our data even more. We used a semi-supervised process to help us do this.

3.1.4 Data pre-processing:

Before we feed the images to our CNNs, we do some work on each one of them. These steps are called face detection, facial land mark detection, and alignment. If there are more than one face in an image, we choose the one that is closest to the centre of the picture. Given the face bounding boxes, we find 68 facial landmarks and use them to align the picture. As a last step, all of the aligned faces are cut and resized to the same size. With our work, we used face alignment instead of some previous work that didn't. This made a big difference when we came up with our accuracy numbers.

3.1.5 Deep training

In the first step, we train a deep neural network to recognise faces by looking at four million images of more than 41,000 people. Using just the CPU, our face recognition model extracts features in just 70ms. It also does very well on the LFW dataset, which is a set of faces. As a part of our facial attribute recognition engine, this model is the main thing. For each job, we built a network that was very well-optimized for speed and accuracy. In some recent studies [28], researchers try to design a network that can do all of its jobs at the same time. They have made some progress, but it's not much. Even though we had separate networks for each task, we were able to make faster and more portable models for each one of them. Running all the models together also takes less time than running the all-in-one model, and we get better results. We should also say that even though the network architecture is different for each task, all networks are trained first for the task of facial recognition with the four million image set that we used to train them.

3.1.6 Feature Extraction Using CNN Models.

This CNN arrangement is similar to AlexNet, pretrained CNN which contains 5 convolutional levels and 3 completely linked levels. Nevertheless, this network applies only 4 convolutional levels and solitary completely linked level. Furthermore, batch regularization, ReLU and pooling levels were positioned amid convolutional levels.

1. Convolutional Layer

This level was utilized to convolve input from descending filters horizontally and vertically alongside input. Dot result of weights and input are calculated for every filter and later, bias term was augmented. Here, there were 4 convolutional levels which were used with precise hyper parameters. Formula for calculating output dimension of convolutional level and quantity of weights per filter are revealed in (1) and (2) correspondingly.

$$\text{Output size} = \frac{W_I - F + 2P}{S} + 1$$

$$\text{.......(1)}$$

$$\text{Weight} = F \times F \times 3$$

$$\text{............(2)}$$

In which is input dimension, is filter dimension, is padding dimension and is quantity of strides.

2. Batch Normalization Layer

This level uses every input channel through mini-group plus standardizes this from subtracting mini-group average and splitting from mini-group normal aberration. Later, level moves input by learnable offset and gauges this from learnable scaling factor γ. This accelerates exercise of CNN and decreases sensitivity to network initialization.

$$\hat{x} = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + s}} \qquad y_i = \gamma \hat{x} + \beta$$

$$\text{.... (3)}$$

3. Rectifier Linear Units (ReLU)

This executes threshold process to rudiments of input which alters -ve amounts to zero plus upholding +ve amounts. This is also denoted as an activation purpose which determines situation of neuron whether it fervours or stay inactive.

$$f(x) = \max(0, x)$$

………… (4)

4. Pooling layer

This level works like down-sampling which splits input as rectangular pooling sections plus calculated values of every section. This also polishes the output and stop local alterations. There are 3 kinds of pooling approaches that are maximum, minimum and normal pooling. Here, supreme pooling level is utilized to sub-samples output of convolutional level. Supreme pooling level calculates maximum values of every section. Formula for calculating output dimension of pooling level are presented in (5).

$$\text{Output size} = \frac{W_I - P_I + 2P}{S} + 1$$

……… (5)

In which is input dimension, is pool dimension, is padding dimension and is quantity of strides.

5. Fully Connected Layer

This level performs like classifier level in termination procedure of CNN since this achieves cataloguing of preceding mined characteristic by functioning skilled weighted links. This augments input by weight matrix plus augments a bias vector. This level uses output of whichever preceding level and output an N dimensional vector which regulates quantity of selectable divisions for cataloguing. Output dimension and preconception of level was fit to 23 since 23 likely subjects were present to be categorized.

A deep feature representation of facial expressions is made from a 2D-CNN. This is how it works. Deep learning models have a complicated structure that learns about different things on each layer of the structure (hierarchical representations of layered features). This layered structure makes it easy to get high-level, medium-level, and low-level features from an instructor's face. There are two types of networks: a sequential network and one called a directed acyclic graph (DAG) [29]. In AlexNet, for example, there are 8 layers and 227 x 227 2-dimensional input. A serial network has layers that are arranged in a certain order, like that. On the other hand, a DAG network has layers in the form of a directed acyclic graph, with multiple layers processing at the same time to get the best results. GoogleNet [30], DenseNet201 [31], ResNet50, ResNet18, ResNet101 [32], and Inceptionv3 [33] are some examples of DAGs. They have a depth of 22, 201, 50, 18, 101, and 44 layers, respectively. The deeper layers have high-level representations of features, but that doesn't mean they're always the most accurate. Instead of just getting features from the last layer, features are taken from convolution, pooling, and regularisation layers as well. We have tested the performance of different layers of deep networks. Conv4 block9 1 bn layer is used to get features from DenseNet201. This is how the features for AlexNet, GoogleNet, Inceptionv3, and ResNet50 are found: drop 7, pool5 drop 7x7 s1, activation 94 relu, and avg pool For ResNet101 and ResNet18, we chose pool 5. The DenseNet architecture was made to make sure that as much information as possible moves between layers in the network. All layers are directly connected to each other, and each layer gets feature maps made by all the layers that came before it. These feature maps then move on to the next layer. Like ResNet, here features are combined by concatenation rather than summation before going into a layer, so they don't get mixed up. Thus, the lth layer has l inputs, which are the feature maps of all the convolutional blocks that came before it. Its own feature maps are passed on to all L − 1 subsequent layers. So in an L-layer network, there are L (L + 1)/2 direct connections unlike traditional architectures which have L number of connections. .e l th layer receives the feature maps xl of all preceding layers, x0, . . . , xl− 1, which is in the following form:

$$x^l = H^l\left(\left[x^0, x^1, \ldots, x^{l-1}\right]\right),$$

………………….(6)

where x0, x1, . . . , xl− 1 are the concatenation of the feature maps in layers 0, . . . , l − 1th layer. Hl (.) is a composite function comprised of three operations which are batch normalization (BN) [34], a rectified linear unit (ReLU) [35], and a 3 ∗ 3 convolution (Conv). In traditional deep CNNs, layers are followed by a pooling layer that reduces feature maps size to half. Consequently, the concatenation operation used in equation (1) would be erroneous due to the change in feature maps. However, down sampling layers are an essential part of convolutional networks. To

facilitate consistent down sampling, DenseNets are designed so as to divide the network into multiple densely connected dense blocks and transition layers are introduced [36].

3.2 Face Emotion Detection based on Linguistic Approach

Another goal of this study is to come up with a new way to solve the problem of recognising or finding a person's face, either with or without using any numerical measures that are linked to that person's face images. A professional expert in criminology or the person who witnessed the crime could both help with the process of identifying the person who did the crime. We want to learn more about how analytic hierarchy process, or AHP, is used in multi-criteria decision-making theory. Face recognition can be broken down into two levels of hierarchy. There is a higher level in the hierarchy where we figure out how much weight each face feature has in the process of classifying it. At the bottom of the hierarchy, we use the AHP method again to turn the linguistic descriptions of concrete facial features into numeric variables that show how important each attribute is for that feature. Besides that, our goal is to look into how the AHP can help people recognise faces better when it is used with other methods that look at how shapes and lines connect in the face. [37].

3.2.1 The Role of the Analytic Hierarchy Process

This section gives a quick overview of the AHP method as it was first proposed in [38]. Using this method, one can figure out the order and importance of any set of features or attributes that are being looked at. The algorithm is shown in this way. First, the structure of the problem is made clear, so we can figure out what it is. The goal is at the top, and then the criteria are written down. At the bottom of the hierarchy, the set of alternatives is found. In our case, we have two goals. As a first step, we want to figure out how much each face feature weighs, which can be used in the process of classifying people (whenever the weights can be applied to prioritise the classifiers). Second, we want to know how many of the facial features belong to the words that describe the set of traits that people have in common. There are a lot of things we want to know about people, like whether their noses are short or long, and how much they belong to each of the three groups.

It is at this point in the algorithm that an expert (or a group of experts) decides how important it is to measure the differences between the elements (alternatives). These evaluations are based on how the elements compare to each other. For n options, the experts' responses are put together in the nn matrix A, where n is the number of options to be looked at(in our case, facial features).

The experts generate the pairwise comparisons' results using the following scale: equal importance (1), weak importance (2), moderate importance (3), moderate plus (4), essential/strong importance (5), strong plus (6), very strong/demonstrated importance (7), very, very strong (8), extreme importance (9). A is called a reciprocal matrix, meaning that it satisfies the following requirements: For each element aij we have

$$a_{ij}=1/a_{ji}, i,j=1,\ldots,n, \text{ and} \quad \ldots\ldots\ldots (7)$$
$$a_{ii}=1. \quad \ldots\ldots\ldots\ldots\ldots\ldots (8)$$

Let us introduce the expression

$$v=(\lambda_{max}-n)/(n-1), \quad \ldots\ldots\ldots\ldots\ldots(9)$$

where $\lambda_{max} \geq n$ is a maximal eigenvalue of the reciprocal matrix A and the value $\mu=v/r$, where r=0,0,0.52,0.89,1.11,1.25,1.35,1.40,1.45,1.49 for n=1,…,10, respectively. These values concern the mean consistency indices of 500 randomly generated reciprocal matrices [36]. For the matrices of higher dimensionality, the methods generating the pertinent values are discussed in [37]. v, μ, and r are called inconsistency index, consistency ratio, and random inconsistency index, respectively. From a practical perspective, it is considered that the consistency ratio should not exceed the value 0.1 to assure the satisfactory level of consistency of results. However, it can be difficult to obtain such level of consistency, especially when the non-numerical, intangible features are compared. The final ranking of the priorities is constructed using the values of the elements of the eigenvector of the matrix A associated with the maximal eigenvalue λ_{max}.

IV.   RESULTS

The experimental results of Face Emotion Detection based on Age Group as well as Linguistic approach is given below:

Table 1: Face emotion Detection based on age using CNN algorithm

Sl. No.	Algorithm	Accuracy
1	CNN	90%

Table 2: Face emotion Detection based on linguistic approach using AHP method

Sl. No.	Algorithm	Accuracy
1	AHP	97%

V. CONCLUSION AND DISCUSSION

Human facial emotion recognition (FER) has been getting a lot of attention from researchers because of its promising applications. Make sure that different facial expressions match up with the emotions they are meant to show Two big parts make up the classic FER: extracting features and figuring out how people feel. Today, Deep Neural Networks, especially the Convolutional Neural Network (CNN), are very popular for use in FER because of their ability to find features in images. Since CNN can recognise and classify emotions based on age, we used it to do this work. Then, we showed a new way to use the analytic hierarchy process to make linguistic descriptions of facial emotions based on the experts' opinions of the faces. This method worked very well when it was combined with a standard classifier.

VI. REFRENCES

[1]. T. Baltrusaitis, D. McDuff, N. Banda, M. Mahmoud, R. El Kaliouby, P. Robinson and R. Picard, Real-time inference of mental states from facial expressions and upper body gestures. In Automatic Face & Gesture Recognition and Workshops (FG 2011), IEEE International Conference on (pp. 909-914). IEEE, 2011.

[2]. S. Piana, A. Stagliano, F. Odone, A. Verri, and A. Camurri, Real time automatic emotion recognition from body gestures. arXiv preprint arXiv:1402.5047, 2014.

[3]. M. Papakostas, E. Spyrou, T. Giannakopoulos, G. Siantikos, D. Sgouropoulos, Ph. Mylonas and F. Makedon, Deep Visual Attributes vs. Hand-Crafted Audio Features on Multidomain Speech Emotion Recognition. Computation, 5(2), 26, 2017.

[4]. S. G. Koolagudi and K. S. Rao. Emotion recognition from speech: a review. International journal of speech technology, 15(2), 99-117, 2012.

[5]. Kumari, J., R. Rajesh, and K. Pooja, Facial expression recognition: A survey. Procedia Computer Science, 2015. 58: p. 486-491.

[6]. Fu, Y., G. Guo, and T.S. Huang, Age synthesis and estimation via faces: A survey. IEEE transactions on pattern analysis and machine intelligence, 2010. 32(11): p. 1955-1976.

[7]. Sahni, S. and S. Saxena, SURVEy: Techniques for Aging Problems in Face Recognition. MIT International Journal of Computer Science and Information Technology, 2014. 4(2): p. 82-88.

[8]. Short, L.J., et al., Validation of a computer modelled forensic facial reconstruction technique using CT data from live subjects: a pilot study. Forensic science international, 2014. 237: p. 147. e1-147. e8.

[9]. Pandya, J.M., D. Rathod, and J.J. Jadav, A survey of face recognition approach. International Journal of Engineering Research and Applications (IJERA), 2013. 3(1): p. 632-635.

[10]. Dehghan, A., et al., View Independent Vehicle Make, Model and Color Recognition Using Convolutional Neural Network. arXiv preprint arXiv:1702.01721, 2017.

[11]. Everingham, M., et al., The Pascal Visual Object Classes Challenge: A Retrospective. International Journal of Computer Vision, 2015. 111(1): p. 98-136.

[12]. Dongshu Wang, Heshan Wang, Jiwen Sun, Jianbin Xin, and Yong Luo Face Recognition in Complex Unconstrained Environment with an Enhanced WWN Algorithm, Intell. Syst. 2021; 30:18–39.

[13]. M. A. CHYAD, H. A. ALSATTAR, B. B. ZAIDAN , A. A. ZAIDAN , GHAILAN A. AL SHAFEEY , 2019, The Landscape of Research on Skin Detectors: Coherent Taxonomy, Open Challenges, Motivations, Recommendations and Statistical Analysis, Future Directions, Faculty of Arts, Computing and Creative Industry, Department of Computing, Universiti Pendidikan, Tanjung Malim 35900, Malaysia.

[14]. Antipov, Grigory; Baccouche, Moez; Dugelay, Jean-Luc (2017). [IEEE 2017 IEEE International Conference on Image Processing (ICIP) - Beijing, China (2017.9.17-2017.9.20)] 2017 IEEE International Conference on Image Processing (ICIP) - Face aging with conditional generative adversarial networks. , (), 2089–2093. doi:10.1109/ICIP.2017.8296650.

[15]. Atallah, Rasha; Kamsin, Amirrudin; Ismail, Maizatul Akmar; Abdelrahman, Sherin; Zerdoumi, Saber (2018). Face Recognition and Age Estimation Implications of Changes in Facial Features: A Critical Review Study. IEEE Access, (), 1–1. doi:10.1109/ACCESS.2018.2836924.

[16]. Bereta, Michal & Karczmarek, Paweł & Pedrycz, Witold & Reformat, Marek. (2013). Local descriptors in application to the aging problem in face recognition. Pattern Recognition. 46. 10.1016/j.patcog.2013.03.010.

[17]. Biswas, Soma & Aggarwal, Gaurav & Ramanathan, Narayanan & Chellappa, Rama. (2008). A Non-generative Approach for Face Recognition Across Aging. BTAS 2008 - IEEE 2nd International Conference on Biometrics: Theory, Applications and Systems. 1 - 6. 10.1109/BTAS.2008.4699331.

[18]. Z. Li, U. Park and A. K. Jain, "A Discriminative Model for Age Invariant Face Recognition," in IEEE Transactions on Information Forensics and Security, vol. 6, no. 3, pp. 1028-1037, Sept. 2011, doi: 10.1109/TIFS.2011.2156787.

[19]. N. Zeng, Z. Wang, H. Zhang, K.-E. Kim, Y. Li, and X. Liu, "An improved particle filter with a novel hybrid proposal distribution for quantitative analysis of gold immunochromatographic strips," IEEE Transactions on Nanotechnology, vol. 18, pp. 819–829, 2019.

[20]. F. Zhi-Peng, Z. Yan-Ning, and H. Hai-Yan, "Survey of deep learning in face recognition," in Proceedings of the 2014 International Conference on Orange Technologies, September 2014.

[21]. H. Yu, Z. Luo, and Y. Tang, "Transfer learning for face identification with deep face model," in Proceedings of the 2016 7th International Conference on Cloud Computing and Big Data (CCBD), November 2016.

[22]. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Communications of the ACM, vol. 60, no. 6, pp. 84–90, 2017.

[23]. Dehghan, A., Masood, S.Z., Shu, G., Ortiz., E.G.: View independent vehicle make, model and color recognition using convolutional neural network. In: arXiv:1702.01721. (2017)

[24]. Everingham, M., E.S.M.A.V.G.L.W.C.K.I.W.J., Zisserman, A.: The pascal visual object classes challenge: A retrospective . In: International Journal of Computer Vision. (2015)

[25]. Eidinger, E., Enbar, R., Hassner., T.: Age and gender estimation of unfiltered faces. In: IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY. (2013)

[26]. Rothe, R., Radu Timofte, L.V.G.: Dex: Deep expectation of apparent age from a single image. In: International Conference on Computer Vision (ICCV),. (2015)

[27]. Levi, G., Hassner., T.: Emotion recognition in the wild via convolutional neural networks and mapped binary patterns. In: Proceedings of the 2015 ACM on International Conference on Multimodal Interaction. ACM. (2015)

[28]. Ranjan, R., Sankaranarayanan, S., Castillo, C.D., Chellappa, R.: An all-in-one convolutional neural network for face analysis. In: arXiv:1611.00851. (2016).

[29]. J. Li, X. Li, and D. He, "A directed acyclic graph network combined with CNN and LSTM for remaining useful life prediction," IEEE Access, vol. 7, pp. 75464–75475, 2019.

[30]. C. Szegedy, W. Wei Liu, Y. Yangqing Jia et al., "Going deeper with convolutions," in Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, June 2015.

[31]. F. Iandola, M. Moskewicz, S. Karayev, R. Girshick, T. Darrell, and K. D. Keutzer, "Implementing efficient convnet descriptor pyramids," 2014, http://arxiv.org/abs/1404.1869.

[32]. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2016.

[33]. C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, June 2016.

[34]. S. Ioffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," 2015, http://arxiv.org/abs/1502.0316.

[35]. X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," in Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, pp. 315–323, Fort Lauderdale, FL, USA, April 2011.

[36]. G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4700–4708, Las Vegas, NV, USA, July 2017.

[37]. Karczmarek, P., Kiersztyn, A., Pedrycz, W. et al. Linguistic Descriptors in Face Recognition. Int. J. Fuzzy Syst. 20, 2668–2676 (2018). https://doi.org/10.1007/s40815-018-0517-0.

[38]. Saaty, T.L., Vargas, L.G.: Models, Methods, Concepts & Applications of the Analytic Hierarchy Process. Springer, New York (2012)

[39]. G. Sathish 砼 S. V. Saravanan's. Narmadha砼S. U. Maheswari 2012. "Personal Authentication System using Hand Vein Biometric." International Journal of ComputerTechnology and Applications砼vol. 3(1) 砼 pp. 383-391砼 Jan-Feb 2012.