A New Deep Learning Model for Face Recognition and Registration in Distance Learning

https://doi.org/10.3991/ijet.v17i12.30377

Ahmed B Salem Salamh, Halil Ibrahim Akyüz(✉)
Kastamonu University, Kastamonu, Turkey
hakyuz@kastamonu.edu.tr

Abstract—The demand for secure, accurate and reliable identification of individuals using facial recognition has attracted considerable interest in education, security and many other sectors, not limited because it is robust, secure and authentic. Recently, the demand for distance learning has increased dramatically. This increase is due to various barriers to learning that arise from enforced conditions such as seclusion and social distancing. Facial feature extraction in distance education is valuable in supporting face authenticity as it prevents the position of participants from changing, especially during the examination phase. In the field of face recognition, there is a mismatch between research and practical application. In this paper, we present a novel but highly efficient Deep Learning model for improving face recognition and registration in distance education. The technique is based on a combination of sequential and residual identity blocking. This makes it possible to evaluate the effectiveness of using deeper blocks than other models. The new model has proven to be able to extract features from faces in a high and accurate manner in compared with other state-of-the-art methods. In registration processing, there are several challenges related to training data limitation, face recognition and verification. We present a new architecture for face recognition and registration. Experiments have shown that our registration model is capable of recognizing almost all faces and registering the corresponding labels.

Keywords—face recognition, deep learning, face identification, distance learning, feature extraction

1 Introduction

The distance learning has become a global issue in the culture of learning and information and becomes immediately important in 2020 in the scenario of COVID -19 pandemic[1], which has greatly changed the way education is delivered in many countries[2][3].The traditional form of educational delivery has evolved into Internet-based, computer-oriented online education[4]. The most important factor here is the degree of social distance and security [5][6]. Online education is one of the terms used to refer to the learning process where the teachers and learners are not physically present in the same place. There are many different terms that have the same meaning such as online learning and distance learning. In an online class, the teacher and learner interact with each other. They appear to be close, but they are not. The teaching and learning process is done through video calls that supplement the online course. The communication between the teacher and the learner can be face to face through video chat or video conferencing which is the ideal form of learning.

One of the problems faced by online lecturers is the authenticity of learners at different stages of online teaching, such as teaching and assessment. Authenticity of learners is ensured by recognition technologies. In many distance education systems, It is unattainable to report if the participant is a real person who needs to be instructed or if the student is actively pursuing the course[7]. This paper focuses on face recognition in online class registration. It is believed that facial recognition is a valuable biometric technology that can help in distance learning and directly support the whole educational process. The contribution of this part is a new model for face recognition and registration. Train and test the whole structure of the proposed model and implement further techniques to control the overall recognition and registration of the system.

1.1 Face recognition

Although great progress has been made in the field of recognition, there are still difficult problems in face recognition and recognition tasks. The limitations are in head pose, scaling, illumination variations, environment segmentation and occlusion. Face recognition is one of the identity verification techniques used in many computer vision applications[8]. The recognition is done by using different feature extraction methods after face detection and alignment process. The most popular technique is the traditional and deep learning technique. The traditional methods are known for extracting handcrafted features such as Multi Descriptor model[9], Local Binary Pattern [10][11], Histogram of Oriented Gradients [12], Gabor filter [13]and Scale Invariant Feature Transform [14] which are used for extracting distinctive features. These techniques were robust and efficient for face recognition. Moreover, these techniques have been modified and developed several times[15] for example, LBP has been modified to be more efficient in face recognition[16].

Convolution neural networks (CNN) is among the common techniques for solving image processing and computer vision problems throughout the last decade[17]. At many levels of abstraction, the strength of a Convolution Neural Ne work (CNN) lies in obtaining a set of distinguishing feature maps. Convolutional neural networks have positively influenced the top of most other implementations[18]. In general, there are many CNN architectures, deep models such as GoogLeNet[19],ResNet[20] and[21]. Moreover, a visual geometry group at Oxford University proposed a convolutional neural network architecture called Deep Face Recognition[22]. Face detection, feature extraction and identification are the 3 main components of face recognition. Face detection extracts single or multiple faces from images using common techniques such as Viola Jones face detector [23] and Multi-task cascaded CNN[24]. The most important phase is feature extraction, which determines the features of the face, while identification is performed to determine who that person belongs to. First, face detection: the goal of this phase is to select a particular face in the image. Two well-known algorithms are used in face recognition: viola-jones face detector and Multi-task cascaded CNN. Our research has shown that the Viola-Jones face detector works well in constrained environments with different poses, illuminations, and occlusions, but that problems are encountered in unconstrained situations with different poses, illuminations, and occlusions. Multi-task Cascade CNN uses a three-stage cascaded structure of specially designed deep convolution networks to detect the face and landmark region in a coarse to fine manner. The cascaded multi-Task CNN was used in this paper because it can solve the low detection rate problem. Second, feature extraction: feature extraction is the primary goal of any pattern recognition algorithm. For face recognition, a Deep Learning Convolution Neural Network is used for feature extraction. This is a simple strategy that does not require a complicated feature extraction method. Third, identification: identification is the recognition of a person's facial image to assign it to a specific class. After feature extraction, the technique compares a face in a photo or video image with an existing face dataset. Fourth, the softmax contribution is composed of the output of the fully linked layer that was in front of it and the extracted features of the real neural network. This result represents the probability function for all class-labels.

2 Proposed deep learning model for face recognition

The proposed architectures have the input image with the size of 224X224x3 pixels and the first layer with the filter size of 7X7. Table 1 and Figure 2, show that the İndepthnet19 architecture with a stride of 2 allows the CNN to learn more relevant features with a total of 64 filters. The following layer has a dimensionality of data of 112 and a filter size of 5X5 with 64 filters. The proposed normal block, as the main component of the network, has a learning architecture with the average pooling of the early layer to complement the final output of the block, as shown in Figure 1.a.

$$y = \max(F(x, \{Wi\})) + \max(x) \tag{1}$$

$$y = \max(F(x, \{Wi\})) + \mathrm{avg}(x) \tag{2}$$

The core network has a learning architecture where the average pooling precedes the early layer to complete the final output of the block, with both average and maximum pooling for the extended block as shown in equation (1) and (2). Finally, each convolutional layer for all models is followed by batch normalization and rectifying linear unit (ReLU) to speed up the processing and stability [25][26] of the model. Finally, global average pooling, a fully connected layer, and a softmax layer were added. To train the model, the cross entropy loss function is applied. For weight updates and optimizing the Loss function, Stochastic Gradient Descent with Momentum is applied.

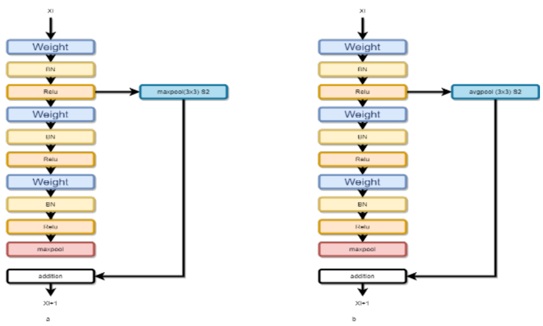


Fig. 1. Learning structure for Indepthnet19, a) proposed normal block as main component of network, b) extended used for deeper network based on block

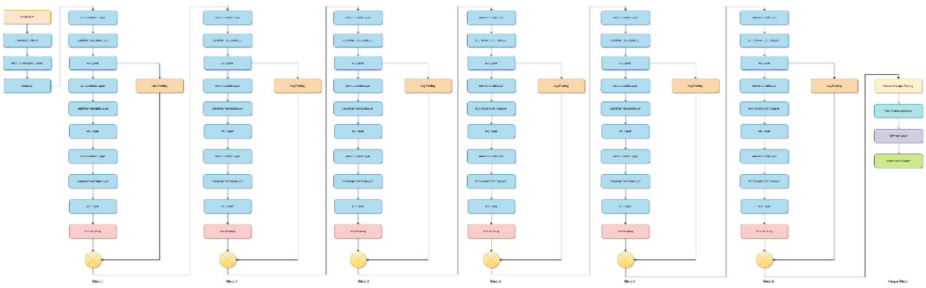


Fig. 2. The configuration block of Indepthnet19 with normal block for the network. We use the down-sample with maxpooling and avgpooling with the same block identity except for the first conv2d layer, which is used as the base layer for the block

Table 1. The architecture of Indepthnet19

Layer	Number of Convolution	Kernal Size	Channels	Stride	Input Size
Input Image					224x224
Conv1	1	7x7	64	2	224x224
Conv1_1	1	5x5	64	1	112x112
Block 1	2	3x3	64	1	112x112
Block 2	3	3x3	128	1	55x55
Block 3	3	3x3	256	1	27x27
Block 4	3	3x3	512	1	13x13
Block 5	3	3x3	1024	1	6x6
Block 6	3	3x3	2048	1	6x6

3 Face recognition datasets

Because we apply the proposed model for face recognition and registration in distance education, the selection of the dataset is crucial for our study. The dataset collected in the class was selected for model validation. The University of Essex has, Computer Vision Science Research Projects introduced four folders of data (Faces94, Faces95, Faces96, Grimace) [27]. The total face database contains 395 faces, consisting of males and females for each subject, and includes 20 images. It contains photos of people from different ethnic backgrounds, mainly first-year university students. Most of the subjects are between 18 and 20 years old, and there are also some older subjects. Some other individuals wear glasses and beards. Face94 contains 152 people, including 20 females, 113 males, and 20 males, for a total of 3060 people. The resolution of the images is 180 x 200 pixels. The variation of an individual is presented to summarise some features. The Face95 database contains 72 male and female subjects, and for each subject there are 20 examples with a total of 1440 individuals. The features include head rotation and some expression variations. Face96 included 151 subjects with 20 3013 samples collected for each subject. The resolution of the images is 196 x 196 pixels. The features of this database are considered large-scale and some expressions and others are included in Table 4, which contains some features of face images. The grimace database contains male and female subjects with 18 subjects and 20 subjects with an image resolution of 180 x 200 pixels. The database contains the main expression variations. Table 2 shows the overall characteristics of the four datasets used in this paper. The experiment was evaluated using 5-fold cross-validation. We augmented the four datasets based on minor left and right rotation without further processing.

Table 2. Summary of facial recognition datasets

Dataset	Number of subject	Number of image	Background	Head turn	Head Scale	Position of face	expression variations	lighting variation
Faces94	152	3060	plain green	very minor variation	none	minor changes	consider expression changes	none
Faces95	72	1440	red curtain	minor variation	large head scale variation	some translation	some expression variation	significant lighting changes
Faces96	72	3013	complex	minor variation	large head scale variation	some translation	some expression variation	significant lighting changes
Grimace	18	360	plain	considerable variation	small head scale variation	some translation	major expression variation	very little

4 Performance evaluation and comparison

To evaluate and compare the performance of the proposed CNN, we used the following indices: the number of positive and negative detections (P, N) and the number of correct and false detections (T, F), total accuracy in equation (3) and the mean accuracy in equation (4).

1. Accuracy (Acc):

$$\text{Acc} = \frac{(TP+TN)}{(TP+TN+FP+FN)} \tag{3}$$

2. Mean accuracy μ is given by:

$$\mu = \frac{\sum_{i=1}^{5} p_i}{5} \tag{4}$$

Where p_i denote the proportion of valid classification.

5 Results and discussion

The public dataset was chosen because it facilitates comparison with other approaches and is related to face recognition and classroom registration problems. It is also commonly used by other researchers. The results were evaluated with a 5-fold cross-validation. Diagnosis of the proposed models is a crucial step for face recognition. Figure 3, shows the training accuracy, validation accuracy, training loss and validation loss for the datasets. From the analysis of the curves, it can be seen that all the datasets have good fit. The face recognition rate of this technique was compared with other state-of-the-art deep learning models. Figure 4, shows that our model outperforms the other methods with an accuracy of 100% for the Face94 and Grimace datasets and achieves 99.86% for Face95. In the case of Face96, the model achieves a good accuracy of 99.54 compared to resnet50 and densenet201.

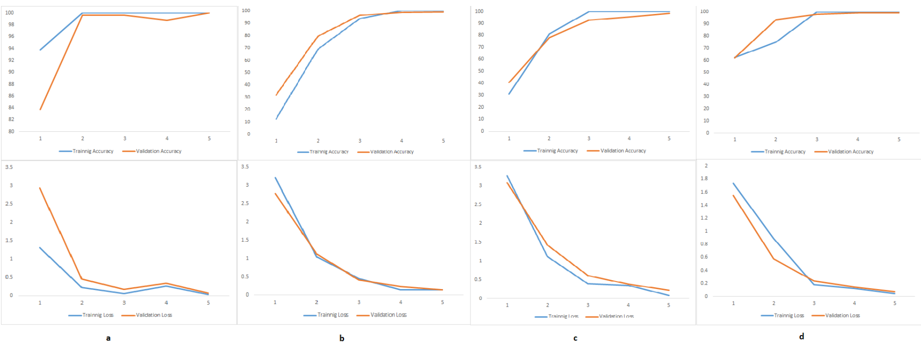
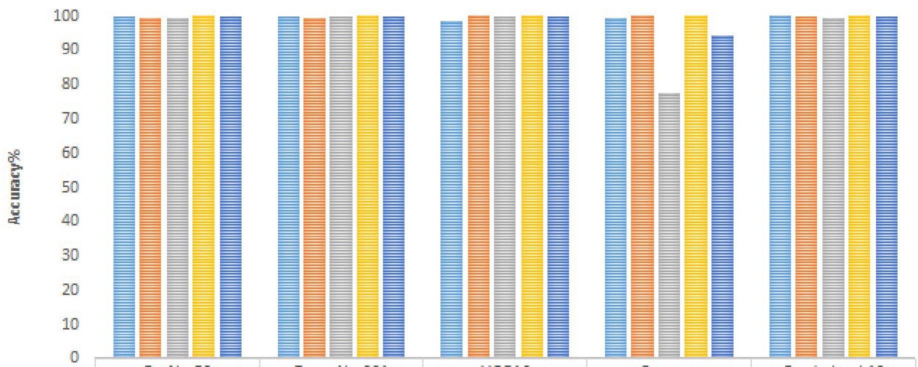


Fig. 3. Shows the training and validation accuracy, training loss and validation loss curves for İndepthnet19, (a) Faces94, (b) Faces95, (c) Faces96, (d) Grimace



	ResNet50	DenseNet201	VGG16	Facenet	Our Indepth19
Face94	99.93	99.97	98.64	99.37	100
Face95	99.44	99.2	100	100	99.86
Face96	99.27	99.64	99.67	77.67	99.54
grimace	100	100	100	100	100
Average	99.66	99.7025	99.5775	94.26	99.85

Face94 Face95 Face96 grimace Average

Fig. 4. Performance comparison of Indepthnet19

To understand the results, four data sets with different challenges were created, as shown in Table 2. For the challenges related to the datasets, faces94 and grimaces with minor head scale, face position, head turn and major expression differences, the proposed model obtained significant results. For face95 with significant illumination changes and large head scale differences, the model shows perfect performance compared to other methods. The dataset face96 has a different complex background with large head scale differences, which slightly affects the performance of the proposed model and other models. To make better use of Indepthnet19, we consider Face94, Grimace and Face95 challenges to perform well. For Facenet, the accuracy in the faces96 dataset is not optimal, which is likely because there are numerous face labels in this dataset that contain different faces. However, the accuracy of the FaceNet results from [26] was only 77.67%. This seems quite feasible since each face label has multiple differences and the FaceNet training technique is using triplet loss.

All the state-of art methods achieved significant performance. The average results indicate that Indepthnet19 proved to be the most effective method and outperformed the other state of art methods, making it more interesting for the challenges related to the four datasets.

6 Distance learning and registration

6.1 Distance learning datasets

All faces are acquired from live detection. Every face image in online classroom was cropped automatically using face detection method see section II. The detection and cropped was based on a single image for each face. The images were prepared using data transformation and augmentation. This technique was used to develop the datasets in online face recognition. Further, steps were introduced on how to train the classifier step by step as it is appearing in Figure 5. In addition, the main concern of data transformation in this paper to process the classes that has a few face images as well as the classes that have more than face image. The training process is valuable and a challenging step of all systems. This method was applied using proposed CNN and support vector machine for face feature extraction. Moreover, the classifier with the support vector machine was retained. This valuable and crucial step revealed that the SVM using best hyper plane that separates classes tended to less error than softmax probability function to identify the classes. The test stage was based on live face recognition on the faces that appeared on screen.

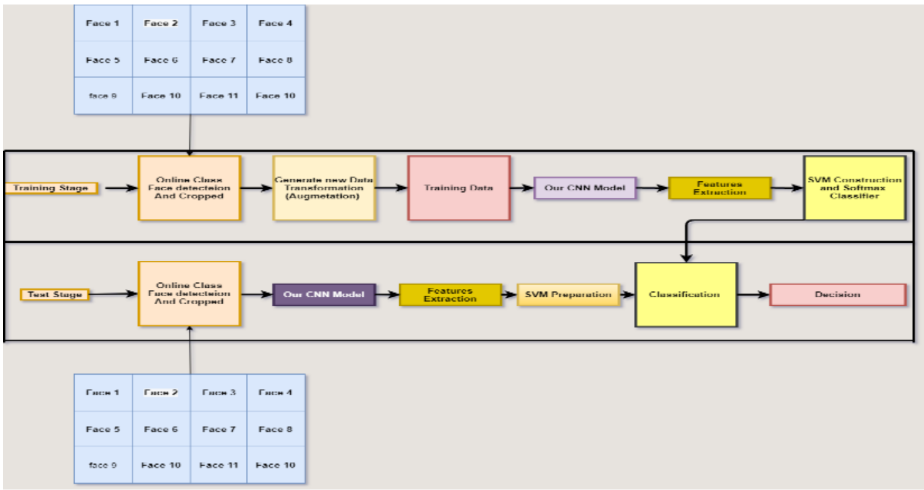


Fig. 5. Show proposed CNN and SVM online class training and testing architecture

6.2 Data augmentation

Data augmentation is useful and provides an excellent method for increasing the accuracy of face recognition by implementing data modifications to brightness, contrast, and saturation [28]. Contrast, Different shooting conditions lead to unexpected photo contrasts that can make recognition difficult. Such variations can be represented by digitally altering the contrast information using the HSV transform. This shows that it is necessary to generate new contrast information by increasing and decreasing the contrast by 20% while preserving the original face image information. Brightness and illumination variations are still a clear challenge in a real-time operation. Transforms that change the brightness of the image were implemented to basically detect such an action. The current researchers increased and decreased the brightness by 20% by developing a new image with a different condition, while keeping the original face image as part of the work in all operations. Saturation, poor lighting conditions, and efficiency, including the camera lens, can cause photo saturation to be extremely high or extremely low. Since a high value leads to undesirable colors in the facial image, the saturation ratios were generated using the S-channel of the HSV color space. To do this, the researchers increased and decreased the saturation by 20%.

7 Online face recognition and registration model

The face recognition model has been selected and validated and has proven to be able to overcome almost all challenges. For the online face recognition and registration model, we need another process to complete the registration using multi descriptor for feature extraction[9] and cosine similarity technique [33]. Multi descriptor used to generate two vectors for face verification as seen in equations (5), (6) and (7). The cosine similarity is calculated using equation (9). M and N are represented by two images, the tested image and the identified image. The two images must be introduced as feature vectors. The value of equation (9) must be between -1 and 1, where 1 represents perfect similarity and -1 represents perfect dissimilarity. The equation (10), (11) for the output of support vector machine and softmax probability. The threshold value should be set to a reasonable value.

$$MD = (x_8 f\left(g_p - g_c\right) + \ (\textstyle\sum_{i=7}^{n=0} \frac{x_{(i+1)}}{k} f\left(g_p - g_c\right)))2.55 \tag{5}$$

$$\mathrm{F(x)} = \begin{cases} New\ value & if\ X\ \geq 0, \\ 0 & otherwise. \end{cases} \tag{6}$$

$$\mathrm{F(x)} = \begin{cases} New\ value & if\ X\ \leq 0, \\ 0 & otherwise. \end{cases} \tag{7}$$

$$\text{similarity (M, N)} = \frac{M.N}{\|M\| * \|N\|} \tag{8}$$

Where M and N represented by two images, tested and predicted face.

$$\text{similarity (M, N)} = \frac{\sum_{i=1}^{n} M_i * N_i}{\sqrt{\sum_{i=1}^{n} M_i^2} * \sqrt{\sum_{i=1}^{n} N_i^2}} \tag{9}$$

$$s_1 = \text{ similarity } (M_1, \mathrm{N}_1) \tag{10}$$

$$s_1 = \text{ similarity } (M_1, \mathrm{N}_2) \tag{11}$$

Where s_1 And s_2 scores using cosine similarity and M_1 and N_1 for the support vector machine output and M_1 and N_2 for the softmax output.

$$Pridected(T) = \max\,(s_1, s_2) \tag{12}$$

Where T is a final predicted face that successfully verified as seen in equation (12).

This model was developed to measure the performance of CNN Softmax and SVM in the real world application and to clarify how it can be used in distance education as seen in Figure 6. The class with better similarity should be considered as the true class, as Equation 10 shows. Based on the recognition and cropping of faces considering the number of faces as reference, the system passes through. The system change the reference number itself if any error occurs during the registration process.

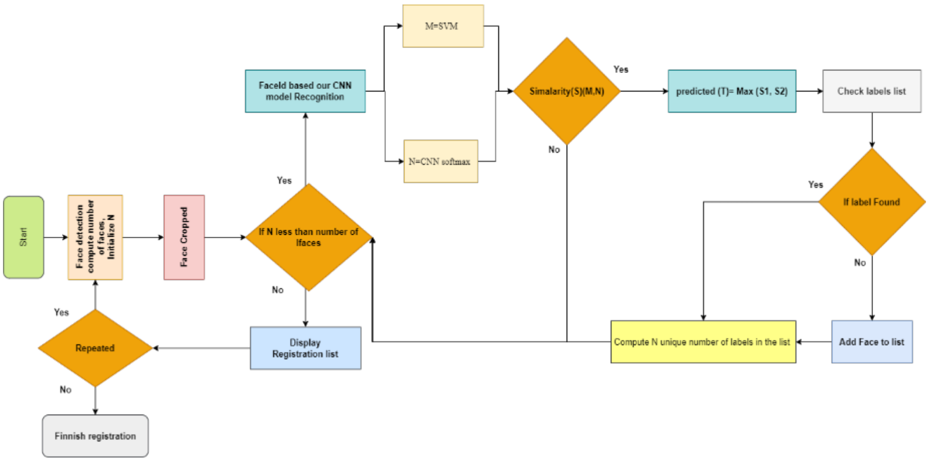


Fig. 6. Shows the complete architecture of online class face recognition system

7.1 Results and discussion

The performance comparisons were made on the basis of two tests. The first test was related to the recognition of faces of all classes. Figure 7, shows the result of the proposed CNN Softmax with the support of SVM which showed significant results in test 1, frontal face with different facial expression and slight right and left rotation with 100% accuracy. This is due to the cosine similarity function that we used to highlight the identification process. Figure 7, in test 2, frontal face with large facial expression and large right and left rotation as well as different background and large head scale. The result for this case shows 88.88% accuracy.

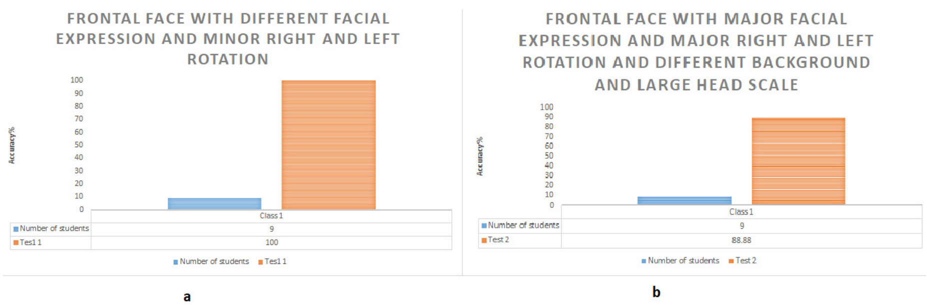


Fig. 7. Shows the complete results of online face recognition, a) test 1, minor variations, b) test 2, major variations

In Experiment 1, the frontal face tasks with different facial expressions and low right and left rotation were safer, more accurate, and more reliable than in Experiment 2. However, Experiment 2, test 2 with 88.88% accuracy means that the model was able to recognize 8 out of 9 faces in this situation, due to the extreme changes in face position, scale, resolution and low training data variation.

The proposed model need not the large data set to learn and this considering to be more effective since processing the dataset based on a fewer face image for training CNN is better than large data set and transfer learning as well as more effective in the face verification using multi descriptor for feature extraction.

The results shed light on the corresponding challenges that should be considered when developing the application of face recognition and registration in distance learning and other tasks. Because, we develop the model based on the similarity between datasets purpose, our methodology should consider to avoid the gap between face recognition models and real-world application in the case of distance learning.

8 Conclusion

In this paper, we presented a unique deep neural network with a high performance model for face recognition called İndepthnet19 motivated by various network design techniques. The datasets we selected for model validation are suitable for face recognition and distance learning requirements. Tests on the four datasets with different face recognition challenges have shown that our İndepthnet19 is efficient in terms of accuracy and the effectiveness of the model is comparable to most other deep face recognition networks. We also presented a new CNN model and technique for face recognition and registration. The application of face recognition in distance learning is a problem that arises from the difficulty of obtaining training data that is available online. We show how to tackle and solve this problem using data transformation and augmentation. The face feature was extracted from the proposed CNN and trained using Softmax and Support Vector Machine, which is suitable for limited datasets and lower face recognition errors. For all datasets, we emphasize the advantage of using Support Vector Machine to improve face recognition in real-time systems, which supports the verification phase. In our work, we presented a face recognition model that can be successfully used in distance learning registration and is suitable to overcome the challenges of face recognition and registration.

Finally, our model could also contribute to other classification tasks. The architecture for face recognition and registration is suitable for any work that needs to be done remotely, or it can be used in some institutions, especially in the current and future phases due to the constraints imposed on gatherings and working in a closed location. Using our model for emotion recognition in distance learning could be beneficial in the future and could give a good impact on learning processes.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

9 References

[1] A. B. Salem Salamh, A. A. Salamah, and H. I. Akyüz, "A Study of a New Technique of the CT Scan View and Disease Classification Protocol Based on Level Challenges in Cases of Coronavirus Disease," *Radiol. Res. Pract.*, vol. 2021, pp. 1–9, Mar. 2021. https://doi.org/10.1155/2021/5554408

[2] N. F. R. S.K. Nazarova, G. A.Jalilova,N.S.Mukhamedova, "Features ofDistance Learning Organization," *Ann. RSCB*, no. 1, pp. 339–347, 2021.

[3] D. Li, "Deep Learning Technology Based on Intelligent Teaching in Social Psychology Courses," *Int. J. Emerg. Technol. Learn.*, vol. 16, no. 24, pp. 40–56, 2021. https://doi.org/10.3991/ijet.v16i24.27255

[4] D. Mu and W. Guo, "Impact of Students' Online Learning Burnout on Learning Performance – the Intermediary Role of Game Evaluation," *Int. J. Emerg. Technol. Learn.*, vol. 17, no. 02, pp. 239–253, Jan. 2022. https://doi.org/10.3991/ijet.v17i02.28555

[5] M. H. Akmal Abdullaev, "INNOVATIVE DISTANCE LEARNING TECHNOLOGIES," *J. Crit. Rev.*, vol. 7, no. 11, 2020. https://doi.org/10.5958/2249-7137.2021.00198.1

[6] K. K. Abduraximovich, "Teaching based on distance education technology as a complete pedagogical problem," *Acad. AN Int. Multidiscip. Res. J.*, vol. 11, no. 1, pp. 1199–1203, 2021. https://doi.org/10.17718/tojde.849872

[7] D. Ozdemir and M. E. Ugur, "A Model Proposal On The Determination Of Student Attendance In Distance Education With Face Recognition Technology," *Turkish Online J. Distance Educ.*, vol. 22, no. 1, pp. 19–32, 2021. https://doi.org/10.17718/tojde.849872

[8] Z. Pei, H. Xu, Y. Zhang, M. Guo, and Y.-H. Yang, "Face Recognition via Deep Learning Using Data Augmentation Based on Orthogonal Experiments," *Electronics*, vol. 8, no. 10, p. 1088, Sep. 2019. https://doi.org/10.3390/electronics8101088

[9] A. B. S. Salamh and H. I. Akyüz, "A Novel Feature Extraction Descriptor for Face Recognition," *Eng. Technol. Appl. Sci. Res.*, vol. 12, no. 1, pp. 8033–8038, Feb. 2022. https://doi.org/10.48084/etasr.4624

[10] L. Zhou, H. Wang, S. Lin, S. Hao, and Z.-M. Lu, "Face recognition based on local binary pattern and improved Pairwise-constrained Multiple Metric Learning," *Multimed. Tools Appl.*, vol. 79, no. 1–2, pp. 675–691, Jan. 2020. https://doi.org/10.1007/s11042-019-08157-0

[11] T. Lindahl, "Study of Local Binary Patterns," *Science And Technology*. p. 3, 2007, [Online]. Available: http://www.ep.liu.se/

[12] C. H. Hima Bindu and K. Manjunatha Chari, "Face Recognition Using Improved Co-HOG Features," 2020, pp. 647–655. https://doi.org/10.1007/978-981-15-3992-3_55

[13] T. M. Abhishree, J. Latha, K. Manikantan, and S. Ramachandran, "Face Recognition Using Gabor Filter Based Feature Extraction with Anisotropic Diffusion as a Pre-processing Technique," *Procedia Comput. Sci.*, vol. 45, pp. 312–321, 2015. https://doi.org/10.1016/j.procs.2015.03.149

[14] T. Lindeberg, "Scale Invariant Feature Transform," *Scholarpedia*, vol. 7, no. 5, p. 10491, 2012. https://doi.org/10.4249/scholarpedia.10491

[15] M. Ghorbani, A. T. Targhi, and M. M. Dehshibi, "HOG and LBP: Towards a robust face recognition system," in *2015 Tenth International Conference on Digital Information Management (ICDIM)*, Oct. 2015, pp. 138–141. https://doi.org/10.1109/ICDIM.2015.7381860

[16] B. Ahuja and V. P. Vishwakarma, "Local Binary Pattern Based Feature Extraction with KELM for Face Identification," in *2020 6th International Conference on Signal Processing and Communication (ICSC)*, Mar. 2020, pp. 91–95. https://doi.org/10.1109/ICSC48311.2020.9182760

[17] X. Wu, R. He, Z. Sun, and T. Tan, "A Light CNN for Deep Face Representation With Noisy Labels," *IEEE Trans. Inf. Forensics Secur.*, vol. 13, no. 11, pp. 2884–2896, Nov. 2018. https://doi.org/10.1109/TIFS.2018.2833032

[18] K. Mrhar, L. Benhiba, S. Bourekkache, and M. Abik, "A Bayesian CNN-LSTM Model for Sentiment Analysis in Massive Open Online Courses MOOCs," *Int. J. Emerg. Technol. Learn.*, vol. 16, no. 23, pp. 216–232, 2021. https://doi.org/10.3991/ijet.v16i23.24457

[19] G. Zeng, Y. He, Z. Yu, X. Yang, R. Yang, and L. Zhang, "Preparation of novel high copper ions removal membranes by embedding organosilane-functionalized multi-walled carbon nanotube," *J. Chem. Technol. Biotechnol.*, vol. 91, no. 8, pp. 2322–2330, Aug. 2016. https://doi.org/10.1002/jctb.4820

[20] V. Sangeetha and K. J. R. Prasad, "Syntheses of Novel Derivatives of 2-Acetylfuro[2,3-a] carbazoles, Benzo[1,2-b]-1,4-thiazepino[2,3-a]carbazoles and 1-Acetyloxycarbazole-2-carbaldehydes.," *ChemInform*, vol. 37, no. 50, Dec. 2006. https://doi.org/10.1002/chin.200650130

[21] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size," Feb. 2016, [Online]. Available: http://arxiv.org/abs/1602.07360

[22] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep Face Recognition," in *Procedings of the British Machine Vision Conference 2015*, 2015, pp. 41.1-41.12. https://doi.org/10.5244/C.29.41

[23] P. Viola and M. J. Jones, "Robust Real-Time Face Detection," *Int. J. Comput. Vis.*, vol. 57, no. 2, pp. 137–154, May 2004. https://doi.org/10.1023/B:VISI.0000013087.49260.fb

[24] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint Face Detection and Alignment Using Multi-task Cascaded Convolutional Networks," *IEEE Signal Process. Lett.*, vol. 23, no. 10, pp. 1499–1503, Oct. 2016. https://doi.org/10.1109/LSP.2016.2603342

[25] Sergey Ioffe;Christian Szegedy, "Batch Normalization : Accelerating Deep Network Training by Reducing Internal Covariate Shift," *Proc. 32nd Int. Conf. Mach. Learn.*, vol. 37, pp. 448–456, 2015.

[26] S. Santurkar, D. Tsipras, A. Ilyas, and A. M. ˛adry, "How Does Batch Normalization Help Optimization," 2018.

[27] L. Spacek, "Libor Spacek's Facial Images Databases," *Computer Vision Science Research Projects*, 2009. https://cmp.felk.cvut.cz/~spacelib/faces/

[28] V. Uchoa, K. Aires, R. Veras, A. Paiva, and L. Britto, "Data Augmentation for Face Recognition with CNN Transfer Learning," in *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, Jul. 2020, pp. 143–148. https://doi.org/10.1109/IWSSIP48289.2020.9145453

10 Authors

Ahmed B Salem Salamh, PHD student, Institute of Science, Material Science and Engineering, Kastamonu University, Kuzey Kent /PO Box, Kastamonu /state/Turkey, postal code 37150, Turkey.

Halil Ibrahim Akyüz, PHD, Associate professor, Faculty of Education, Computer and Teaching Technologies Education, Kastamonu University, street address Kuzey Kent /PO Box, Kastamonu /state/Turkey, postal code 37150, Turkey.

Article submitted 2022-02-23. Resubmitted 2022-03-28. Final acceptance 2022-03-28. Final version published as submitted by the authors.