Early stroke disease prediction with facial features using convolutional neural network model

Ali Ahmad, Muhammad Usama, Yasir Niaz Khan
Computer Science and IT Department, The University of Lahore, Lahore, Pakistan

Article Info

Article history:

Received Jul 27, 2021
Revised Nov 9, 2023
Accepted Nov 30, 2023

Keywords:

Artificial intelligence
Deep learning
Medical disease
Neural network
Stroke disease prediction

ABSTRACT

Past researcher has proposed computed tomography (CT) and magnetic resonance image (MRI) scan images as the most efficient ways to diagnose stroke disease. These methods are not only hectic and take much time but are also costly. This paper proposes a new approach to diagnosing this disease and gives a time and cost-efficient solution. We have offered a two-step solution to diagnose stroke disease in a patient using only the patient's facial image. In the first step, we gathered a dataset of several stroke patients and normal persons. Then we applied several pre-processing operations, including red, green and blue (RGB) to grayscale conversion, scaling/ resizing, and normalization on dataset images before training them. In the second step, we trained the cropped images of their face regions and trained them using a convolutional neural network (CNN). We have successfully achieved an efficiency of 98%. The accuracy, precision, recall, and f-measure of the results were measured at 98%, 97%, 99%, and 98% respectively.

Corresponding Author:

Ali Ahmad
Computer Science and IT Department, The University of Lahore
Km. 1 Defence Road, near Bhuptian Chowk, Lahore, Pakistan
Email: alijakhar129@gmail.com

1. INTRODUCTION

The stroke is a medical emergency, and the brain of the person suffering from this medical condition stops working due to a lack of blood supply, ultimately leading to the death of brain cells. Brain cells start dying within a few minutes after the person experiences a stroke attack. The sooner the patient gets treated, the higher the chance of recovery and delay in treatment can lead to complications [1]. In 2018, one in every six deaths from cardiovascular disease was due to stroke [2]. Heart attack and stroke are leading causes of death globally. The World Health Organization (WHO) estimates that 7.3 million deaths globally were due to coronary heart disease, and 6.2 million were due to stroke in 2008 [3]. There are several methods to diagnose this disease, and two are more popular than the rest. The first is a computed tomography (CT) scan that uses a series of x-rays to create a detailed image of your brain. The second one is magnetic resonance image (MRI) which uses powerful radio waves and magnets to create a detailed view of your brain [4]. It is done in an enclosed space, and the loud noises made by the magnets can make some people feel fearful of being on a closely enclosed surface while having an MRI scan [5]. In recent years, the MRI has become more popular than (CT) scans [6].

The part of the body that gets effected by this stroke disease most of the time is the patient's face. More than 80% of stroke patients hospitalized with a confirmed stroke presented with some paresis, most common was of the arms 75.5%, while a majority reported he face paresis 54.6% and legs 68.6%, location of paresis was equally distributed between right and left legs [7]. Our research revolves around extracting the features of a patient face. If a person suffers a stroke attack, a few signs can be seen in the patient: the mouth appears to be falling by gravity, eyebrows get raised, and the face get tilted towards the either left or the right side. However, the most important among those is the inability of the patient to change its facial expressions. Thus, there are noticeable changes on the face of a patient suffering from stroke disease. Figure 1 will help you to understand. This research aims to present a simple, quick, and easy solution, unlike currently used CT scan and MRI methods. And we have also tried to provide the fact that facial image of a patient is enough to diagnose stroke disease in a patient.

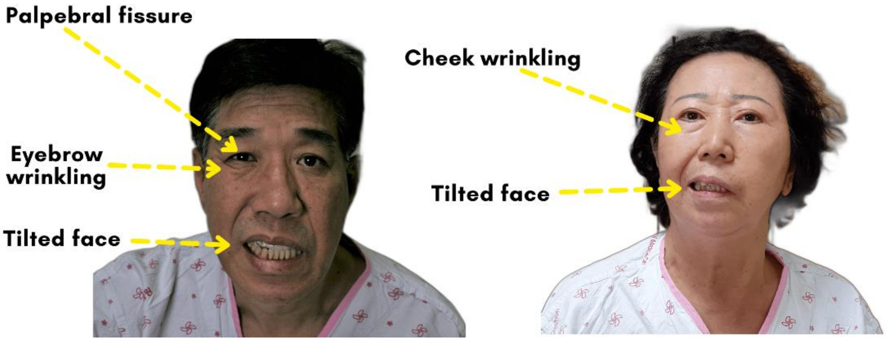


Figure 1. Shows stroke impact on face

2. RELATED WORK

Sabin Umirzakova present in his research paper to detect the initial symptoms of stroke disease by using facial features like the forehead, eyeballs movement, jaw dropping, and changes occurring in cheeks, and has achieved 91% accuracy [8]. In the past, researchers used CT scan images to identify stroke disease. One of the researchers applied different image processing techniques to CT scan images and achieved an accuracy of 90% [9]. Another research by JT Marbun to classify stroke using CT scan images. He used convolutional neural network (CNN) for classification of stroke images. Still, before training the neural network, he applied some pre-processing on the dataset to turn images into black and white. The preprocessing techniques he used included gray scaling, image scaling, and contrast limited adaptive histogram equalization (CLAHE). After their pre-processing steps, he trained the CNN model and achieved 90% accuracy [10]. Another researcher Bhagyashree Rajendra published his research paper on the classification and segmentation of stroke using MRI images. He used CNN and deep learning models in his research and got 96-97% accuracy on the classification model 85-87% accuracy on the segmentation model [11]. Another researcher HemaRajini published his research paper in 2013. He proposed to identify stroke disease. His method consists of five stages: pre-processing, tracking the brain’s midline, extracting features, segmentation, and classification, support vector machine (SVM) and k-nearest neighbor (KNN) have obtained a classification accuracy of 98% and 97%, respectively [12].

3. METHODOLOGY

Facial paresis is one of the most commonly occurring disorders in patients after stroke [13]. Our methodology is entirely different from others. Unlike others, who have been using traditional and hectic ways like medical training initiative (MTI) and CT scans to diagnose stroke, we have used face images of stroke patients to diagnose the disease. We collected images of stroke patients and normal people and classified them into two categories, “stroke faces” and “normal faces” The methodology is divided into three steps dataset collection, applying to pre-process on images, and developing our custom CNN model. Figure 2 shows the complete steps or general architecture of our methodology.

3.1. Dataset acquisition

The dataset consists of images of stroke patients and normal persons. There are 2,827 images in our dataset, out of which 35% are females, 65% are male images, 40% are stroke images, and 60% are normal face images. The dataset characteristics are shown in Table 1 and Figure 3 shown stroke patient and normal patient dataset sample.

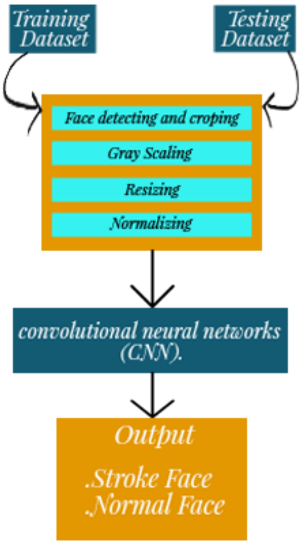


Figure 2. Shows general architecture

Table 1. dataset characteristics

Type	Number of images
Stroke persons	1,121
Number of patients	123
Male stroke	777
Female stroke	334
Normal persons	1,706
Male normal	1,053
Female normal	653
Total images	2,827

Figure 3. Shows dataset samples

3.2. Pre-processsing

The training phase becomes difficult if much irrelevant and redundant information is present or noisy and unreliable. Several steps are involved in pre-processing: face detection and cropping, gray scaling, resizing, and image normalization. Data pre-processing is the most important part of any automatic learning process. In some cases, it focuses on correcting the deficiencies that may negatively impact the process of learning, e.g. omission, noise, and outliers [14].

3.2.1. Face detection and cropping

The first step is to detect faces in the images and select their face region. We need cropped images of the face region in the training step, so we cropped the face region in the images for training because we have to work with face region images. That is why we need face in the images. There are various face detection techniques nowadays, but the best face detection technique is multi-task cascaded convolutional networks (MTCNN) [15], so we have used MTCNN for detection and cropping purposes. Figure 4 shown cropped faces of the stroke patient faces and normal patient faces after face detection using MTCNN.

Figure 4. Shows cropped face samples

3.2.2. Gray scaling

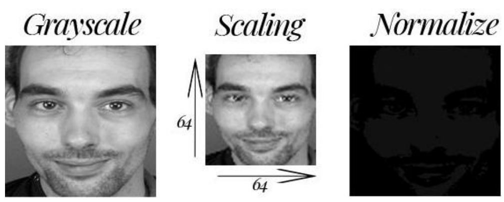
Gray scaling, often termed as "grayscale conversion," involves transforming a color image into shades of gray, where each pixel represents a shade from white to black, thereby eliminating chromatic information while retaining luminance details [16] Gray scaling is the process of converting the RGB image into a grayscale image because the grayscale image has only one layer, whereas the RGB has three layers red, green, and blue. So, it is easy to deal with grayscale images. Therefore, the conversion of an RGB image into a grayscale image is converting RGB values (24-bit) into grayscale values (8-bit) [17]. Figure 5(a) shows the grayscale image.

3.2.3. Resizing and scaling

Resizing process reduces the image’s size (pixel amount) in the CNN. The period and computational power also depend on the input size of the neural network. If the neural network’s input size increases, the duration, and computational power also increase. We have reduced the 150×150 images to 64×64 images in this research. Figure 5(b) shows the scaling image.

3.2.4. Normalization

Image normalization is a preprocessing technique extensively used in image processing and computer vision tasks. Introduced as a method to standardize the range of pixel intensities in images, normalization ensures a consistent scale across image data, promoting easier convergence for algorithms, especially deep learning models [18]. Normalization is the process of dividing each pixel value of the image by 255 to set the value of the image between 0 and 1 because neural networks process inputs using small weights values, and the values with large values can slow the learning process [19], and large values also make disturbance during backpropagation process. Therefore, normalization does not affect the original image in terms of visualization. Figure 5(c) shows the normalized form of the image.



(a) (b) (c)

Figure 5. Shows pre-processing samples (a) grayscale, (b) scaling, and (c) normalize

3.3. Convolutional neural network (CNN)

Image classification is a challenging problem of image processing, computer vision, and machine learning [20]. CNN is designed to learn spatial hierarchies of features efficiently through back propagation by using different building blocks, to name a few; convolutional layers, pooling layers, and fully connected layers [21]. CNN is mainly used for solving complex image-derived pattern recognition tasks and, with its precise yet straightforward architecture, gives us a simplified and easy method to get started with artificial neural networks (ANNs) [22] and CNNs are a class of deep learning algorithms that have become particularly powerful for tasks related to image recognition and classification. Introduced by LeCun et al [23]. CNNs were inspired by the visual cortex of animals. These networks consist of multiple layers of convolutions with nonlinear activation functions, pooling operations, and fully connected layers. Through these mechanisms, they can automatically and adaptively learn spatial hierarchies of features from input images. Their efficiency in handling image data has rendered CNNs as a cornerstone technique in computer vision applications. As demonstrated by Simonyan and Zisserman [24], the depth of the network, where multiple layers are employed, plays a pivotal role in improving recognition capabilities. We have created our convolutional neural network model. There are two main processes of stroke faces and normal faces classification, the first is training and the second one is testing. For the training, we have used an 80% dataset, and during the training, it was also performed validation testing on the 10% dataset.

There are two steps performed (feedforward and backpropagation). Suppose there is no feedback from the output of neurons towards the input throughout the network. In that case, the network is referred to as a 'feedforward neural network', otherwise, if there is any connection from the outputs to the inputs, the network is called a 'recurrent neural network'. Backpropagation is the most popular and most used one of the training of feedforward neural networks [25]. Backpropagation is used to trace the error by counting the weights from the output layer and sending them back to the hidden layers, so the neural network obtains weights with minimum error. I have used 20% data in the testing phase to test the CNN model. CNN only has a feedforward process. The sample architecture of our CNN model as shown in Figure 6 and the hyper-parameters we have used in our CNN model are also shown in Table 2.

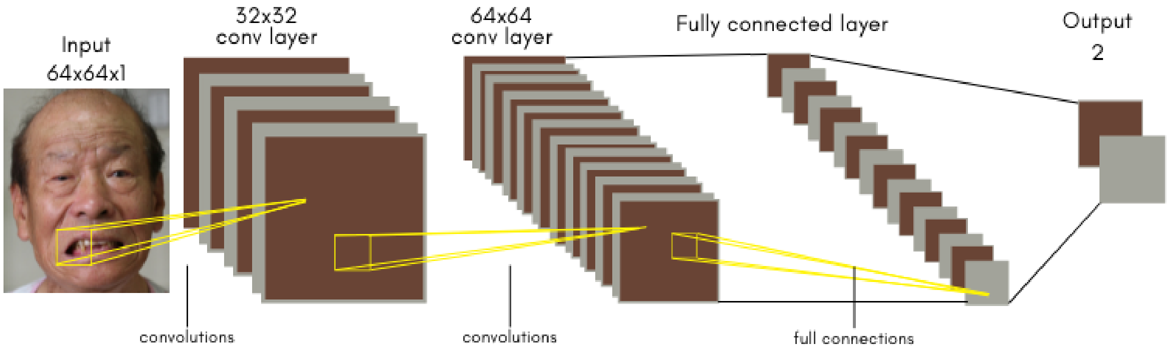


Figure 6. Shows sample architecture of our CNN model

Table 2. CNN model hyper-parameters

Parameters	Description
Convolutional layers	2 convolutional layers with kernel size: 5×5, filter: 32 2 convolutional layers with kernel size: 3×3, filter: 64
Pooling layers	2 max-pool layers with 3×3
Output nodes	2, 0 for normal and 1 for stroke
Learning rate	0.001
Optimizer	Adam
Batch size	32
Epochs	40
Fully connected layer nodes	250
Dropout	0.45

4. RESULT AND DISCUSSION

After training, we have achieved 99% training accuracy, and 98% validation accuracy shown in Figure 7 and we have also gotten 2% training loss and 5% validation loss shown in Figure 8. The performance analysis of the CNN model is done by testing the test dataset. We have successfully achieved 98% accuracy which is the best accuracy among most of the solutions mentioned above in the related work section [8]–[12]. The confusion matrix is shown in Figure 9 and other evaluation metrics results are shown in Table 3.

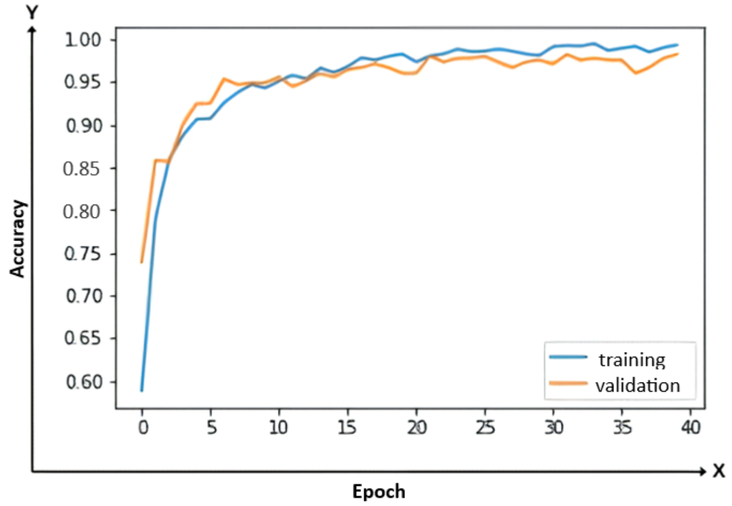


Figure 7. Training and validation accuracy graph

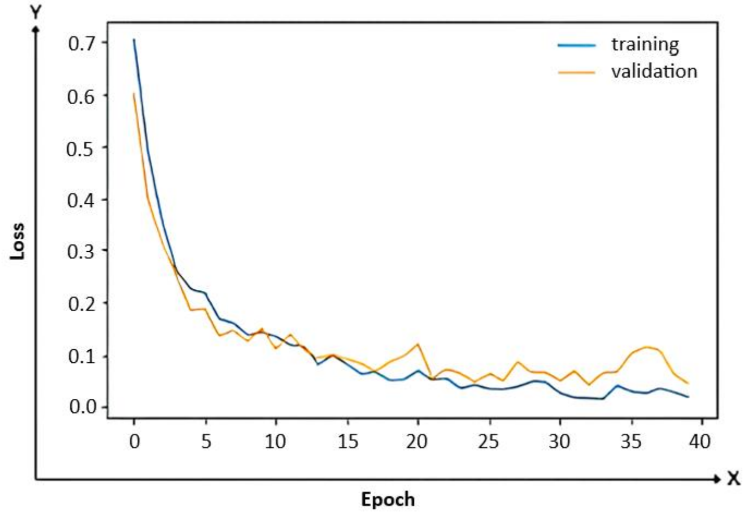


Figure 8. Training and validation loss graph

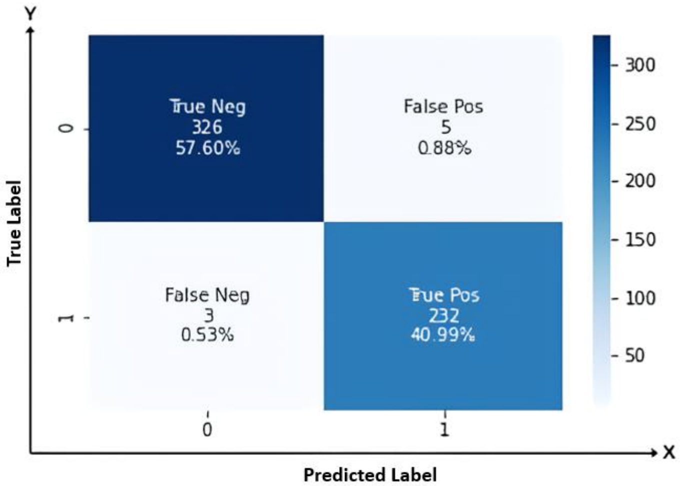


Figure 9. Shows confusion matrix

Table 3. Evaluation metrics results

Metrics	Results
Precision	97%
Recall	99%
F-measure	98%
Sensitivity	98%
Specificity	99%

5. CONCLUSION

Various techniques are used nowadays to detect stroke disease, but the most ignored approach is diagnosing it using facial images. Stroke disease diagnosis using facial images is not used to diagnose this disease and is considered an inefficient way to do so and other methods that are hectic and expensive are in place to diagnose stroke. This paper proposes the facial image of the patient as sufficient evidence to diagnose stroke disease in a patient. Facial images are an important yet ignored factor in diagnosing stroke, and they can serve as an effective alternative method to diagnose stroke disease. Other diagnosing methods that are being widely used nowadays, like CT scans and MRI, are hectic and expensive, so our proposed method is the best choice for diagnosis at the initial or early stages. With the diagnosis using our proposed methods, the patient can then proceed with a CT scan and MRI. We built a custom CNN model of our own, and 20% of images of the dataset were used as test data and the overall accuracy achieved is 98%.

REFERENCES

[1] S. Gupta, A. Mishra, and R. Menaka, "Ischemic stroke detection using image processing and ANN," in *Proceedings of 2014 IEEE International Conference on Advanced Communication, Control and Computing Technologies, ICACCCT 2014*, May 2015, pp. 1416–1420, doi: 10.1109/ICACCCT.2014.7019334.

[2] C. for D. C. and Prevention, "About the underlying cause of death, 1999-2019," [Online]. Available: https://wonder.cdc.gov/ucd-icd10.html.

[3] M. Shanthi, P. Pekka, and N. Bo, "Global atlas on cardiovascular disease prevention and control," *World Health Organization*, pp. 3–18, 2011, [Online]. Available: https://apps.who.int/iris/handle/10665/44701.

[4] M. Clinic, "Stroke disease diagnosis," [Online]. Available: https://www.omegapds.com/mri-advantages-and-disadvantages/.

[5] OMeg, "MRI advantages, and disadvantages," [Online]. Available: https://www.omegapds.com/mri-advantages-and-disadvantages/.

[6] H. R. Jäger, "Diagnosis of stroke with advanced CT and MR imaging," *British Medical Bulletin*, vol. 56, no. 2, pp. 318–333, Jan. 2000, doi: 10.1258/0007142001903247.

[7] S. S. Rathore, W. D. Rosamond, L. S. Cooper, L. E. Chambless, H. A. Tyroler, and A. R. Hinn, "Characterization of stroke signs and symptoms: findings from the Atherosclerosis risk in communities study," *Circulation*, vol. 103, no. suppl_1, pp. 1362–1362, Mar. 2001, doi: 10.1161/circ.103.suppl_1.9998-59.

[8] S. Umirzakova and T. K. Whangbo, "Study on detect stroke symptoms using face features," in *9th International Conference on Information and Communication Technology Convergence: ICT Convergence Powered by Smart Intelligence, ICTC 2018*, Oct. 2018, pp. 429–431, doi: 10.1109/ICTC.2018.8539440.

[9] I. Software, S. J. Said, J. A. E. Noor, and Y. Yueniwati, "Identification of Ischemic Stroke stages in CT-scan brain images using identification of Ischemic Stroke stages in CT scan brain images using imagej software," vol. 3, no. July 2014, pp. 24–31, 2020.

[10] J. T. Marbun, Seniman, and U. Andayani, "Classification of stroke disease using convolutional neural network," *Journal of Physics: Conference Series*, vol. 978, no. 1, p. 12092, Mar. 2018, doi: 10.1088/1742-6596/978/1/012092.

[11] B. R. Gaidhani, R. Rajamenakshi, and S. Sonavane, "Brain stroke detection using convolutional neural network and deep learning models," in *2019 2nd International Conference on Intelligent Communication and Computational Techniques, ICCT 2019*, Sep. 2019, pp. 242–249, doi: 10.1109/ICCT46177.2019.8969052.

[12] N. Hema Rajini and R. Bhavani, "Computer aided detection of ischemic stroke using segmentation and texture features," *Measurement: Journal of the International Measurement Confederation*, vol. 46, no. 6, pp. 1865–1874, Jul. 2013, doi: 10.1016/j.measurement.2013.01.010.

[13] P. Konecny, M. Elfmark, and K. Urbanek, "Facial paresis after stroke and its impact on patients' facial movement and mental status," *Journal of Rehabilitation Medicine*, vol. 43, no. 1, pp. 73–75, 2011, doi: 10.2340/16501977-0645.

[14] S. Tabik, D. Peralta, A. Herrera-Poyatos, and F. Herrera, "A snapshot of image pre-processing for convolutional neural networks: case study of MNIST," *International Journal of Computational Intelligence Systems*, vol. 10, no. 1, pp. 555–568, 2017, doi: 10.2991/ijcis.2017.10.1.38.

[15] "Face recognition techniques: a survey," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 10, no. 2, pp. 697–700, Apr. 2021, doi: 10.30534/ijatcse/2021/341022021.

[16] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004, doi: 10.1109/TIP.2003.819861.

[17] C. Saravanan, "Color image to grayscale image conversion," in *2010 2nd International Conference on Computer Engineering and Applications, ICCEA 2010*, 2010, vol. 2, pp. 196–199, doi: 10.1109/ICCEA.2010.192.

[18] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *32nd International Conference on Machine Learning, ICML 2015*, vol. 1, pp. 448–456, 2015, [Online]. Available: https://ieeexplore.ieee.org/document/7486899.

[19] "Scale images pixels data for deep learning," [Online]. Available: https://machinelearningmastery.com.

[20] M. M. Krishna, M. Neelima, M. Harshali, and M. V. G. Rao, "Image classification using deep learning," *International Journal of Engineering and Technology(UAE)*, vol. 7, no. 2.7, pp. 614–617, Mar. 2018, doi: 10.14419/ijet.v7i2.7.10892.

[21] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology," *Insights into Imaging*, vol. 9, no. 4, pp. 611–629, Jun. 2018, doi: 10.1007/s13244-018-0639-9.

[22] K. O'Shea and R. Nash, "An Introduction to Convolutional Neural Networks," *International Journal for Research in Applied Science and Engineering Technology*, vol. 10, no. 12, pp. 943–947, Nov. 2015, doi: 10.22214/ijraset.2022.47789.

[23] Y. L. Cun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2323, 1998, doi: 10.1109/5.726791.

[24] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 2015.

[25] M. H. Sazli, "A brief review of feed-forward neural networks," *Communications, Faculty Of Science, University of Ankara*, pp. 11–17, 2006, doi: 10.1501/0003168.

BIOGRAPHIES OF AUTHORS

Ali Ahmad holds a bachelor's of computer science degree from The University of Lahore, Pakistan in 2020. He is currently working as a Machine Learning engineer nationally and internationally. He participated in several international conferences. His research area includes face recognition systems, and computer vision-based deep neural networks. He can be contacted at email: alijakhar129@gmail.com.

Muhammad Usama is currently working as a software engineer. He completed his bachelor's in computer science from The University of Lahore. He is working on python for two years. He worked on various projects that sharpen his data science and image processing skills. Proficient in Pandas, OpenCV, Matplotlib, and Django framework. He can be contacted at email: usamabhatti4061@gmail.com

Dr. Yasir Niaz Khan is currently working as Assistant Professor at the Department of Computer Science at The University of Lahore and is the Founder & Head of the Research Group of Robotics and IoT. He has taught computer science courses at various universities in Pakistan. He completed his Ph.D. from the University of Tuebingen, Germany, in the Robotics Research Group. His work was mainly on Visual Terrain Classification onboard wheeled and flying robots. He remained at the Kaiserslautern University of Technology Germany for a few years, teaching robotics and AI group and Image Understanding and Pattern Recognition (IUPR) Group at DFKI. He can be contacted at email: yasir.niaz@gmail.com