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

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.

Keywords:
Artificial intelligence
Deep learning
Medical disease
Neural network
Stroke disease prediction

This is an open access article under the CC BY-SA license.

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 Bhagyarajendra 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.
Early stroke disease prediction with facial features using convolutional neural network model (Ali Ahmad)

3.2. Pre-processing

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

---

Table 1. Dataset characteristics

<table>
<thead>
<tr>
<th>Type</th>
<th>Number of images</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stroke persons</td>
<td>1,121</td>
</tr>
<tr>
<td>Number of patients</td>
<td>123</td>
</tr>
<tr>
<td>Male stroke</td>
<td>777</td>
</tr>
<tr>
<td>Female stroke</td>
<td>334</td>
</tr>
<tr>
<td>Normal persons</td>
<td>1,706</td>
</tr>
<tr>
<td>Male normal</td>
<td>1,053</td>
</tr>
<tr>
<td>Female normal</td>
<td>653</td>
</tr>
<tr>
<td>Total images</td>
<td>2,827</td>
</tr>
</tbody>
</table>

---

Figure 2. Shows general architecture

Figure 3. Shows dataset samples
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.

![Cropped Stroke Patient Faces](image1)

![Cropped Normal Patient Faces](image2)

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.

![Grayscale](image3) ![Scaling](image4) ![Normalize](image5)

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

Ali Ahmad

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 6. Shows sample architecture of our CNN model

<table>
<thead>
<tr>
<th>Table 2. CNN model hyper-parameters</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convolutional layers</td>
<td>2 convolutional layers with kernel size: 5x5, filter: 32</td>
</tr>
<tr>
<td>Pooling layers</td>
<td>2 convolutional layers with kernel size: 3x3, filter: 64</td>
</tr>
<tr>
<td>Output nodes</td>
<td>2 for normal and 1 for stroke</td>
</tr>
<tr>
<td>Learning rate</td>
<td>0.001</td>
</tr>
<tr>
<td>Optimizer</td>
<td>Adam</td>
</tr>
<tr>
<td>Batch size</td>
<td>32</td>
</tr>
<tr>
<td>Epochs</td>
<td>40</td>
</tr>
<tr>
<td>Fully connected layer nodes</td>
<td>250</td>
</tr>
<tr>
<td>Dropout</td>
<td>0.45</td>
</tr>
</tbody>
</table>
Figure 7. Training and validation accuracy graph

Figure 8. Training and validation loss graph

Figure 9. Shows confusion matrix
Table 3. Evaluation metrics results

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precision</td>
<td>97%</td>
</tr>
<tr>
<td>Recall</td>
<td>99%</td>
</tr>
<tr>
<td>F-measure</td>
<td>98%</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>98%</td>
</tr>
<tr>
<td>Specificity</td>
<td>99%</td>
</tr>
</tbody>
</table>

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


Early stroke disease prediction with facial features using convolutional neural network model (Ali Ahmad)


**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