GrenzeInternational Journal of Engineering and Technology, Jan Issue

Stroke Prediction using MRI Images

Dr. Mahesh Kaluti[1], Ramya S[2], Sahana Nayak[3], Sanjana R[4] and Shrunga SD[5]
[1-5]PES College of Engineering/CSE Engineering Department, Mandya, India
Email: mahesh.rkcet@gmail.com, ramyasshetty2000@gmail.com, sahananayak1903@gmail.com, sanjanaramalinganna@gmail.com, shrungadhanu2000@gmail.com

Abstract— **Early prediction of stroke plays an important role as it can reduce brain damage and also reduce the consequences caused by the stroke. Stroke is the second leading cause of death all over the world and one of the fatal diseases for people above the age 60 years. So, after considering all these risks we wanted to develop an automated system that could be helpful to predict the stroke in the early stages. Medical imaging is a well-developing technology in recent times and by considering the accuracy that the proposed system should provide we used MRI images as the dataset. In this study, we divided the whole process into six distinctive stages which included three major processes which are Data Collection, Pre-processing of the data, Performance Analysis Method of Classification. And In pre-processing, Gabor filters were used to improve the quality of the MRI images. Adaptive histogram equalization is the(AHE) is the technique used for enhancement..**

Index Terms— **Brain Stroke, Complexity, MRI, Gabor Filters.**

I. Introduction

Everyone's health is seen as a vital component of their lives. Because the brain is such an important organ, it requires a steady supply of blood, similar to how our hearts do. A brain stroke is caused by a blockage in the arteries that provide blood to the brain. Ischemic stroke (IS), haemorrhagic stroke (HE), and transient ischemic attack (TIA) are the three forms of stroke that can occur (TIA). One of the most common types of stroke is ischemic stroke. According to the American Heart Association (AHA), ischemic stroke accounts for 87 percent of all strokes. Ischemic stroke develops when a clot or obstruction in blood arteries persists. Although only 10-15% of strokes are expected to be haemorrhagic, the death rate is substantial when compared to ischemic stroke. In comparison to other types of stroke, TIA is a transitory obstruction that lasts only a few minutes. Biomedical imaging techniques can result in a visual representation of the interior of the human body. In the medical industry, a variety of imaging techniques are used, including x-rays, ultrasounds, computed tomography scans (CT), and magnetic resonance imaging scans (MRI). In comparison to other forms of medical imaging, the MRI is critical in providing analogous images of the brain with comparable resolution in multiple projections. And it is possible to prescribe an effective treatment using this method. As a result, the suggested system will go through a series of testing steps with unusual MRI scans, and the results will be positive.

II. Litreture survey

A. Machine Learning for Predicting Delayed Onset Trauma Following Ischemic Stroke

By: Anthony Ma, Gus Liu

Grenze ID:01.GIJET.9.1.636

Canonical profiles of the most frequent individuals that doctors are expected to encounter after an initial stroke attack are created in this study using a combination of supervised outcome data and unsupervised learning techniques like K-Means. Utilizing information from differential diagnoses and medical interviews, the risk of death is predicted. Critical qualities were found using principal component analysis (PCA). To classify individuals into canonical patient profiles, K-Means is used. Based on how similar they are, the objects will be grouped into K categories. A cluster patient profile serves as a representation of each patient's profile data. Support vector machines, or SVMs, are a tool for classifying data and solving regression issues. When the output variable has a continuous or real value, regression is utilised. When predicting the likelihood of a patient's mortality, SVM splits the data it has collected into two categories after conducting an analysis. The most representative profile for new patients can be determined, and a plan can be developed to reduce the risk of the most serious problems. We produced a typical profile of patients who are likely to get a stroke using both trained and unsupervised data. The most representative profile for new patients can be determined, and a plan can be developed to reduce the risk of the most serious problems. For various datasets and methodologies, different predictions are made. Therefore, the forecast is less precise and efficient. A technique called PCA can be used to reduce a huge number of variables to a manageable amount while still keeping most of the data from the larger set.

B. Use of Combination of PCA and ANFIS in Infarction Volume Growth Rate Prediction in Ischemic Stroke

By: Rahma Ali, UvaisQidwai, Saadat K. Ilyas

A procedure called Decompressive Hemicraniectomy necessitates the patient to go through numerous CT scans in order to quantify the size of the stroke-affected area, also known as the infarction volume. A strategy was presented that can predict the pace of infarction volume increase based on just one CT scan and a few clinical parameters. The prediction of infarction volume at a specific time is the main topic covered here. In order to predict infarction volume, the recommended method has proven to be more accurate. It combines Principal Component Analysis (PCA) and Adaptive Neuro-Fuzzy 0 percent Plagiarized 100 percent Unique Inference System (ANFIS). A technique called the Adaptive Neuro-Fuzzy Inference System (ANFIS) combines artificial neural networks (ANN) and fuzzy inference systems (FIS) to study infarction growth patterns and forecast infarction growth rate and volume for stroke patients with significant artery occlusion in their anterior circulation at a certain period. By reducing the dimensionality of a dataset while retaining as much information as feasible, PCA is a feature reduction approach. The suggested method estimates the second infarction growth rate from a smaller data set via feature reduction, producing results that are more accurate. It is inaccurate and ineffective.

C. Fuzzy Data to Crisp Estimates: Helping the Neurosurgeon Making Better Treatment Choices for Stroke Patients

By: Qidwai, Uvais

When the infraction volume reaches critical levels and requires immediate surgery, the infraction growth rate (IGR) has been established as a key factor. This book discusses some of the most challenging treatment choices that neurosurgeons must make. The infarction volume is mapped using a machine learning framework based on the Adaptive Neuro-Fuzzy Inference System (ANFIS). Using an artificial neural network as its first phase, ANFIS inputs data and groups them according to their shared characteristics. To develop expert decision surfaces that calculate specific outcomes, these clusters are combined using fuzzy rules. Using clinical data acquired from monitoring devices, an open source CT scan system predicts IGR. By figuring out what influences an infarction volume's behaviour, it can predict it with accuracy. 130 patients make up the dataset, making it a less effective model.

D. Machine Learning for Outcome Prediction of Acute Ischemic Stroke

By: Hamed Asadi, Richard Dowling, Bernard Yan, Peter Mitchell

After intra-arterial therapy, the acute anterior circulation was examined with a computer-aided diagnostic system. The objective is to forecast how an endovascular unsupervised feature perception enhancement algorithm will perform using an ischemic stroke on the basis of a number of variables, it produces predictions. We employed Support Vector Machine's (SVM) supervised learning methods. An ANN is a network of nodes that takes its cue from the ease with which brain neurons function. Both classification and regression problems are solved using SVM. The network is lowered to a specified level of error, indicating that the training is complete. The value is created. The identification and validation of predictive factors is possible with a logistic regression model. When the output variable is a real or continuous value, the regression problem is used. Using artificial neural networks and support vector techniques, a supervised machine is developed that can categorise these predictions into potential positive and negative outcomes. The algorithms were trained, verified, and assessed using data that was partitioned randomly. a set of standards for determining whether patients will benefit from intervention while avoiding possibly harmful therapeutic side effects. It proves that invasive stroke therapies are effective. The capacity to produce output is unaffected by the fault tolerance corruption of one or more cells. The ideal network structure is found through experience and trial & error. The network's error level has been reduced to a predetermined level, signifying that training is finished. The results for this value are not optimal. To improve their performance, they need big training datasets.

III. MOTIVATION

Cardiovascular diseases (CVDs) are the leading cause of death worldwide, according to the World Health Organization (WHO), claiming over 17.9 million lives each year. Heart attacks and strokes account for about 85% of all cardiovascular disease deaths. The number of strokes in low- and middle-income nations has more than doubled in the previous four decades. Given its tremendous impact on countries' socioeconomic development, this escalating catastrophe has gotten comparatively little attention to yet. A stroke can affect anyone, even children, but the great majority of instances involve adults nearing the end of their working lives. According to the latest numbers from the National Stroke Association, stroke is currently the sixth highest cause of death in the United States and is a major source of adult disability. Every year, it is estimated that 0.8 million people in the United States experience a stroke, with 0.13 million of them dying as a result. This means that one American dies from a stroke every four minutes.

One of the contributing factors for the remarkable increase in stroke-related deaths is a lack of understanding of stroke indications. Stroke is a condition that is often overlooked, and because of a lack of sensitivity to the signs and symptoms of stroke, many people do not seek treatment, resulting in a delay in treatment.

IV. PROBLEM STATEMENT

Stroke is the second leading cause of death all over the worldand individuals and national healthcare systems both face a significant health burden. Some of the complex and multifactorial risk factors for stroke include hypertension, cardiac disease, diabetes, dysregulation of glucose metabolism, arrhythmia, and lifestyle habits. As a result, the purpose of our study is to apply CNN principles to huge current data sets in order to accurately predict whether a particular MRI image contains a stroke as well as to determine the type or classification of stroke.

V. PROPOSED SYSTEM

1.Test image is given as input to the CNN model.
2.These input images are uploaded through the web interface.
3.Computers are unable to recognize or evaluate images in the same way that humans do. So, we have to figure out how to turn these images into numbers. Using NumPy we can convert these images to the array.
4.The array contains RGB values of each pixel of an image ranging from 0-256.
5.The trained model detects the stroke type whether it is Ischemic Stroke, Hemorrhagic Stroke or Normal.

VI. METHODOLOGY

The proposed methodology describes a series of techniques that help to detect brain stroke.The whole process is divided into six different stages.At every stage, the image is processed at various levels to obtain detailed information. AtFirst,Magnetic Resonance Imaging(MRI) of the brain is collected as data sets that will be used for training and testing the model.Secondly,pre-processing of the collected imagesis performed to enhance the quality of the MRI image and make it suitable for further processing.It also helps to improve some of the features of MRI images such as reducing the noises, improving the visual appearance. Thenext part is to filter the images to delete noises. The contrast in the image is also improved.Then,the required then classified according to the trained dataset. features are obtained from the segmented image.And later this image is then classified according to trained dataset.

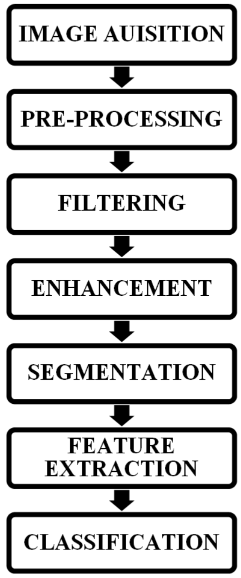

Chart-1: Flow Chart of Model

A. Image Acquisition

The initial stage of any system is the Image Acquisition stage. If the required image is not obtained properly, then the planned tasks will not be achieved properly. This process involves compression, display of data sets, storage, and printing.

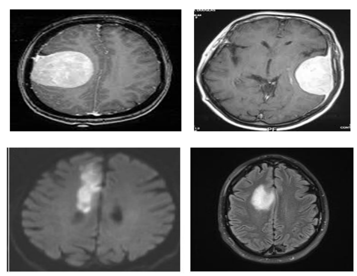

Fig 1: Datasets

B. Pre-Processing

It is a process that is used to improve the quality of the collected images. These involve processes like grayscale conversion, removal of noise, cropping, scaling, etc.

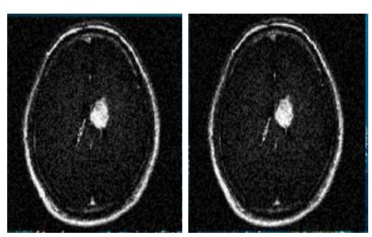

Fig 2: Pre-processing

C. Filtering

Filtering is a technique used to modify or enhance the image. Edge Enhancement, sharpening is some of the other image processing tasks that are done along with filtering.
In this process, Gabor filters are used to enhance the images and also analyse the texture of the image.

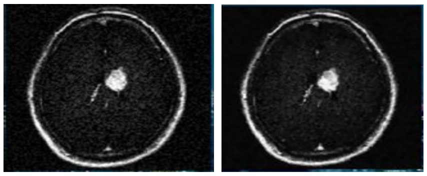

Fig 3: Filtering

D. Enhancement

Enhancement of the image is the process of altering and calibrating the images which were processed earlier so that the results are suitable for further processing. The practise of enhancing an image before processing aims to increase its quality and informational richness.

Common techniques include FCC, density slicing, spatial filtering, and contrast augmentation.

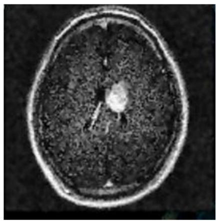

Fig 4(a): After the AHE process

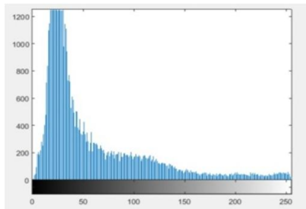

Fig 4(b):Adaptive Histogram

E. Segmentation

This segmentation process aims to divide the image into many subparts and analyse the minute details. The image is segmented into different classes where every class has a degree of membership.The Fuzzy C Means technique is used for clustering. Based on the histogram from the previous step FCM is applied to each cluster formed.

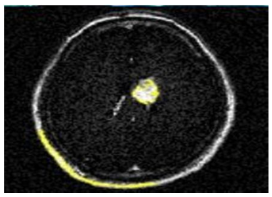

Fig 5: Segmented image

F. Feature Extarction

Feature Extraction is a type of dimensionality reduction that reduces the resources to define whole information. Extraction processes analyse the objects and images to export some of the prominent features which represent the classes of objects. The prominent characteristics in this case were extracted using the image analysis technique Gray-Level Co-Occurrence Matrix (GLCM).

G. Classification

The classification algorithm reviews the numerical properties and organizes the data into categories. The term "support vector machine" refers to a range of supervised methods for analysing data, recognising patterns, and classifying them.

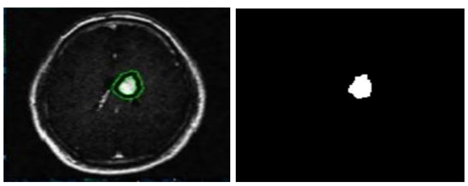

Fig 6: Affected region located

VII. SYSTEM ARCHITECTURE AND DESIGN

System Architecture is designed to identify the overall hypermedia structure for the applications. Applications architecture is defined within the development environment in which the application is implemented.
The entire process is divided into three phases which are, Pre-processing phase, Model creation phase, and Execution phase. In the first phase, the datasets are collected and the images are processed. The important features are retrieved in the second step, and the data is divided into training and testing data. The model is then trained and tested using training and test data. The model is ready if the outcome is satisfactory.And if the accuracy of the model is not satisfactory the model is again trained using training data and the process continues until the desired model is obtained. At the last phase, which is the execution phase a new MRI scan is fed to the system, and the system is going to predict the output.

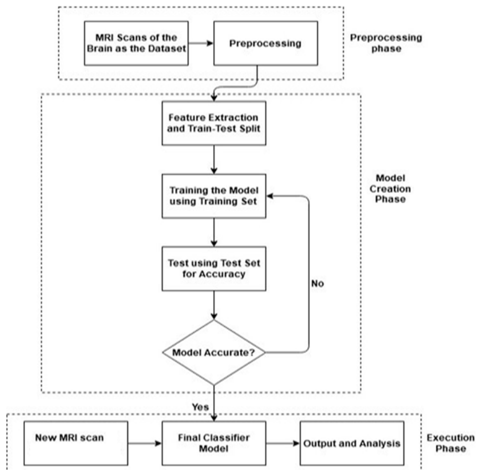

Fig7: System Architecture

VIII. ALGORITHM DETAILS

Convolutional Neural Networks

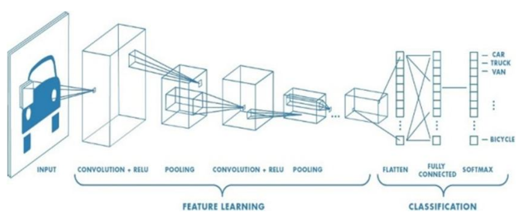

Figure 8 : Neural network with many convolutional layers

Convolution layer

The convolution layer is the first to take features out of an input image. Convolution learns visual features from small input data squares, preserving the link between pixels. It is a mathematical process that requires two inputs, such as an image matrix and a filter or kernel.

- An image matrix (volume) of dimension **($h \times w \times d$)**
- A filter **($f_h \times f_w \times d$)**
- Outputs a volume dimension **$(h - f_h + 1) \times (w - f_w + 1) \times 1$**

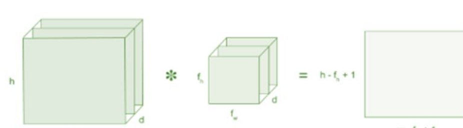

Figure9: Image matrix multiplies kernel or filter matrix

ReLU Layer

Rectified Linear Unit is the abbreviation for a non-linear operation.
The output is ***f(x) = max(0,x).***

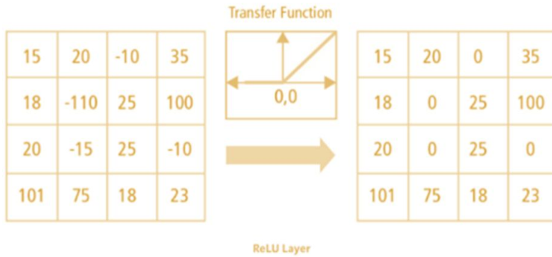

Figure10: ReLU Operation

Pooling Layer

When the photos are too huge, the section on pooling layers would lower the number of parameters. Spatial pooling, also known as subsampling or down sampling, decreases the dimensionality of each map while preserving crucial information. Multiple forms of spatial pooling are possible:

1.Max Pooling
2.Average Pooling
3.Sum Pooling

The largest element in the corrected feature map is selected using max pooling. The average pooling could be taken instead of just the largest element. Sum pooling refers to the total of all feature map elements.

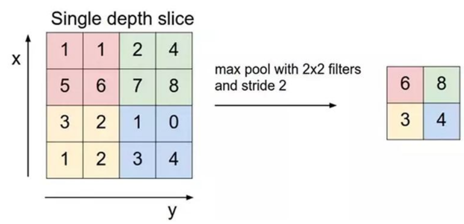

Figure11: Max Pooling

Fully connected layer

Neural networks that are fed forward are the Fully Connected Layer. The final several layers of the network are known as Fully Connected Layers. The output from the last pooling or convolutional layer is passed into the fully connected layer, where it is flattened before being applied.

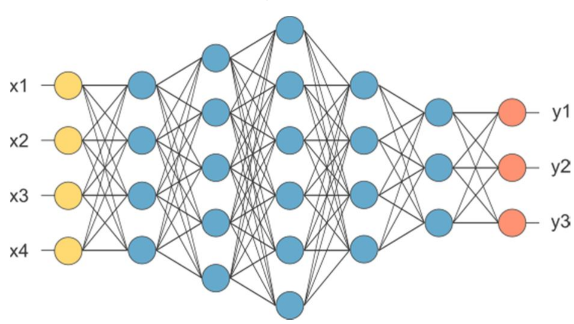

Figure12: Fully connected layer

IX. HARDWARE AND SOFTWARE REQUIREMENTS

Hardware

- System: intel i3/i5 2.4 GHz.
- Hard Disk:500 GB
- Ram:4/8 GB

Software

- Operating system: Windows XP/ Windows7/8/10.
- Software Tool: Open CV
- Coding Language: Python
- Toolbox: Image processing toolbox.

X. CONCLUSIONS

Stroke causes damage to multiple body components and causes unpredictable death. So, this will be a problem for not only the individual but also for the health care system. If a stroke is identified early enough, it is possible to receive sufficient treatment and recover. As a result, the purpose of this study is to develop a stroke prediction system that will help doctors anticipate strokes more accurately and quickly. The proposed machine learning algorithm can help determine whether or not someone is having a stroke also it displays the type of stroke.
This paper hence helps to predict the stroke risk. In the future, we will continue to collect more information on stroke and we will try to decrease the death rate due to stroke in our country.
likely to get a stroke using both trained and unsupervised data. The most representative profile for new patients can be determined, and a plan can be developed to reduce the risk of the most serious problems. For various datasets and methodologies, different predictions are made. Therefore, the forecast is less precise and efficient. A technique called PCA can be used to reduce a huge number of variables to a manageable amount while still keeping most of the data from the larger set.

REFERENCES

[1] [1] Mayank Chawla, Saurabh Sharma, Jayanthi Sivaswamy, Kishore L.T, “A method for automatic detection and classification of stroke from brain CT images”.

[2] [2] Jia Xie, Zhu Wang, Zhiwen Yu, Bin Guo, “Enabling Efficient Stroke Prediction by Exploring Sleep Related Features.

[3] [3] Luis Garcia-Temza, Jose L. Risco-Martin, Jose L. Ayala, Gemma ReigRosello, Juan M. Camarasltas, “Comparison of Different Machine Learning Approaches to Model Subtype Classification and Risk Prediction.

[4] [4] SFariborzKhorvash, Rasul Norouzi, “Prediction and Control of Stroke by Data Mining”.

[5] [5] Mahsa Khodadadi, HeidaraliShayanfar, KevianMaghooli and Amir HooshangMazinan, “Prediction of Stroke Probability Occurrence Based on Fuzzy Cognitive Maps”.

[6] [6] Pattanapong Chantamit-o-pas and Madhu Goyal, “Prediction of Stroke Disease using Deep Learning Model”.