Encephalon Disease Detection and Classification Using S-Transform and Sine Cosine Algorithm Based Deep Convolutional Network

By

Birhanu Abera Tafes

A Thesis Submitted to the

Department of Electronics and Communication Engineering

School of Electrical Engineering and Computing

Presented in Partial Fulfillment of the Requirement for the Degree of Master of Science in Electronics and Communication Engineering

Office of Graduate Studies

Adama Science and Technology University

Adama, Ethiopia
July 2020

Encephalon Disease Detection and Classification Using S-Transform and Sine Cosine Algorithm Based Deep Convolutional Network

By

Birhanu Abera Tafes

Advisor

Dr.Satyasis Mirshra (Ph.D)

A Thesis Submitted to the

Department of Electronics and Communication Engineering Program

School of Electrical Engineering and Computing

Presented in Partial Fulfillment of the Requirement for the Degree of Master of Science in Electronics and Communication Engineering

Office of Graduate Studies

Adama Science and Technology University

Adama, Ethiopia
July 2020

APPROVAL OF BOARD OF EXAMINERS

We, the undersigned, members of the Board of Examiners of the final open defense by Birhanu Abera Tafes have read and evaluated his thesis entitled "Encephalon Disease Detection and Classification Using S-Transform and Sine Cosine Algorithm Based Deep Convolutional Network" and examined the candidate. This is, therefore, to certify that the thesis has been accepted in partial fulfillment of the requirement of the Degree of Masters in Electronics and Communication Engineering.

Name of Student	Signature	Date
Advisor /Supervisor/	Signature	Date
Chair Person	Signature	Date
Internal Examiner	Signature	Date
External Examiner	Signature	Date
Department Head	Signature	Date
School Dean	Signature	Date
Postgraduate Dean	Signature	Date

Declaration

I hereby declare that this MSc thesis is my original work and has not been presented for a degree in any other university, and all sources of material used for this thesis have been duly acknowledged.

Name: _____

Signature: _____

This thesis has been submitted for examination with my approval as thesis advisor.

Name: _____

Signature: _____

Date of submission: _____

ADVISOR APPROVAL SHEET

To: Department of Electronics and Communication Engineering

Subject: Thesis Submission

This is to certify that the thesis entitled “Encephalon Disease Detection and Classification Using S-Transform and Sine Cosine Algorithm Based Deep Convolutional Network” submitted in partial fulfillment of the requirements of the Master of Science Degree in Electronics and Communication Engineering, postgraduate in Electronics and Communication Engineering, and has been carried out by Birhanu Abera Tafes Id No. PGR/18073/11 under my supervision. Therefore, I recommend that the student has fulfilled the requirements and hereby he can submit the thesis to the department.

Name of Advisor	Signature	Date

ACKNOWLEDGMENT

First and foremost, I would like to thank the almighty God for the strength and his blessing in completing this thesis.

Next, I would like to convey my sincere gratitude to my advisor Dr.Satyasis Mirshra (PhD) for his advice, continuous support, motivation, enthusiasm, patience, constructive comments and immense knowledge throughout the process of researching and writing this thesis.

I would like to thanks to Dr. Feyisa Debbo (PhD) for his academic guidance, motivation, encouragement and valuable suggestions.

I am also indebted to Mr. Tadesse Hailu, Head of the Electronics and Communication Engineering Department, for his guidance and valuable advice.

Finally, it is my great deep appreciation to my parents for their unlimited support and continuous encouragement throughout my years of study. This accomplishment would not have been possible without them. Sincere thanks to all.

Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYM

AI	Artificial Intelligence
ANN	Artificial Neural Network
BWT	Berkeley Wavelet Transformation
CA	Cellular Automata
CNS	Central Nervous System
CT	Computed Tomography
CWT	Continuous Wavelet Transform
dB	Decibel
DCNN	Deep Convolutional Neural Network
DL	Deep Learning
DNN	Deep Neural Network
DOST	Discrete Orthonormal Stockwell Transform
DST	Discrete Stockwell Transform
DWT	Discrete Wavelet Transform
FCA	Fuzzy C Means Algorithm
FCM	Fuzzy C-Mean
FFCA	Fast Fuzzy C Means Algorithm
FFT	Fast Fourier Transform
FIS	Fuzzy Interference System
FT	Fourier Transform
GLCM	Gray Level Co-occurrence Matrix
IDM	Inverse Difference Moment
KNN	K-Nearest Neighbor
LBP	Local Binary Patterns
LDA	Linear Discriminant Analysis
LLRBFNN	Local Linear Radial Basis Function Neural Networks
LIPS	Local Independent Projection-Based Classification
LMS	Least Mean Square Algorithm
ML	Machine Learning
MLP	Multilayer Perceptron
MRF	Margo Random Field

MRI	Magnetic Resonance Image
NVSS	National Vital Statistics Systems
PCA	Principle Component Analysis
PNN	Probabilistic Neural Network
PET	Positron Emission Tomography
RBFN	Radial Basis Function Neural Networks
RBF	Radial Basis Function
ReLU	Rectified Linear Unit
SVM	Support Vector Machine
STFT	Short Time Fourier Transform
SCA	Sine Cosine Algorithm
SMO	Sequential Minimal Optimization
WHO	World Health Organization
WT	Wavelet Transform

ABSTRACT

Magnetic Resonance Imaging (MRI) is playing an important role in the research of neuroscience for studying brain images. It is the one of the most frequently used diagnosis tool to detect and classify abnormalities in the brain. Automatic classification and detection is a difficult and complex task for a radiologist or clinical practitioner for identification and extraction of infected tumor areas from the MRI (Magnetic Resonance Imaging).Further, classifying the type of tumor from Magnetic Resonance (MR) images also a vital part of the diagnostic system. Factors like size, shape, and position of tumor vary from different patient's brain. Many efforts have been made for image detection and classification, but getting accurate automated techniques taken higher computational time. Motivated by the above difficulties, the research proposes Stockwell transform based segmentation technique called as Discrete Orthonormal S-Transform (DOST) to improve the performance of detection process. The DOST also utilized for feature extraction of the image. Further, a SCA (Sine Cosine Algorithm) based DCNN (Deep Convolutional Neural Network) model has been developed for classification of brain tumors into malignant (cancerous) and benign (noncancerous) category. The SCA has been utilized for weight optimization in the fully connected layer of the DCNN model. Also the different category of hidden neuron functions at the hidden layer has been tested with the new hybrid SCA-DCCN model and comparison results are presented. In this thesis an effort will be made to list and cover previous work of different researchers to improve the accuracy of diagnosis process.

CHAPTER-1

INTRODUCTION

1.1 Back ground

Brain Cancer is one of the main causes of death. As per the reports of World Health Organization (WHO), there are 9.6 million peoples died of cancer worldwide in 2018 and 17,760 adults are died from brain tumors in 2019 [1]. In the human body, Brain is one of the utmost complex structures and brain tumor is uninhibited division of cells creating an irregular group and that group of cells can affect the normal brain activity and destroy the healthy cells [2, 3]. The low-grade (grade I and II) termed as benign (non-cancerous) and high-grade (grade III and IV) are termed as malignant (cancerous) tumors [4-6]. The automatic medical imaging techniques become popular due to magnetic resonance imaging (MRI) [7-9].

Overall, a person's likelihood of developing primary brain or spinal cord tumor type of tumor in their lifetime is less than 1%. In fact, brain and other nervous system cancer is the 10^{th} foremost cause of death for men and women. The death rate from primary cancerous brain and CNS tumors is projected as 60%. According to the last report from National Vital Statistics Systems (NVSS), the mortality from 1975 to 2016 was higher among men and more in older individuals. People with age 65+ years had a significant increase in mortality for all tumors, while people aged <20 years had no significant changes in mortality. There have been significant increases in mortality in the elderly (age 65+ years), especially those age 75–84 years [10-11]. The rate of survival in the age group less than 15 is more than 74% and age between 15 to 39 is about 71%. [11].

The incidence of brain tumors in the last 20 years is increasing in all ages; however, it has grown more than 40 percent in adult person [12]. Global findings indicate a wide variation in the incidence of tumors of the nervous system in such a way that the standardization of age in different countries is between 0.01 and 12.7 in males and 0.01 and 10.7 in women, per 100,000 people. The lowest incidence is in Africa while its highest level is in northern Europe1. As a consequence, the difference in the diagnosis of the disease and the registration and reporting system in different countries can be due to these disparities [13]. The mortality rate of the neural system cancers is estimated at 3.4 per 100,000 everywhere

in the world. The increasing trend registered in prevalence among the elderly can be due to an increase in their life expectancy. In developed countries, in comparison with other countries of the world, the factors such as increased life expectancy, urbanization, and lifestyle changes are associated with an increase in the incidence and death of the brain tumor [14].

The involute structure of the human encephalon perplexes the diagnosis of the tumor in the encephalon region. Especially in encephalon imaging, the MRI method provides a unique presence in visualization of soft tissues with spatial resolution and contrast resolution. Among the studies predicated on image processing, transform-predicated approaches and thresholding methods are mundane. Salem and Alfonso utilized Expeditious Fourier Transform (FFT) and Minimal Redundancy-Maximal Pertinence (MRMR) techniques are utilized for automatic relegation of MRI encephalon tumor images [15]. Remya et al. utilized discrete wavelet transform (DWT) together with Fuzzy C-Denotes method for segmentation of tumors from MRI images [16]. While thresholding and clustering are mundane methods for tumor segmentation [17-19], it is withal possible to cumulate these with transform-predicated methods [20].

Machine learning (ML) [21, 22] is a center of manmade consciousness [23]. AI has additionally been broadly applied in medical diagnosis [9], and computer vision [25-28]. With the improvements of neural networks [19, 30], the awareness of deep learning (DL) has been proposed. DL can discover new methodologies are also known as automatic features acquisition strategies [31-34]. A convolution neural network (CNN) represents a most eminent DL structure utilized commonly for description and segmentation of brain tumors.

This proposed research work introduces an efficient model for brain tumor classification, where, the real Magnetic Resonance (MR) images are classified into non-cancerous (benign) brain tumor and cancerous (malignant) brain tumor. In this research work we propose a SCA based DCNN model for brain tumor classification of benign and malignant tumors and also we have proposed Discrete orthonormal S Transform to decompose the MRI image into different levels of approximate and detailed coefficients.

1.2 Brain Tumor Magnetic Resonance Image Overview

The intricate structure of the human encephalon perplexes the diagnosis of the tumor in the encephalon region. Encephalon tumor is incited because of a mundane boom of cells [35]. It is divided into two categories that are malignant or benign [36-39]. Encephalon tumors sodalities [39] relegate the encephalon tumor in four grades where grade I and II are referred to as benign and the remaining III and IV are labeled as malignant [40-44].

GLIOMAS are the encephalon tumors with the highest mortality rate and prevalence [45]. These neoplasms can be graded into Low Grade Gliomas (LGG) and High Grade Gliomas (HGG), with the former being less truculent and infiltrative than the latter [45-49].The technique magnetic resonance imaging (MRI) is of good quality among other techniques as it gives better contrast details about the encephalon tissues from a variety of image sequences [50-54]. The medical image analysis field has made CNNs very popular. The machine learning methods have been utilizing kernels like Gaussian or Haar-like for getting the occurrence detail, but CNNs pick up the kernel set predicated on the training data as provided [55].

The case studies of the brain tumour has been taken from the Harvard medical school website

- http://www.med.harvard.edu/aanlib/cases/case33/case.html

(i) Case study-1

A 73 year old dextral man sought medical attention because of a grand mal seizure and progressive arduousness with verbalization. Encephalon biopsy in 3/95 revealed grade II astrocytoma; with PET.Case contributed by Drs. C. D. Sturm and R. Bucholz, St. Louis University.

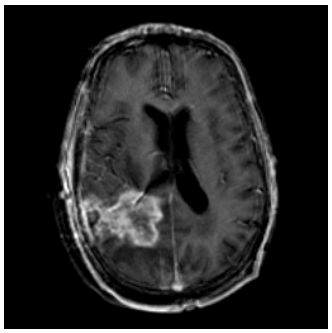

Fig. 1.1 Astrocytoma

(ii) Case study-2

A 35 year old man commenced having headaches eight months anteriorly. A low grade glioma was found, and he was treated with radiation. Thallium images show an anteriorborder of high uptake, consistent with a diminutive region of tumour recurrence.

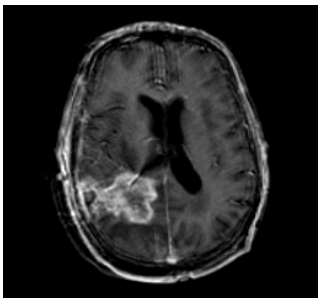

Fig. 1.2 Astrocytoma

(iii) Case study-3

A 62 year old man suffering a first seizure, a tonic-clonic convulsion with focal onset, and became unresponsive to verbal commands. MR images show a lesion involving the right second frontal convolution and another in the cerebellum, near the fourth ventricle, additionally visible on the sagittal image map.

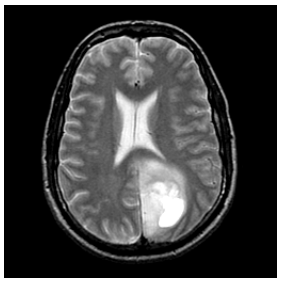

Fig. 1.3 Metastatic Carcinoma of the colonatic Carcinoma of the colon of the colon

(iv) Case study-4

A 51 year old woman sought medical attention because of gradually increasing right hemiparesis (weakness) and hemianopia (visual loss).

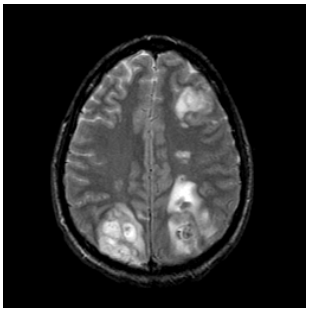

Fig. 1.4 Anaplastic Astrocytoma

(v) Case study-5

The patient was a 22 year old man who was admitted for resection of Ewing's sarcoma (peripheral/primitive neuroepithelialtumor- PNET) and had a right homonomous hemianopia, a left inferior quadrantanopia, right lower extremity hyperreflexia, and right extensor plantar response.

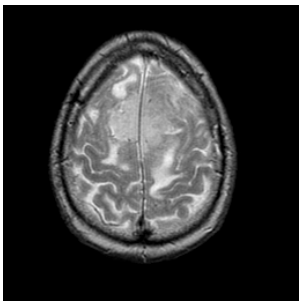

Fig. 1.5 Sarcoma

(vi) Case study-6

The patient was a 75 year old man who had an 8 - 10 month history of progressive difficulty walking. He had noted some left lower extremity weakness and some difficulty with memory and concentration.

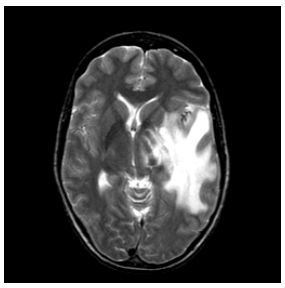

Fig. 1.6 Meningioma

(vii) Case study-7

A 42 year old woman with a long history of tobacco use began having headaches. The MR demonstrates the tumor as an area of high signal intensity on proton density (PD) and T2-weighted (T2) images in a large left temporal region. Contrast enhancement shows the lesion to contain a cystic component here.

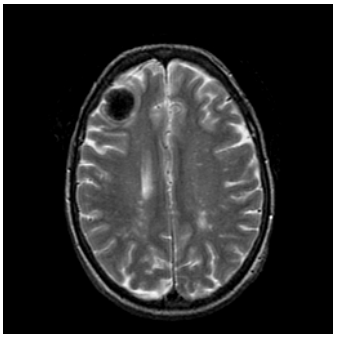

Fig. 1.7 Metastatic bronchogenic carcinoma

1.3 Signal Processing Transform Techniques

1.3.1Wavelet Transform

The Fourier transform is commonly used to decompose a signal into its frequency components in signal and image processing [56]. Fourier transform of a one-dimensional function is defined as:

$$H(f) = F\{h(t)\} = \int_{-\infty}^{\infty} h(t)e^{-i2\pi ft}\,dt \tag{1.1}$$

The inverse Fourier transform of *H(f)* is defined as:

$$h(t) = F^{-1}\{H(f)\} = \int_{-\infty}^{\infty} H(f)e^{j2\pi ft}\,df \tag{1.2}$$

In order to resolve this global issue one may use the short-time Fourier transform (STFT) [56]. The STFT is given by:

$$STFT\{h(t)\} = G(\tau, f) = \int_{-\infty}^{\infty} h(t)\omega(t-\tau)e^{-j2\pi ft}\,dt \tag{1.3}$$

And the inverse STFT is:

$$h(t) = \frac{1}{\omega(0)}F^{-1}G(\tau, f) = \frac{1}{\omega(0)}\int_{-\infty}^{\infty} G(\tau, f)e^{j2\pi ft}\,df \tag{1.4}$$

Where τ is a translation variable that allows the window $\omega(t)$ to translate in time. The wavelet transforms successfully overcome the shortcomings of the STFT [56].

The continuous wavelet transform for a continuous-domain input is given as:

$$W(\tau, s) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} h(t) \varphi \left(\frac{t-\tau}{s} \right) dt \quad (1.5)$$

Where τ is a translation variable while s is a scaling parameter.

If an image is decomposed by wavelets, the overlap in the frequency domain becomes non-avoidable. Even though the term "scale" can be approximately interpreted as "frequency" due to its ability in adjusting the size of the basis function, there is no straightforward way to turn this scale information into proper frequency information. In response to this restriction, the Stockwell transform was published [56].

The Stockwell Transform (ST) improves the time-frequency resolution in the multi-resolution analysis domain [57]. It maintains the absolutely-referenced frequency and phase information.

1.3.2 The 1-D Stockwell Transform

Stockwell Transform gives a full spatial-frequency decomposition of a signal. The Stockwell transform of an input signal is defined as the Fourier transform (FT) of the product between the input signal and a Gaussian window function [56]. It is based on the continuous wavelet transform, and the STFT [57]. Combining these two transforms, the 1D S transform is given by:

$$S(\tau, f) = S\{h(t)\} = \int_{-\infty}^{\infty} h(t) \frac{|f|}{\sqrt{2\pi}} e^{\frac{-(\tau-t)^2}{2} f^2} e^{-j2\pi f t} dt \quad (1.6)$$

S (τ, f) decomposes the input signal *h(t)* into temporal (*τ*) and frequency (*f*) components. The value of τ represents the center of the window function, and thus, by picking all possible values for τ, the S transform coefficients will cover the whole temporal axis and create full resolutions for each designated frequency [56]. Therefore ,for different *f* values there will be different resolutions [56]. As opposed to the wavelet transform, the kernel of the S transform is not shifted in time. Only the modulating window (Gaussian) is shifted. Therefore, the phase of the signal does not change with respect to the origin as it does for the wavelet transform. Stockwell calls this phase "absolutely referenced", because at t = 0, the phase of the time-frequency signal is zero [57].

Because the Stockwell coefficients *S(τ, f)* are complex, formula 1.6 can be generalized as:

$$S(\tau, f) = A(\tau, f) e^{j\emptyset(\tau, f)} \quad (1.7)$$

Where *A(τ, f)* is the "amplitude S-spectrum" and $\emptyset(\tau, f)$ is the "phase S-spectrum" which allows the definition of a broadband generalization of instantaneous frequency.

Using the convolution theorem in the Fourier space, the S transform can also be written as:

$$S(\tau, f) = \int_{-\infty}^{\infty} F(\alpha + f) e^{\frac{-2\pi^2\alpha^2}{f^2}} e^{j2\pi\alpha\tau}\, d\alpha \tag{1.8}$$

Where $F(\alpha + f)$ is the Fourier transform of $h(t)e^{-j2\pi ft}$.

This allows us to use the Fast Fourier Transform (FFT) while computing the S-transform. The relation between *S (τ, f)* and the Fourier transform of $h(t)$ is expressed as:

$$H(f) = \int_{-\infty}^{\infty} S(\tau, f) d\tau \tag{1.9}$$

Where $H(f)$is the Fourier transform of $h(t)$.

The original function $h(t)$can be recovered by calculating the inverse Fourier transform of $H(f)$as:

$$h(t) = F^{-1}\{H(f)\} = \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} S(\tau, f) d\tau \right\} e^{j2\pi ft} df \tag{1.10}$$

For an input $h(m)$, *m = 0,1,…,N − 1*, the discrete Stockwell transform (DST) of the input can be defined as:

$$S(i, n) = \sum_{m=0}^{N-1} H(m + n) e^{\frac{-2\pi^2 m^2}{n^2}} e^{\frac{j2\pi mi}{N}} \tag{1.11}$$

where *i = 0,1,…,N-1*.

For a fixed value of i, the DST coefficients for different *n* can be regarded as the inverse Fourier transform of the term $H(m + n)e^{\frac{-2\pi^2 m^2}{n^2}}$

1.3.3 Fuzzy C Means Algorithm

In FCM, let $X = \{x_1, x_2, x_3 \ldots\ldots x_N\}$ signifies the data with N data samples. Fuzzy C Means [58]-[61] method is based on minimization of the following objective function:

$$J_m = \sum_{k=1}^{N} \sum_{i=1}^{C} u_{ik}^{m} \left\| x_k - v_i \right\|^2 \tag{1.12}$$

Where m=2, fuzziness coefficient, u_{ik} is the degree of membership of x_k in cluster i, x_k is the i_{th} of n-dimensional measured data, v_i is the n-dimensional center of the cluster. $\|.\|$ is a norm metric and 'm' is a constant. The parameter "m" decides the "fuzziness of the consequential partition". By taking the derivative of the equation and make it equal to zero by using "Lagrange method", we have following equations as

$$v_i = \frac{\sum_{k=1}^{N} u_{ik}^m . x_i}{\sum_{k=1}^{N} u_{ik}^m}, \quad u_{ik} = \sum_{k=1}^{C} \left[\frac{\|x_k - v_i\|}{\|x_k - v_k\|} \right]^{-2/m-1} \tag{1.13}$$

Many variations to the FCM have been proposed because conventional FCM couldn't perform well in the presence of noise and intensity inhomogeneity.

1.3.4 Fast Fuzzy C Means Algorithm

According to Fast fuzzy c means algorithm [62], Let $X = [x_1, x_2 ..... x_n]$ be a N sample data set and assume that each sample x_k is represented by a set of p features and U is the hard partition matrices whose general term is given by $u_{ik} = 1$ if $x_k \in X_i$, and 0 otherwise. To get partition matrix, the HCM (Hard c Means) algorithm is chosen which minimizes the objective function

$$J = \sum_{k=1}^{N} \sum_{i=1}^{L} u_{ik}^m \|x_k - C_i\|^2 \tag{1.14}$$

Where "L" is the number of clusters and C_i is the cluster center and "m" is the Fuzzifier exponent and $u_{ik} \in [0,1]$.

Ruspini et al.[63] proposed fuzzy partition, where samples may partially belong to "several clusters through the conception of partial membership degrees".

Minimization of equation (1.14) is obtained by an optimization technique that successively updates the cluster centers C_i and partition matrix U by using the formula

$$C_i = \frac{\sum_{k=1}^{N} u_{ik}^m x_k}{\sum_{k=1}^{N} u_{ik}^m} \tag{1.15}$$

$$\text{And } u_{ik} = \frac{1}{\sum_{j=1}^{L} \frac{\|x_k - C_i\|}{\left(\|x_k - C_j\|\right)^{2/(m-1)}}} \tag{1.16}$$

1.3.4.1 Centroids Initialization

As defined, iterative fuzzy clustering methods do not ensure a unique final partition because different results are obtained with different initializations of V (or U)[64].

However, most of the practitioners initialize in a random manner for the initialization of V [65].

For each p-dimensional sample x_k, we define its relative mean by

$$m_k = \frac{1}{p}\sum_{j=1}^{p} x_{kj} \tag{1.17}$$

so that we obtain the n-dimensional vector $m = (m_1, m_2, ...m_n)$.

Postulating that the clusters are equipollently distributed, we uniformly split the n-dimensional vector m into c groups. In other terms, we set $c+1$ indices, say $l_0, l_1..l_c$, such that the c differences $(l_i - l_{i-1})$ are roughly equipollent. More formally, each index is given by

$$l_i = i * \lfloor n/c \rfloor \tag{1.18}$$

Where $\lfloor . \rfloor$ is the floor function. We iteratively build c subsets S_i of n as follows

$$S_i = \{l_{i-1} + 1, ...l_i\}$$

We obtain the subset of indices in each cluster by applying the inverse function:

$$C_i = \sigma^{-1}(S_i) \tag{1.19}$$

Finally, each "cluster center" is computed using:

$$v_i = \frac{1}{|C_i|}\sum_{j \in C_i} x_j \tag{1.20}$$

Where $|C_i|$ is the cardinality of C_i.

1.3.5 Fuzzy clustering with spatial constraints (FCM_S) and its variants

Ahmed et al. [66] proposed a modification to the standard FCM, in that modified objective function of FCM_S is defined as follows:

$$J_m = \sum_{i=1}^{c}\sum_{k=1}^{N} u_{ik}^{m} \|x_k - v_i\|^2 + \frac{\alpha}{N_R}\sum_{i=1}^{c}\sum_{k=1}^{N} u_{ik}^{m} \sum_{r \in N_k} \|x_r - v_i\|^2 \tag{1.21}$$

where x_k is the gray value of the k^{th} pixel, v_i represents the prototype value of the i^{th} cluster, u_{ik} represents the fuzzy membership of the k^{th} pixel with respect to cluster i, The parameter α is used to control the effect of the neighbours term. By definition, each

sample point x_k satisfies the constraint that $\sum_{i=1}^{c} u_{ik} = 1$. Two necessary but not sufficient conditions for J_m to be at its local extreme will be obtained as follows:

$$u_{ik} = \frac{\left(\left\| x_k - v_i \right\|^2 + \frac{\alpha}{N_R} \sum_{r \in N_k} \left\| x_k - v_i \right\|^2 \right)^{-\frac{1}{m-1}}}{\sum_{i=1}^{c} \left(\left\| x_k - v_i \right\|^2 + \frac{\alpha}{N_R} \sum_{r \in N_k} \left\| x_k - v_i \right\|^2 \right)^{-\frac{1}{m-1}}} \tag{1.22}$$

$$v_i = \frac{\sum_{k=1}^{N} u_{ik}^m \left(x_k + \frac{\alpha}{N_R} \sum_{r \in N_k} x_r \right)}{(1+\alpha) \sum_{k=1}^{N} u_{ik}^m} \tag{1.23}$$

A inadequacy of (1.22) and (1.23) is that calculating the neighbour term will take considerably more time in each iteration step. In order to reduce the computation, Chen and Zhang [67] proposed a variant of FCM_S, the low-complexity objective function can be written as follows:

$$J_m = \sum_{i=1}^{c} \sum_{k \in N_k} u_{ik}^m \left\| x_k - v_i \right\|^2 + \alpha \sum_{i=1}^{c} \sum_{k=1}^{N} u_{ik}^m \sum_{k \in N_k} \left\| \overline{x_k} - v_i \right\|^2 \tag{1.24}$$

Where $\overline{x_k}$ is a means of neighbouring pixels lying within a window around x_k. An iterative algorithm for minimizing (6) with respect to u_{ik} and v_i can similarly in FCM_S be derived, as:

$$u_{ik} = \frac{\left(\left\| x_k - v_i \right\|^2 + \alpha \left\| \overline{x_k} - v_i \right\|^2 \right)^{-\frac{1}{m-1}}}{\sum_{i=1}^{c} \left(\left\| x_k - v_i \right\|^2 + \alpha \left\| \overline{x_k} - v_i \right\|^2 \right)^{-\frac{1}{m-1}}} \tag{1.25}$$

$$v_i = \frac{\sum_{k=1}^{N} u_{ik}^m \left(x_k + \alpha \overline{x_k} \right)}{(1+\alpha) \sum_{k=1}^{N} u_{ik}^m} \tag{1.26}$$

When α is set to zero, the algorithm is equivalent to the original FCM.

1.3.6 Enhanced Fuzzy C Means Algorithm

L.Szilágyi et al. [68] proposed the En FCM algorithm to expedite the segmentation process for gray level image. In order to expedite FCM_S [69], a linearly-weighted sum image ξ is in advance composed from the pristine image and its local neighbour average image in terms of:

$$\xi_k = \frac{1}{\alpha}\left(x_k + \frac{\alpha}{N_R}\sum_{j\in N_k} x_j \right) \tag{1.27}$$

Where ξ_k denote the gray value of the k[th] pixel of the image ξ, x_j represents the neighbours of x_k, N_k stands for the set of neighbours falling into a window around x_k. Concretely, the objective function used for fast segmenting the newly-generated image ξ is defined as:

$$J_s = \sum_{i=1}^{c}\sum_{l=1}^{q}\gamma_l u_{il}^m (\xi_l - v_i)^2 \tag{1.28}$$

where v_i represents the prototype of the i[th] cluster, u_{il} represents the fuzzy membership of gray value and γ_l is the number of the pixels having the gray value equal to l, where l=1,.., q. Naturally, we have

$$\sum_{i=1}^{q}\gamma_l = N \tag{1.29}$$

And under the constraint that $\sum_{i=1}^{c} u_{il} = 1$ for any l, minimize J_s, specifically, taking the first derivatives of J_s with respect to u_{il} and v_i, and zeroing them, will be obtained as follows

$$u_{il} = \frac{(\xi_l - v_i)^{-\frac{2}{m-1}}}{\sum_{j=1}^{c}(\xi_l - v_i)^{-\frac{2}{m-1}}} \tag{1.30}$$

$$v_i = \frac{\sum_{i=1}^{q}\gamma_l u_{il}^m \xi_l}{\sum_{l=1}^{q}\gamma_l u_{il}^m} \tag{1.31}$$

En FCM provides commensurable segmenting quality in considerably expeditious manner for the encephalon image used there to FCM_S, but the segmenting quality depends on window size , the parameter α and the filtering method.

1.3.7 KNN Algorithm

K Nearest Neighbor [70] is a non-parametric lazy learning algorithm. That is a pretty concise statement. To use KNN [71-72] for classification, we need to have data points for training and also a new unlabeled data for testing.

One of their striking results is to obtain a fairly tight error bound to the Nearest Neighbor rule. The bound is

$$P^* \le P \le P^*\left(2 - \frac{c}{c-1}P^*\right) \tag{1.32}$$

Where P^* is the Bayes error rate, c is the number of classes and P is the error rate of Nearest Neighbor.

In k-Nearest Neighbor Rule, typically the object is classified based on the labels of its *k* nearest neighbours by majority vote. After converting each image to a vector of fixed-length with real numbers, we used the most common distance function for KNN which is Euclidean distance:

$$d(x,y) = \sum_{i}^{k} \sqrt{(x_i - y_i)(x_i - y_i)'} \tag{1.33}$$

Where x and y are histograms in X.

1.3.8 K- Means Algorithm and process of implementation

K-Means[73-74] clustering is a partition-based cluster analysis method. The *k*-means clustering algorithm considered as datasets having "n" data points are partitioned into "k" groups or clusters[75-76].The " k centroids" will modify their regulated areas and will not change their positions anymore [77]. The "objective function" is shown below; and thus, the algorithm aims to reduce the squared error in this function:

$$J_m = \sum_{j=1}^{K} \sum_{i \in S_I} \left\| x_i - c_j \right\|^2 \tag{1.34}$$

The distance indicator of the n data set points is x_i and their represented centroids c_j are $\left\| x_i - c_j \right\|^2$.

The algorithm should run at different times to decrease this impact though it will choose a different set of centroid search time, making it very hard to compare its initial indications [76].

The K-Means [77] Clustering Algorithm starts by picking the number K of centres and randomly assigning the data points x_i to S_i subsets containing N_j data points that minimize the cost function. It then uses a simple re-estimation procedure to end up with a partition of the data points into clusters containing *N* data points that minimize the sum squared clustering function. The clustering process terminates when no more data points switch from one cluster to another based on minimization of the following objective function:

$$J_m = \sum_{j=1}^{K} \sum_{i \in S_I} \left\| x_i - c_j \right\|^2$$

(1.35)

Where $c_j = \frac{1}{N_j} \sum_{i \in s_j} x_i$

1.4 Statement of the Problem

Detection of solid tumors, whether it is benign or malignant, is often difficult in brain MR images. Identifying the exact size and coverage of tumors is also challenging in brain MRI because the original medical image has the problem like noise, low contrast, and bad resolution and so on. So, it takes minutes to get MR images of a subject and more time to view the images generated on a screen and carry out visual assessment.

Visual assessment of the MR images is subjective, often time consuming and hardly repetitive which might give rise to inaccurate diagnosis. Once , MRI shows that there is a tumor in the encephalon the most mundane way to determine the type of encephalon tumor is to optically canvass the result from a sample of tissue after biopsy or surgery. There is no mechanism that detects tumors and classifies the tumors as malignant or benign in MR. This could be avoided if there exist a tool that could be used for accurate detection of abnormalities in the brain automatically by clever analysis of the MR images non-invasively. This calls for the development of methodology which could be used for effective analysis of the MR images that could robustly detect and classify brain tumors.

1.5 Objective

1.5.1 General objective

- The general objective of this proposed thesis is brain tumor detection and classification using discrete orthonormal stockwell transform(DOST) and sine cosine algorithm based deep convolutional neural network(DCNN)

1.5.2 Specific objective

- ✓ To understand feature extraction techniques
- ✓ Develop an algorithm that can extract useful, robust and non-redundant textural features
- ✓ Detect and classify the tumor in the brain MRI image
- ✓ Detect and segment tumors from single parameters MR images.
- ✓ Investigating the application of the Discrete orthonormal Stockwell (S) transform in extraction of useful information from brain MRI images
- ✓ Indicate possible clinical implications

1.6 Significance of the study

The increasing incidence of brain tumors increases the number of images that need to be reviewed by oncologists/radiologists. In addition, particularly in a low resource setting, the high cost of examinations and lack of trained experts hinder patients from receiving the service. In this regard, automating brain tumor detection could deliver many potential benefits. In a screening setting, it allows the examination of a large number of images objectively in less time than current observer driven techniques which are subjective. It can also be an important diagnostic aid and can reduce the workload of trained graders, thereby reducing costs inside clinics So that this study provides help to radiologist, medicos and surgeons in diagnoses of disease in very short time and with the high precision.

1.7 Thesis Organization

Chapter 1 presents a brief background to the research work by describing the different signal processing transform techniques, segmentation and classification of brain tumor research work and its advantages and disadvantages. It includes a critical review on research work reported on different segmentation and classification task followed by statement of the problem, objective of the research and also provides the thesis organization.

Chapter 2 presents literature review of the segmentation techniques and classification techniques of brain tumor form magnetic resonance images.

Chapter 3 presents a recently developed DOST(Discrete orthonormal S-Transform) for image segmentation, sine cosine algorithm(SCA), Deep convolutional neural network(DCNN), and pseudo for the hybrid SCA-DCNN model for classification.

Chapter 4 presents data collection, results and discussion on classification and segmentation techniques. The segmentation results from DOST, S-Transform and Feature extraction are presented. Further the comparison results with different hidden neuron activation functions in the SCA-DCNN model and conventional DCNN Model has been presented.

Chapter 5 presents the conclusion for the entire thesis.

CHAPTER-2

LITERATURE REVIEW ON SEGMENTATION AND CLASSIFICATION

2.1 Introduction

In recent years, major effort has been made to develop MRI applications which can help radiologists in the detection and characterization of malignant and benign abnormalities. The study of cancerous tumor has attracted the attention of many researchers around the world. Many of researches are being published yearly that argue issues related to the brain tumor and the different methods for its early detection and classification.

The detection and classification of the brain tumors have been presented by the researchers through different classifiers such as SVM, PNN, RBFNN etc. and found classification results in terms of accuracy and computational time for the cancerous and noncancerous brain tumors.

2.2 Literature review on FCM based segmentation along with Classification

In [78], the Fuzzy C-Denotes (FCM) segmentation is applied to dissever the tumor and non-tumor region of encephalon and achieved a precision rate of 96.97% in the analysis of DNN predicated encephalon tumor relegation. In [79], a novel bio-physio mechanical tumor growth modeling is presented to analyze the likelihood to tacitly segment tumor-bearing brain images based on atlas based registration. In [80], new multi-fractal (MultiFD) feature extraction and improved AdaBoost classification schemes are used to detect and segment the brain tumor. In [81], local independent projection-based classification (LIPC) method is used to classify the voxel of the brain In [82] , a seeded tumor segmentation method with new Cellular Automata (CA) technique is presented and achieved low accuracy. In [83] multimodal brain tumor segmentation scheme is proposed to achieve high performance than the existing method. In [84], Margo Random Field (MRF) segmentation, deformable model, geometric deformable models are analyzed. In [85], decision Tree, Bagging C predicated wrapper approach is utilized to obtain the decision rules. Additionally, simplify the decision rules by utilizing hybrid feature cull. In [86], the fuzzy predicated control theory is utilized for encephalon tumor segmentation and relegation method and it takes more processing to relegate the tumors. For this purport, methods predicated on machine learning are utilized. Shree and Kumar used Berkeley Wavelet Transformation (BWT) and Support Vector Machine (SVM) to detect mundane and aberrant tissues from MRI

encephalon images [87]. In the study by Kumar and Vijaya kumar the principle component analysis (PCA) and Radial Substructure Function (RBF) are utilized for segmentation and relegation of encephalon tumor and obtained better results [88]. Ullah et al. [89] used ANN to relegate MRI encephalon images by utilizing Haar wavelet and statistical moments. For a homogeneous purport, Muneer and Joseph utilized ANN together with (PCA) to abbreviate the dimensionality of the feature vector [90].

Vidyarthi, A., & Mittal [91] proposed threshold segmentation and morphological operations. To preserve the background and correctly identification of the tumor region. Li et al [92] proposed framework employs local binary patterns (LBPs) to merge results from the classifier ensemble. Moreover, the efficient extreme learning machine with a very simple structure is employed as the classifier. Liu and Liu [93] proposed apperception utilizing gray level co-occurrence matrix (GLCM) in order to efficaciously extract the feature information of human viruses (HV) microscopic images. Parveen and Singh [94] proposed an incipient hybrid technique predicated on the fortification vector machine (SVM) and fuzzy c betokens for encephalon tumor relegation. Amasyali and Ersoy [95] proposed ensemble classifier in order to ameliorate the precision and execution time.

2.3 Literature review on Deep learning based classification

Deep learning is one of the machine learning methods that utilizes neural network [96], image relegation [97], object detection [98, 99] and verbalization apperception [100, 101]. Convolutional neural networks (CNN) is a component of DL having convolution, pooling, or rectified linear unit (ReLU) [102] . Deep learning is a highly efficacious method in solving medical image analysis quandaries including lung cancer diagnosis [103], tibial cartilage segmentation [104] and encephalon tumor detection [105-108]. Utilizing deep learning methods, automatic segmentation has been prosperously performed on astronomically immense magnitudes of MRI images [105-108]. Among more recent solutions to this quandary, Sajid et al. [109] proposed a CNN architecture that takes local and contextual information into account and obtained sensitivity and specificity values for glioma detection are 0.86 and 0.91, respectively with a precision of 95.62%[110]. Zikic et al.[111] utilized a shallow CNN with one plenarily-connected (FC) layer and a softmax layer. Urban et al. [112] evaluated the utilization of 3D

filters to receive more sizably voluminous patches for more sizably voluminous context view over the image [113] In additament to their two-pathway network, Havaei et al. [114] built a cascade of two networks and performed a two-stage training, by training with balanced classes and then refining it with proportions near the originals. Lyksborg et al. [115] utilize a binary CNN to identify the consummate tumor. Then, a cellular automata smooths the segmentation, afore a multi-class CNN discriminates the sub-regions of tumor. Rao et al. [116] extracted patches in each plane of each voxel and trained a CNN in each MRI sequence. Dvorák and Menze [117] divided the encephalon tumor regions segmentation tasks into binary sub-tasks and proposed structured prognostications utilizing a CNN as learning method. In this work, inspired by the groundbreaking work of Simonyan and Zisserman [118] investigated the potential of utilizing deep architectures with minute convolutional kernels for segmentation of gliomas in MRI images. Simonyan and Zisserman [119]. proposed the utilization of minute 3 x3 kernels for the utilization of the intensity normalization

CHAPTER-3

BRAIN TUMOR IMAGE SEGMENTATION AND CLASSIFICATION UTILIZING DISCRETE ORTHONORMAL S-TRANSFORM (DOST) AND SCA-DCNN

3.1 Introduction

This chapter presents the detailed mathematical calculations involved in segmentation based on DOST and SCA-DCNN. The DOST segmentation is proposed as the energy concentration provides more information of internal structure of the image than the other conventional techniques. The SCA algorithm is utilized for the weight optimization of DCNN fully connected layer. The detailed mathematical calculation of SCA and DCNN along with research flow diagram is presented.

3.2 Methodology

The proposed methodology based on discrete orthonormal S-Transform(DOST) segmentation and the Deep CNN learning architecture for classification where the classifier is identifying the encephalon tumors in encephalon MRIs. The proposed methodology for relegating the encephalon tumors in encephalon MRIs is as follows:

Step 1: Brain MRIs Dataset collection

Step 2: Image segmentation using DOST

Step 3: Feature extraction using Discrete Orthonormal S- transform (DOST)

Step 4**:** Classification using SCA-DCNN

Step 5: Comparison Results

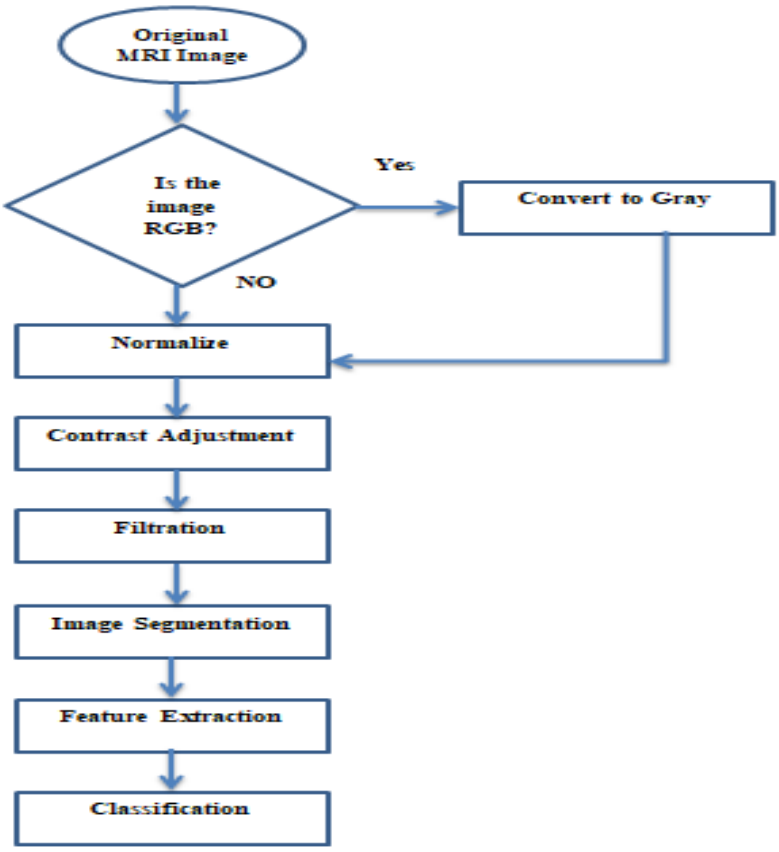

Fig. 3.1 The block diagram of research methodology

3.3 Research flow diagram

The research work is organized with the steps such as (i) The magnetic resonance images has been first amassed and segmented by the novel DOST[120,121] algorithm and the features has been extracted from the images utilizing statistical feature extraction technique . Further (ii) the segmented image has been given as input to the proposed SCA-DCNN for the relegation of encephalon tumors. In the third stage (iii), the weights on fully connected layer are updated utilizing SCA (Sine Cosine Algorithm) algorithm to update the weights of the extreme learning machine model. In the 4th stage (iv), The relegation comparison results from the proposed SCA-DCNN and DCNN is performed.

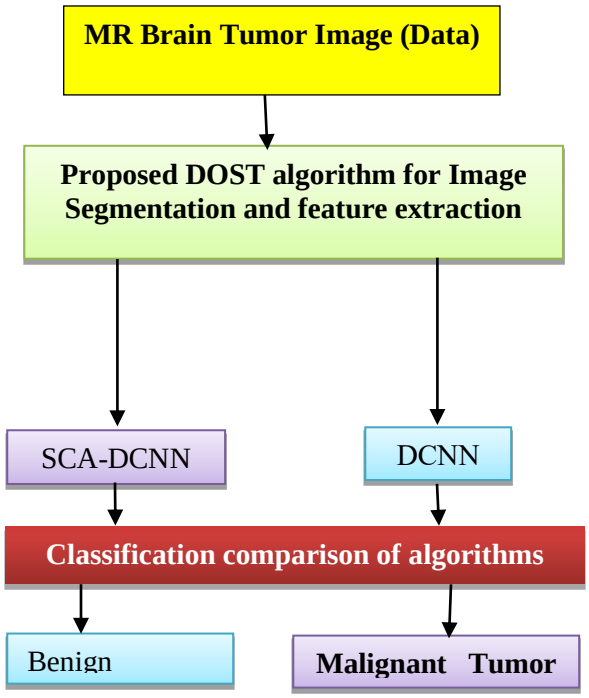

Fig. 3.2 Research work flow diagram

3.4 Detailed analysis of Discrete orthonormal S-Transform

3.4.1 The 2-D Stockwell Transform

For a 2-D continuous-domain function $h(x,y)$, the "2-D S-Transform" is defined with a "2-D Gaussian modulate of cosinusoidal functions" as:

$$S(X,Y,k_x,k_y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)\frac{|k_x||k_y|}{2\pi} e^{-\frac{(X-x)^2 {k_x}^2+(Y-y)^2 {k_y}^2}{2}} e^{-j2\pi(k_x x+k_y y)}dxdy \qquad (3.1)$$

The Gaussian kernel changes shape with respect to spatial frequencies and in the k_x and k_y in the x and y directions, respectively.

Similar to the 2-D Fourier transform, the "2D S-Transform" is a separable transform over different dimensions [121], Therefore, calculation can be pursued first over one dimension and then over the other dimension. Using the 2D convolution theorem in the Fourier space, the 2-D S transform can be written as:

$$S(X,Y,k_x,k_y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(\alpha+k_x,\beta+k_y) e^{-\frac{-2\pi^2\alpha^2}{{k_x}^2}} e^{\frac{-2\pi^2\beta^2}{{k_y}^2}} e^{j2\pi(\alpha X+\beta Y)}d\alpha d\beta \qquad (3.2)$$

Integration of $S(X,Y,k_x,k_y)$ over the variables x and y gives the "2-D Fourier spectrum"

$$H(k_x,k_y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(X,Y,k_x,k_y)dXdY \qquad (3.3)$$

Then the 2-D inverse Fourier transform can be applied to $H(k_x, k_y)$ to recover the original function.

For an input 2-D signal (image), the 2-D Stockwell transform is a complex function of *x*, *y*, k_x and k_y which gives convenience and the freedom to manipulate data over spatial and frequency domains. The transformed data has four-dimensional data set which gives a big challenge in possessing, analyzing and visualizing the coefficients.

Therefore, we only analyze, process, compute or store relevant components of the (X, Y, k_x, k_y) . We will describe one of the strategies to deal with this challenge later in this section.

The discrete 2-D Stockwell coefficients of an image $h(p, q)$can be expressed explicitly as:

$$S(p,q,n,m) = \sum_{n'=0}^{N-1} \sum_{m'=0}^{M-1} H(n'+n, m'+m) e^{-\frac{-2\pi^2 n'^2}{n^2}} e^{\frac{j2\pi n' p}{N}} e^{-\frac{-2\pi^2 m'^2}{m^2}} e^{\frac{j2\pi m' q}{M}} \quad (3.4)$$

where $p = 0,1,\ldots,N-1$ and $q = 0,1,\ldots,M-1$.

It is shown in the continuous S transform that integration of $S(X, Y, k_x, k_y)$ over the variables x and y gives the 2-D Fourier spectrum. Similarly summation of $S(p, q, n, m)$ over the variables p and q gives the 2-D discreet Fourier spectrum.

$$\frac{1}{M}\sum_{q=0}^{M-1} \frac{1}{N}\sum_{p=0}^{N-1} S(p,q,n,m) = H(n,m) \quad (3.5)$$

Where $H(n, m)$ are the discrete 2-D Fourier coefficients.

The original image can be reconstructed using:

$$h(p,q) = \left(\frac{1}{M}\right)^2 \sum_{q'=0}^{M-1} \sum_{m=0}^{M-1} \left(\frac{1}{N}\right)^2 \sum_{p'=0}^{N-1} \sum_{n=0}^{N-1} S(p', q', n, m) e^{\frac{j2\pi np}{N}} e^{\frac{j2\pi mq}{M}} \quad (3.6)$$

3.4.2 Discrete Orthonormal Stockwell Transform

In S-Transform, for a signal of length *N*, there are N^2 Stockwell coefficients and therefore the computing of all N^2 coefficients of the Stockwell transform has computational complexity. If the dimension for the input signal is higher, the computational complexity is also higher. The ST gets exponentially more expensive for higher-dimensional signals [120].

The Discrete Orthonormal Stockwell Transform (DOST) is the best solution to reduce the computational cost without changing its multi resolution nature and the absolutely-referenced frequency and phase information.

The DOST is a pared-down version of the fully redundant Stockwell transform. In the sense of multiresolution, less temporal resolution is required for a lower frequency band based on Nyquist criterion or the sampling theorem [120]. The Nyquist criterion indicates that low frequency band

and high frequency band have very different sampling requirements. Lower frequencies have longer periods and high frequencies have smaller periods. So lower frequencies have lower sampling rates and higher frequencies have higher sampling rates. Hence, the DOST subsamples the low frequencies and takes the advantage of this sample spacing paradigm, and distributes its coefficients accordingly for higher frequencies. It does so by constructing a set of N orthogonal unit-length basis vectors, each of which targets a particular region in the time-frequency domain. Which region is dictated by a set of parameters: v specifies the center of each frequency band (voice), β is the width of that band, and τ specifies the location in time. Using these parameters, the *kth* basis vector is defined as.

$$D[k][v,\beta,\tau]=\frac{1}{\sqrt{\beta}}\sum_{f=v-\beta/2}^{v+\beta/2-1}\left(e^{-i2\pi\frac{k}{N}f}\right)\left(e^{-i2\pi\frac{\tau}{\beta}f}\right)\left(e^{-i2\pi\tau}\right) \quad (3.7)$$

for k=0, 1,……., N-1, which can be summed analytically to:

$$D[k][v,\beta,\tau]=i\pi\tau\frac{e^{-i2\alpha\left(v-\beta/2-1/2\right)}-e^{-i2\alpha\left(v+\beta/2-1/2\right)}}{2\sqrt{\beta}sin\alpha} \quad (3.8)$$

Where $\alpha=\pi\left(\frac{k}{N}-\frac{\tau}{\beta}\right)$ can be regarded the center of the temporal window.

3.5 Feature Extraction

A significant part of the medical image processing is feature extraction, since it exposes the mathematical properties of various areas. The main objective of the feature is to reduce the data involved in the calculation, as it already contains the relevant details of the image. In this understanding, there are a number of classic features derived from medical picture in extended use. So, this step is to find a feature set of brain tumor discern non-tumor from tumor or benign from malignant. Features are extracted from the segmented images to distinguish them into benign and malignant ones.

If we assume $(i,)$ is the $(i,)'$ th entry of size $W=m\times n$ MRI image matrix, and then the statistical measure image becomes $I(x,y)$.

Mean: Shows average intensity level of an image..

$$I(x,y)=\frac{1}{mn}\sum_{(i,j)\in W}g(i,j) \quad (3.9)$$

Variance: Utilized to find how each pixel varies from the neighboring pixel and measures how far a set of numbers is spread out.

$$I(x,y) = \frac{1}{mn-1} \sum_{(i,j)\in W} \left(g(i,j) - \frac{1}{mn-1} \sum_{(i,j)\in W} g(i,j) \right)^2 \tag{3.10}$$

Standard deviation: Used to characterize the amount of 2D dispersion around the mean.

$$I(x,y) = \sqrt{\frac{1}{mn-1} \sum_{(i,j)\in W} \left(g(i,j) - \frac{1}{mn-1} \sum_{(i,j)\in W} g(i,j) \right)^2} \tag{3.11}$$

Skewness: the symmetry property of the image

Kurtosis: the flatness of the histogram:

Entropy: Measures the uniformity/variability of an image histogram.

$$I(x,y) = -\sum_{(i,j)\in W} g(i,j)\big(\log\big(g(i,j)\big)\big) \tag{3.12}$$

Correlation: Correlation is measurement of the similarity between two signals/sequences

$$g_x(i) = \textstyle\sum_{j=1}^{n} g(i,j) \tag{3.13}$$

$$g_y(j) = \sum_{i=1}^{n} g(i,j) \tag{3.14}$$

Where n is the number of distinct gray levels of the image

$$I(x,y) = \frac{\sum_{(i,j)\in W}(i \times j)g(i,j) - \mu_x\mu_y}{\sigma_x\sigma_y} \tag{3.15}$$

Where $\mu_x, \mu_y, \sigma_x \; and \; \sigma_y$ are the means and standard deviations of $g_x \; and \; g_y$

Contrast: Contrast is the difference in luminance or color that makes an object distinguishable from other objects within the same field of view.

$$I(x,y) = \sum_{k=0}^{L-1} k^2 \times \sum_{(i,j)\in W} \delta(|i-j|,k)g(i,j) \tag{3.16}$$

Where L is the number of distinct gray levels of the image and

$$\delta(x,y) = \begin{cases} 1, & x = y \\ 0, & x \neq y \end{cases}$$

Energy: The quantifiable amount of the extent of pixel pair repetition, this parameter is used for measuring the similarity of the image.

Inverse Difference Moment (IDM): The measurement of local homogeneity of an image and it could be used to detect whether the image is textured or not.

3.6 Deep Convolutional Neural Network

3.6.1 Convolution Neural Network (CNN) Basic Structure

Convolutional neural network is similar to artificial neural network, as both of them are made up of self-optimized neurons, which are imported by inputs and perform a nonlinear transformation [121]. Compared with the artificial neural network, convolution neural network is widely used in pattern recognition on images, for it encodes image specific features into the network architecture, making the network more suitable for image-based feature learning [122].

There are five basic elements within the convolution neural network: input layer, convolutional layer, non-linear layer, pooling layer and fully connected layer.

3.6.1.1 Convolution Layer

The convolution layer is the foremost part of "convolutional neural network (CNN)", which does the heavy calculation of convolution neural network operation. The convolution neural network extracts different features of the inputs and the convolution layer extracts low-level features of the image, such as edges, lines, and corners. The key points of the convolution layers are the usage of learnable kernels. The kernels are generally small in spatial dimensionality, but they can spread along the entire input. When an input comes into a convolutional layer, the convolution layer convolves each filter across the input and then it produces a 2 dimensional activation map.

3.6.1.2 Non-linear Layer

The convolution neural network applies a non-linear transformation on the input, whose purpose is to identify the features within each hidden layer.

3.6.1.3 Pooling Layer

The major purpose of the pooling layer is also to reduce the dimensionality, the number of the parameters as well as the computational complexity. Besides, it helps to make the features robust against noise and distortion.

3.6.1.4 Fully Connected Layer

After several repeats of the previous layers, the data comes to the final layer of the convolution neural network, which is the fully connected layer. Within the fully connected layer, the neurons

are directly connected to the neurons in the two adjacent layers. The aim of this layer is to sum up the weights of the features coming from the previous layers and indicates the probability of each class. Convolution is performed on the artificial neural networks (ANN) that gives us the Convolutional neural network [123]. Neurons having weights and biases, which can be learned, forms a CNN. Three main layers utilized in building up the convolutional neural network architecture are- Convolutional layer, pooling layer, and plenarily connected layer. A convolutional neural network (CNN) is so denominated as it contains one or more convolutional layers. In the input images, certain local features are detected through convolutional layers [124]. A pooling layer is the next layer after convolution layer in CNN whose main purport is decrementing the size of the representation spatially, which avails in controlling the over fitting [125]. A number of plenarily connected (FC) layers will follow other layers as mentioned. Onclusively, CNN contains a Softmax layer that engenders the desired outputs.

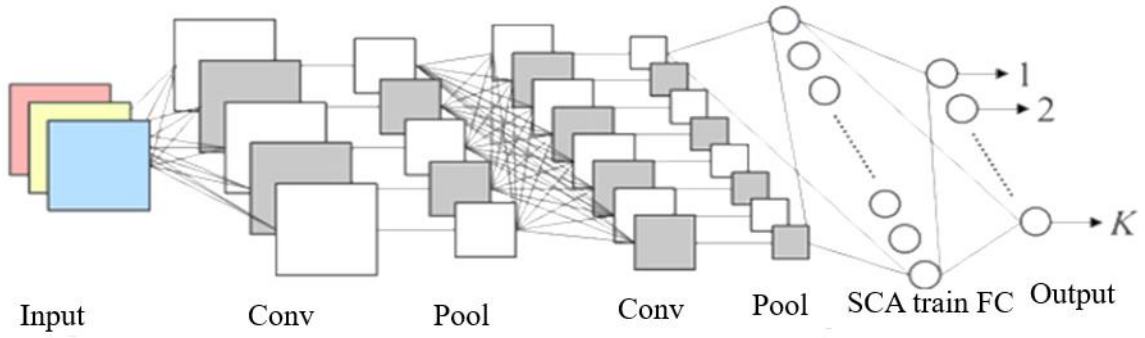

Fig. 3.3 Architecture of DCNN [123]

3.7 SINE COSINE ALGORITHM

SCA algorithm is a population-based optimization algorithm using a mathematical model based on sine and cosine functions. The performance metrics are used to qualitatively to observe and confirm the performance of SCA. The SCA algorithm can highly be effective in solving real problems with constrained and unknown search spaces.

This research work proposes a novel population-based optimization algorithm called Sine Cosine Algorithm (SCA) for solving the weight optimization of Deep CNN.

According to sine cosine algorithm the position equation is updated as,

$$X_i^{n+1} = \begin{cases} X_i^n + \alpha_1 \times \sin(\alpha_2) \times \left| \alpha_3 p^{gbest} - X_i^n \right|, & \alpha_4 < 0.5 \\ X_i^n + \alpha_1 \times \cos(\alpha_2) \times \left| \alpha_3 p^{gbest} - X_i^n \right|, & \alpha_4 \geq 0.5 \end{cases} \quad (3.17)$$

Where α_1, α_2, α_3, α_4 are the random variables and α_1 is given by

$$\alpha_1 = a\left(1 - \frac{n}{K}\right) \quad (3.18)$$

Where n is the current iteration, K is the maximum number of iterations.

The current position X_i^n and X_i^{n+1} updated position has been mentioned in eqn(3.17). The parameter α_1 determines the next position regions, α_2 evaluates the direction of movement from $x_i(n)$ .The parameter α_3 controls the current movement, and the parameter α_4 equally changes between the sine and cosine functions.

To have faster speed of convergence of the parameter α_1 is improved as

$$\alpha_m = \exp\left(\frac{1}{1+(\alpha_1)}\right) \quad (3.19)$$

And the corresponding updated position equation is given as

$$X_{ij}^{n+1} = \begin{cases} X_{ij}^n + \alpha_m \times \sin(\alpha_2) \times \left| \alpha_3 p^{gbest} - X_i^n \right|, & \alpha_4 < 0.5 \\ X_{ij}^n + \alpha_{m1} \times \cos(\alpha_2) \times \left| \alpha_3 p^{gbest} - X_i^n \right|, & \alpha_4 \geq 0.5 \end{cases} \quad (3.20)$$

Further the weights are mentioned as $W = [w_{i0} + w_{i1}x_1 + ...w_{iN}x_N]$ and the weigthts are mapped and updated using

$$W_{ij}^{n+1} = \begin{cases} W_{ij}^n + \alpha_m \times \sin(\alpha_2) \times \left| \alpha_3 p^{gbest} - W_i^n \right|, & \alpha_4 < 0.5 \\ W_{ij}^n + \alpha_{m1} \times \cos(\alpha_2) \times \left| \alpha_3 p^{gbest} - W_i^n \right|, & \alpha_4 \geq 0.5 \end{cases} \quad (3.21)$$

3.7.1 Pseudo code of the SCA Algorithm for DCNN

Pseudo code: SCA for weight optimization of DCNN Model.

1. Initialize the parameters of SCA parameters $\alpha_1, \alpha_2, \alpha_3, \alpha_m, K$
2. Initialize the position equation and map with the weights
3.Initialize the weights as W_{ij}^n
4.%starting of loop

for l=1:K

update(W_{ij}^n) as

$$W_{ij}^{n+1} = \begin{cases} W_{ij}^n + \alpha_m \times \sin(\alpha_2) \times \left| \alpha_3 p^{gbest} - W_i^n \right|, & \alpha_4 < 0.5 \\ W_{ij}^n + \alpha_{m1} \times \cos(\alpha_2) \times \left| \alpha_3 p^{gbest} - W_i^n \right|, & \alpha_4 \geq 0.5 \end{cases}$$

End

update W_{ij}^{n+1} to obtain minimum weight values

end of for loop

6.Stopping criteria: Continue till optimization gets minimum error values

7. If not converges, repeat until nearly zero error satisfied.

CHAPTER-4

RESULTS AND DISCUSSION

4.1 Introduction

This chapter will present the results of the work in this thesis. Different experiments will be carried out under different parameters and image processing characteristics. The results will be tabulated, discussed, and evaluated in this chapter.

4.2 Data set

The brain tumor data set was taken from "Harvard Medical School website (http://med. harvard.edu/AANLIB/)". The proposed methodology was implemented on a MATLAB R2018a platform. A total of 150 MRI images with 70 benign and 80 malignant tumors were used. Unfortunately, the dataset didn't include control subject (normal cases).

4.3 Pre-processing Results

It is a technique used to enhance image data by enhancing certain features that are necessary for further processing. Which include techniques for eliminating noise, normalizing the intensities, eliminating reflections on the image, and masking portions of the image. However, the only way to prevent these unwanted image features is to obtain a high-quality image. Nevertheless, in reality, digital images are error prone during the various stages of image processing that result in varying pixel values that do not represent the true intensities of the real scene.

Pre-processing is very helpful in a number of these circumstances because it helps to remove irrelevant details to a particular study Many widely used image pre-processing tools include image enhancement, cropping, filtering and restoration. Image enhancement enhances the visual quality of an image and offers much clearer sources for more image processing.

The methods used for enhancing image quality include density slicing and contrast stretching. Image cropping focuses the region of interest and removes the irrelevant parts of the image. Filtering either emphasizes or eliminates certain features and includes smoothing, sharpening, and edge enhancement. Filtering could be performed in spatial domain or in a transformed domain (such as frequency domain)

Mostly used filters include mean filter, Gaussian and Wiener filters. Image filtering has numerous applications in biomedical imaging and image processing. Restoration is different from the other pre-processing applications mentioned above because it is more objective whereas

the others are subjective. Restoration techniques aim to restore images corrupted by some form of degradation (a good example is motion degradation).

Restoration methods include simulation of the degradation process (mostly using convolutions) at the beginning and application of the reverse process (using such as de-convolution) at the end.

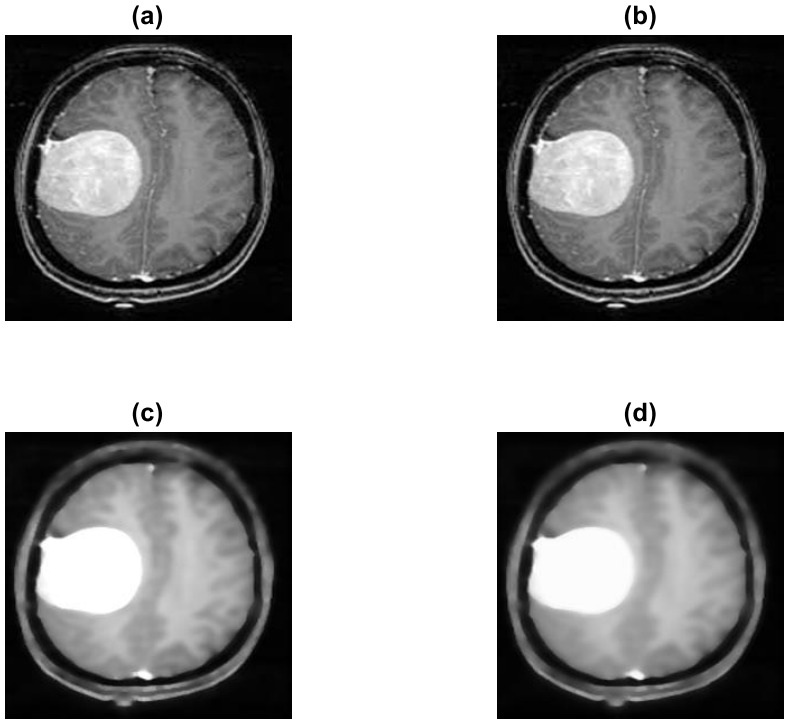

Fig. 4.1 Preprocessing result -(a) Original image (b) Normalized gray image (c) Contrast enhanced (d) Filtered

4.4 Image segmentation results

Convolution is performed on the artificial neural networks (ANN) that gives us the Convolutional neural network [123]. Neurons having weights and biases, which can be learned, forms a CNN. Three main layers utilized in building up the Image segmentation is the non-frivolous task of dissevering the different mundane encephalon in encephalon MR images [126]. In this thesis, we used discrete orthonormal s transform technique to segment the image into 5 sections as it had good results in our precedent work and additionally for comparison purposes Fig. 4.7 shows the results of segmenting a sample image utilizing DOST.

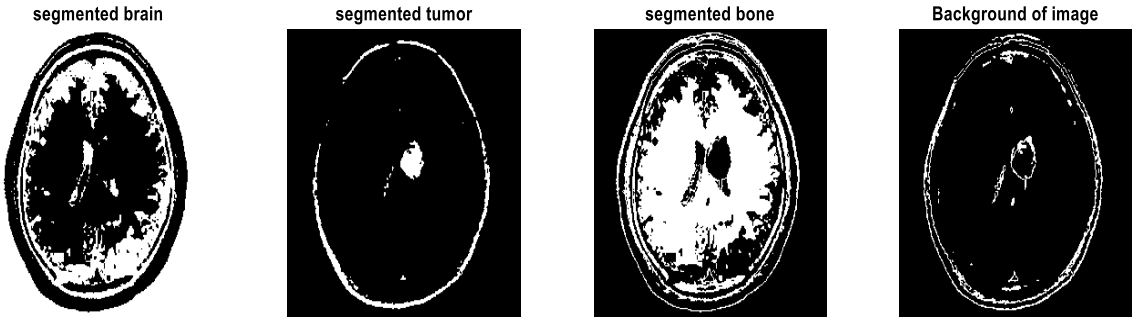

Fig. 4.2 Segmentation by using K Means Algorithm

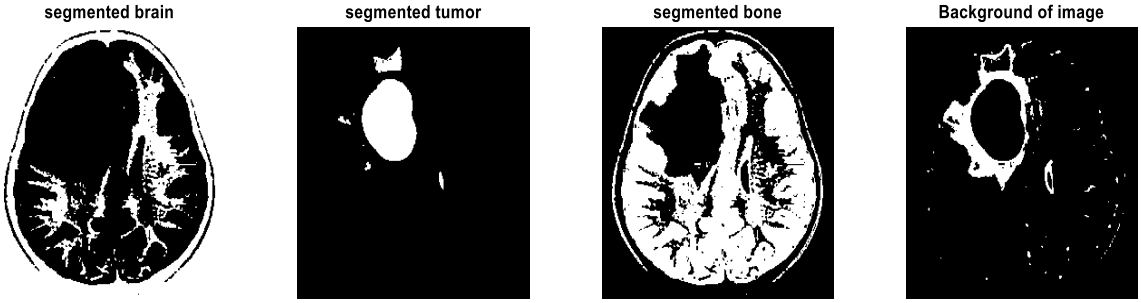

Fig. 4.3: Segmentation by using Fuzzy C Means Algorithm

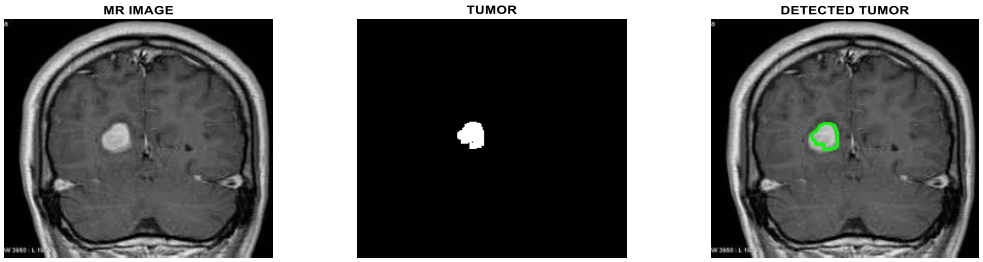

Fig. 4.4 Image segmentation using Fast Fuzzy C Means

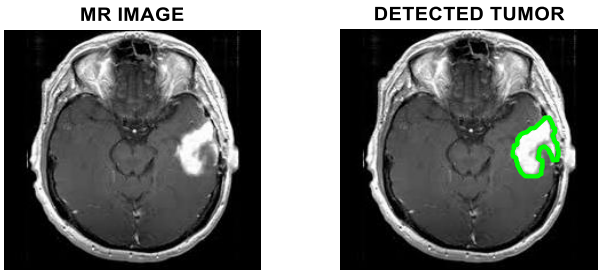

Fig. 4.5 Segmentation by using KNN Algorithm

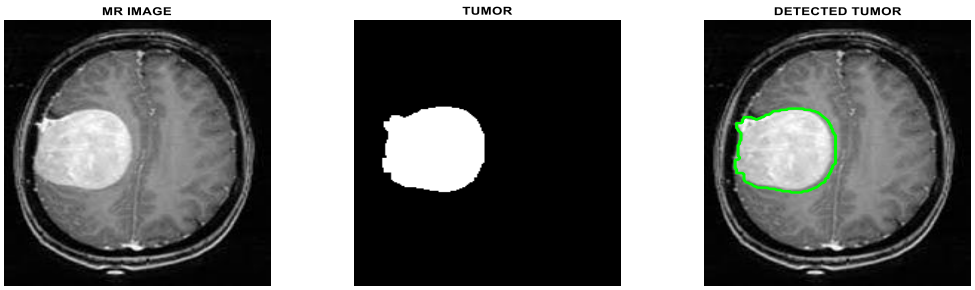

Fig. 4.6 Image segmentation using wavelet transform

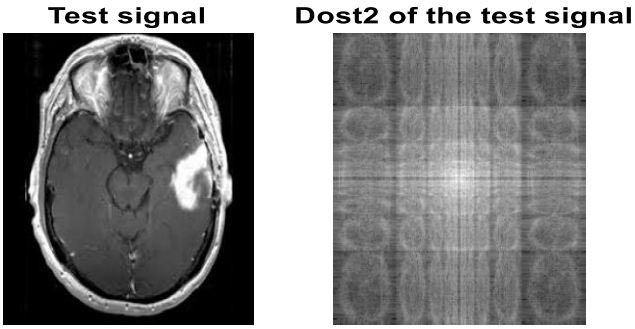

Fig. 4.7 MRI image and Dost of MRI image

The above figure shows the density of the tumor in hazed white concentration where the effect of the disease is minute. The wavelet transform is showing the location of the tumor, but could not able to perform the concentration of energy. Even though the computational time is slightly more In DOSI, it is preferable for the purpose of application in segmentation process.

Table 4.1 : Segmentation Accuracy and computational time

Model	Computational time	Accuracy
Wavelet Transform	11. 214534	97.93
DOST	12.213146	98.12
KNN	37.324524	91.31
K-MEANS	29.214433	85.23
Fuzzy C Means	24.233232	93.56
Fast Fuzzy C Means	21.123242	95.82

4.5 Statistical Feature extraction

After segmenting the Encephalon MRI into different sections features of the segmented tumor is extracted utilizing DOST. DOST have the advantage of extracting the most germane features at different directions and scales as they provide localized time-frequency information of MR image.

Table 4.2: Feature extraction

Feature	Benign Tumor	Malignant Tumor
Contrast	46.2267	50.999
Correlation	8.1281	8.8093
Energy	1.1144	1.2154
Homogeneity	7.3827	8.0726
Mean	0.0024	0.0031
Standard Deviation	0.6875	0.75
Entropy	34.8699	38.4521
RMS	0.6875	0.75
Variance	0.0429	0.0468
Smoothness	9.1866	9.6425
Kurtosis	92.3887	96.643
Skewness	2.7475	2.7801
IDM	1.6466	1.1015

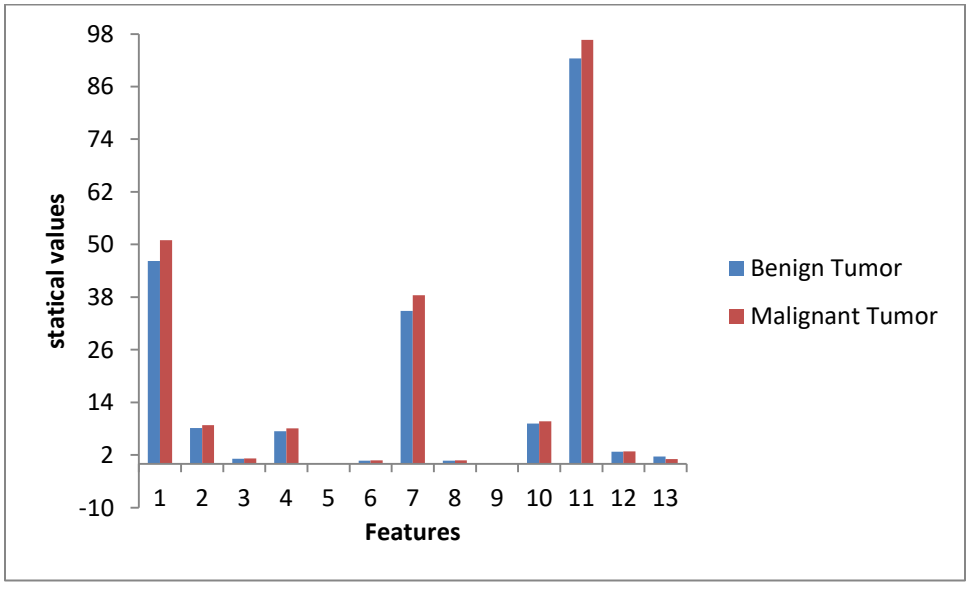

Fig. 4.8 Feature extraction

4.6 Classification Results

After the features are extracted and selected, the classification step using DNN is performed on the resulted feature vector. Classification is performed by using SCA optimization technique for building the DCNN of hidden layers structure. Also for evaluating the performance of the selected classifier, SCA has been employed for optimizing the weights of DCNN.

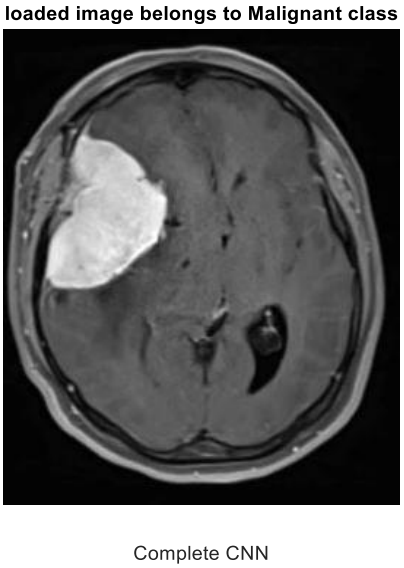

Fig. 4.9 Malignant

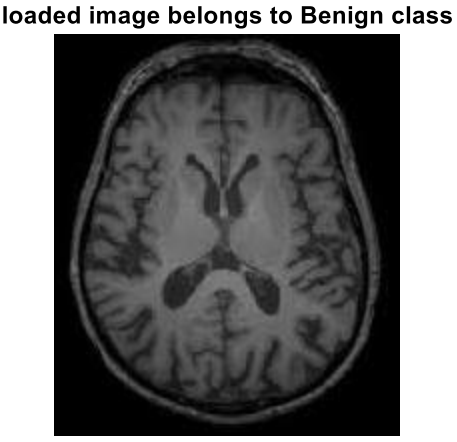

Fig. 4.10 Benign

4.7 Sine Cosine Algorithm

While considering Tanh kernel function, the error settles between 0.4 and 0.6, But the Tansig function settles between 0.2 and 0.4 but takes more time for convergence. The proposed softmax and logsig function converges zero at less time as shown in fig 4.11.

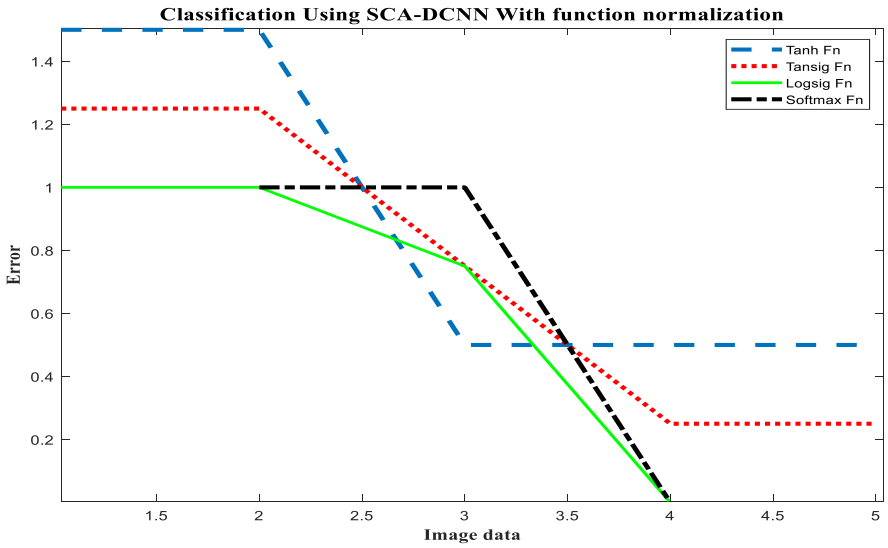

Fig 4.11 SCA-DCNN with function normalization

While considering Tansig kernel function, the error settles between 0.2 and 0.4, But the Tanh function settles at zero but takes more time for convergence. The proposed softmax and logsig function converges at less time as shown in fig 4.12.

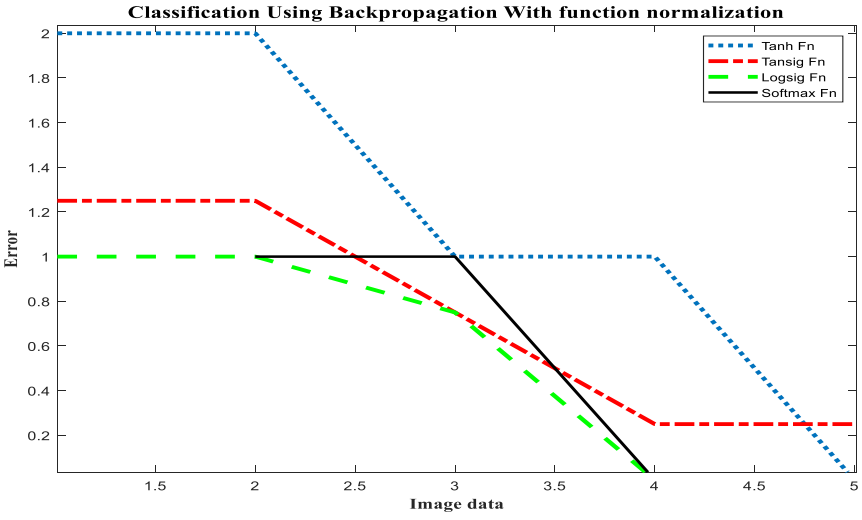

Fig. 4.12 Backpropagtion with function normalization.

While considering tanh kernel function, the error settles at 0.5, But the Logsig function settles at zero but takes more time for convergence. The proposed softmax function takes less time for convergence as shown in fig 4.13.

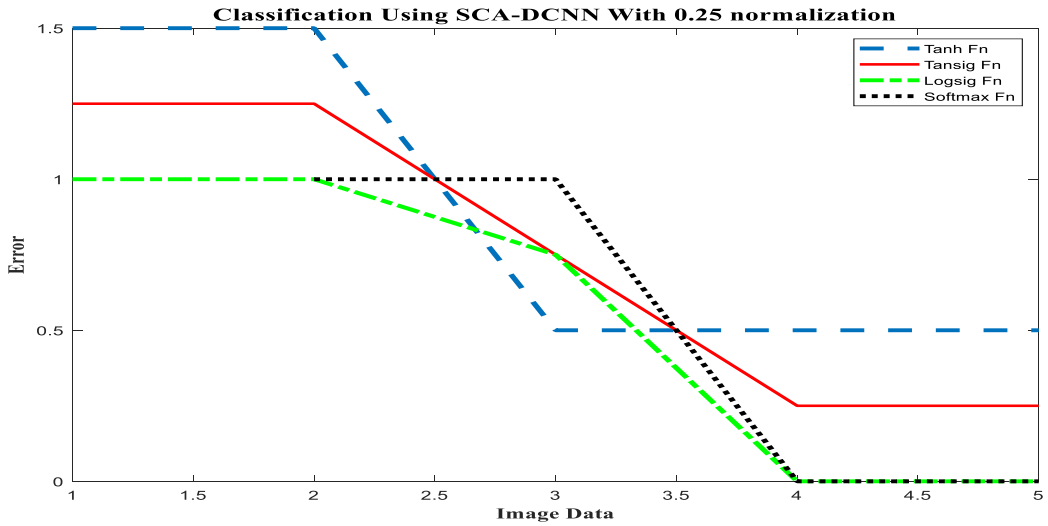

Fig.4.13 SCA-DCNN with function 0.25 normalization

The softmax function settles at zero but takes more time for convergence. The proposed logsig function takes less time for convergence as shown in fig 4.14.

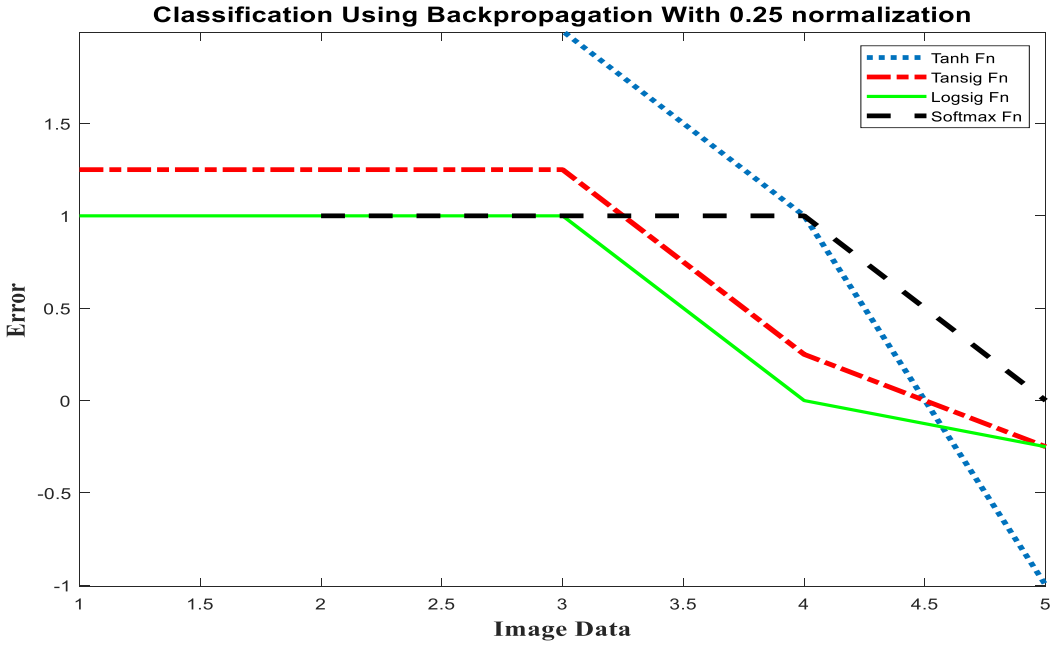

Fig. 4.14 Backpropagation with function 0.25normalization

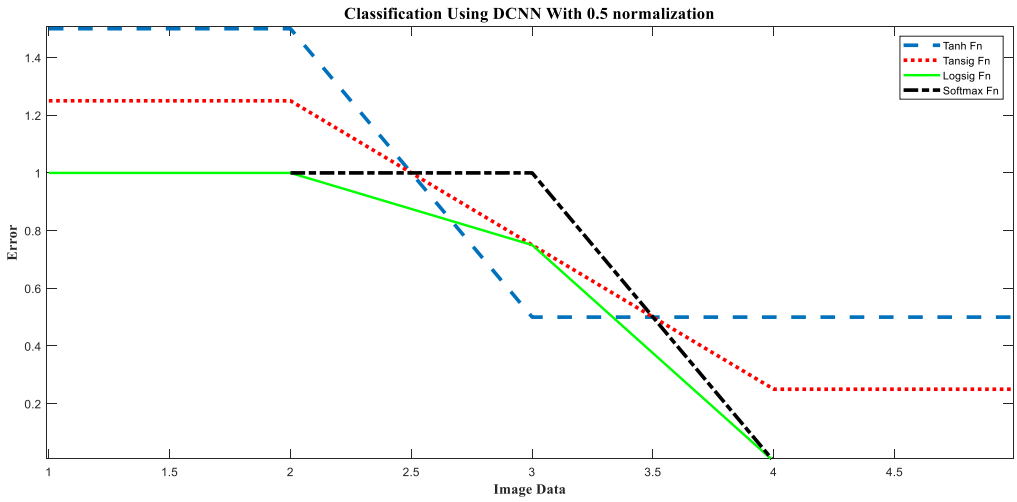

Fig. 4.15 DCNN with function 0.5 normalization

While considering Tanh kernel function, the error settles between 0.4 and 0.6, But the Logsig function settles at zero but takes more time for convergence. The proposed softmax function takes less time for convergence as shown in fig 4.15

Table 4.3 Performance evaluation

	Elapsed time (seconds)		**Accuracy in %**	
Activation Function	**DCNN+SCA**	**DCNN+ Back propagation**	**DCNN+SCA**	**DCNN+ Back propagation**
Softmax	2.0224	2.1639	98.28	97.56
tanh	2.2800	2.9633	91.12	90.17
tansig	2.0072	2.1960	95.40	92.16
logsig	2.0677	2.1308	96.37	93.15

The proposed DCNN with SCA model has been assigned for the classification and the results were compared with the conventional logsig, tanh, softmax and tansig functions.

From the table 4.3 we see using logsig, tanh, softmax and tansig functions the computational time achieved as 2.0677, 2.2800, 2.0224 and 2.072 respectively. But the softmax function in SCA-DCCN achieved 2.0224 sec, which shows the robustness of the proposed SCA-DCCN model.

CHAPTER-5

CONCLUSION AND RECOMMENDATION

5.1 Conclusion

The literature survey has been done with more than 40 research papers. From the literature survey there are several methods has been proposed for classification of brain tumor and the computational time and accuracy has been presented. In this thesis we have proposed improved and novel biologically inspired classifiers for automatic detection of brain tumors and results has been compared with other already proposed methods.

The first and second chapter presents the literature survey and signal processing algorithms. In the third chapter, the automatic detection and classification using the proposed DCNN model with SCA training on fully connected layer is proposed. Images are segmented and features are extracted using DOST algorithm at the first step. There are 13 features have been considered for the classification task. The feature such as contrast, correlation, energy, Homogeneity, Mean, standard deviation, entropy, rms, variance, smoothness, kurtosis and idm are considered for the classification task.

The proposed DCNN with SCA model has been assigned for the classification and the results were compared with the conventional logsig, tanh, softmax and tansig functions.

By using logsig, tanh, softmax and tansig functions the computational time achieved as 2.0677, 2.2800, 2.0224 and 2.072, respectively. But the softmax function in SCA-DCCN achieved 2.0224 sec, which shows the robustness of the proposed SCA-DCCN model. Further classification accuracy of 98.12 % achieved with SCA DCNN with softmax function, approach. From the result, it is found that the proposed SCA-DCNN model provides better classification result and the computations time obtained as less as compared to other conventional methods

5.2 Recommendation.

This thesis works on the 2-dimensional MRI images, there is huge volume of data encapsulated in the 3-dimensional images plus the video, so there is a need for applying deep learning on these kinds of data. However, this will restrict the use of the network as the depth of the MRI data will change.Moreover, in this thesis, the number of sample images taken for training and testing was restricted. Incorporating more samples could boost the classification efficiency. These issues await further investigations.

Acknowledgment

This research thesis was founded by Adama science and Technology University under the grant number of ASTU/SM-R/.Adama, Ethiopia.

REFERENCES

1. Cancer.Net Editorial Board. (2019). Brain Tumor: Statistics. Retrieved from https://www.cancer.net/cancer-types/braintumor/statistics
2. R.Kavitha, L.Chitra, and R.kanaga,"Brain tumor segmentation using genetic algorithm with SVM classifier," Int J Adv Res Electr Electron Instrum Eng 2016; 5(3):1468-1471.
3. T.Logeswari, M.Karnan,"An improved implementation of brain tumor detection using segmentation based on hierarchical self-organizing map," Int J Comput Theory Eng 2010; 2(4):591-595.
4. G.Khambhata Kruti, R.Panchal Sandip," Multiclass classification of brain tumor in MR images," Int J Innov Res Comput Commun Eng 2016;4(5): 8982-8992.
5. G.Kaur, and J.Rani, "MRI brain tumor segmentation methods-a review,". Int J Comput Eng Technol (IJCET) 2016;6(3):760-764.
6. V.Das, and J.Rajan, "Techniques for MRI brain tumor detection," a survey. Int J Res Comput Appl Inf Tech 2016; 4(3):53-56.
7. EL.Zacharaki, S.Wang, S.Chawla, D.Soo Yoo, R.Wolf, and E.Melhem, "Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme,"Magn Reson Med 2009; 62:1609-1618.
8. G.Litjens, T.Kooi, B.Bejnordi, A.Setio, F.Ciompi, M.Ghafoorian," A survey on deep learning in medical image analysis," Med Image Anal 2017; 42:60-88.
9. L.Singh, G.Chetty, and D.Sharma," A novel machine learning approach for detecting the brain abnormalities from MRI structural images," In: IAPR international conference on pattern recognition in bioinformatics. Berlin Heidelberg: Springer; 2012. p. 94-105
10. H.Gittleman,C.Kromer,Q.Ostrom ,R.Blanda ,J.Russell, C.Kruchko, and J.Barnholtz-Sloan,"Is mortality due to primary malignant brain and other central nervous system tumors decreasing?," J Neurooncol 2017; 133: 265-275
11. S.Bab, E.Abdifard, S.Elyasianfar, P.Mohammadi, E.Mohammadi, N.Izadi, and M.Heidari,"Trend of the incidence of brain cancer in Iran and it's 6 geographical regions during 2000-2005. Pharmacophore 2018; 9: 41-52.
12. Statistics adapted from the American Cancer Society's publication, Cancer Facts & Figures 2020, the ACS website (January 2020), the CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2012-2016 (January 2020), and the National Cancer Institute website (January 2020).

13. E.Maile, I.Barnes, A.Finlayson, S.Sayeed, R.Ali,"Nervous system and intracranial tumor incidence by ethnicity in England", 2001-2007: a descriptive epidemiological study. PLoS One 2016; 11: e0154347.
14. F.Khodamoradi, M.Ghoncheh, R.Pakzad, H.Gandomani, H.Salehiniya ,"The incidence and mortality of brain and central nervous system cancer and their relationship with human development index in the world," WCRJ 2017; 4: e985.
15. M.Alfonse, and A.Salem, "An automatic classification of brain tumors through MRI using support vector machine, "Egyptian Computer Science Journal, 40(3), 11-21,2016
16. R.Remya, G.K.Parimala, and S.Sundaravadivelu,"Enhanced DWT Filtering Technique for Brain Tumor Detection,"IETE Journal of Research, 1-10,2019
17. M.Wu, C.Lin, and C.Chang," Brain Tumor Detection Using Color-Based K-Means Clustering Segmentation," Third International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP 2007), 2, 245- 250,2007
18. P.Natarajan, N.Krishnan,N.Kenkre,S.Nancy, and B.Singh,"Tumor detection using threshold operation in MRI brain images," IEEE International Conference on Computational Intelligence and Computing Research, 1-4,2012
19. A. Mustaqeem, A.Javed, & T.Fatima, "An Efficient Brain Tumor Detection Algorithm Using Watershed & Thresholding Based Segmentation," International Journal of Image, Graphics and Signal Processing, 4(10), 34-39,2012
20. B. Deepa, and M. G, Sumithra,"An intensity factorized thresholding based segmentation technique with gradient discrete wavelet fusion for diagnosing stroke and tumor in brain MRI. Multidimensional Systems and Signal Processing", 30(4), 2081-2112. ,2019
21. J. Amin, M. Sharif, M. Raza, and M. Yasmin, "Detection of Brain Tumor based on Features Fusion and Machine Learning," Journal of Ambient Intelligence and Humanized Computing, pp. 1-17, 2018.
22. J. Amin, M. Sharif, M. Raza, T. Saba, and M. A. Anjum, "Brain tumor detection using statistical and machine learning method," Computer methods and programs in biomedicine, vol. 177, pp. 69-79, 2019.
23. J. Liu, Y. Pan, M. Li, Z. Chen, L. Tang, C. Lu, and J. Wang, "Applications of deep learning to MRI images: A survey," Big Data Mining and Analytics, vol. 1, pp. 1-18, 2018
24. T. Saba, S. T. F. Bokhari, M. Sharif, M. Yasmin, and M. Raza, "Fundus image classification methods for the detection of glaucoma: A review," Microscopy research and technique, vol. 81, pp. 1105-1121, 2018

25. S. M. Naqi, M. Sharif, and M. Yasmin, "Multistage segmentation model and SVM-ensemble for precise lung nodule detection," International journal of computer assisted radiology and surgery, vol. 13, pp. 1083-1095, 2018.
26. A. Liaqat, M. A. Khan, J. H. Shah, M. Sharif, M. Yasmin, and S. L. Fernandes, "Automated ulcer and bleeding classification from WCE images using multiple features fusion and selection," Journal of Mechanics in Medicine and Biology, vol. 18, p. 1850038, 2018.
27. M. Sharif, M. A. Khan, M. Faisal, M. Yasmin, and S. L. Fernandes, "A framework for offline signature verification system: Best features selection approach," Pattern Recognition Letters, 2018.
28. J. H. Shah, M. Sharif, M. Yasmin, and S. L. Fernandes, "Facial expressions classification and false label reduction using LDA and threefold SVM," Pattern Recognition Letters, 2017
29. M. Yasmin, M. Sharif, and S. Mohsin, "Neural networks in medical imaging applications: A survey," World Applied Sciences Journal, vol. 22, pp. 85-96, 2013.
30. M. A. Khan, T. Akram, M. Sharif, M. Y. Javed, N. Muhammad, and M. Yasmin, "An implementation of optimized framework for action classification using multilayers neural network on selected fused features," Pattern Analysis and Applications, vol. 22, pp. 1377-1397, 2019
31. M. Fayyaz, M. Yasmin, M. Sharif, J. H. Shah, M. Raza, and T. Iqbal, "Person re-identification with features-based clustering and deep features," Neural Computing and Applications, vol.17.pp. 1-22, 2019.
32. S. M. Naqi, M. Sharif, and A. Jaffar, "Lung nodule detection and classification based on geometric fit in parametric form and deep learning," Neural Computing and Applications, pp. 1-19, 2018.
33. M. Sharif, U. Tanvir, E. U. Munir, M. A. Khan, and M. Yasmin, "Brain tumor segmentation and classification by improved binomial thresholding and multi-features selection," Journal of Ambient Intelligence and Humanized Computing, pp. 1-20, 2018
34. J. Amin, M. Sharif, M. Yasmin, and S. L. Fernandes, "A distinctive approach in brain tumor detection and classification using MRI," Pattern Recognition Letters, 2017
35. S. Sajjad, A. Hanan Abdullah, M. Sharif, and S. Mohsin, "Psychotherapy through video game to target illness related problematic behaviors of children with brain tumor," Current Medical Imaging Reviews, vol. 10, pp. 62-72, 2014.
36. J. Amin, M. Sharif, M. Yasmin, and S. L. Fernandes, "A distinctive approach in brain tumor detection and classification using MRI," Pattern Recognition Letters, 2017.

37. V. Rajinikanth, S. C. Satapathy, S. L. Fernandes, and S. Nachiappan, "Entropy based segmentation of tumor from brain MR images–a study with teaching learning based optimization," Pattern Recognition Letters, vol. 94, pp. 87-95, 2017.
38. V. Rajinikanth, S. L. Fernandes, B. Bhushan, and N. R. Sunder, "Segmentation and analysis of brain tumor using Tsallis entropy and regularised level set," in Proceedings of 2nd International Conference on Micro-Electronics, Electromagnetics and Telecommunications, 2018, pp. 313-321.
39. L. M. DeAngelis, "Brain tumors," New England Journal of Medicine, vol. 344, pp. 114-123, 2001.
40. J. Naik and S. Patel, "Tumor detection and classification using decision tree in brain MRI," International Journal of Computer Science and Network Security (IJCSNS), vol. 14, p. 87, 2014.
41. U. R. Acharya, S. L. Fernandes, J. E. WeiKoh, E. J. Ciaccio, M. K. M. Fabell, U. J. Tanik, V. Rajinikanth, and C. H. Yeong, "Automated Detection of Alzheimer's Disease Using Brain MRI Images–A Study with Various Feature Extraction Techniques," Journal of Medical Systems, vol. 43, p. 302, 2019.
42. S. L. Fernandes, U. J. Tanik, V. Rajinikanth, and K. A. Karthik, "A reliable framework for accurate brain image examination and treatment planning based on early diagnosis support for clinicians," Neural Computing and Applications, pp. 1-12, 2019.
43. R. A. Gatenby, O. Grove, and R. J. Gillies, "Quantitative imaging in cancer evolution and ecology," Radiology, vol. 269, pp. 8-14, 2013.
44. J. Amin, M. Sharif, M. A. Anjum, M. Raza, and S. A. C. Bukhari, "Convolutional neural network with batch normalization for glioma and stroke lesion detection using MRI," Cognitive Systems Research, vol. 59, pp. 304-311, 2020.
45. S. Bauer, "A survey of MRI-based medical image analysis for brain tumor studies," Phys. Med. Biol., vol. 58, no. 13, pp. 97–129, 2013.
46. D. N. Louis, "The 2007 who classification of tumours of the central nervous system," Acta Neuropathologica, vol. 114, no. 2, pp. 97–109, 2007.
47. E. G. Van Meir, "Exciting new advances in neuro-oncology: The avenue to a cure for malignant glioma," CA, Cancer J. Clinicians, vol. 60, no. 3, pp. 166–193, 2010.
48. G. Tabatabai,"Molecular diagnostics of gliomas: The clinical perspective," Acta Neuropathologica, vol. 120, no. 5, pp. 585–592, 2010.
49. B. Menze,"The multimodal brain tumor image segmentation benchmark (BRATS)," IEEE Trans. Med. Imag., vol. 34, no. 10, pp. 1993–2024, Oct. 2015.
50. N. J. Tustison, "N4ITK: Improved n3 bias correction," IEEE Trans. Med. Imag., vol. 29, no. 6, pp. 1310–1320, Jun. 2010.

51. L. G. Nyúl, J. K. Udupa, and X. Zhang, “New variants of a method of MRI scale standardization,” IEEE Trans. Med. Imag., vol. 19, no. 2, pp. 143–150, Feb. 2000.
52. S.Franchini, A.Gentile, F.Sorbello, G.Vassallo, and S.Vitabile, "ConformalALU: A Conformal Geometric Algebra Coprocessor for Medical Image Processing", IEEE Transactions On Computers, VOL. 64, NO. 4, APRIL 2015.
53. J.Shang-Ling ,Z.Shichen, X.Weilun, Y.Fangxiaoqi, F.Mingjian, and W.Dongmei, "Brain MRI Tumor Segmentation with 3D Intracranial Structure Deformation Features", IEEE Intelligent Systems,V.31 n.2, p.66-76, March 2016.
54. S.Bauer., “ A Survey of MRI-based medical image analysis for brain tumor studies”, Phys. Med. Biol., vol. 58, no. 13, pp. 97-129, 2013.
55. T.Nima, S.Jae, R.Suryakanth, R. Gurudu, R. Todd Hurst, "Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning,” IEEE Transactions On Medical Imaging, vol. 35, no. 5, may 2016.
56. Y. Wang, “Efficient Stockwell Transform with Applications to Image Processing”, University of Waterloo, Waterloo, Ontario, Canada, 2011.
57. J. Ladan, “An Analysis of Stockwell Transforms, With Applications to Image Processing”, University of Waterloo, Ontario, Canada, 2014
58. M. Singh Choudhry and Rajiv Kapoor,“Performance Analysis of Fuzzy C-Means Clustering Methods for MRI Image Segmentation,” Twelfth International Multi-Conference on Information Processing-2016 (IMCIP-2016), Science Direct, Procedia Computer Science 89 (2016) 749 – 758
59. S.Chen, and D.Zhang, “Robust Image Segmentation using FCM with Spatial Constraints Based on New Kernel-Induced Distance Measure”, IEEE Transaction on Systems, Man and Cybernetics-Part B, vol. 34(4), August (2004).
60. S. R. Kannana, S. Ramathilagamb, R. Devia and E. Hinesc, “Strong Fuzzy C-Means in Medical Image Data Analysis”, The Journal of Systems and Software, pp. 2425–2438, 2012.
61. A. Zexuan Ji, Yong Xiab, Qiang Chena, Quansen Suna, Deshen Xiaa and David Dagan Feng, “Fuzzy C-Means Clustering with Weighted Image Patch for Image Segmentation”, Applied Soft Computing, vol. 12, pp. 1659–1667, 2012.
62. H.Capitaine and C.Frélicot, “A fast fuzzy c-means algorithm for color image segmentation” EUSFLAT-LFA 2011 July 2011 Aix-les-Bains, France.
63. E. H. Ruspini. “A new approach to clustering. Information and control”, 15(1):22–32, 1969
64. A. Jain, M. Murty, and P. Flynn. Data clustering: A review. ACM Computing Surveys, 31:264–323, 1999.

65. S.Thirumuruganathan, “A Detailed Introduction to K-Nearest Neighbor (KNN) Algorithm book” , International Journal of Advanced and Applied Sciences, 6(7), 2010.
66. M.Ahmed, S.Yamany, N.Mohamed, A.Farag,and T.Moriarty T, "A modifed fuzzy c-means algorithm for bias feld estimation and segmentation of MRI data,” IEEE Trans Med Imaging 21(3):193–199,2002
67. H.RamziA,D.Mohamad,and M.Masud, “An Enhanced k-Means Clustering Algorithm for Pattern Discovery in Healthcare Data”, International Journal of Distributed Sensor Networks Volume 2015.
68. S. Chen and D. Zhang. “Robust image segmentation using fcm with spatial constraints based on new kernel-induced distance measure”. IEEE Transactions on Systems, Man and Cybernetics, Part B., 34(4):1907–1916, 2004.
69. S Krinidis, V Chatzis, ”A robust fuzzy local information cmeans clustering algorithm,” IEEE Trans Image Process 19(5):1328– 1337, 2010
70. José L. A. Rosa Nelson F. F. Ebecken, “Data Mining for Data Classification Based on the KNN-Fuzzy Method Supported by Genetic Algorithm”, Part of the Lecture Notes in Computer Science book series (LNCS, volume 2565), pp 126-133 on Image Processing, 19:1328–1337, 2010.
71. R. J. Ramteke, Khachane Monali Y. “Automatic Medical Image Classification and Abnormality Detection Using KNearest Neighbour”, International Journal of Advanced Computer Research (ISSN (print): 2249-7277 ISSN (online): 2277-7970), Vol-2, Issue-6, December-2012.
72. N. Suguna, and K. Thanushkodi, “An Improved k-Nearest Neighbor Classification Using Genetic Algorithm”, IJCSI International Journal of Computer Science Issues, Vol. 7, Issue 4, No 2, July 2010, ISSN (Online): 1694-0784.
73. S. K. Khan and A. Ahmad. “Cluster center initialization algorithm for K-Means clustering”, Pattern Recognition Letters, Vol.25, pp1293–1302, 2004.
74. A. K. Jain and R. C. Dubes, “Algorithms for Clustering Data”, Prentice Hall, Upper Saddle River, NJ, USA, 1988.
75. J.B.Mac Queen, “Some methods for classification and analysis of multivariate observations,” in Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability, pp.281–297, University of California Press,1967.
76. J.A.Hartigan and M.A.Wong, “Algorithm AS136: a k-means clustering algorithm,” Journal of the Royal Statistical Society, SeriesC, vol.28, no.1, pp.100–108, 1979.
77. L. Bottou and Y. Bengio, “Convergence properties of the k means algorithms,” in Advances in Neural Information Processing Systems,pp.585– 592,MITPress,NewYork,NY,USA,1995.

78. H.Mohsen, “Classification using Deep Learning Neural Networks for Brain Tumors”, Future Computing and Informatics, pp 1-4 ,2017
79. S.Bauer, “Multiscale Modeling for Image Analysis of Brain Tumor Studies”, IEEE Transactions on Biomedical Engineering, 59(1),2012
80. A.Islam, “Multi-fractal Texture Estimation for Detection and Segmentation of Brain Tumors”, IEEE, 61(1),2013
81. M.Huang , “Brain Tumor Segmentation Based on Local Independent Projection based Classification”, IEEE Transactions on Biomedical Engineering, IEEE, 33(3), 2013
82. AndacHamamci, “Tumor-Cut: Segmentation of Brain Tumors on Contrast Enhanced MR Images for Radiosurgery Applications”, IEEE Transactions on Medical Imaging, 31(3),2012
83. B.H. Menze , “The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)”, IEEE Transactions on Medical Imaging, 2014
84. J.Liu, “A Survey of MRI-Based Brain Tumor Segmentation Methods”, TSINGHUA Science and Technology, 19(6) ,2011
85. S.Huda, “A Hybrid Feature Selection with Ensemble Classification for Imbalanced Healthcare Data: A Case Study for Brain Tumor Diagnosis”, IEEE Access, 4,2017
86. R. Karuppathal and V. Palanisamy, “Fuzzy based automatic detection and classification approach for MRI-brain tumor”, ARPN Journal of Engineering and Applied Sciences, 9(12), 2014
87. N.V.Shree, and T. Kumar, “Identification and classification of brain tumor MRI images with feature extraction using DWT and probabilistic neural network,” Brain Informatics, 5(1), 23-30,2018
88. P.Kumar,and B.Vijayakumar,” Brain Tumour MR Image Segmentation and Classification Using by PCA and RBF Kernel Based Support Vector Machine,” Middle-East Journal of Scientific Research, 23(9), 2106-2116,2015
89. Z.Ullah, S.H.Lee, and M.Fayaz,“Enhanced feature extraction technique for brain MRI classification based on Haar wavelet and statistical moments, “International Journal of Advanced and Applied Sciences, 6(7), 89-98,2019
90. K.V.A.Muneer, and K.P.Joseph, “Automation of MR Brain Image Classification for Malignancy Detection,” Journal of Mechanics in Medicine and Biology, 19(1),Art. no. 1940002 ,2019
91. A.Vidyarthi, and N.Mittal, “A hybrid model for extraction of brain tumor in MR images,” In Signal Processing and Communication (ICSC), International Conference on (pp. 202-206). IEEE,2013

92. W.Li, C.Chen,H.Su, and & Q.Du,"Local Binary Patterns and Extreme Learning Machine for Hyperspectral Imagery Classification," IEEE Trans. Geoscience and Remote Sensing, 53(7), 3681-3693,2015
93. Q.Liu, and X.Liu, "Feature Extraction of Human Viruses Microscopic Images Using Gray Level Co-occurrence Matrix", IEEE ,International Conference on Computer Sciences and Applications (CSA), 2013 (pp. 619-622),2013
94. A.Singh, "Detection of brain tumor in MRI images, using combination of fuzzy c-means and SVM," 2nd IEEE International Conferenceon Signal Processing and Integrated Networks (SPIN), 2015 (pp. 98- 102),2015
95. M.F.Amasyali, and O.K. Ersoy, "Comparison of single and ensemble classifiers in terms of accuracy and execution time", IEEE International Symposium on Innovations in Intelligent Systems and Applications (INISTA), (pp. 470-474),2011
96. S.Şeker, , B.Diri, , and Balık, H.Derin Öğrenme Yöntemleri ve Uygulamaları Hakkında Bir İnceleme (In Turkish). Gazi Mühendislik Bilimleri Dergisi, 3(3), 47-64,2017
97. T.Chan, K.Jia, S.Gao, J.Lu, Z.Zeng, and Y.Ma, "A Simple Deep Learning Baseline for Image Classification," IEEE Transactions on Image Processing, 24(12), 5017-5032,2015
98. W. Ouyang,"Deepid-net: Deformable deep convolutional neural networks for object detection," Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, 2403-2412,2015
99. J.Han, D.Zhang, G.Cheng, N.Liu, and D.Xu, "Advanced Deep-Learning Techniques for Salient and Category-Specific Object Detection," A Survey. IEEE Signal Processing Magazine, 35(1), 84-100,2018
100. L.Deng, G.Hinton, and B.Kingsbury, "New types of deep neural network learning for speech recognition and related applications," An overview. IEEE International Conference on Acoustics, Speech and Signal Processing, 8599- 8603,2013
101. L.Deng,and J.C.Platt,"Ensemble deep learning for speech recognition,"15th Annual Conference of the International Speech Communication Association,2014
102. R. Girshick, J.Donahue, T.Darrell, and J.Malik,"Rich feature hierarchies for accurate object detection and semantic segmentation,"Proceedings of the IEEE conference on computer vision and pattern recognition, 580-587,2014
103. W.Sun,B.Zheng, and W.Qian," Automatic feature learning using multichannel ROI based on deep structured algorithms for computerized lung cancer diagnosis," Computers in Biology and Medicine, 89, 530-539,2017

104. A.Prasoon, K. Petersen, C.Igel, F.Lauze, E.Dam, and M.Nielsen, " Deep Feature Learning for Knee Cartilage Segmentation Using a Triplanar Convolutional Neural Network, "International Conference on Medical Image Computing and Computer-Assisted Intervention, 246-253,2013

105. M. Havaei, A.Davy, D.Warde-Farley, A. Biard, A.Courville, Y. Bengio, C. Pal, P.Jodoin and H.Larochelle,"Brain tumor segmentation with Deep Neural Networks. Medical Image Analysis,"35, 18-3, 2016.

106. A.Işın, C.Direkoğlu,and M.Şah, "Review of MRI-based Brain Tumor Image Segmentation Using Deep Learning Methods," Procedia Computer Science, 102, 317-324,2016

107. S.P.Justin, J.P.Andre, A.L.Bennett, and D.Fabbri, "Deep learning for brain tumor classification," Proc. SPIE Medical Imaging : Biomedical Applications in Molecular, Structural, and Functional Imaging, 10137,2017

108. G.Urban, M.Bendszus, F.A.Hamprecht, and J.Kleesiek,"Multi-modal Brain Tumor Segmentation using Deep Convolutional Neural Networks,"MICCAI BraTS (Brain Tumor Segmentation) Challenge. Proceedings, winning contribution, 31-35,2014

109. S.Sajid,S.Hussain, and A.Sarwar, "Brain Tumor Detection and Segmentation in MR Images Using Deep Learning," Arabian Journal for Science and Engineering, 44(11), 9249-9261,2019

110. F.Özyurt, E.Sert, E.Avci, and E.Dogantekin, "Brain tumor detection based on Convolutional Neural Network with neutrosophic expert maximum fuzzy sure entropy," Measurement, 147, p. 106830,2019

111. D.Zikic, "Segmentation of brain tumor tissues with convolutional neural networks," MICCAI Multimodal Brain Tumor Segmentation Challenge (BraTS), pp. 36–39, 2014.

112. G. Urban et al., "Multi-modal brain tumor segmentation using deep convolutional neural networks," MICCAI Multimodal Brain Tumor Segmentation Challenge (BraTS), pp. 1–5, 2014.

113. A. Davy, "Brain tumor segmentation with deep neural networks," MICCAI Multimodal Brain Tumor Segmentation Challenge (BraTS), pp. 31–35, 2014.

114. M. Havaei, "Brain tumor segmentation with deep neural networks ," International Journal of Advanced and Applied Sciences, 6(7), 89-98,2015

115. M. Lyksborg, "An ensemble of 2d convolutional neural networks for tumor segmentation," in Image Analysis. New York: Springer, 2015, pp. 201–211.

116. V. Rao, M. Sharifi, and A. Jaiswal, "Brain tumor segmentation with deep learning," MICCAI Multimodal Brain Tumor Segmentation Challenge (BraTS), pp. 56–59, 2015.

117. P. Dvorák and B. Menze, “Structured prediction with convolutional neural networks for multimodal brain tumor segmentation,” MICCAI Multimodal Brain Tumor Segmentation Challenge (BraTS), pp. 13–24, 2015.

118. K. Simonyan and A. Zisserman, Very deep convolutional networks for large-scale image recognition ,2014

119. L. G. Nyúl, J. K. Udupa, and X. Zhang, “New variants of a method of MRI scale standardization,” IEEE Trans. Med. Imag., vol. 19, no. 2, pp. 143–150, Feb. 2000.

120. S.Shahla, and C.Nasrollah, “Palm print authentication based on discrete orthonormal S-Transform” Applied Soft Computing vol.21,pp.341–351, 2014

121. R. G. Stockwell, L. Mansinha, and R. P. Lowe, “Localization of the complex spectrum the S transform,” IEEE Transactions on Signal Processing, vol. 44, no. 4, pp. 998–1001, Apr. 1996

122. D.Shen, W.Guorong, and S.Heung-Il, "Deep learning in medical image analysis," Annual Review of Biomedical Engineering, 2017.

123. N.Tajbakhsh, J.Y. Shin, R.G.Suryakanth, and R.Todd Hurst, "Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?", IEEE Transactions On Medical Imaging, vol. 35, no. 5, may 2016.

124. A.S. Razavian, H. Azizpour, J. Sullivan, and S. Carlsson, “CNN features off-the-shelf: An astounding baseline for recognition,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops, pp. 512–519, 2014.

125. Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521,no. 7553, pp. 436–444, 2015.

126. N.Gordillo, E.Montseny, P.Sobrevilla,”State of the art survey on MRI brain tumor segmentation,” Magn Reson Image, 31(8):1426-1438, 2013

127. S.Sonu, S.L.Mohan,”Automatic brain tumor detection and classification using svm classifier,” In: Proceedings of ISER 2nd international conference, pp. 55-59, July 2015

APPENDIX

MATLAB CODE

Matlab code for Segmentation using DOST

```
clear all
close all
clc
% read the test-signal

[filename,pathname] =
uigetfile({'*.*';'*.bmp';'*.tif';'*.gif';'*.jpg';'*.png'},'Pick an Image
File');
test_signal= imread([pathname,filename]);
figure, imshow(test_signal); title('Brain MRI Image');
test=imresize(test_signal,[256 256]);
 test1  =rgb2gray(test);
% compute the dost coefficients of the signal "test1"
dost2_test_signal = dost2(test1);
% % reconstruct the signal using the idost2
% reconstructed_test_signal = idost2(dost2_test_signal);
figure
subplot(1,2,1)
Imagesc(test_signal)
title('test signal','FontSize',20);
  axis off
subplot(1,2,2)
imagesc(log(abs(dost2_test_signal)))
% imagesc([-.5 .5], [-.5 .5], log(abs(dost2_test_signal)))
title(' Dost2 of the test signal','FontSize',20);
 axis off
% subplot(1,3,3)
% imagesc(real(reconstructed_test_signal))
% title('reconstructed test signal')
% axis tight
colormap gray
```

Matlab code for classification

```
outputfolder=('Tumor data');
rootfolder=fullfile(outputfolder,'Tumor data');%from this we take only two
catagories not all images so
catagories={'Malignant','Benign'};
imds=imageDatastore(fullfile(rootfolder,catagories),'LabelSource','foldername
s');%it is two argumentthe location of the images and names and vlaues
tbl=countEachLabel(imds);%so the number of catagories is not the same make it
the same
% %  %which catagory is least number of images
% minSetCount=min(tbl)%%then it least we focus on second column,make look
minSetCount=min(tbl{:,2});%we are accessing the leat number for all
% % %now updet image data store and in the updeted catagory each  images
% % %so call it
imds=splitEachLabel(imds,minSetCount,'randomize');%select each image
randomize
countEachLabel(imds);%count each label again we get all catagories the same
 Malignant=find(imds.Labels=='Malignant',1);%sequance is 1
 Benign=find(imds.Labels=='Benign',1);%sequance is 1
```

```
figure
subplot(2,2,1);
imshow(readimage(imds,Malignant));
subplot(2,2,2);
imshow(readimage(imds,Benign));
%now pretrained convolutional nueral network
%resonet 50 is tranined in image net data set this is 1000 image catagories
and one mil trained images
net=resnet50();
% %visualaized architecture
figure
plot(net)
title('Architecture of ResNet_50')%it is complex architecture difficulte to
resize
% %using set fucntion to resize the network
set(gca, 'YLim',[ 150 170]);
 net.Layers(1);%the input layer property
net.Layers(end);%which is the laset layer name is classf.fc1000
numel(net.Layers(end).ClassNames); % num of class
[trainingSet,testSet]=splitEachLabel(imds,0.7,'randomize');%30 per for
testing rest for validation
imageSize=net.Layers(1).InputSize;%IMAGE input size
augmentedTrainingSet=augmentedImageDatastore(imageSize,trainingSet,...
    'ColorPreprocessing','gray2rgb');%to resize the image and to convert any
image into rgb omage
 augmentedTestSet=augmentedImageDatastore(imageSize,testSet,...
     'ColorPreprocessing','gray2rgb');
w1=net.Layers(2).Weights;%get the weights of second convolutional layer and
store to w1
 w1=mat2gray(w1);
 figure
 montage(w1)
 title('first convolutional layer weight')
 featureLayer='fc1000';
 trainingFeatures=activations(net,augmentedTrainingSet,...
     featureLayer,'MiniBatchSize',20,'OutputAs','columns');
 trainingLable=trainingSet.Labels;
classifier=fitcecoc(trainingFeatures,trainingLable,'Learner','Linear','Coding
','onevsall',...
    'ObservationsIn','column')
testFeatures=activations(net,augmentedTestSet,...
featureLayer,'MiniBatchSize',20,'OutputAs','columns');% gives features of
test image
% %measure accuracy ogf trained classifier
predictLable=predict(classifier,testFeatures,'ObservationsIn','column');
 testLable=testSet.Labels; %actual label
% %we used conf.matric to evaluate the performance of classifier
 confMat=confusionmat(testLable,predictLable);
% %when we evaluate using a percentage value using
confMat=bsxfun(@rdivide,confMat,sum(confMat,1));
Accuracy=mean(diag(confMat));%this measur accuracy i.e the classifier for
@augmentedTrainingSet is accurace
disp(['Accuracy = ',num2str(Accuracy),'%'])
plotconfusion(testLable,predictLable);
%now we are going to classify images make comment for figure above
%two catagories
% newImage=imread(fullfile('test1.jpeg'));
```

```
[filename,pathname] = uigetfile('*.*');
newImage = imread([pathname,filename]);
ds=augmentedImageDatastore(imageSize,newImage,...
    'ColorPreprocessing','gray2rgb');
imageFeatures=activations(net,ds,...
    featureLayer,'MiniBatchSize',32,'OutputAs','columns');
label=predict(classifier,imageFeatures,'ObservationsIn','column');%pridicted
label
y=sprintf('loaded image belongs to %s class',label);
figure,imshow(newImage); title(y);xlabel('Complete CNN')
disp(' Done!');
```

Matlab code for SCA Algorithm

```
clc
clear all
x = zeros(5,5,5);
W1 = 2*rand(20, 25) - 1;
W2 = 2*rand(20, 20) - 1;
W3 = 2*rand(20, 20) - 1;
W4 = 2*rand( 5, 20) - 1;
alpha = 0.01;
N = 5;
 for epoch=1:100
for k = 1:N
    x = zeros(5,5,5);
x(:, :, 1) = imread('2_output.jpg')./max(imread('2_output.jpg'));
x(:, :, 2) = imread('3_output.jpg')./max(imread('3_output.jpg'));
x(:, :, 3) = imread('7_output.jpg')./max(imread('7_output.jpg'));
x(:, :, 4) = imread('8_output.jpg')./max(imread('8_output.jpg'));
x(:, :, 5) = imread('10_output.jpg')./max(imread('10_output.jpg'));
D = (imread('2_output.jpg')./max(imread('2_output.jpg')))';
x = reshape(x(:, :, k), 25, 1);

t=2;
Max_iteration=1000;
% Destination_position=zeros(1,5);
  %  while t<=Max_iteration
     a = 2;
    r1=a-t*((a)/Max_iteration); % r1 decreases linearly from a to 0
    % Update the position of solutions with respect to destination
        for i=1:20 % in i-th solution
        for j=1:25 % in j-th dimension
            % Update r2, r3, and r4
            r2=(2*pi)*rand();
            r3=2*rand;
            r4=rand();

            if r4<0.5
%
                W1(i,j)= W1(i,j)+(r1*sin(r2)*abs(r3*N-W1(i,j)));
            else
%
                W1(i,j)= W1(i,j)+(r1*cos(r2)*abs(r3*N-W1(i,j)));
             end
        end
        end
```

```
v1 = W1*x;
y1 = ReLU(v1);
 for i=1:20 % in i-th solution
        for j=1:20 % in j-th dimension
            % Update r2, r3, and r4
            r2=(2*pi)*rand();
            r3=2*rand;
            r4=rand();
            if r4<0.5
%
                W2(i,j)= W2(i,j)+(r1*sin(r2)*abs(r3*N-W2(i,j)));
            else
%
                W2(i,j)= W2(i,j)+(r1*cos(r2)*abs(r3*N-W2(i,j)));
             end
        end
 end
v2 = W2*y1;
y2 = ReLU(v2);
 for i=1:20 % in i-th solution
        for j=1:20 % in j-th dimension
            % Update r2, r3, and r4
            r2=(2*pi)*rand();
            r3=2*rand;
            r4=rand();

            if r4<0.5
%
                W3(i,j)= W3(i,j)+(r1*sin(r2)*abs(r3*N-W3(i,j)));
            else
%
                W3(i,j)= W3(i,j)+(r1*cos(r2)*abs(r3*N-W3(i,j)));
             end
        end
 end
v3 = W3*y2;
y3 = ReLU(v3);

    % Update the position of solutions with respect to destination
        for i=1:5 % in i-th solution
        for j=1:20 % in j-th dimension
            % Update r2, r3, and r4
            r2=(2*pi)*rand();
            r3=2*rand;
            r4=rand();

            if r4<0.5
%
                W4(i,j)= W4(i,j)+(r1*sin(r2)*abs(r3-W4(i,j)));
            else
%
                W4(i,j)= W4(i,j)+(r1*cos(r2)*abs(r3*-W4(i,j)));
            end
            end
% v = W4*y3;
% %
```

```
% %    y = tanh(v);
%    y = tansig(v);
% %   y = logsig(v);
%    %y = Softmax(v);
% d   = D(k, :)';
% d=double(d);
% e   = d - y;
% delta = e;
% % plot(sort(e,'descend'))
%
end
end
%  delta;
% plot(sort(e,'descend'))
 end
v = W4*y3;
%
  y = tanh(v);
%    y = tansig(v);
%   y = logsig(v);
%    y = Softmax(v);
d   = D(k, :)';
d=double(d);
e   = d - y;
delta = e;
% plot(sort(e,'descend'))
plot(sort(e,'descend'))
%    subplot(121);plot(sort(e,'descend'))
  v1 = W4*y3;
%
%      y1 = tanh(v1);
   y1 = 0.25*tansig(v1);
%    y1 = logsig(v);
%    y1 = Softmax(v1);
d1   = D(k, :)';
d1=double(d1);
e1   = d1 - y1;
delta = e1;
 hold on
plot(sort(e1,'descend'),'r')
% subplot(122);plot(sort(e1,'descend'))
v2 = W4*y3;
%
%      y1 = tanh(v1);
%     y1 = tansig(v);
   y2 = 0.25*logsig(v2);
%    y1 = Softmax(v1);
d2   = D(k, :)';
d2=double(d2);
e2   = d2 - y2;
delta = e2;
 hold on
plot(sort(e2,'descend'),'g')
v3 = W4*y3;
%
%      y1 = tanh(v1);
%     y1 = tansig(v);
```

```
%    y2 = 0.25*logsig(v);
  y31 = 4*Softmax(v3);
d3  = D(k, :)';
d3=double(d3);
e3  = d3 - y31;
delta = e3;
 hold on
plot(sort(e3,'descend'),'k')
 hold off
```

Matlab code for feature extraction

```
ddd=dost2(test_signal);
z=(abs(ddd));
ST_feat = z;
G = pca(ST_feat);
whos ST_feat
whos G
[glcm,SI] = graycomatrix(z,'Offset',[1,0]);%'Symmetric',true);
%g= graycomatrix(ddd);
g = graycomatrix(z);
stats = graycoprops(glcm,'Contrast Correlation Energy Homogeneity');
Contrast = stats.Contrast;
Correlation = stats.Correlation;
Energy = stats.Energy;
Homogeneity = stats.Homogeneity;
Mean =mean2(G);
Standard_Deviation =std2(double(G));
Entropy = entropy(G);
RMS = mean2(rms(G));
%Skewness = skewness(img)
Variance =mean2(var(double(seg_img)));
a = sum(double(G(:)));
Smoothness = 1-(1/(1+a));
Kurtosis = kurtosis(double(G(:)));
Skewness = skewness(double(G(:)));
% Inverse Difference Movement
m = size(G,1);
n = size(G,2);
in_diff = 0;
for i = 1:m
    for j = 1:n
        temp = G(i,j)./(1+(i-j).^2);
        in_diff = in_diff+temp;
    end
end
IDM = double(in_diff);
feat =
[Contrast,Correlation,Energy,Homogeneity,Mean,Standard_Deviation,Entropy,RMS,
Variance, Smoothness, Kurtosis, Skewness, IDM];
```