DETECTING AND CLASSIFYING FETAL BRAIN ABNORMALITIES USING DECISION TREE ALGORITHM

Mrs.P.J.Mercy[1] and M.Utchimahali@usha [2]
1 Associate Professor, 2 III MCA Student
Department of Computer Applications, Sarah Tucker College, Tirunelveli-7.

ABSTRACT: **Detecting and classifying fetal brain abnormalities from magnetic resonance imaging (MRI) is important, as approximately 3 in 1000 women are pregnant with a fetal of abnormal brain. Early detection of fetal brain abnormalities using machine learning techniques can improve the quality of diagnosis and treatment planning. The literature has shown that most of the work made to classify brain abnormalities in very early age is for preterm infants and neonates not fetal. However, research papers that studied fetal brain MRI images have mapped these images with the neonates MRI images to classify an abnormal behaviour in newborns not fetal. In this work, a pipeline process is proposed for fetal brain classification (FBC) which uses machine learning techniques. The main contribution of this work is the classification of fetal brain abnormalities in early stage, before the fetal is born. The proposed algorithm is capable of detecting and classifying a variety of abnormalities from MRI images with a wide range of fetal gestational age (GA) (from 16 to 39 weeks) using a flexible and simple method with low computational cost. The novel proposed method consists of four phases; pre-processing, segmentation, feature extraction and classification. In the pre-process, the input image is converted into gray scale and apply the wiener filter for image enhancement. After pre-processing, the adaptive segmentation is applied for segmenting the image. Then DWT and statistical features are extracted from segmented image. Finally, the input image is classified using Decision Tree algorithm**.

Index Terms - **Leaf Classification, Image processing techniques, Morphological operators, Machine Learning algorithms**

1. INTRODUCTION

Magnetic resonance imaging (MRI) is an important tool for medical investigation of the brain and is considerably enhanced in uterus fetal brain imaging. Scanning the fetal brain is essential as approximately 3 in 1000 pregnancies have fetuses with different types of brain abnormalities. Moreover, several series of neuropathological variations happen, some related to serious clinical morbidities. Recently, fetal MRI has been considered as a neoteric, non-invasive tool for monitoring fetal brain development. Fetal MRI confers high contrast resolution for the tissue. Therefore, it confers detailed fetal structural information of the brain. Machine learning techniques can be used to detect and classify brain abnormalities and tumors from fetal MRI images at an early stage without any surgical interventions. Early detection of the fetal brain abnormalities will indicate how the pregnancy will be managed, possible treatments that can be taken, and help parents understand and be prepared for dealing with the abnormality

Moreover, the early discovery of these abnormalities can improve the quality of diagnosis and follow-up planning. Machine learning is a well-known field in engineering and computer science. It is a great tool used in pattern recognition and image processing. It can be applied for medical images to help physicians in rendering medical diagnoses. Also, it prevents human-based diagnostic error and the effort during an examination. It is can be used as well for fetal brain MRI images for early detection of abnormalities, diagnosis, and classification of certain diseases. Most of the work that used fetal MRI images focused on segmentation. However, few works used fetal MRI images for classification. To our knowledge, papers that used fetal brain images for classification are limited. Although these papers used fetal brain MRI images for classification, they used these images to predict an abnormal neuro-behavior in small for gestational age (SGA) neonates, not fetuses. This means that they classify the SGA abnormality after the fetuses are born and become neonates. The authors of these papers assumed the hypothesis that fetal MRI textural patterns are related to new-born neurobehavioral.

Machine learning is an exciting field of research in computer science and engineering. It is considered a branch of artificial intelligence because it enables the extraction of meaningful patterns from examples, which is a component of human intelligence. The appeal of having a computer that performs repetitive and well-defined tasks is clear: computers will perform a given task consistently and tirelessly; however, this is less true for humans. More recently, machines have demonstrated the capability to learn and even master tasks that were thought to be too complex for machines, showing that machine learning algorithms are potentially useful components of computer-aided diagnosis and decision support systems. Even more exciting is the finding that in some cases, computers seem to be able to "see" patterns that are beyond human perception.

Brain tumors are the most common brain disease that affects the central nervous system (CNS), the brain, and spinal cord. During the last two decades, the computer-aided diagnosis system (CAD) has been employed to improve the accuracy of the diagnostic ability of radiologists in detecting, segmenting, and identifying the type of brain tumor [2].Magnetic Resonance Imaging (MRI) is the most common modality that is exploited to diagnose abnormal tissue growth in the brain. Usually, radiologists depend on manual diagnosis to identify brain tumors [3]. In manual diagnosis, the amount of MRI images to be analyzed is large enough to make readings based on visual interpretation expensive, inaccurate, and intensive. Besides; the human eye is sensitive to any changes in the MRI image, this sensitivity may be decreased with the increase in the number of images. Furthermore, manual diagnosis is time-consuming, and it depends on the subjective decisions of the radiologists which is hard to quantify, thus it may lead to misclassification [6]. Radiologists used invasive techniques such as biopsy and spinal tap methods for a brain tumor diagnosis. Both approaches, biopsy and spinal tap, are painful, time-consuming, risky, and expensive. Consequently, developing an accurate, robust, practical, and automatic diagnosis system for this purpose is significant and essential to overcome the drawbacks of manual diagnosis and invasive techniques.

2. LITERATURE REVIEW

[1] They addressed this issue by automatic localization of the fetal anatomy, in particular, the brain which is a structure of interest for many fetal MRI studies. We first extract superpixels followed by the computation of a histogram of features for each superpixel using bag of words based on dense scale invariant feature transform (DSIFT) descriptors. We construct a graph of superpixels and train a random forest classifier to distinguish between brain and non-brain superpixelsIn

[2] the proposed system, at first the self-organising map neural network trains the features extracted from the discrete wavelet transform blend wavelets and the resultant filter factors are consequently trained by the K-nearest neighbour and the testing process is also accomplished in two stages. The proposed two-tier classification system classifies the brain tumours in double training process which gives preferable performance over the traditional classification method. The proposed system has been validated with the support of real data sets and the experimental results showed enhanced performance

In[3] this study, they apply machine-learning methods to compare whole-brain functional connectivity in preterm infants at term-equivalent age and healthy term-born neonates in order to test the hypothesis that preterm birth results in specific alterations to functional connectivity by term-equivalent age.

[4]Machine learning is a technique for recognizing patterns that can be applied to medical images. Although it is a powerful tool that can help in rendering medical diagnoses, it can be misapplied. Machine learning typically begins with the machine learning algorithm system computing the image features that are believed to be of importance in making the prediction or diagnosis of interest. The machine learning algorithm system then identifies the best combination of these image features for classifying the image or computing some metric for the given image region. There are several methods that can be used, each with different strengths and weaknesses. There are open-source versions of most of these machine learning methods that make them easy to try and apply to images. Several metrics for measuring the performance of an algorithm exist; however, one must be aware of the possible associated pitfalls that can result in misleading metrics. More recently, deep learning has started to be used; this method has the benefit that it does not require image feature identification and calculation as a first step; rather, features are identified as part of the learning process. Machine learning has been used in medical imaging and will have a greater influence in the future. Those working in medical imaging must be aware of how machine learning worworks

[5]Recent resting-state functional MRI investigations have demonstrated that much of the large-scale functional network architecture supporting motor, sensory and cognitive functions in older pediatric and adult populations is present in term- and prematurely-born infants. Application of new analytical approaches can help translate the improved understanding of early functional connectivity provided through these studies into predictive models of neuro developmental outcome. One approach to achieving this goal is multivariate pattern analysis, a machine-learning, pattern classification approach well-suited for high-dimensional neuro imaging data. It has previously been adapted to predict brain maturity in children and adolescents using structural and resting state-functional MRI data.

[6] They present a technique for real-time adaptive thresholding using the integral image of the input. Our technique is an extension of a previous method. However, our solution is more robust to illumination changes in the image. Additionally, our method is simple and easy to implement.Our technique is suitable for processing live video streams at a real-time frame-rate, making it a valuable tool for interactive applications such as augmented reality. Source code is available online.

[7]The purpose of this study was to describe fetal brain anomalies identified during nuchal translucency (NT) examination and their clinical management. The current report highlights the evolving ability to detect CNS malformation during NT scan. FTS novel findings expand our embryological understanding of early fetal development. While severe anomalies can be managed by early termination, other defects necessitate comprehensive evaluation and repetitive imaging.

[8] We tested the hypothesis whether texture analysis (TA) from MR images could identify patterns associated with an abnormal neurobehavior in small for gestational age (SGA) neonates. Ultra sound .Ultra sound and MRI were performed on 91 SGA fetuses at 37 weeks of GA. Frontal lobe, basal ganglia, mesencephalon and cerebellum were delineated from fetal MRIs. SGA neonates underwent NBAS test and were classified asabnormal if $1 area was ,5th centile and as normal if all areas were .5th centile. Textural features associated with neurodevelopment were selected and machine learning was used to model a predictive algorithm. Fetal brain MRI textural patterns were associated with neonatal neurodevelopment. Brain MRI TA could be a useful tool to predict abnormal neurodevelopment in SGA.

[9] In this paper, we present a fully automatic brain tumor segmentation method based on Deep Neural Networks (DNNs). The proposed networks are tailored to glioblastomas (both low and high grade) pictured in MR images. By their very nature, these tumors can appear anywhere in the brain and have almost any kind of shape, size, and contrast. These reasons motivate our exploration of a machine learning solution that exploits a flexible, high capacity DNN while being extremely efficient. Here, we give a description of different model choices that we've found to be necessary for obtaining competitive performance. We explore in particular different architectures based on Convolutional Neural Networks (CNN), i.e. DNNs specifically adapted to image data

[10] they propose a novel multikernel support vector machine classification framework by using the connectivity features gathered from WM connectivity networks, which are generated via multiscale regions of interest (ROIs) and multiple diffusion statistics such as fractional anisotropy, mean diffusivity, and average fiber length

3. METHODOLOGY

In this work a novel pipeline is proposed to classify fetal brain abnormalities. It consists of four steps segmentation, enhancement, feature extraction and classification. The proposed method has several advantages. First, it is used to detect and predict several fetal brain abnormalities before the fetal is born. These abnormalities includes; Cerebellar Hypoplasia, Colpocephaly, Dandy-Walker Variant/Malformation, Agenesis of the Corpus Callosum, Agenesis of the Septi Pellucidi, Megacisterna Magna, and Polymicrogyria. Second, it can be used for a wide range of fetal GA. Moreover, the dataset used to test the performance of the proposed method is based on 20 images. Finally, the segmentation step used to crop fetal brain in the proposed method is a semi-automatic procedure and not manual.

This project has following modules

- **Data Acquisition**
- **Segmentation Phase**
- **Watershed segmentation**
- **Feature Extraction Phase**
- **Discrete Wavelet Transform**
- **GLCM and other statistical descriptive features**
- **Classification Phase**
  - **DECISION TREE**

Data Acquisition :

The dataset used in the proposed method consists of 20 fetal images (113 normal and 114 abnormal brain MRI images) between 16-39 weeks fetal GA. Details on the dataset can be found in [15]. Stacks of T2 weighted images in axial, coronal, and sagittal planes are obtained with half Fourier single shot RARE technique. Because of the fetal motion throughout the examination, each acquisition serves as the scout for the subsequent acquisition.

Segmentation Phase:

Firstly, the fetal brain is extracted from the cerebrospinal fluid (CSF) and the amniotic fluid that surrounds the fetal head. For this step, an optimal intensity threshold **T** is estimated using adaptive local threshold method [16] .This threshold is used to separate the pixels of an image into two classes foreground and background. A binary image of the fetal brain **Ibinary** will be generated from the original image **I** after this step using (1):

Next, some morphological operations as shown in fig.1 are applied to remove maternal tissues that are connected weakly with the outer part of the fetal skull to form a new binary image called **Iclose** . Afterwards, the **Iclose** image's boundaries are traced and classified into parent boundaries and children boundaries where; the parent boundaries are the outer boundary of the skull and the children boundaries are the holes inside the skull which are completely enclosed by the parents' boundary. The binary image obtained after defining parent and children boundaries is called **Iregion.**. Next, Watershed segmentation is applied on **Iregion** to remove any other connection from the maternal tissue and skull. A new image is then formed and called **Iwatershed.**

Watershed segmentation

Watershed segmentation technique is a region-based segmentation approach. It uses Watershed transform to find out the "Watershed ridge lines" or "catchment basins" from a binary image by considering it as a surface, where; the black pixels are considered as the low elevation and the white pixel are the high elevation.

Following this step, the area of all the objects remaining from **Iwatershed** is measured using region properties method. They are then sorted in descending order and the largest area is extracted to form a new binary image called **IABS**. After that, the **IABS** is segmented from the original image **I** using a mask named **Imask** to produce the final segmented whole fetal brain image **Ifinal** using equation (2). Images after these segmentation steps are shown in fig 2.
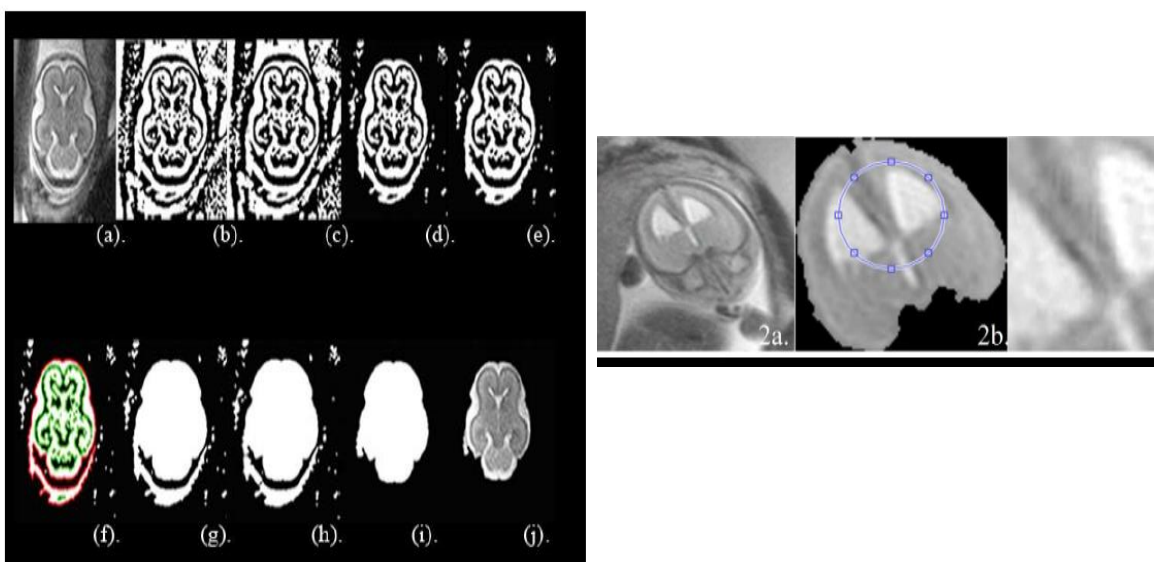
After the whole fetal brain is segmented and detected, a minor ROI is delineated and segmented from the whole brain images based on the brain shape and the changes in the intensity or in the gray level. That is why a blue circle located in the top of the image is chosen as boundary for the minor ROI as shown in fig 3.This minor ROI allocates the region where abnormalities exist in the whole segmented fetal brain image. Figures 3(1a) and (2a) show the whole fetal brain images before segmentation for normal and abnormal cases respectively. Figures 3(1b) and (2b) show the fetal brain after segmentation for the normal and abnormal cases respectively. The delineated minor ROI for the normal fetal brain is shown as well in Figure 3(1b) and the delineated minor ROI for the abnormal fetal brain is shown in Figure 3(2b).

Contrast Enhancement Phase:

In this step the minor ROI image is enhanced using a combination of local and global stretching contrast enhancement methods [19]. This step is applied to enhance the quality, contrast, and preserve the brightness in the fetal brain region. This step is also made to improve the accuracy of classification done later in the proposed approach. The segmented minor ROI after contrast enhancement for normal fetal brain is shown in fig 3(1c) and the segmented minor ROI after contrast enhancement for abnormal fetal brain is shown in fig3

Feature Extraction Phase
a) Discrete Wavelet Transform

DWT is used to provide frequencies representation of signals or images. The Wavelet analysis decomposes the images using a set of orthogonal basis functions (orthonormal). In one level DWT, decomposition is performed on an input image by convolving with one low pass and a high pass filter followed by dyadic decimation . Multi-decomposition levels can be done, by passing the image that has been convolved with the low pass filter into several low and high pass filters. At each level four components are generated which include; an approximation component and three details components (horizontal, vertical, and diagonal). The approximation component is used for the next decomposition level.

Firstly in this phase, 4 -decomposition levels of 2D-DWT are applied on the enhanced minor ROI image to extract some discriminate features. The wavelet basis function used is the discrete approximation of Meyer wavelet (dmey) with mode periodization. After getting the 4-decomposition levels, a features matrix is generated containing; the approximate, horizontal, vertical, and diagonal coefficients of the 4th decomposition level.

b) GLCM and other statistical descriptive features

This feature matrix is then used for texture analysis using grey level covariance matrix (GLCM) algorithm [22] to extract texture features such as; contrast, correlation, energy and homogeneity in the images. Moreover, other statistical descriptive features were extracted from the minor ROI enhanced image. These features include; the mean, standard deviation, variance, root mean square (RMS), smoothness, skewness, entropy , Inverse Difference Moment (IDM) [23], and Kurtosis. The final extracted features are therefore 13 features (4 texture features and 9 descriptive statistical features).

4) Classification Phase

In this phase, the 13 feature extracted is used to train and test machine learning classifiers. These classifiers include decision tree. 5-cross validation technique was used for the validation of the classifiers.

DECISION TREE

The core algorithm for building decision trees called **ID3** by J. R. Quinlan which employs a top-down, greedy search through the space of possible branches with no backtracking. ID3 uses Entropy and Information Gain to construct a decision tree.
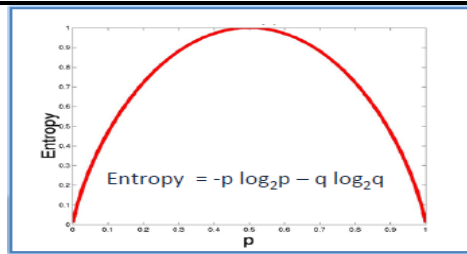
Entropy

A decision tree is built top-down from a root node and involves partitioning the data into subsets that contain instances with similar values (homogenous). ID3 algorithm uses entropy to calculate the homogeneity of a sample. If the sample is completely homogeneous the entropy is zero and if the sample is an equally divided it has entropy of one.

Information Gain

The information gain is based on the decrease in entropy after a dataset is split on an attribute. Constructing a decision tree is all about finding attribute that returns the highest information gain (i.e., the most homogeneous branches).

Entropy

$$E(T, X) = \sum_{c \in X} P(c)E(c)$$

$$Entropy = -0.5 \log_2 0.5 - 0.5 \log_2 0.5 = 1$$

Information Gain

$$Gain(T, X) = Entropy(T) - Entropy(T, X)$$

DECISION TREE WORKING FLOW

Decision tree builds classification or regression models in the form of a tree structure. It breaks down a dataset into smaller and smaller subsets while at the same time an associated decision tree is incrementally developed. The final result is a tree with **decision nodes** and **leaf nodes**. A decision node (e.g., Outlook) has two or more branches (e.g., Sunny, Overcast and Rainy). Leaf node (e.g., Play) represents a classification or decision. The topmost decision node in a tree which corresponds to the best predictor called **root node**. Decision trees can handle both categorical and numerical data.

4. RESULTS

5. CONCLUSION

A pipeline method for detecting and predicting the fetal brain abnormalities was proposed. The main contribution of this paper is the classification of fetal brain abnormalities using a simple and low computational cost method. The proposed method successfully classified fetal brain abnormalities with images of different gestational age (from 16 weeks to 39 weeks). Moreover, the technique was capable of predicting different abnormalities not only one type of abnormality. Most of the work done for classification of MRI images of very young gestational age is for preterm and neonatal. The only work that used fetal MRI images for classification have mapped these images with their neonatal MRI images to predict small for gestational age (SGA) abnormality in new-borns not in fetal. Therefore, the proposed method was compared with other methods for classifying preterm and neonatal. The results showed that the proposed method outperform most of these methods. The results are promising; however future work will focus on improving the performance of the proposed algorithm and increasing the size of the dataset. This work will encourage several researchers to start on working on fetal brain classification.

6. REFERENCE

[1] Automatic Brain Localization In Fetal Mri Using Superpixel Graphs [Amir Alansary ,Matthew Lee, Kevin Keraudren, Bernhard Kainz,Christina Malamateniou,Mary Rutherford,Joseph V. Hajnal,Ben Glocker,Daniel Rueckert 30 December 2015]

[2]Brain Tumour Classification Using Two-Tier Classifier With Adaptive Segmentation Technique [V. Anitha S. Murugavalli 2016]

[3]Machine-Learning To Characterise Neonatal Functional Connectivity In The Preterm Brain[G Ball Et Al. Neuroimage 2016]

[4] Machine Learning For Medical Imaging [Bradley J.Erickson , Panagiotis Korfiatis, Zeynettin Akkus, Timothy L. Klinefeb 17 2017]

[5] Prediction Of Brain Maturity In Infants Using Machine-Learning Algorithms, [Christopher D. Smyser, Nico U.F. Dosenbach, [...], And Jeffrey J. Neil 2016]

[6] Adaptive Thresholding Using The Integral Image [Derek Bradley, Gerhard Roth 30 Jan 2011]

[7] Fetal Brain Anomalies Detection During The First Trimester: Expanding The Scope Of Antenatal Sonography[Eldad Katorza Et Al. J Matern Fetal Neonatal Med 2018]

[8] Automatic Quantitative Mri Texture Analysis In Small-For-Gestational-Age Fetuses Discriminates Abnormal Neonatal Neurobehavior [Magdalena Sanz-Cortes Et Al. Plos One. 2013]

[9]Brain Tumor Segmentation With Deep Neural Networks [Mohammad Havaeia,1, Axel Davyb, David Warde-Farleyc, Antoine Biardc,D, Aaron Courvillec, Yoshua Bengioc, Chris Palc,E,Pierre-Marc Jodoina, Hugo Larochellea,F Received 27 April 2015, Revised 2 March 2016, Accepted 11 May 2016, Available Online 19 May 2016.]

[10]Identification Of Infants At High-Risk For Autism Spectrum Disorder Using Multiparameter Multiscale White Matter Connectivity Networks [Yan Jin Et Al. Hum Brain Mapp2015 Dec]