Detection of Thyroid Cancer Using Deep Learning Method

Keerthana S M[1], Aravind Kumar A[2], Avinash A[3], Aravindhan K[4]
[1]Assistant Professor, Dept of computer science engineering
[2, 3, 4] Dept of computer science engineering
[1, 2, 3, 4] SVS College Of Engineering, Coimbatore.

Abstract- *Thyroid cancer is the leading cause of deaths worldwide. Both researchers and doctors are facing the challenges of fighting with thyroid cancer. In that thyroid disease is a major cause of formation in medical diagnosis and in the prediction, onset to which it is a difficult axiom in the medical research. Thyroid gland is one of the most important organs in our body. The secretions of thyroid hormones are culpable in controlling the metabolism. Hyperthyroidism and hypothyroidism are one of the two common cancer of the thyroid that releases thyroid hormones in regulating the rate of body's metabolism. Early detection of thyroid cancer is the top priority for saving the lives of many. Typically, visual examination and manual techniques are used for these types of a thyroid cancer diagnosis. This manual interpretation of medical images demands high time consumption and is highly prone to mistakes. Thus, in this project we apply deep learning algorithms to detect the thyroid cancer and its presence without the need of several consultations from different doctors. This leads to earlier prediction of the presence of the cancer and allows us to take prior actions immediately to avoid further consequences in an effective and cheap manner avoiding human error rate. Thus, this project is used to predict the thyroid cancer effectively.*

I. INTRODUCTION

Thyroid cancer occurs in the cells of the thyroid — a butterfly-shaped gland located at the base of your neck, just below your Adam's apple. Your thyroid produces hormones that regulate your heart rate, blood pressure, body temperature and weight. Thyroid cancer might not cause any symptoms at first. But as it grows, it can cause pain and swelling in your neck. Several types of thyroid cancer exist. Some grow very slowly, while others can be very aggressive. Most cases of thyroid cancer can be cured with treatment. Thyroid cancer rates seem to be increasing. Some doctors think this is because new technology is allowing them to find small thyroid cancers that may not have been found in the past.

II. LITERATURE SURVEY

Hui Zhou, Kun Wang, Jie Tian, Member, IEEE," Online Transfer Learning for Differential Diagnosis of Benign and Malignant Thyroid Nodules with Ultrasound Images", IEEE Transactions on Biomedical Engineering. We aimed to propose a highly automatic and objective model named online transfer learning (OTL) for the differential diagnosis of benign and malignant thyroid nodules from ultrasound (US) images. Methods: The OTL mothed combined the strategy of transfer learning and online learning. Two datasets (1750 thyroid nodules with 1078 benign and 672 malignant nodules, and 3852 thyroid nodules with 3213 benign and 639 malignant nodules) were collected to develop the model. The diagnostic accuracy was also compared with VGG-16 based transfer learning model and different input images based model. Analysis of receiver operating characteristic (ROC) curves were performed to calculate optimal area under it (AUC) for benign and malignant nodules. Results: AUC, sensitivity and specificity of OTL were 0.98 (95% confidence interval [CI]: 0.97-0.99), 98.7% (95% confidence interval [CI]: 97.8%-99.6%) and 98.8% (95% confidence interval [CI]: 97.9%-99.7%) in the final online learning step, which was significantly better than other deep learning models ($P < 0.01$).OTL achieved the most accurate differential diagnosis of benign and malignant thyroid nodules comparing with transfer learning and multi-ROI based model

Beaumon, P. Onoma, M. Rimlinger , D. Broggio, P. Caldeira Ideias and D.Franck,"Age-specific experimental and computational calibration of thyroid in vivo monitoring",IEEE Transactions on Radiation and Plasma Medical Sciences Age-specific thyroid phantoms corresponding to 5, 10, 15 years-old and the adult case have been designed and manufactured with a 3D printer. Reference measurements of the counting efficiency have been carried out for thyroid in vivo monitoring of 131 I with all these phantoms. These measurements where performed for the emergency mobile units of IRSN. The full efficiency curve, between 29 and 1000 keV, was then obtained by Monte-Carlo calculations and validated by comparison of a large set of measurements. The obtained efficiency curves are consistent and show that the relative difference in efficiency between the adult and the children case are energy dependent. The developed thyroid phantoms enabled to obtain age specific calibration factors for emergency in vivo monitoring of children. Taking into account the size of thyroid for uptake measurement might be also useful in nuclear medicine department.

Indeed, the treatment of benign thyroid disease, like Grave's disease, requires a personalized dosimetry and hence personalized thyroid retention function .

Mohammad Mehrmohammadi*, Member, IEEE, Pengfei Song, Student Member, IEEE, Duane D. Meixner, Robert T. Fazzio, Shigao Chen, Member, IEEE, James F. Greenleaf, Life Fellow, IEEE, Mostafa Fatemi, Fellow, IEEE, and Azra Alizad, Member, IEEE,"Comb-Push Ultrasound Shear Elastography (CUSE) for Evaluation of Thyroid Nodules: Preliminary In Vivo Results", IEEE Transactions on medical imaging.In clinical practice, an overwhelming majority of biopsied thyroid nodules are benign. Therefore, there is a need for a complementary and noninvasive imaging tool to provide clinically relevant diagnostic information about thyroid nodules to reduce the rate of unnecessary biopsies. The goal of this study was to evaluate the feasibility of utilizing comb-push ultrasound shear elastography (CUSE) to measure the mechanical properties (i.e., stiffness) of thyroid nodules and use this information to help classify nodules as benign or malignant. CUSE is a fast and robust 2-D shear elastography technique in which multiple laterally distributed acoustic radiation force beams are utilized simultaneously to produce shear waves. Unlike other shear elasticity imaging modalities, CUSE does not suffer from limited field of view (FOV) due to shear wave attenuation and can provide a large FOV at high frame rates. To evaluate the utility of CUSE in thyroid imaging, a preliminary study was performed on a group of five healthy volunteers and 10 patients with ultrasound-detected thyroid nodules prior to fine needle aspiration biopsy. The measured shear wave speeds in normal thyroid tissue and thyroid nodules were converted to Young's modulus (E), indicating a measure of tissue stiffness. Our results indicate an increase in E for thyroid nodules compared to normal thyroid tissue. This increase was significantly higher in malignant nodules compared to benign.

Pinle Qin, Kuan Wu, Yishan Hu, Jianchao Zeng, Xiangfei Chai," Diagnosis of benign and malignant thyroid nodules using combined conventional ultrasound and ultrasound elasticity imaging", IEEE Journal of Biomedical and Health Informatics. Ultrasonography is one of the main imaging methods for diagnosing thyroid nodules. Automatic differentiation between benign and malignant nodules in ultrasound images can great assist inexperienced clinicians in their diagnosis. The core of problem is the effective utilization of the features of ultrasound images. In this study, we propose a method that is based on the combination of conventional ultrasound and ultrasound elasticity images based on a convolutional neural network and introduces richer feature information for the classification of benign and malignant thyroid nodules. First, the conventional network model performs pretraining on ImageNet and transfers the feature parameters to the ultrasound image domain by transfer learning so that depth features may be extracted and small samples may be processed. Then, we combine the depth features of conventional ultrasound and ultrasound elasticity images to form a hybrid feature space. Finally, the classification is completed on the hybrid feature space, and an end-to-end CNN model is implemented.

We proposed a method for feature extraction and fusion of conventional ultrasound and ultrasound elasticity images for the differentiation of benign and malignant thyroid nodules. Considering the clinical practicality of ultrasound imaging, we used two different data sources, that is, conventional ultrasound and ultrasound elasticity images, to distinguish benign and malignant thyroid nodules.

III. PROPOSED SYSTEM:

In this project we apply deep learning algorithms to detect thyroid cancer. We will applying deep learning algorithm such as **VGG16**, etc.., to predict the highest accuracy. Detect its presence without the need of a number of consultations from different doctors. This leads to earlier prediction of the presence of the thyroid cancer and allows us to take prior actions immediately to avoid further consequences in an effective and cheap manner avoiding human error rate. Optimization and loss minimization techniques will be applied to increase the accuracy of model. Thus this project is used to predict the thyroid cancer effectively.

ADVANTAGES:

- An early prediction of thyroid cancer.
- Easy and cheap to determine the presence of thyroid cancer.
- Eliminates the human error rate.
- Saves time and efficient.

A. Dataset Handling and Pre-Processing

The images from the dataset are read one after the other using the imread function from the opencv library in Python. The output from imread is added to a larger matix. The matrix that holds the images is then saved in the pickle format. This file is then loaded and used while training the CNN. Similarly, another matrix is created to hold the labels for all the images.

B. Convolutional Neural Network

Convolutional Neural Networks are based on the idea that it is sufficient to have a local understanding of the image and it is not necessary that every pixel be connected to every other pixel like in the case of Fully Connected Networks. As a result, the number of parameters in CNN are lesser and do not depend on the dimensions of the input. Fully Connected Networks usually have an error percentage of 8-12% while on the other hand Convolutional neural networks only have an error percentage of 3-5%.

In this study, we have used a model which has 10 convolutional layers with small receptive fields (3x3), 5 max-pool layers (2x2 size) and three fully connected layers, with the final layer has a soft-max activation function. Fig 2 shows our network architecture. In this model we have used the ReLU (Rectified Linear Unit) activation function in the convolutional layers to increase the training speed without any change in the accuracy compared to non linear functions like Sigmoid and tanh. It also helps to alleviate problems of vanishing gradients, which is the issue where the lower layers of the network train very slowly because the gradient decreases exponentially through the layers. The ReLu layer applies the function f(x)=max(0,x) to all the values in the input volume, thus changing the negative values *to 0.*

C. Loss minimization

Training a model simply means learning (determining) good values for all the weights and the bias from labeled examples. In supervised learning, a machine learning algorithm builds a model by examining many examples and attempting to find a model that minimizes loss; this process is called loss minimization

Cross-Entropy

Cross-entropy loss, or log loss, measures the performance of a classification model whose output is a probability value between 0 and 1. Cross-entropy loss increases as the predicted probability diverges from the actual label. So predicting a probability of .012 when the actual observation label is 1 would be bad and result in a high loss value. A perfect model would have a log loss of 0.

The CE Loss is defined as:

$$CE = -\sum_{i}^{C} t_i log(s_i)$$

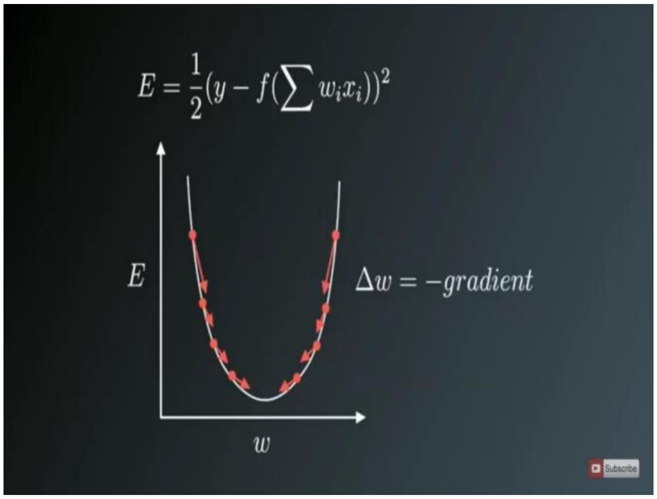
D. *Data Optimizer*

In this project, Gradient descent for optimizing the model. Gradient Descent is the most important technique and the foundation of how we train and optimize Intelligent Systems.Find the Minima, control the variance and then update the Model's parameters and finally lead us to Convergence. θ=θ−η⋅∇J (θ)— is the formula of the parameter updates, where 'η'is the learning rate,'∇J (θ)'is theGradient of Loss function-J **(**θ**)** w.r.t parameters-'θ'.

It is the most popular Optimization algorithms used in optimizing a Neural Network. Now gradient descent is majorly used to do Weights updates in a Neural Network

Model, i.e. update and tune the Model's parameters in a direction so that we can minimize the Loss function. Now we all know a Neural Network trains via a famous technique called Back propagation **,** in which we first propagate forward calculating the dot product of Inputs signals and their corresponding Weights and then apply a activation functionto those sum of products, which transforms the input signal to an output signal and also is important to model complex Non-linear functions and introduces Non-linearities to the Model which enables the Model to learn almost any arbitrary functional mappings.

After this we propagate backwards in the Network carrying Error terms and updating Weights values using Gradient Descent, in which we calculate the gradient of Error(E) functionwith respect to the Weights (W) or the parameters , and update the parameters (here Weights) in the opposite direction of the Gradient of the Loss function w.r.t to the Model's parameters.



The image on above shows the process of Weight updates in the opposite direction of the Gradient Vector of Error w.r.t to the Weights of the Network. The U-Shaped curve is the Gradient (slope). As one can notice if the Weight**(**W**)** values are too small or too large then we have large Errors, so want to update and optimize the weights such that is neither too small nor too large, so we descent downwards opposite to the Gradients until we find a local minimum.

E..EXPERIMENTAL RESULT

In this section, training strategies and test results were presented. Before the training phase all images are resized for the target network model And test results follows:-

PREDICTION	*ACCURACY*
Thyroid cancer	*97%*
Thyroid normal	*97%*

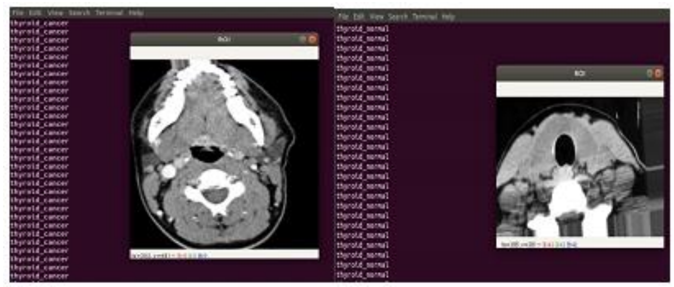
IV. FINAL RESULTS

In this project we have designed a thyroid cancer prediction system which can be used to predict the presence of cancer without human intervention**.**

To begin with, testing of the trained model, we can split our project into modules of implementation that is done. These datasets are then preprocessed to form an equal aspect ratio so that it can be made ready for training with the model. Dataset collection involves the process of collecting different thyroid cancer dataset.The datasets are separated into different categories to undergo preprocessing. After this the final implementation is done where the training process takes place and the results are obtained.

In order to run the code, first we need to go inside the project folder and select the location to run from there such that the location of the code is the main thing to be considered. Then as the code is run, the execution begins where the datasets are trained using the Deeper Google net architecture and the achieved an accuracy of 100% and has been printed that can be seen in the following figure

Then the recognize code is written which is used to predict the presence of thyroid cancer by importing the model generated after the training process. This can be seen in the following figures



Thus, from the above results and discussion, it is clear that we have efficiently made a project for predicting the thyroid cancer and achieved an accuracy of about 100% which is more efficient and can be easily applied in real time process. Thus, we have successfully implemented the scope of the project

V. CONCLUSION

This project is used to find the presence of thyroid cancer and provide prior measures to avoid the disease. This also help in providing efficient treatment in a most cheap way and eventually reduce the time required for finding the thyroid cancer in the current state. it is done manually which consumes more time and also involves human error rate. So, reduces the time required for manual classification and eliminates the human error rate by this project.

REFERENCES

[1] Beaumon, P. Onoma, M. Rimlinger, D. Broggio, P. Caldeira Ideias and D.Franck,"Age-specific experimental and computational calibration of thyroid in vivo monitoring",IEEE Transactions on Radiation and Plasma Medical Sciences,Vol.24697311,2018.

[2] Behnaz Jarrahi1,2, (Member, IEEE), Roger Gassert3, (Senior Member, IEEE),

[3] Johann Wanek4 , Lars Michels1 , Ulrich Mehnert4 , And Spyros S. Kollias1,"Design And Application Of a New Automated Fluidic Visceralstimulation Device For Human Fmri Studies Of Interoception",IEEE. Translations And Content Mining,Volume 4, 2016.
Jennifer E. Rosen∗, Hyunsuk Suh, Nicholas J. Giordano, Ousama M. A'amar,
Eladio Rodriguez-Diaz, Irving I. Bigio, and Stephanie L. Lee,"Preoperative Discrimination of Benign from Malignant Disease in Thyroid Nodules With
Indeterminate Cytology Using Elastic Light-Scattering Spectroscopy",IEEE Transactions On Biomedical Engineering, Vol. 61, No. 8, August 2014.

[4] Hui Zhou, Kun Wang, Jie Tian, Member, IEEE, "Online Transfer Learning for Differential Diagnosis of Benign and Malignant Thyroid Nodules with Ultrasound Images", IEEE Transactions on Biomedical Engineering,Vol.0018-9294,2019

[5] Julien Rouyer, Member, IEEE, Tony Cueva, Student Member, IEEE, Tamy Yamamoto, Alberto Portal, and Roberto Lavarello, Senior Member, IEEE,"In vivo

[6] Estimation of Attenuation and Backscatter Coefficients from Human Thyroids",IEEE Transactions on Ultrasonics , Ferroelectrics, and Frequency Control,Vol.0885-3010 (c) 2015.