DEEP LEARNING BASED THYROID TUMOR PREDICTION

Muthulakshmi M.
Department of ECE
Amrita School of Engineering
Amrita Vishwa Vidyapeetham
chennai, India
m_muthulakshmi@ch.amrita.edu

Harsha Vardhan A.
Department of CSE
Amrita School of computing
Amrita Vishwa Vidyapeetham
chennai, India
akulaharsha1435@gmail.com

Veda Sampreetha M.
Department of CSE
Amrita School of computing
Amrita Vishwa Vidyapeetham
chennai, India
veda.maganti@gmail.com

Hanuma Siva Sairam A.
Department of CSE
Amrita school of Computing
Amrita Vishwa Vidyapeetham
chennai, India
adapasairam090904@gmail.com

Syfullah Sd.
Department of CSE
Amrita school of Computing
Amrita Vishwa Vidyapeetham
chennai, India
syedsyfullah4@gmail.com

Sriram K.
Department of CSE
Amrita school of Computing
Amrita Vishwa Vidyapeetham
chennai, India
kdvsriram@gmail.com

***Abstract*—Thyroid tumors and disorders are serious health concerns and the necessity for accurate diagnostic procedures are required. In this paper, we are going to build a model that predicts thyroid tumors using ultrasonic images and classifies them into three categories – benign, malignant, and normal thyroid. Using a large dataset of ultrasonic images, we employ deep learning techniques for Restnet101-based feature extraction and pattern recognition. Through examining and cross–validating, we improve the model's performance and aim for high accuracy in categorizing between benign, malignant, and normal thyroid conditions and predicting the nodules. We implemented Resnet 101 which is a Convolutional neural network along with a training function to stop the iterations when there is no change in validation and training values to avoid overfitting and to reduce the time complexity. Our findings demonstrate that the model can efficiently aid doctors in making accurate diagnostic decisions, minimizing the need for invasive treatments, and improving patient outcomes. We Successfully deployed a model with an accuracy of 0.93 and Customised CNN having 4 layers with an accuracy of 0.873.**

***Index Terms*—Deep learning, ResNet 101, Data Augmentation, Image transformation, DLRT**

I. INTRODUCTION

Thyroid cancer is a significant malignant that develops in the cells of the thyroid gland and is very rare. A butterfly-shaped organ in the neck that produces hormones is termed as Thyroid that develops various bodily functions. Papillary and follicular are different types of thyroid cancers, followed by medullary and anaplastic thyroid cancers which vary in aggressiveness. Thyroid cancer is detected by the appearance of a lump or a nodule in the neck. Depending on the type and the stage of the cancer, the treatment is done which includes iodine therapy, chemotherapy, etc. In the recent survey in September 2021, over 52000 people affected with thyroid cancer in the US.

Deep Learning involves the usage of massive datasets, and deep neural networks which are trained to learn and recognize complex patterns and representations. Deep Learning models are modeled following the structure and function of the human brain learning models, such as deep neural networks process information through multiple layers of neurons or nodes, which does different tasks such as image and speech recognition, autonomous decision-making, and NLP. It has the ability to extract hierarchical features from data that revolutionize fields such as language processing, computer vision, and data analysis making it a potent tool for handling intricate and data-rich problems across various domains.

In this study, we are using ResNet101, shown in Fig.1. for detecting thyroid cancer. ResNet which is known as Residual Network, is a sophisticated deep learning architecture and is widely used in medical image analysis, which also includes detection of thyroid cancer. The use of residual connections, allows deep neural networks to train extremely without risk of any disappearing gradients. ResNet models are trained on medical images such as ultrasonic images, which enables the recognition of intricate patterns and abnormalities that help medical professionals detect thyroid cancer early.

In the context of training deep learning models in which neural networks are included there are techniques that are commonly used Image Transformation and Data Augmentation which increase model performance by enhancing or manipulating training data. The appearance of an image is altered without changing its semantic content in Image Transformation. Cropping, flipping, color adjustment, and rotation are some of the common transformations done to images.

A convolutional neural network(CNN) is designed for processing and analyzing visual data. Its architecture uses convolutional layers to capture spatial hierarchies of features in an input image. These layers apply kernels or filters that scan the input image and search for patterns like shapes, edges, and textures. Convolutional layers are followed by pooling layers, which downsample the spatial dimensions to decrease computational complexity and concentrate on the most essential features. CNNs are more effective in tasks like object detection, facial recognition, and image classification. It is capable of automatically extracting features from complex data.

CNNs use backpropagation and optimization techniques to tailor the network's weights and biases to decrease the error between expected and actual outputs. CNN is robust in handling multiple real-world image recognition techniques due to its iterative learning process, which allows it to adapt and generalize to diverse visual patterns. CNNs are well suited for tasks in computer vision due to their hierarchical feature extraction and shared parameter architecture.

II. Literature Survey

[1]Thyroid detection is done using a deep learning approach which is through computer-aided diagnosis for the detection of malignant nodules using ultrasonic images, in this method images are pre-processed to remove the artifacts, GoogleNet model which is a random forest classifier that is used to categorize the picture data according to "benign" and "malignant" scenarios after a deep learning model is employed for superior feature extraction.

TI-RADS model has been used for thyroid nodule classification which is accurate but is of high cost, and not only that TI-RADS method is time-consuming and is not so robust so it is less preferred for thyroid classification. [2]Many works have been proposed like texture-based features combining to support vector machines for the classification of thyroid nodules.Many researchers who studied thyroid nodules have come to the conclusion that, in comparison to clinical parameters like vascularity, form, and margins, many non-clinical variables—such as tissue stiffness scores, texture, etc. have an impact on the relevance of improved classification accuracy. So, they come up with a solution through which all these can be overcome and have given an effective result. In this process their method was Deep convolutional neural network which is proven to be effective in image classification, segmentation and retrieval[3]. Additionally, the DCNN model has the benefit of being resistant to distortion, such as changes in shape brought on by the camera lens, and having relatively cheap feature extraction costs because the input picture is transformed using the convolutional layer's coefficients.

The goal was to create a comprehensive system for classifying thyroid ultrasonic images by utilizing the refined GoogleNet model. The collected images were processed to extract features and remove artifacts, and sample augmentation[4] was used to fine-tune the GoogleNet model. They created 22 convolutional layers to train and validate the created data from. The accuracy of their model was 96.34% and 93.90% was obtained as specificity. Therefore, optimizing the current deep-learning network (DNN) has the benefit of using fewer training data to create a domain-specific DNN that can efficiently extract high-level features from thyroid ultrasound pictures[5] and accurately categorize benign and malignant thyroid images.
This work has the goal of automating the identification of thyroid nodules[6], in ultrasound scans, which addresses the increasing concern about thyroid cancer rates among women in North America. The paper highlights the limitations of detection methods that rely on experiences of radiologists. To overcome this the study introduces a Mask R CNN model[7]. This is a CNN architecture adjusted using a personalized dataset.Through a two-step process this model efficiently detects thyroid nodules by generating region proposals and predicting object categories leading to improved accuracy in bounding box predictions[8] and creating masks for each detected nodule. The study shows encouraging results in terms of accuracy (mAP), surpassing both classic Mask R CNN and Faster R CNN.

The ultrasonic scans of one hundred individuals with thyroid nodules were gathered for this investigation. These scans were divided into groups according to their content, size, margin, and echogenicity. Proficient annotation markers precisely designated contour masks and bounding boxes for annotations. During training various parameters such, as learning rates, momentum, and optimization using stochastic gradient descent[9] was fine-tuned. The model underwent 300 epochs of training starting with ImageNet weights while employing data augmentation techniques to prevent overfitting.The Mask R CNN model that was proposed has achieved a precision (mAP) score of 0.82 surpassing both the Faster R CNN and traditional Mask R CNN models. The improvement, in detecting thyroid nodules can be attributed to the regularization of the loss function.

[10] In this study, they provided a technique that uses a U Net-based learning model to identify thyroid nodules[11]. The goal is to address the time-consuming nature of detection processes and overcome challenges related to feature extraction. By comparing their proposed algorithm with deep networks like VGG19[12], Inception V3 and DenseNet 161 they have observed several advantages including improved accuracy in segmenting nodules, better quality the edges, and faster network operation time. U Net network achieves a segmentation accuracy rate of 97.85% which closely aligns with delineated nodule areas. Additionally, it outperforms the networks by 3% in terms of thyroid segmentation

accuracy. Overall their approach based on U Net presents an enhancement, in thyroid nodule segmentation accuracy especially when dealing with training data. It brings support to diagnosis and treatment processes.

[13] The ability to definitively diagnose "benign" and "malignant" thyroid nodules have been widely recognized. Nonetheless, its use is constrained by specimen collection and the user's expertise.Thus, Radiomics—which is based on machine learning and collects and evaluates many computed picture characteristics from medical images—is the sole basis for this study. These radiomics[14] are mostly and extensively used in the analysis of CT and MR images, and the results are impressive and effective. However, the use of these radiomics is very rarely discussed in the detection and classification of thyroid nodules.The model, dubbed DLRT (Deep Learning Radiomics of Thyroid)[15], is based on a transfer learning technique that uses a CNN.

They used a real-time database to gather ultrasound pictures of patients who had undergone a full year of treatment, totaling around 2300 ultrasound photos. The patients also underwent a test using ultrasound equipment and a standard style of thyroid examination.With 1003 "benign" and 642 "malignant" nodule pictures, 1629 patients were enrolled to train and validate the DLRT model. Many parameters in the DLRT model were improved, and after this validation, the ultrasound pictures were tested. The CNN architecture and transfer learning techniques were eventually adopted by the DLRT model, which was built on the Keras library.

The DLRT model compared to the CNN and Transfer Learning models, it was found that, although having the same network architecture as CNN, the main distinction was that the DLRT model was trained using ultrasound pictures; only the last layer of the DLRT model shared parameters with the transfer learning model. They have compared the accuracy of DLRT, basic CNN and [16]Transfer Learning (TL). DLRT had a training accuracy of approximately 0.96, internal validation of approximately 0.95, and external validation of approximately 0.97. TL had the second-highest accuracy, with a training accuracy of approximately 0.87, internal validation of approximately 0.85, and external validation of approximately 0.87. Basic CNN had the lowest accuracy, with a training accuracy of approximately 0.82, internal validation of approximately 0.81, and external validation of approximately 0.82. This necessitates further opportunity to improve thyroid nodule detection.

III. Methodology

Fig. 1 describes the flow of work that has been done to predict the thyroid cancer from given ultrasound images.

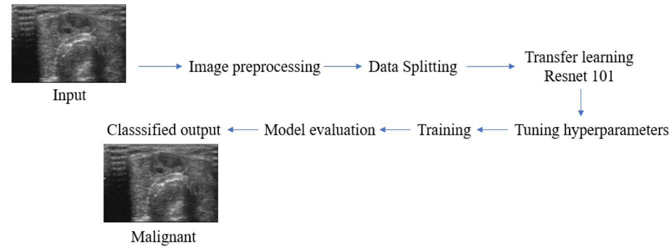

Fig. 1. Workflow

A. Data collection

Collected data from Algeria's hospitals, The dataset contains three classes:

Benign, Malignant, Normal

Benign with 1472 images, Malignant with 1895 and 171 Normal ultrasound images

B. Data Preprocessing

- Images that are collected should be resized to 256x256 pixels to make it similar to the input the size requirement of Resnet 101
- For consistent scaling pixel values should be Normalized and image preprocessing may include noise reduction, histogram equalization

C. Data Splitting

Splitting data into training, testing and validation. In our case, we split 72% of the data for training, 12% for validation, and 16% of data for testing.

D. Data Augmentation

The data that is split for training undergoes augmentation process. To increase the diversity of data set artificially which is used for training different augmentation processes are done to the existing data. To reduce overfitting and unseen data augmentation helps to generalize the model [17].

E. Transfer Learning

From Py-Torch or Tensor Flow load the pre – trained ResNet-101 model, This model is pre trained for large data sets.

F. Architecture of ResNet – 101

The architecture of Resnet 101 is shown in Fig. 2 and the description as follows:

- Residual Block – These are fundamental building blocks of ResNet 101[18]. There consist 2 paths of each block:
  1. Identity path
  2. Residual path

  The identity path pass the input simply to the next layer. Input information is being preserved. A series of convolutional and activation is performed in the residual path on the input Element wise Sum of identity path and the residual path is the output of residual block.
- Bottleneck Architecture – In some Residual blocks resnet 101 uses bottleneck architecture reduces computational complexity[19]. To reduce the dimension it includes 1x1

convolutional layer, 3x3 convolution layer and another 1x1 convolution layer to expand the dimension.

- Poling – A global average pooling(GAP) is used by ResNet 101 at the end of the network, This computes the average value of each feature map across the entire spatial dimensions

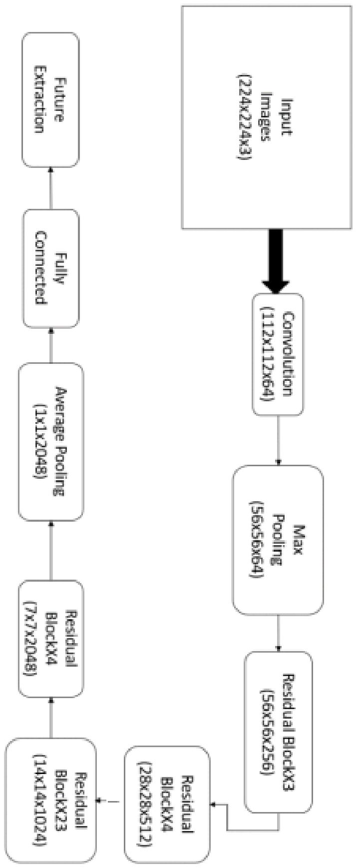

Fig. 2. Architecture of Resnet101

G. Training Function

- We had a training function by adding an early stop to avoid extra processing time. Training will stop if validation loss does not improve for 2 consecutive epochs. Best model weights, loss threshold,an a counter for consecutive epochs are included as variables in the function.
- Trains and validates the model for the given number of epochs, while training it calculate the loss and accuracy for each epoch and continuously change the weights during training.
- If there is no improvement in validation loss counter for consecutive epochs get increased, If the counter gets on the increase up to a specified value it triggers an early stopping and exiting the function.
- As a result it gives the best weights and measures classification metrics.

H. Fine Tuning and Training

- Fine Tune[20] will train only newly added layers for a smaller number of epochs and we used a low learning rate for fine-tuning.
- Performed 20 epochs which results precision, recall f1 score, and confusion matrix.

I. Workflow of CNN

Upto data splitting the model is similar to resnet 101. After splitting the data into test, train, and for validation, initialization of CNN[21] starts. Our CNN model has 4 layers and each layer is of 3x3 sized kernel which generates a single output by performing element-wise multiplication with a relu activation function and maxpooling size of 2x2. Layer 1 is builtthe with 64 filters that can generate an output of 64 different feature maps and layer 2 and layer 3 with 128, layer 4 and with 512 filters all with a 5x5 size kernel.
Adam optimizer is used to change the weights during the epochs with learning rates as 0.0005. Finally adds a dense layer to the CNN which squashes values between 0 and 1 with softmax activation function.

IV. Model Evaluation and Result

For CNN: Running for 10 epochs got a final accuracy of 0.873. The model is saved as my_disease.h5.The model is tested with the test data set that contains images of benign, malignant, and normal thyroid. By giving any of these images as input the model predicts which type it belongs to and prints its name and image as output.

Fig. 3 Fig. 4 and Fig. 5 Represent the predicted outputs by CNN as normal Thyroid, malignant, and benign respectively.

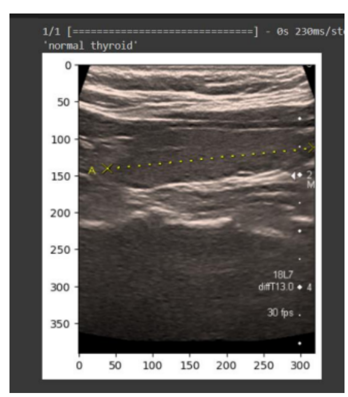

Fig. 3. Normal Thyroid

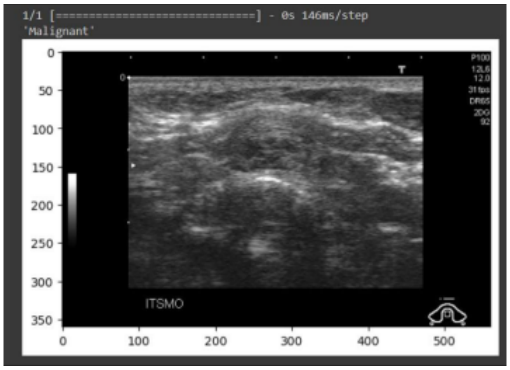

Fig. 4. Malignant

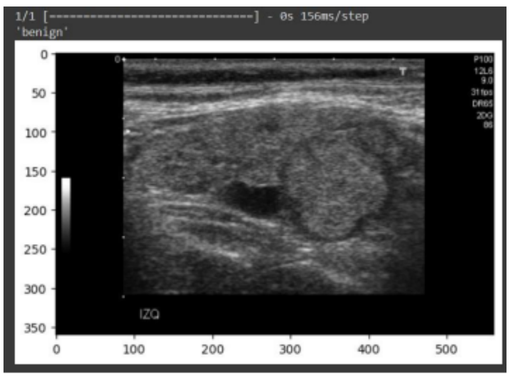

Fig. 5. Benign

For Resnet 101: Graphical Representation of Training accuracy and validation accuracy – Training loss and validation loss.
Fig. 6 and Fig. 7 graphically represent the training and validation of losses and accuracy respectively and stop the training at the 8th epoch as the readings are not changing for successive iterations.

Fig. 8 describes the confusion matrix where we tested the model with nearly 245 sample images where it predicted 130 benign as benign 5 benign as 4 malignant and 1 normal similar inference for Malignant and normal.

Fig. 6. Graphical representation of loss

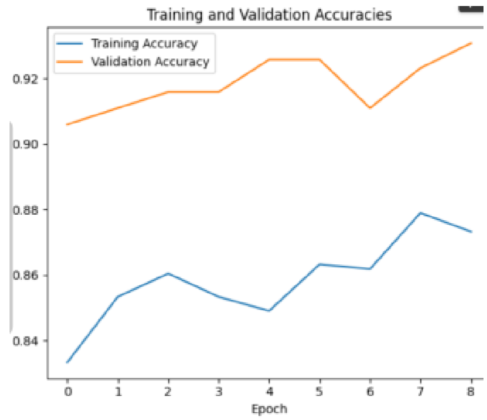

Fig. 7. Graphical representation of accuracy

TABLE I
RESULT

	Precision	Recall	F1 Score
Micro avg	0.94	0.91	0.92
Weighted avg	0.93	0.93	0.92

Accuracy : 0.933

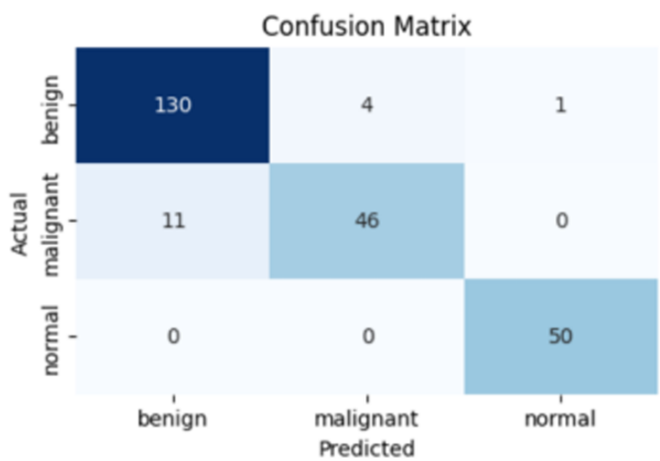

Fig. 8. Confusion Matrix

V. Conclusion

This work presents a groundbreaking approach to early Thyroid nodule prediction, based on Resnet 101 and a CNN model. By testing on data sets including ultrasound images of thyroid nodules with 3 classes of benign, malignant, and normal thyroid, it was demonstrated the model's performance and efficiency. Build a 4-layered custom cnn model with relu activation function having 3x3 filters. By testing on data sets including ultrasound images of thyroid nodules with 3 classes of benign, malignant, and normal thyroid, an accuracy of 0.87 was obtained. Further, we attained an accuracy of 0.933 with resnet 101.. Hence this research study can be considered as a proof of concept for the adoption of Resnet 101 for thyroid tumor Prediction using ultrasound images.

References

[1] Chi, Jianning, Ekta Walia, Paul Babyn, Jimmy Wang, Gary Groot, and Mark Eramian. "Thyroid nodule classification in ultrasound images by fine-tuning deep convolutional neural network." Journal of digital imaging 30 (2017): 477-486.

[2] Aksoy, Suleyman Hilmi, Ozlem Uygun, Isil Yurdaisik, Lora Ates, and Selami Aydin. "The relationship between ultrasound-based TIRADS and BETHESDA categories in patients undergoing thyroid biopsy." Clinical and Experimental Medicine 22, no. 4 (2022): 661-666.

[3] Zhang, Dingwen, Jiajia Zhang, Qiang Zhang, Jungong Han, Shu Zhang, and Junwei Han. "Automatic pancreas segmentation based on lightweight DCNN modules and spatial prior propagation." Pattern Recognition 114 (2021): 107762.

[4] Shorten, Connor, and Taghi M. Khoshgoftaar. "A survey on image data augmentation for deep learning." Journal of big data 6, no. 1 (2019): 1-48.

[5] Liu, Lingqiao, Chunhua Shen, and Anton Van den Hengel. "The treasure beneath convolutional layers: Cross convolutional-layer pooling for image classification." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4749-4757. 2015.

[6] Abdolali, Fatemeh, Jeevesh Kapur, Jacob L. Jaremko, Michelle Noga, Abhilash R. Hareendranathan, and Kumaradevan Punithakumar. "Automated thyroid nodule detection from ultrasound imaging using deep convolutional neural networks." Computers in Biology and Medicine 122 (2020): 103871.

[7] Bharati, Puja, and Ankita Pramanik. "Deep learning techniques—R-CNN to mask R-CNN: a survey." Computational Intelligence in Pattern Recognition: Proceedings of CIPR 2019 (2020): 657-668.

[8] Lempitsky, Victor, Pushmeet Kohli, Carsten Rother, and Toby Sharp. "Image segmentation with a bounding box prior." In 2009 IEEE 12th international conference on computer vision, pp. 277-284. IEEE, 2009.

[9] Netrapalli, Praneeth. "Stochastic gradient descent and its variants in machine learning." Journal of the Indian Institute of Science 99, no. 2 (2019): 201-213.

[10] Chu, Chen, Jihui Zheng, and Yong Zhou. "Ultrasonic thyroid nodule detection method based on U-Net network." Computer Methods and Programs in Biomedicine 199 (2021): 105906.

[11] Burman, Kenneth D., and Leonard Wartofsky. "Thyroid nodules." New England Journal of Medicine 373, no. 24 (2015): 2347-2356.

[12] Hsieh, Ya-Chu, Chiun-Li Chin, Chia-Shin Wei, I-Miao Chen, Pin-Yu Yeh, and Ru-Jiun Tseng. "Combining VGG16, Mask R-CNN and Inception V3 to identify the benign and malignant of breast microcalcification clusters." In 2020 International Conference on Fuzzy Theory and Its Applications (iFUZZY), pp. 1-4. IEEE, 2020.

[13] Zhou, Hui, Yinhua Jin, Lei Dai, Meiwu Zhang, Yuqin Qiu, Jie Tian, and Jianjun Zheng. "Differential diagnosis of benign and malignant thyroid nodules using deep learning radiomics of thyroid ultrasound images." European Journal of Radiology 127 (2020): 108992.

[14] Zhang, Bin, Xin He, Fusheng Ouyang, Dongsheng Gu, Yuhao Dong, Lu Zhang, Xiaokai Mo, Wenhui Huang, Jie Tian, and Shuixing Zhang. "Radiomic machine-learning classifiers for prognostic biomarkers of advanced nasopharyngeal carcinoma." Cancer letters 403 (2017): 21-27.

[15] Zhou, Hui, Kun Wang, and Jie Tian. "Deep learning radiomics for non-invasive diagnosis of benign and malignant thyroid nodules using ultrasound images." In Medical Imaging 2020: Ultrasonic Imaging and Tomography, vol. 11319, pp. 37-44. SPIE, 2020.

[16] Ajilisa, O. A., V. P. Jagathy Raj, and M. K. Sabu. "A Deep Learning Framework for the Characterization of Thyroid Nodules from Ultrasound Images Using Improved Inception Network and Multi-Level Transfer Learning." Diagnostics 13, no. 14 (2023): 2463.

[17] Shin, Hoo-Chang, Le Lu, and Ronald M. Summers. "Natural language processing for large-scale medical image analysis using deep learning." Deep learning for medical image analysis (2017): 405-421.

[18] Sarwinda, Devvi, Radifa Hilya Paradisa, Alhadi Bustamam, and Pinkie Anggia. "Deep learning in image classification using residual network (ResNet) variants for detection of colorectal cancer." Procedia Computer Science 179 (2021): 423-431.

[19] Ghosal, Palash, Lokesh Nandanwar, Swati Kanchan, Ashok Bhadra, Jayasree Chakraborty, and Debashis Nandi. "Brain tumor classification using ResNet-101 based squeeze and excitation deep neural network." In 2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP), pp. 1-6. IEEE, 2019.

[20] Moussa, Olfa, Hajer Khachnaoui, Ramzi Guetari, and Nawres Khlifa. "Thyroid nodules classification and diagnosis in ultrasound images using fine-tuning deep convolutional neural network." International Journal of Imaging Systems and Technology 30, no. 1 (2020): 185-195.

[21] Palaniappan, S., S. Varshaa Sai Sripriya, Amalladinna Rama Lalitha Pranathi, and M. Muthulakshmi. "Diagnosis of Acute Respiratory Syndromes from X-Rays using Customised CNN Architecture." In 2022 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET), pp. 1-5.IEEE,2022.