Proceedings of the Seventh International Conference on Research in Intelligent and Computing in Engineering pp. 129–133 DOI: 10.15439/2022R12 ISSN 2300-5963 ACSIS, Vol. 33

A Study on Thyroid Nodule Image Classification System Using Small Amount of Training Samples

1st Quoc Tuan Hoang
Faculty of Mechanical Engineering
Hung Yen University of Technology and Education
Hung Yen, Vietnam
hoangquoctuan@utehy.edu.vn

2nd Xuan Hien Pham
Department of Mechanical Engineering
University of Transport and Communications
Hanoi, Vietnam
hienpx@utc.edu.vn

3rd Xuan Thang Trinh
Faculty of Mechanical Engineering
Hung Yen University of Technology and Education
Hung Yen, Vietnam
xttrinh@utehy.edu.vn

***Abstract*—To reduce errors caused by traditional diagnostic methods that rely heavily on physician experience, the diagnostic systems based on computer-aided have been researched and developed to assist physicians in diagnosing thyroid disease. Therefore, performance of the computer systems plays an important role to improve the quality of diagnostic processes. Although there has been a number of publish related to this issue, those studies still have limitations in which needing large data sets for training classification models is considered the most concerning limitation of previous studies. To solve this limitations, in this work, a classification method using artificial intelligence with a small amount of data to analyze thyroid ultrasound images was proposed. Thus we can save time and effort for data collection and the classification model processing time. Through baseline tests with an open data set, especially thyroid digital image database (TDID), the proposed method has improved the limitations of previous methods.**

***Index Terms*—malignant thyroid nodule, deep learning, artificial intelligence, ultrasound image**

I. Introduction

Cancer is an extremely dangerous disease, characterized by the ability to invade and spread to other parts of the body. If cancer is not detected and treated in time, the patient will die quickly. Cancer cells grow and divide rapidly and have the ability to invade and spread to other organs. As noted, several cancers have a high incidence today, such as Breast cancer in women is 24.5%, the highest among all cancers in women. Lung cancer is also common cancer, accounting for about 12% of all cancers worldwide. In Vietnam, lung cancer ranks first among the 10 most common cancers in both males and females and is the leading cause of death. Thyroid cancer is a common disease with no obvious symptoms. It occurs in both men and women. This kind of cancer is often found in the area of the head, face and neck. Thyroid cancer cells can metastasize to several parts of the body such as the liver, lungs, brain, and bones, thereby making a higher mortality rate.

On the human body, the thyroid gland is an important organ and it is located in the human neck. The thyroid gland produces and secretes two important hormones, including thyroxine and triiodothyronine. They are responsible for regulating metabolism in the human body.

For these reasons, it is absolutely necessary and important to diagnose and treat thyroid disease [2-6]. In previous reports, they indicated that the appearance of thyroid cancer nodules was a common serious problem in the thyroid region. Thyroid nodules are formed by abnormal growths appearing in the thyroid area of the human body. Many factors can cause thyroid nodules, for example overgrowth of thyroid tissue, iodine deficiency or thyroid cancer. Based on their characteristics, thyroid nodules are usually classified into two types to be benign cases (which are noncancerous nodules) and malignant cases (which can cause thyroid cancer) [7]. In both cases, benign and malignant, the presence of thyroid nodules can cause health effects on the patient. The thyroid gland may be affected with the appearance of nodules. Benign cases have little effect on the patient's health, but affect the patient's appearance, making it difficult for the patient to breathe and swallow. Malignant cases can cause thyroid cancer. The diagnosis and treatment of malignancies become very important.

The traditional diagnosis method is mainly based on the doctor's knowledge and experience when examining the external manifestations such as the appearance of tumors with hard characteristics; clear margins; smooth or rough surface; the presence of cervical lymphadenopathy; large and solid tumor which is fixed in front of the neck; hoarseness; possibly difficulty in breathing, difficulty in swallowing or entanglement due to tumor compression, infiltrated or ulcerated with bleeding neck's skin. However, the major limitation of traditional diagnostic method is its performance is highly dependent on personal knowledge and the experience of the physician. For that reason, diagnostic performance varies and can take longer, these methods are laborious, invasive, and expensive.

In this century of growing technology, imaging-based diagnostic techniques have been widely used, especially computer-aided diagnostic (CAD) systems that have been developed to help physicians in the diagnostic and therapeutic process [8-10, 20]. In contrast, to the conventional thyroid diagnostic methods mentioned above, CAD methods used ultrasound images of thyroid nodules as input and provide information about nodules (benign or malignant) [2-6], thus limiting the unevenness of different doctor's qualifications at different hospital levels and leading to faster initial diagnosis. Technological advances such as back propagation algorithms, neural networks and graphics processing units (GPUs), and techniques based on deep learning have recently been applied to solve many problems in medical image processing systems [7,8,9,10, 11,12,]. Regarding detecting/classifying thyroid nodules, the "deep learning" method has achieved much success. As the name implies, methods based on "deep learning", such as convolutional neural networks (CNN), automatically learn image features to train and produce results with high performance. Therefore, a deep learning-based approach can produce superior diagnostic results.

Deep Learning has high accuracy with a flexible neural network architecture, which can be easily changed to fit many different problems in the training process. High automation, self-adjusting, self-optimizing and the capability of performing parallel computing make deep learning-based methods produce a good performance and can handle large amounts of training data. However, there are still some limitations of previous studies. For instance, the large amount of training data is needed. Therefore, it takes longer time and more effort to collect these data. Additionally, the image acquisition devices are expensive and moreover, it requires co-operation and agreement between doctors and patients [7]. The training process consumes long time. The good computer's hardware is required to train model. Therefore, it is inapplicable to apply to some problems in which a large amount of training data is not available, such as a new kind of disease indicated in [13,14,15]. In this work, to solve the above-mentioned problem of previous studies, we propose a thyroid nodule classification method that can be trained on a small number of training data, we can save time and effort for both data collection and training classification models. Therefrom, we design a deep-learning network that can learn to not only recognize images of a class but also distinguish images between classes.

II. Proposed Method

A. Network architecture

The overview of proposed network architecture is described in Fig. 1. As mentioned in section I, we design a network that can distinguish images from different class labels, and recognize images of the same class. Therefore, our proposed network accepts two input images, as shown in Fig. 1.

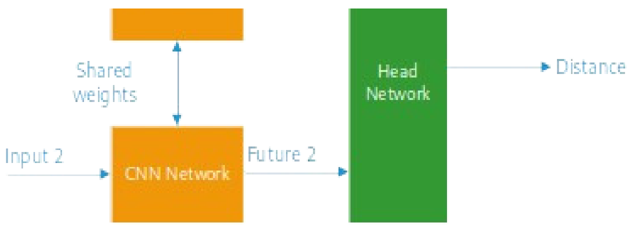

Fig 1. Overview of the proposed network.

In the training phase, our network receives two input images and uses a CNN network to extract the image feature vectors that are the best representation of the two input images. After that, the two feature vectors are transformed into another feature domain using a head network. The detailed descriptions of the CNN and head network are present in the next sections.

In the testing phase, a new input image is inputted to the network and our network will measure the distance from the input images to the training samples in the training dataset to find the best-matched samples with the input image. We used same network architecture with pre-trained weight for image feature extraction for both input 1 and input 2.

B. Related Work

Convolutional Neural Networks (CNNs), a type of artificial neural network, have become essential in various computer vision operations, and are receiving more attention across different published studies. For example, artificial intelligence association into brain magnetic resonance imaging (MRI) in cancer diagnosis [20], audiogram classification method [21], and cancerous region detection in the prostate [22]. The basic layers in a CNN include: convolution layer, pooling layer, fully connected layer, which are changed in number and arrangement to create suitable training models for different problems.

Convolutional Layer is the most important component in CNN, also the place to express the idea of building local connections instead of connecting all pixels. These local links are calculated by convolution between pixel values in a local image region with filters of small size. This filter is shifted through each image area in turn until it completes scanning the entire image, creating a new image that is less than or equal to the size of the input image. After feeding an image to the convolution layer, we get the output as a series of images corresponding to the filters used to perform the convolution. The weights of these filters are initialized randomly for the first time and will improve throughout training. The convolutional layer outputs a series of feature maps that enter the next convolutional or pooling layer.

Pooling Layer is another major computational component in CNN called Pooling, usually placed after the Convolution layer and ReLu layer to reduce the output image size while preserving the important information of the input image. Reducing the data size has the effect of reducing the number of parameters as well as increasing the computational efficiency. The sampling layer also uses a sliding window to scan all regions of the image similar to the Convolution layer and performs sampling instead of convolution, it mean we will choose to save a single representative value. for the entire information area. There are two commonly used Sampling methods, Max Pooling and Average Pooling. Thus, for each input image put through sampling, we obtain a corresponding output image, whose size is significantly reduced but still retains the necessary features for the later calculation process.

Fully-connected Layer is designed completely like a traditional Neuron network, it mean all pixels are fully connected to the Node in the next layer. Compared with the traditional neural network, the input image of this layer has been greatly reduced in size, while still ensuring important information for identification. Therefore, the recognition calculation using the feedforward model is no longer complicated and time consuming as in the traditional neural network. The final dense layer outputs the probabilities for classification by applying a softmax method.

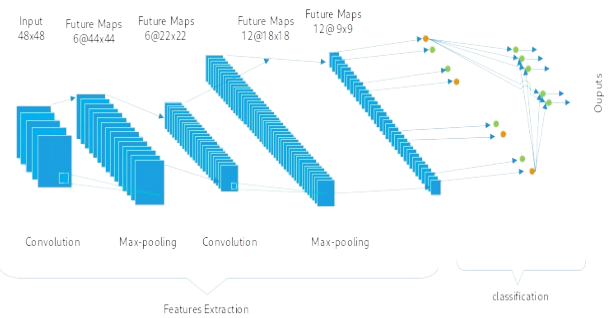

Fig. 2. Conventional CNN network.

Figure 2 shows the CNN network is used to extract image features for an input image. In our study, we use a pre-trained

CNN network trained on the ImageNet dataset to extract image features of input images. The reason is that the conventional CNN network contains a huge number of training parameters. Therefore, it is unable to train this kind of network using a small amount of training data. We called the CNN network in Figure 1 the black-bone in our study. The black-bone can be any CNN network that can be used for image feature extraction such as VGG, Residual, and Inception networks.

C. The head network

The head network based on an MLP network is shown in Fig. 3. The main part of our proposed network is the head network that has responsibility for learning the relationship between images within a class, as well as, among different classes. For that problem, we first transform the input image features to another domain using a conventional neural network. In the new image feature domain, the feature vectors of images that are in the same class are similar, while the feature vectors of different classes are made to be different. In order to obtain that purpose, we then measure the distance between two feature vectors from two input images (after CNN and domain transformation layers). That distance is forced to be 1 if the two input images are from the same class, and forced to be 0 if the input images are from different classes.

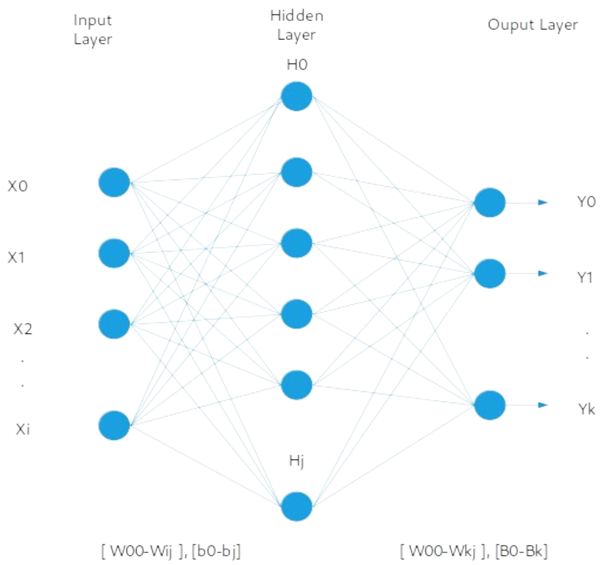

Fig 3. The head network is based on an MLP network.

The MLP network can contain several layers, each layer contains a number of neurons. In our experiments, we use three layers with number of neurons are 256 - 512 - 2. As shown in Figure 3.

III. DATA SET

In this study, the realization of the proposed method is demonstrated using a thyroid ultrasound dataset that has been published under the name Thyroid Digital Imaging Database (TDID). In 2015, at the University of Colombia Nacional de Colombia, Pedraza et al collected and published this dataset [16]. The TDID dataset contains thyroid ultrasound images of 298 patients. Ultrasound images of each patient's thyroid region, which can be single or multiple images, were collected in RGB format with 560 pixels × 360 pixels in image size. As a result, a dataset consisting of 450 thyroid nodule images was extracted for our experiments.

Fig. 4 shows some example thyroid nodule images in the TDID dataset. As we can see that it is hard to distinguish the normal and disease thyroid image by human perception. That is why the computer aid diagnosis system is helpful. It is little difference between the two kinds of images, that is, the disease thyroid nodule images contain nodules with brighter nodule regions with the effect of the calcification phenomenon.

(a)

(b)

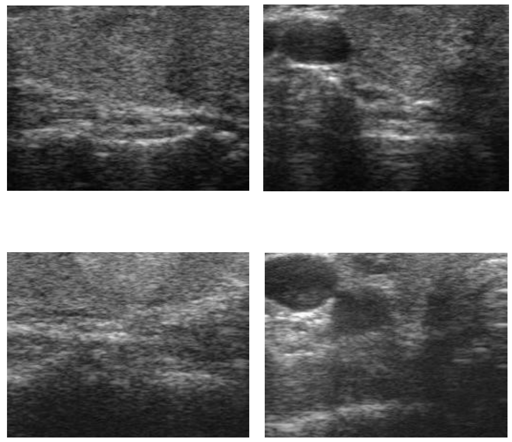

Fig. 4. Example of (a) normal, and (b) disease images in TDID dataset [16].

IV. EXPERIMENTAL RESULTS

As mentioned in our previous section, we are dealing with the classification problem of the medical image that has a small amount of data for training. For our purpose, we divided the TDID dataset (mentioned in Section III), into five parts to perform a five-fold cross-validation method. However, we use data of a single part as the training data, while the images of the other four parts are used as the testing data. This procedure is different from previous studies that use data from four parts for training, and the data of the remaining part is used for testing. Because of different training and testing data, the classification performance of conventional CNN-based methods is different as shown in previous studies. For comparison purposes, we also evaluate the classification performance of two popular CNN-based classification networks, including the Inception and Residual network architectures [17].

In order to evaluate the classification accuracy of the proposed network and the conventional CNN network, we use the classification accuracy metric as shown in (1).

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN) \tag{1}$$

In (1), the TP is the measure of true positive samples that is the case when an image with the disease was correctly classified as a with-disease image. The TN indicates the true negative samples which is the case when the normal image was correctly classified as a normal one. The FP stands for false-positive, the FN stands for false-negative. For a fair comparison with the Residual and Inception networks, we also use the Residual and Inception network as the black-bone of our proposed method.

The experimental results are given in Table 1 and Table 2 for the cases of using a Residual network, and Inception network architecture.

TABLE I. CLASSIFICATION PERFORMANCE OF OUR PROPOSED NETWORK AND PREVIOUS CNN-BASED NETWORK USING RESIDUAL NETWORK ARCHITECTURE (UNIT: %)

Method	Part 1	Part 2	Part 3	Part 4	Part 5	Average
Residual Network	79.619	84.091	75.414	70.055	78.177	77.472
Proposed Network with Residual Black-bone	83.152	84.091	85.912	81.319	79.282	82.751

TABLE II. CLASSIFICATION PERFORMANCE OF OUR PROPOSED NETWORK AND PREVIOUS CNN-BASED NETWORK USING INCEPTION NETWORK ARCHITECTURE (UNIT: %)

Method	Part 1	Part 2	Part 3	Part 4	Part 5	Average
Inception Network	85.054	82.102	85.635	84.066	84.254	84.222
Proposed Network with Inception Black-bone	86.141	84.375	86.188	84.890	85.083	85.335

For reference, we performed experiments with the residual and inception networks and performances are given in Table I and II. To measure the performance with the residual and inception network, we use fine-tuning the network by replaced the number of output neuron by 2 (original output neuron of these networks is 1000) and kept all convolution layers are same. [18-19].

As we can have observed from Table I, the Residual network produced an average classification accuracy of 77,47%. Using our proposed method, we enhanced that accuracy to 82,75%. This result indicates that our proposed network with the Residual network as black-bone outperforms the conventional residual network for the classification problem. A similar result was obtained in Table 2 using the Inception network architectures. As we can see in Table 2, our proposed network with the Inception architectures also outperform the conventional Inception network by producing an average classification accuracy of 85,33% that is higher than 84,22% produced by using the inception network.

As shown in Tables 1 and Tables 2, the classification performance of Inception Network–the based architecture was better than that of Residual Network-based architecture. The reason is that the Inception network-based architecture uses multiple convolution kernels to extract image features. As a result, it is more efficient to extract both large and small features than the Residual network-based architecture.

However, our proposed method outperforms all the two popular CNN networks (Residual and Inception network). It indicates that our proposed method is more efficient than the conventional CNN network when working with the lack of training data problem.

Although our proposed method produced higher classification accuracy than the convention networks network, it takes a longer processing time than the convention CNN network as it must measure the similarity from an input image to all images in the training set. However, it is acceptable with the medical image processing system in which accuracy plays a more important meaning than the processing time. In addition, with the support of graphical processing units (GPUs), the processing time is much reduced when we use strong computer specifications.

V. CONCLUSION

In this paper, we focus on solving a specific problem of medical image classification problem which is when a large amount of training data is not available. In order to solve this problem, we design a deep learning-based classification method based on an image-retrieval approach. In detail, we construct a system that has the ability to recognize images of the same class while distinguishing images between classes. Through experiments, we showed that our proposed method outperforms two conventional CNN-based methods (Residual-based and Inception-based CNN network architecture).

REFERENCES

[1] TDID dataset Pedraza, L.; Vargas, C.; Narvaez, F.; Duran, O.; Munoz, E.; Romero, E. An open access thyroid ultrasoundimage database. In Proceedings of the 10th International Symposium on Medical Information Processing and Analysis, Cartagena de Indias, Colombia, 28 January 2015; pp. 1–6.

[2] Koundal, D.; Gupta, S.; Signh, S. Computer aided thyroid nodule detection system using medical ultrasound images. Biomed. Signal Process. Control 2018, 40, 117–130.

[3] Tessler, F.N.; Middleton, W.D.; Grant, E.G.; Hoang, J.K.; Berland, L.L.; Teefey, S.A.; Cronan, J.J.; Beland, M.D.; Desser, T.S.; Frates, M.C.; et al. ACR thyroid imaging, reporting and data system (TI-RADS): White paper of the ACR TI-RADS committee. J. Am. Coll. Radiol. 2017, 14, 587–595.

[4] Ma, J.; Wu, F.; Zhu, J.; Xu, D.; Kong, D. A pretrained convolutional neural network based method for thyroid nodule diagnosis. Ultrasonics 2017, 73, 221–230.

[5] Pedraza, L.; Vargas, C.; Narvaez, F.; Duran, O.; Munoz, E.; Romero, E. An open access thyroid ultrasoundimage database. In Proceedings of the 10th International Symposium on Medical Information Processing and Analysis, Cartagena de Indias, Colombia, 28 January 2015; pp. 1–6.

[6] Zhu, Y.; Fu, Z.; Fei, J. An image augmentation method using convolutional network for thyroid nodule classification by transfer learning. In Proceedings of the 3rd IEEE International Conference on Computer and Communication, Chengdu, China, 13–16 December 2017; pp. 1819–1823.

[7] Nguyen, D.T.; Pham, D.T.; Batchuluun, G.; Yoon, H.S.; Park, K.R. Artificial intelligence-based thyroid nodule classification using information from spatial and frequency domains. J. Clin. Med. 2019, 8, 1976.

[8] Vuong, Q.H.; Ho, M.T.; Vuong, T.T.; La, V.P.; Ho, M.T.; Nghiem, K.C.P.; Tran, B.X.; Giang, H.H.; Giang, T.V.; Latkin, C.; et al. Artificial intelligence vs. natural stupidity: Evaluating AI readiness for the Vietnamese medical information system. J. Clin. Med. 2019, 8, 168.

[9] Havaei, M.; Davy, A.; Warde-Farley, D.; Biard, A.; Courville, A.; Bengio, Y.; Pal, C.; Jodoin, P.-M.; Larochelle, H. Brain tumor segmentation with deep neural networks. Med. Image Anal. 2017, 35, 18–31.

[10] Bhandary, A. et al. Deep-learning framework to detect lung abnormality – A study with chest x-ray and lung CT scan images. Pattern Recogn. Lett. 2020, 129, 271–278.

[11] Kamnitsas, K.; Ledig, C.; Newcombe, V.F.J.; Simpson, J.P.; Kane, A.D.; Menon, D.K.; Rueckert, D.; Glocker, B. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. Med. Image Anal. 2017, 36, 61–78.

[12] Zhu, Y.; Fu, Z.; Fei, J. An image augmentation method using convolutional network for thyroid nodule classification by transfer learning. In Proceedings of the 3rd IEEE International Conference on Computer and Communication, Chengdu, China, 13–16 December 2017; pp. 1819–1823.

[13] Vuong, Q.H.; Ho, M.T.; Vuong, T.T.; La, V.P.; Ho, M.T.; Nghiem, K.C.P.; Tran, B.X.; Giang, H.H.; Giang,T.V.; Latkin, C.; et al. Artificial intelligence vs. natural stupidity: Evaluating AI readiness for the Vietnamese medical information system. J. Clin. Med. 2019, 8, 168.Sensors 2020, 20, 1822 21 of 23

[14] Havaei, M.; Davy, A.; Warde-Farley, D.; Biard, A.; Courville, A.; Bengio, Y.; Pal, C.; Jodoin, P.-M.; Larochelle, H. Brain tumor segmentation with deep neural networks. Med. Image Anal. 2017, 35, 18–31.

[15] Bhandary, A. et al. Deep-learning framework to detect lung abnormality – A study with chest x-ray and lung CT scan images. Pattern Recogn. Lett. 2020, 129, 271–278.

[16] Pedraza, L.; Vargas, C.; Narvaez, F.; Duran, O.; Munoz, E.; Romero, E. An open access thyroid ultrasoundimage database. In Proceedings of the 10th International Symposium on Medical Information Processing and Analysis, Cartagena de Indias, Colombia, 28 January 2015; pp. 1–6.

[17] Introduction to Convolutional Neural Networks (CNN), Oct. 2022, [Online]. Available: https://www.analyticsvidhya.com/blog/2021/05/convolutional-neural-networks-cnn/

[18] Nikhila, Ponugoti, et al. "Lightweight residual network for the classification of thyroid nodules." arXiv preprint arXiv:1911.08303 (2019).

[19] Guan, Qing, et al. "Deep learning based classification of ultrasound images for thyroid nodules: a large scale of pilot study." Annals of translational medicine 7.7 (2019).

[20] Duong-Trung, Nghia, Dung Ngoc Le Ha, and Hiep Xuan Huynh. "Classification-Segmentation Pipeline for MRI via Transfer Learning and Residual Networks." RICE. 2021.

[21] Kassjański, Michał, Marcin Kulawiak, and Tomasz Przewoźny. "Development of an AI-based audiogram classification method for patient referral." 2022 17th Conference on Computer Science and Intelligence Systems (FedCSIS). IEEE, 2022.

[22] Treigys, Povilas, et al. "Detecting Cancerous Regions in DCE MRI using Functional Data, XGboost and Neural Networks." Annals of Computer Science and Information Systems 32 (2022): 23-30.