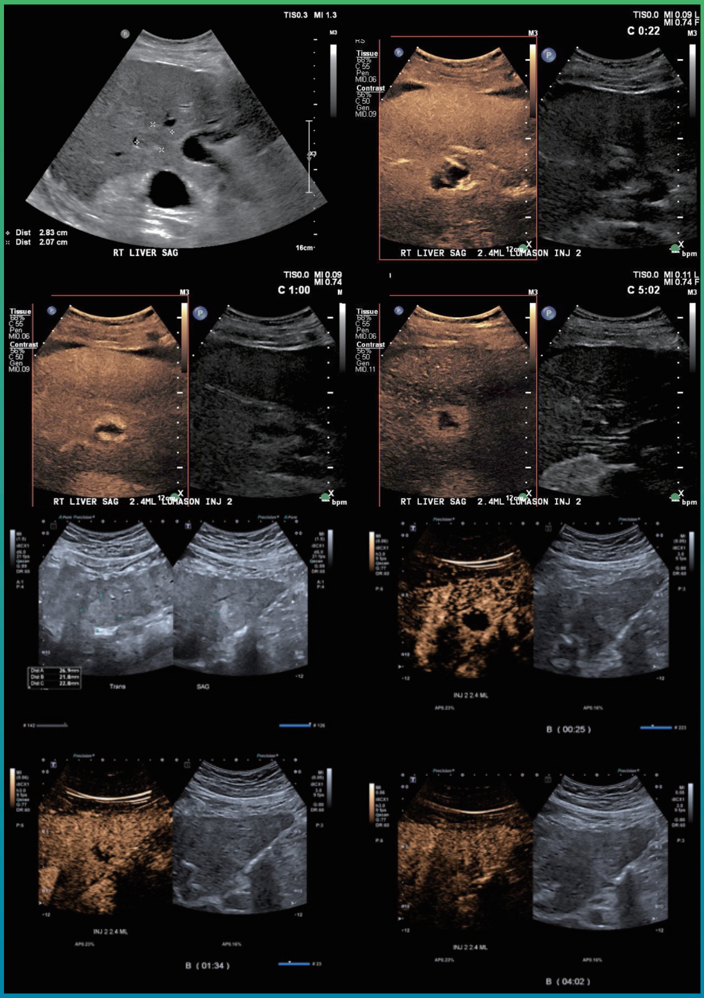
AUDT

Advanced Ultrasound

in Diagnosis and Therapy

December, 2023
Volume: 7
Issue: 4
Pages: 313–422



The official journal of the
CAUSA
China-America Ultrasound Scholar Alliance

https://www.AUDT.org
ISSN: 2576-2508 (Print)
ISSN: 2576-2516 (Online)

CONTENTS

Review Articles

Original Research

Case Reports

ORIGINAL RESEARCH

A Non-Invasive Follicular Thyroid Cancer Risk Prediction System Based on Deep Hybrid Multi-feature Fusion Network

Yalin Wu, PhD [a,1], Qiaoli Ge, MM [a,1], Linyang Yan, PhD [a,1], Desheng Sun, MD [a,]*

[a] *Department of Ultrasound, Peking University Shenzhen Hospital, Shenzhen, China*

Received April 2, 2023; revision received May 14, 2023; accepted June 4, 2023

Objective: A non-invasive assessment of the risk of benign and malignant follicular thyroid cancer is invaluable in the choice of treatment options. The extraction and fusion of multidimensional features from ultrasound images of follicular thyroid cancer is decisive in improving the accuracy of identifying benign and malignant thyroid cancer. This paper presents a non-invasive preoperative benign and malignant risk assessment system for follicular thyroid cancer, based on the proposed deep feature extraction and fusion of ultrasound images of follicular thyroid cancer.

Methods: First, this study uses a convolution neural network (CNN) to obtain a global feature map of the image, and the fusion of global features cropped to local features to identify tumor images. Secondly, this tumour image is also extracted by googleNet and ResNet respectively to extract features and recognize the image. Finally, we employ an averaging algorithm to obtain the final recognition results.

Results: The experimental results show that the method proposed in this study achieved 89.95% accuracy, 88.46% sensitivity, 91.30% specificity and an AUC value of 96.69% in the local dataset obtained from Peking University Shenzhen Hospital, all of which are far superior to other models.

Conclusion: In this study, a non-invasive risk prediction system is proposed for ultrasound images of thyroid follicular tumours. We solve the problem of unbalanced sample distribution by means of an image enhancement algorithm. In order to obtain enough features to differentiate ultrasound images, a three-branched feature extraction network was designed in this study, and a balance of sensitivity and specificity is ensured by an averaging algorithm.

Key words: Follicular thyroid cancer; Ultrasound image; Risk prediction system; Hybrid multi-feature fusion; Convolutional neural network

Advanced Ultrasound in Diagnosis and Therapy 2023; 04: 373-380 DOI: 10.37015/AUDT.2023.230023

Follicular thyroid carcinoma (FTC) is the second most common thyroid malignancy, with approximately 10%-15% of the total incidence of thyroid cancer [1,2]. However, while FTC is relatively rare compared to papillary thyroid carcinoma (PTC), the most common type of thyroid cancer, it is prone to hematogenous metastases to the lung and bone in the early stages of the disease, and some patients are not diagnosed until after the tumour has metastasised, resulting in a higher mortality rate compared to PTC [3].

Ultrasound is the imaging method of choice for thyroid disease with significant implications for the preoperative determination of benign and malignant thyroid nodules. Due to the fact that follicular thyroid

[1] *Yalin Wu, Qiaoli Ge and Linyang Yan contributed equally to this study.*

[*] *Corresponding author: Department of Ultrasonography, Peking University Shenzhen Hospital, 1120 Lianhua Road, Shenzhen, Guangdong, China*
e-mail: szdssun@163.com

tumours include FTC and benign follicular thyroid adenoma (FTA), however, FTC and FTA are not easily distinguished by ultrasound sonography alone [4,5]. In addition, ultrasound-guided fine needle aspiration biopsy (FNAB) is currently the "gold standard" for the preoperative diagnosis of benign and malignant thyroid nodules. However, since both FTC and FTA originate from thyroid follicular cells and therefore have very similar cytomorphological features, the differential value of fine needle aspiration pathology is limited [6,7]. Follicular thyroid tumours that do not show obvious signs of malignancy currently still require surgical excision and are judged to be benign or malignant based on the presence of envelope, vascularity, extra-thyroidal tissue invasion and distant metastases [8]. However, studies have shown that only about 20%-30% of follicular thyroid tumours are proven to be malignant after surgery [9,10], meaning that most patients undergo diagnostic thyroid lobectomy despite the benign nature of the tumour, causing heavy emotional stress and financial burden to the patient. For this reason, it is urgent to have a non-invasive and effective method to predict the risk of benign and malignant thyroid follicular tumours and to guide the clinical diagnosis and treatment process.

Ultrasound images are an important branch of medical imaging, which like CT and MRI images also take the content of the images as an important basis for diagnosis. Given that ultrasound images contain much critical information beyond the human eye's perception of the underlying physiology of a tumour [11,12], the application of machine learning to the recognition of ultrasound images and the development of a powerful, non-invasive automated tumour risk prediction system has become a research hotspot in ultrasound-based precision oncology [13]. Extraction of features from ultrasound images, construction of classifiers based on the extracted features and their application to the prediction of benign and malignant thyroid nodules have yielded some interesting results. In general, there are two main types of feature extraction methods for images, namely automatic feature extraction and manual feature extraction [13,14]. Currently, the two approaches [13,14] are applied to feature extraction and identification of ultrasound images of follicular thyroid tumours often give conflicting results compared to experienced radiologists [15,16]. The reasons are as follows: similarity of benign and malignant image features requiring deeper feature extraction; insufficient sample size; and highly unbalanced distribution of benign and malignant.

In order to solve the problems and to obtain reliable results that can be used in clinical applications, the deep hybrid multi-feature fusion network (DHMFN) has proposed in this study. The proposed DHMFN uses a convolutional neural network (CNN) to obtain a global feature map of the image, and the fusion of global features cropped to local features to identify tumor images. The tumour image is also extracted by googleNet and ResNet respectively to extract features and recognize the image. we employ an averaging algorithm algorithm to obtain the final recognition results.

The rest of this paper is organized as follows: In Section 2, several classical feature extraction algorithms are presented.Then, we describe our proposed DHMFN in Section 3 which includes image collection, preprocessing, feature extraction and classification. Experiments have been carried out on the proposed scheme and the results are discussed in Section 4. Section 5 concludes the paper with summary and future research directions.

Related Work

Selected image feature extraction methods relevant to this study are presented in Section 2.

Convolutional neural network (CNN)

One neural network specifically designed for the image recognition problem is the CNN [17,18]. The CNN mimics the human process of recognising images in its many layers: ingestion of pixels by the pupil; rudimentary processing by certain cells in the cerebral cortex to find shape edges and orientation; as well as abstractly determining shapes (e.g. circles and squares) and further abstracting decisions (e.g. judging objects as balloons.) CNNs typically contain five layers: an input layer, a convolutional layer, an ensemble layer, a fully connected layer and an output layer (Fig. 1).

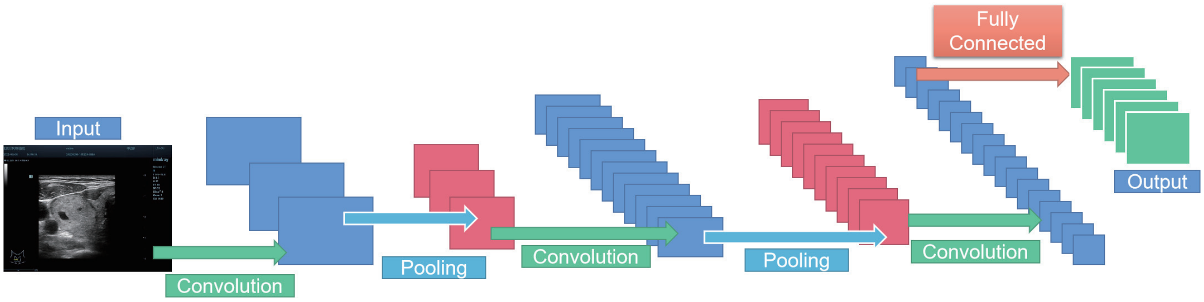


Figure 1 CNN model.

Inception v3

In 2014, Google proposed an improved CNN model, which is called the GoogLeNet network. Compared to traditional CNNs, GoogLeNet networks with the Inception network structure not only reduce the number of network parameters, but also increase the network depth. A module of Inception [19,20] generally involves three different sizes of convolution and a maximum pooling. A module of Inception generally involves three different sizes of convolution and a maximum pooling. The output of the network for the previous layer, following the convolution operation, the channels are aggregated and the fusion is then performed non-linearly. Through this the expressiveness of the network and its adaptability to different scales is improved and over-fitting in this way can be prevented. By 2016, the GoogLeNet network already had been developed with Inception v1 (2014), Inception v2 (2015), Inception v3 (2015), Inception v4 (2016) and Inception-ResNe (2016). The Inception v3 network structure uses a small convolutional kernel which has a 3 × 3 convolution split into 3 × 1 and 1 × 3 convolutions using a grid of 35 × 35, 17 × 17 and 8 × 8 regions (Fig. 2).

ResNet-50

ResNet is an abbreviated form of Residual Network, which is a special form of CNN similar to the GoogLeNet network [21]. ResNet can be considered to be stripped down from the VGG neural network, which has less filters and a simpler structure than the VGG neural network. The ResNet is initially made up of 34 weighted layers which use an algorithm called "shortcut connection" which converts a normal network into a residual network, thus solving the problem of gradient disappearance in normal CNNs. Currently, the more popular network structure for ResNet is ResNet-50, ResNet-50 has an architecture based on the model depicted above, but with one important difference. The 50-layer ResNet uses a bottleneck design for the building block. A bottleneck residual block uses 1 × 1 convolutions, known as a "bottleneck", which reduces the number of parameters and matrix multiplications. This enables much faster training of each layer. It uses a stack of three layers rather than two layers (Fig. 3) .

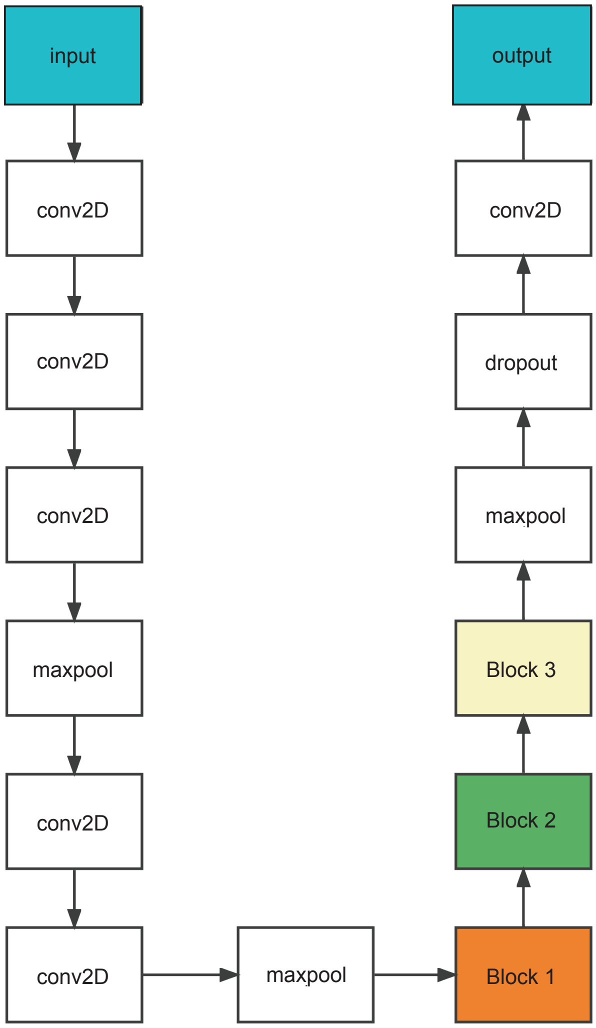


Figure 2 Inception module.

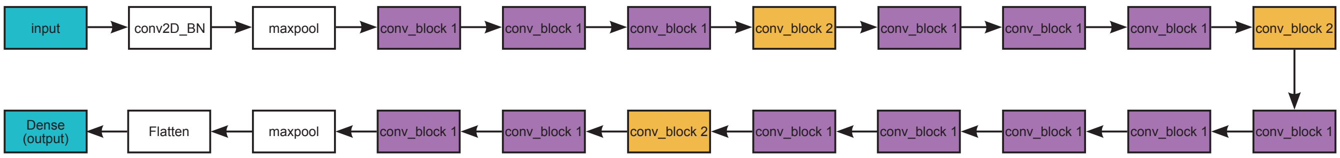


Figure 3 ResNet-50 module.

Proposed Method

The main steps in machine learning include the following: data collection, data pre-processing, feature extraction, model training and evaluation [22,23].

Data collection

This study retrospectively collected preoperative ultrasound sonograms and general clinical data of thyroidectomy cases with postoperative pathologically confirmed FTC or FTA performed at Peking University Shenzhen Hospital from January 1, 2015 to December 31, 2020 and followed up the nodules in the FTA group until December 31, 2022. Also, this study excludes unsuitable data based on the following criteria: (1) Pre-operative ultrasound images are not available or the quality of the pre-operative ultrasound images acquired is poor and cannot be utilised, for example: poor clarity, containing numbers, measurement lines, and so on. (2) Nodes in the FTA group that metastasised during follow-up.

The Ethics Committee (EC) of Peking University Shenzhen Hospital approved the retrospective study and waived the informed consent requirement for the study population.

A total of 269 nodules are eventually included in 263 patients, 57 males and 212 females, aged 13 to 77 years, with a mean age of (38.4 ± 11.9) years. 229 cases of 230 lesions in FTA and 34 cases of 39 lesions in FTC were included in the 263 patients. A total of 493 preoperative ultrasound images of follicular thyroid tumours were included, with a total of 417 images in the FTA group and 77 images in the FTC group (Fig. 4).

Image pre-processing

The impact of image pre-processing on the generalisation of neural networks is very important. Image pre-processing can help the model to better capture image features and thus better generalise to new images [22]. In particular, the ratio of benign to malignant samples was extremely uneven in this study, with malignant samples accounting for only 16% of the entire sample. In order to solve the problem of unbalanced sample distribution, this study incorporates image augmentation algorithm processing in the image pre-processing stage. In the image augmentation strategy, each malignant image is processed in this study by 'Rotating', 'Flipping', 'Adding Gaussian Noise', 'Brighing' and 'Adding Gaussian Blur', respectively (Fig. 5).

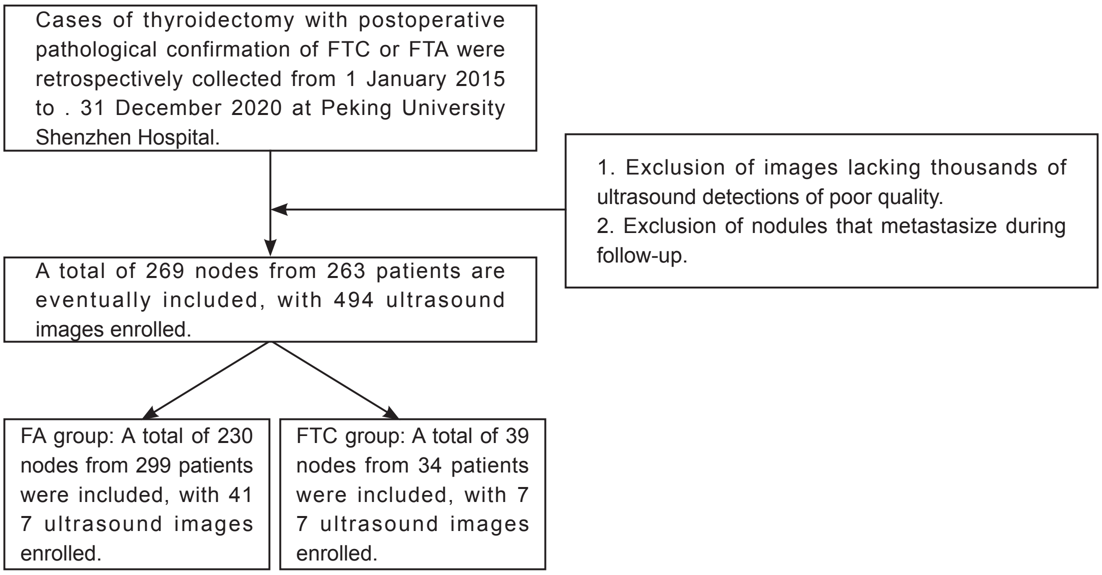


Figure 4 Inclusion diagram of patient population.

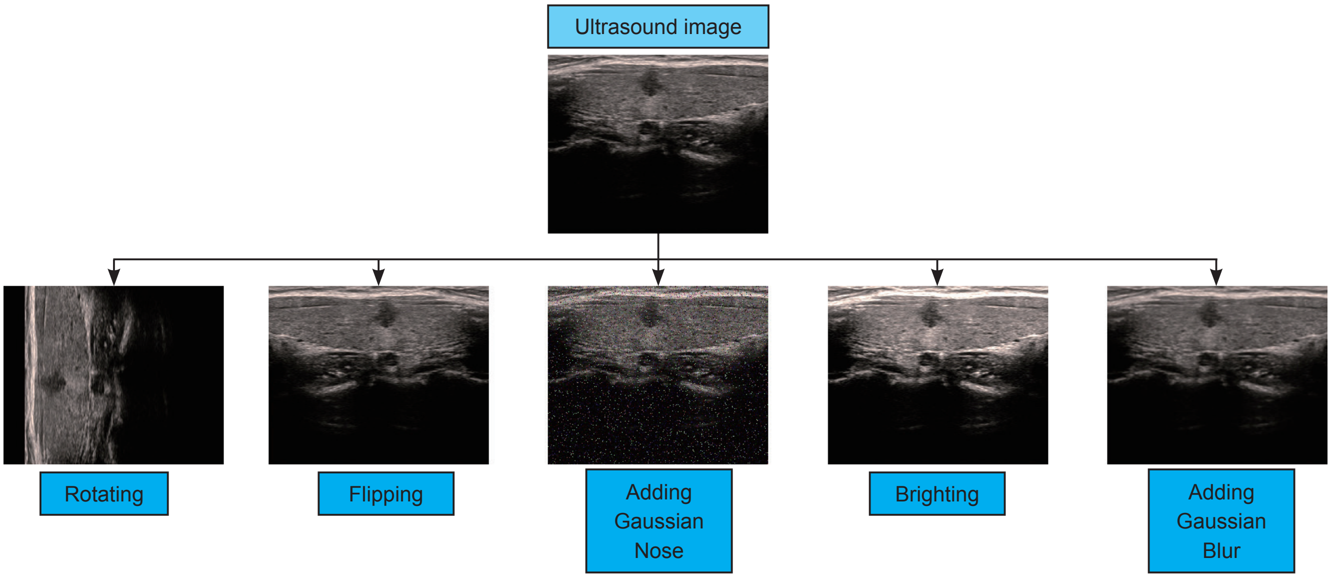


Figure 5 Image augmentation strategy.

Deep hybrid multi-feature extraction

The contribution of feature extraction to the overall machine learning process is crucial and can even affect the performance of the entire risk prediction system. The ultrasound images of benign and malignant follicular thyroid tumours are very similar (Fig. 6) and the practitioners having difficulty distinguishing between them by the naked eye. As seen in many previous studies [14,15], conventional machine algorithms have also been ineffective in distinguishing between benign and malignant ultrasound images of follicular thyroid tumours. An explanation for the problem is that the feature similarity between benign and malignant ultrasound images obtained by current feature extraction algorithms is extremely high and is not sufficient to distinguish between the classes of images.

This study proposes a deep hybrid multi-feature extraction method in order to obtain features that are sufficient to allow the images to be correctly classified.

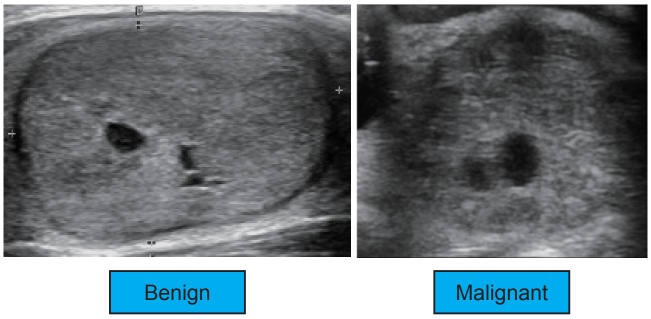


Figure 6 Ultrasound images of follicular tumours of the thyroid.

The proposed feature extraction scheme is composed of three branches: transfer learning features, cropped features and residual network features (Fig. 7).

Inception V1 is used in this study to obtain transfer learning features and to derive predictions of benign and malignant thyroid follicular tumours (Fig. 8). And ResNet-50 has taken on the acquisition of the residual features and the prediction of the results.

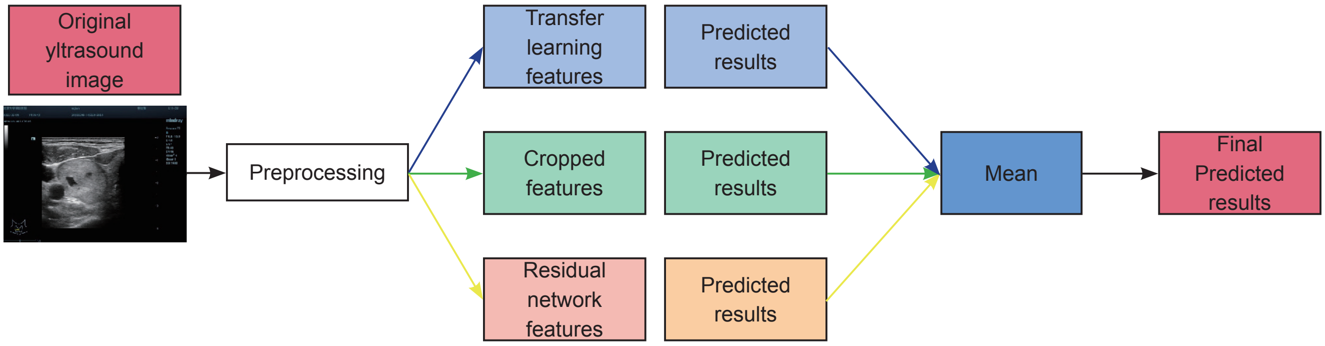


Figure 7 Proposed deep hybrid multi-feature extraction.

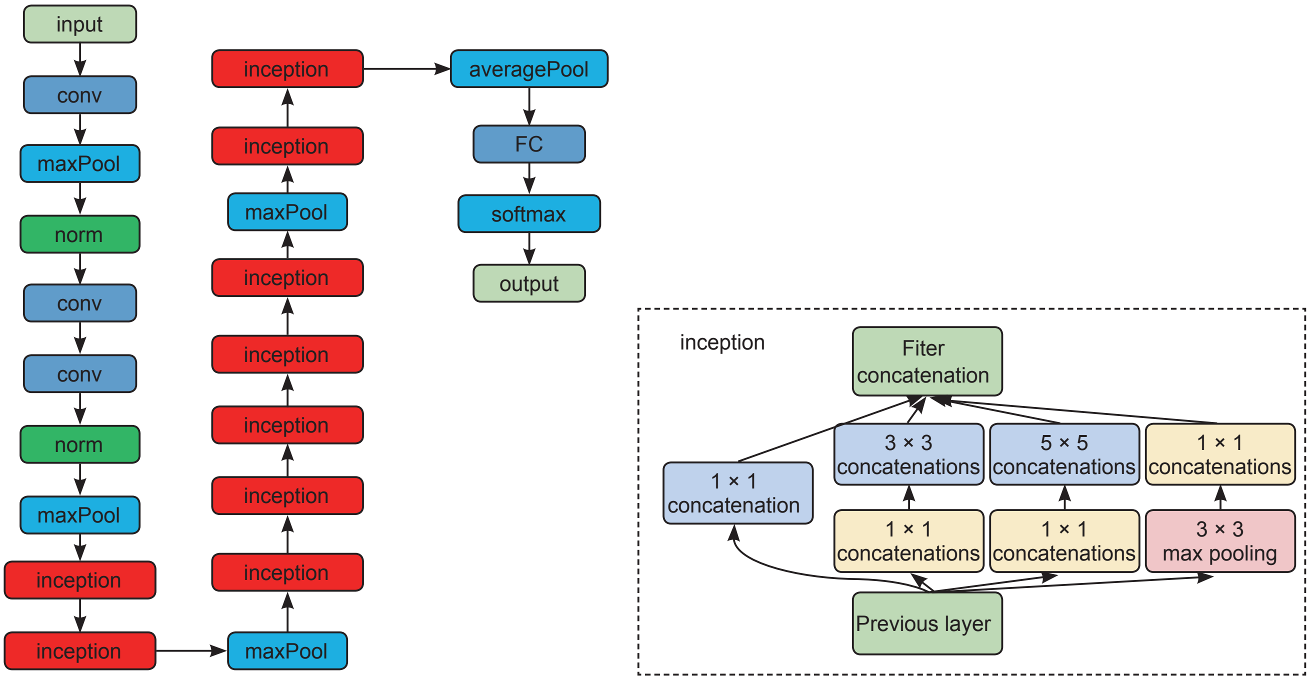


Figure 8 Proposed Inception V1 module.

Among the traditional manual image feature extraction algorithms, local feature extraction schemes have been emerging as research hotspots. Local features for images retain more information about the image, and this research uses the concept of image local features to form the proposed cropped feature strategy. A feature map is acquired by a CNN, and then the feature map is divided into N parts more Equation 1.

$$N = \left(1 + \frac{(1-a)h}{N_h}\right) * \left(1 + \frac{(1-a)w}{N_w}\right) + 1 \qquad (1)$$

where h and w are the height and width of the feature map respectively. a as a variable that dynamically determines the scale of the cropped feature map, in this study a has the value of 0.5. N_h and N_w are equal to 1 (1-α) h and 1 (1-α) w. And are the size of the crop window. An example of feature cropping for a of 0.5 (Fig. 9). The 5 cropped sub-features are extracted separately and the predicted values are derived, and finally the average of these 5 predicted values is taken as the final prediction result of the branch. The final predicted value for the whole system is calculated according to Equation 2.

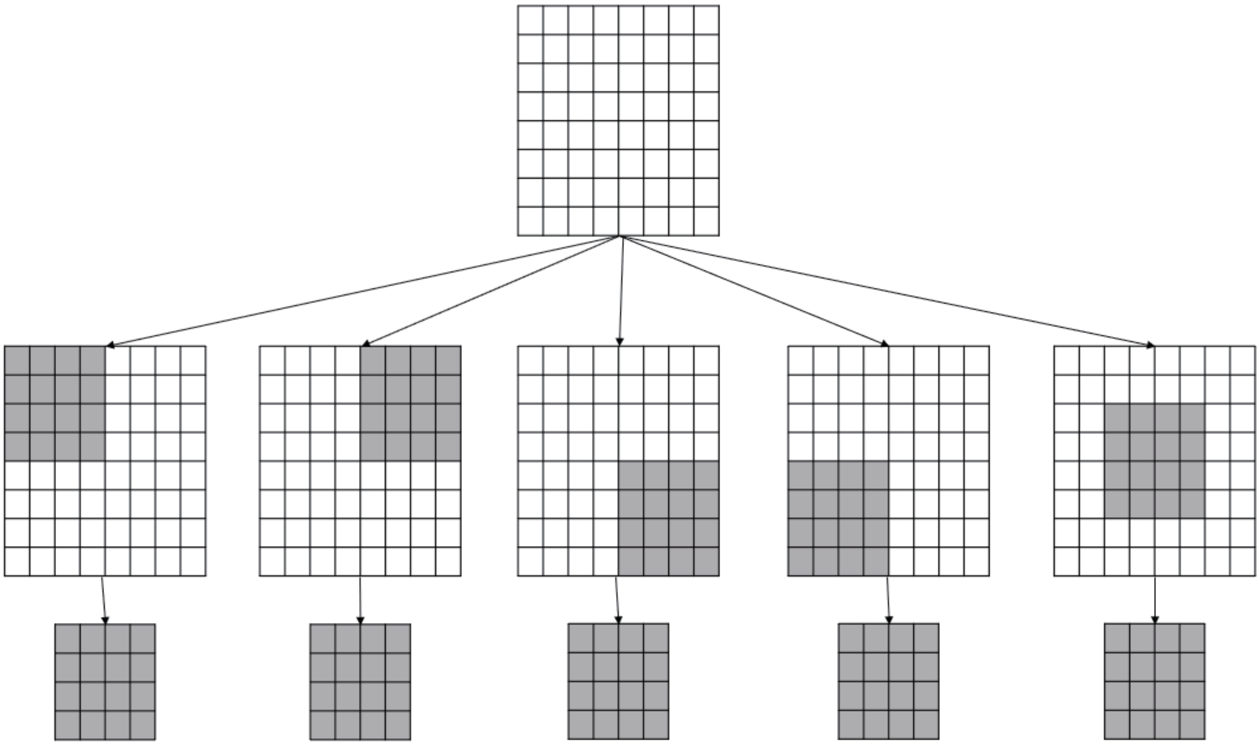


Figure 9 Example of feature cropping.

$$F_p = \frac{T_p + C_p + R_p}{3} \qquad (2)$$

where F_p is the final predicted value. F_p, C_p, R_p are the predictions of each of the transfer, clipping and residual feature branches respectively.

Experiments and Discussion

In this section, in order to validate the performance of the proposed frameworks for ultrasound images of follicular tumours of the thyroid, we have carried out experiments on the ultrasound image database at Peking University Shenzhen Hospital and compared with other state-of-the-art methods which are LeNet, AlexNet, ZFNet, VGGNet, Inception v1, ResNet-50, SENet, CroppingNet. The hardware system platform adopted for the experiments in this study was a DELL G15 5520 laptop with a 12th Gen Intel(R) Core (TM) i7–12700H (20 CPUs), 2.3 GHz, 16384 MB RAM, and three graphics cards: two discrete graphics cards (NVIDIA GeForce RTX 3060 Laptop GPU 6023 MB, NVIDIA GeForce RTX 3080 Ti, 12108 MB) and one integrated graphics card (Intel(R) Iris(R) Xe Graphics, 128 MB). The development system used for this study was Python 3.9.12, with Jupyter Notebook as the editor and Windows 10 Professional 64-bit as the operating system. The training set and test sets are divided according to the following criteria (Fig. 10).

The evaluation criteria for this study are accuracy, sensitivity, specificity and AUC. Our proposed method obtains the optimal AUC of 96.69% and also shows a good balance in terms of sensitivity and specificity(Table 1 and Fig. 11). Although the ZFNet obtained 100% specificity, its sensitivity was only about 3% accurate, due to the small sample size of positives. Furthermore, the results of Inception V1 are good, but the sensitivity and AUC are not as good as the proposed method. Sensitivity performance is very important in practical clinical applications.

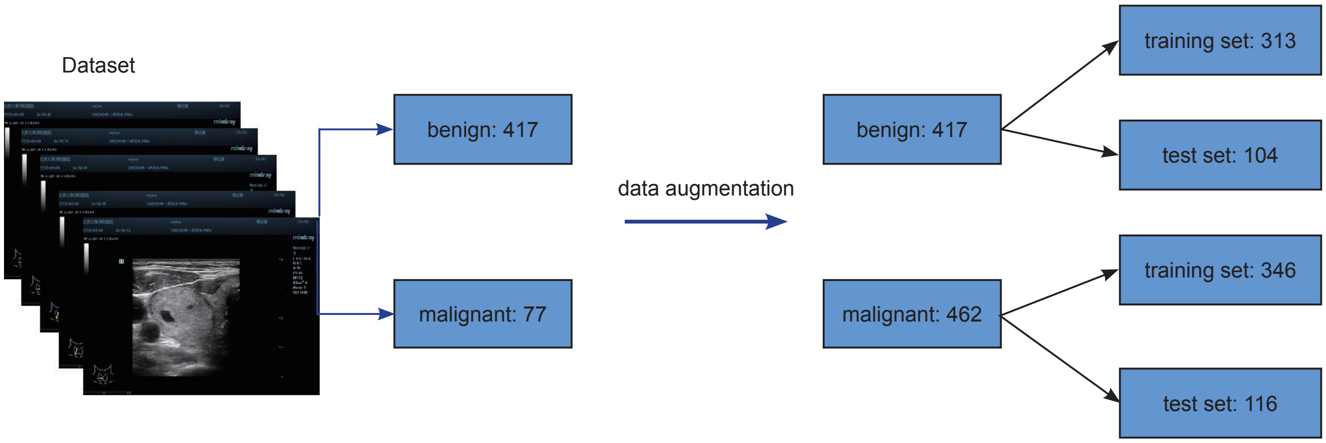


Figure 10 Division of the training and test sets.

Table 1 Performance comparison results

Method	Accuracy(%)	Sensitivity(%)	Specificity(%)	AUC(%)
LeNet	72.60	70.19	74.78	80.99
AlexNet	76.71	75.96	77.39	86.36
ZFNet	53.88	2.88	100.00	70.61
VGGNet	72.60	60.58	83.48	80.14
Inception v1	90.41	86.54	93.91	94.38
ResNet	84.93	83.65	86.09	91.30
SENet	56.62	98.08	19.13	72.03
Cropping	84.93	88.46	81.74	92.07
Proposed method	89.95	88.46	91.30	96.69

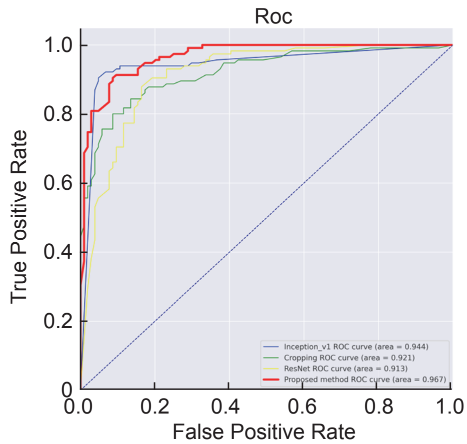


Figure 11 Results for AUC.

Conclusion and Future Work

In this study, a non-invasive risk prediction system is proposed for ultrasound images of thyroid follicular tumours. The core of the system is a proposed deep hybrid multi-feature Fusion Network to address the problems of similar image features, low sample size and unbalanced distribution. We solve the problem of unbalanced sample distribution by means of an image augmentation algorithm. In order to obtain enough features to differentiate ultrasound images, a three-branched feature extraction network was designed in this study, and a balance of sensitivity and specificity is ensured by an averaging algorithm.

In future,we will merge the three branches of features and optimise the parameters to reduce the excessive consumption of resources by the risk prediction system.

Acknowledgment

This work is supported by grants from National Natural Science Foundation of China (No. 81871358, No. 81901767), Natural Science Foundation of Guangdong Province (No. 2022A1515010986), Shenzhen Science and Technology Innovation Commission (No. JCYJ20200109140212277, NO. KCXFZ20200201101048740, No. JCYJ20210324110211031, No. JCYJ20210324131402008, JSGG20191129103020960), Construction of High-Level Hospitals in Guangdong Province (No. GD2019260), and Sanming Project of Medicine in Shenzhen (No. SZSM202111011).

Conflict of interest

The authors have no conflict of interest to declare.

References

[1] Grani G, Lamartina L, Durante C, Filetti S, Cooper DS. Follicular thyroid cancer and Hürthle cell carcinoma: Challenges in diagnosis, treatment, and clinical management. *Lancet Diabetes Endocrinol* 2018;6:500-514.

[2] Tallini G, Tuttle RM, Ghossein RA. The history of the follicular variant of papillary thyroid carcinoma. *J Clin Endocrinol Metab* 2017;102:15-22.

[3] Pstrag N, Ziemnicka K, Bluyssen H, Wesoly J. Thyroid cancers of follicular origin in a genomic light: In-depth overview of common and unique molecular marker candidates. *Mol Cancer* 2018;17:116.

[4] Baldini E, Sorrenti S, Tartaglia F, Catania A, Palmieri A, Pironi D, et al. New perspectives in the diagnosis of thyroid follicular lesions. *Int J Surg* 2017;41 Suppl 1:S7-S12.

[5] Xu B, Tallini G, Scognamiglio T, Roman BR, Tuttle RM, Ghossein RA. Outcome of large noninvasive follicular thyroid neoplasm with papillary-like nuclear features. *Thyroid* 2017;27:512-517.

[6] de Jong MC, McNamara J, McGlashan N, Winter L, and Mihai R. Risk of malignancy in thyroid nodules selected for fine needle aspiration biopsy based on ultrasound risk stratification. *British Journal of Surgery* 2022;109(Supplement_2),znac056.

[7] Tian T, Chen Y, Xiang Y, Liu L, Liu B. Remarkable response of pulmonary metastases rather than remnant thyroid in 131I therapy of follicular thyroid cancer. *Clin Nucl Med* 2019;44:327-329.

[8] García-Burillo A, Monturiol-Duran JA, Iglesias-Felip C, Villasboas-Rosciolesi D, Castell-Conesa J. Follicular thyroid carcinoma metastases on round ligament of liver. *Clin Nucl Med* 2021;46:326-328.

[9] Ohori NP, Nishino M. Follicular neoplasm of thyroid revisited: current differential diagnosis and the impact of molecular testing. *Adv Anat Pathol* 2023;30:11-23.

[10] Xu B, Ghossein RA. Advances in thyroid pathology: high grade follicular cell-derived thyroid carcinoma and anaplastic thyroid carcinoma. *Adv Anat Pathol* 2023;30:3-10.

[11] Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer* 2012;48:441-446.

[12] Erickson BJ, Korfiatis P, Akkus Z, Kline TL. Machine learning for medical imaging. *Radiographics* 2017;37:505-515.

[13] Yu J, Deng Y, Liu T, Zhou J, Jia X, Xiao T, et al. Lymph node metastasis prediction of papillary thyroid carcinoma based on transfer learning radiomics. *Nat Commun* 2020;11:4807.

[14] Shin I, Kim YJ, Han K, Lee E, Kim HJ, Shin JH, et al. Application of machine learning to ultrasound images to differentiate follicular neoplasms of the thyroid gland. *Ultrasonography* 2020;39:257-265.

[15] Lim KJ, Choi CS, Yoon DY, Chang SK, Kim KK, Han H, et al. Computer-aided diagnosis for the differentiation of malignant from benign thyroid nodules on ultrasonography. *Acad Radiol* 2008;15:853-858.

[16] Wu H, Deng Z, Zhang B, Liu Q, Chen J. Classifier model based on machine learning algorithms: application to differential diagnosis of suspicious thyroid nodules via sonography. *AJR Am J Roentgenol* 2016;207:859-864.

[17] Matiz S, Barner KE. Inductive conformal predictor for convolutional neural networks: Applications to active learning for image classification. *Pattern Recognition* 2019;90:172-182.

[18] Wang Y, Chen Y, Yang N, Zheng L, Dey N, Ashour AS, et al. Classification of mice hepatic granuloma microscopic images based on a deep convolutional neural network. *Applied Soft Computing* 2019;74:40-50.

[19] Meng N, Lam EY, Tsia KK, So HK. Large-scale multi-class image-based cell classification with deep learning. *IEEE J Biomed Health Inform* 2019;23:2091-2098.

[20] Wang C, Chen D, Hao L, Liu X, Zeng Y, Chen J, et al. Pulmonary image classification based on inception-v3 transfer learning model. *IEEE Access* 2019;7:146533-146541.

[21] Yu L, Chen H, Dou Q, Qin J, Heng PA. Automated melanoma recognition in dermoscopy images via very deep residual networks. *IEEE Trans Med Imaging* 2017;36:994-1004.

[22] Yan L, Shi Y, Wei M, Wu Y. Multi-feature fusing local directional ternary pattern for facial expressions signal recognition based on video communication system. *Alexandria Engineering Journal* 2023;63:307-320.

[23] Wu Y, Zhang Q, Hu Y, Sun-Woo K, Zhang X, Zhu H, et al. Novel binary logistic regression model based on feature transformation of XGBoost for type 2 diabetes mellitus prediction in healthcare systems. *Future Generation Computer Systems* 2022;129:1-2.