CLASSIFICATION OF THYROID NODULES IN ULTRASOUND IMAGES USING DEEP MODEL BASED TRANSFER LEARNING AND HYBRID FEATURES

Tianjiao Liu[1], Shuaining Xie[1], Jing Yu[2], Lijuan Niu[3], Weidong Sun[1]

[1]Dept. of Electronic Engineering, Tsinghua University, Beijing 100084, China [2]Colg. of Computer Science & Technology, Beijing Univ. of Technology, Beijing 100124, China
[3]Cancer Hospital of Chinese Academy of Medical Sciences, Beijing 100021, China

INTRODUCTION

Ultrasonography has become the most widely used modality for detecting and diagnosing thyroid cancer. Computer aided diagnosis can give diagnosis suggestions, and increase the diagnosis accuracy when lack of experts.

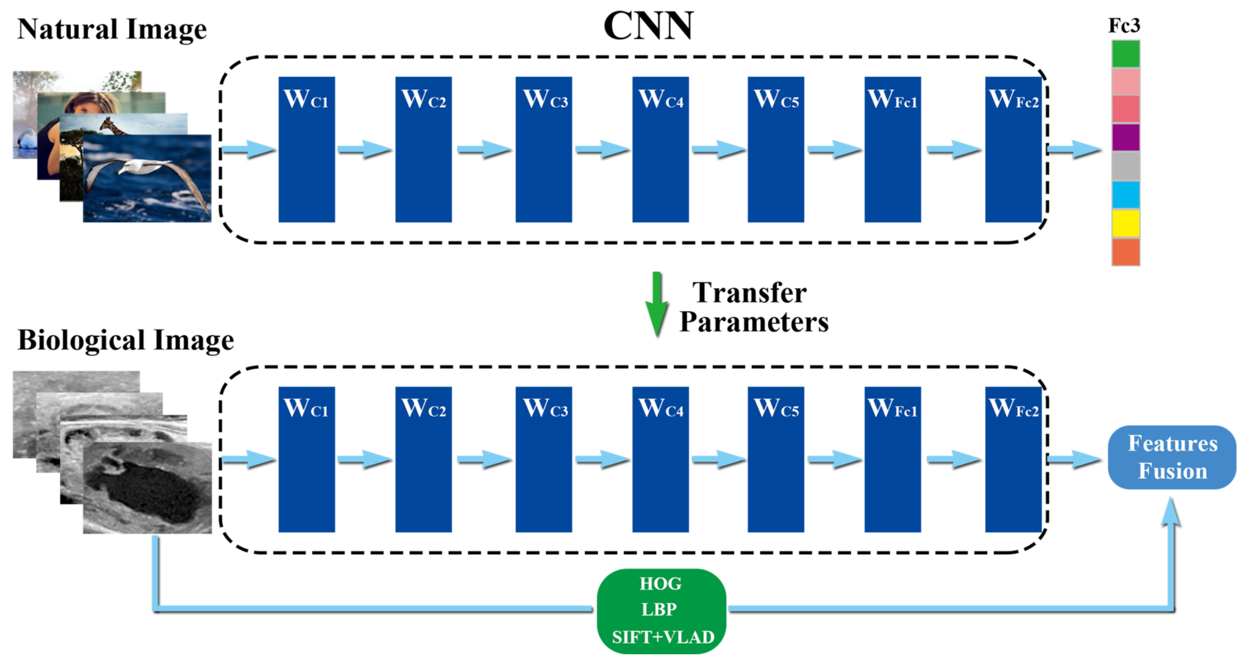
The problem

- Traditional hand-crafted features perform quite unsatisfactorily on the classification task for its intrinsic simplicity and locality.
- CNNs can provide high-level features, but large datasets are usually unavailable in medical field.

The proposed solution

To capture appropriate features and handle the small sample problem, we propose a feature extraction method based on CNNs and transfer learning.

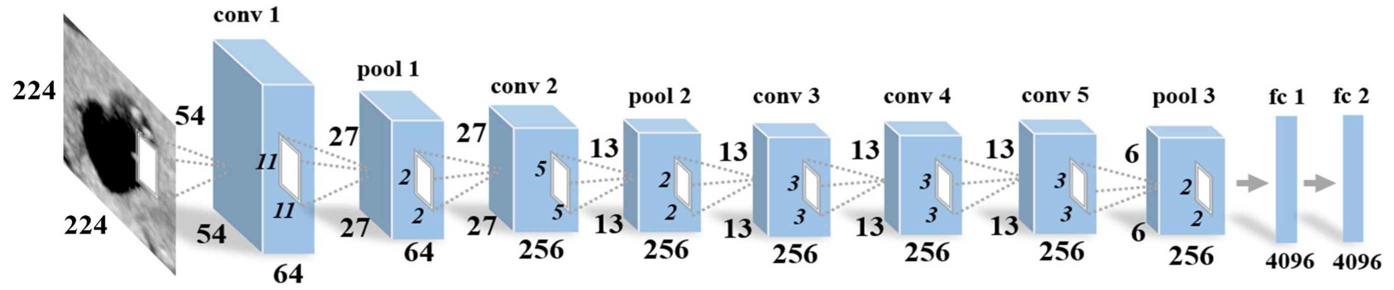
- We combine hand-crafted features with semantic deep features extracted from the pre-trained CNN model.
- A positive-sample-first majority voting and a feature-selected based strategy are employed for the hybrid classification.



Flowchart of our proposed method

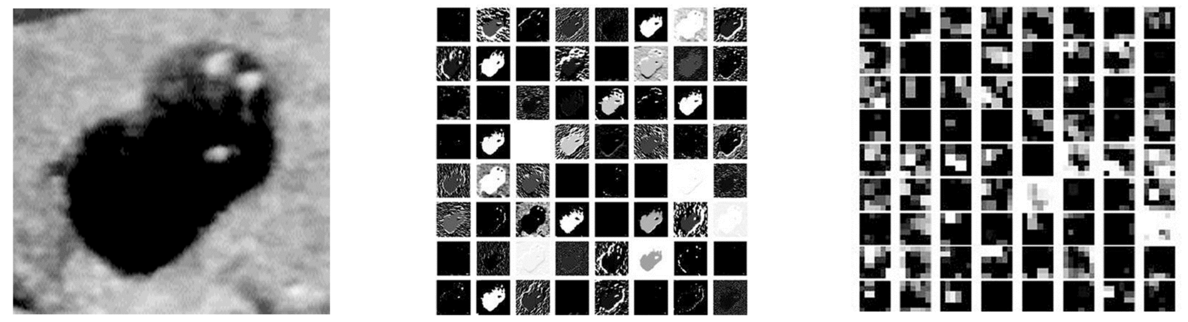
PRE-TRAINED MODEL FOR TRANSFER LEARNING

We employ the VGG-F model trained with ImageNet database. VGG-F model is consist of 5 convolutional and 2 fully-connected layers. The fc layers have 4096-dimentional outputs that can be used as feature descriptor for classification.



Architecture of the VGG model

A tumor image response to a certain convolution layer: lower layer extracts edge, directional and intensity features, while in high layer, various features compositely emerge.



Input image   Outputs of the conv1   Outputs of the conv5

FEATURE FUSION STRATEGY

We integrated HOG, SIFT and LBP features with the high-level features extracted from CNNs. Two kinds of feature fusion strategies are used in our proposed method.

Feature-selected strategy

The feature selection standard based on sorting the differences of benign samples and malignant samples.

$$diff_k = \left| \frac{1}{N_{MB}} \sum_{i=1}^{N_{MB}} v_{ik} - \frac{1}{N_{MM}} \sum_{i=1}^{N_{MM}} v_{ik} \right| (k = 1, \cdots, N)$$

Where, N_{MB} and N_{MM} are the number of benign and malignant nodules in the training set, v_{ik} is the kth dimensional feature of the ith image. The top 1000 features with the largest $diff_k$ will be chosen.

Positive-sample-first majority voting strategy

For a feature extraction method k, a classifier h_k can be trained on the dataset. The final predicted classification result for sample $\boldsymbol{x}$ is expressed as,

$$h(x) = mode(h_1(x), ..., h_D(x))$$

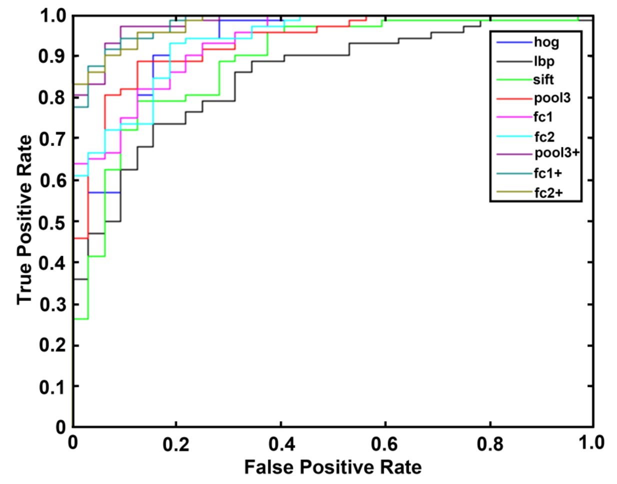
Once the same votes of benign and malignance occurs, we regard the results as malignance.

RESULTS

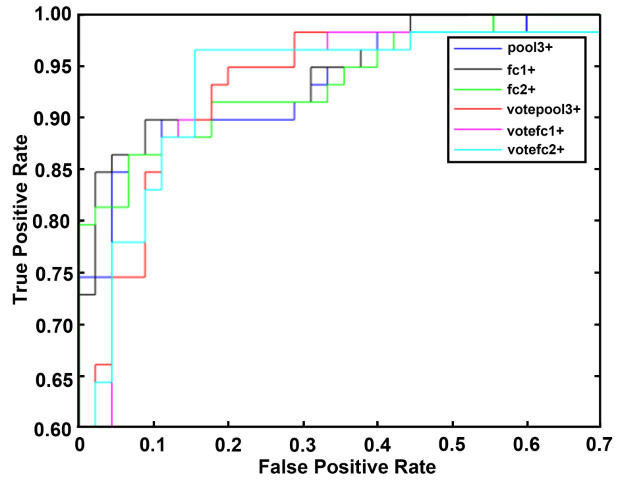
- 1037 thyroid nodule ultrasound images, including 651 benign and 386 malignant images, are used in our experiments and all of the images are demarcated by doctors.
- The classification was performed using SVM classifier with 10-fold cross validation.

Comparison results of different features

Features	Accuracy	Sensitivity	Specificity	AUC
HOG	0.829	0.699	0.906	0.837
LBP	0.780	0.562	0.909	0.793
SIFT+VLAD	0.824	0.726	0.882	0.841
Pool3	0.856	0.773	0.905	0.917
Fc1	0.853	0.774	0.911	0.940
Fc2	0.860	0.788	0.911	0.946
Pool3+	0.931	0.908	0.945	0.977
Fc1+	0.923	0.895	0.939	0.982
Fc2+	0.920	0.887	0.940	0.976
Vote-Pool3+	0.918	0.902	0.928	0.963
Vote-Fc1+	0.913	0.885	0.929	0.959
Vote-Fc2+	0.917	0.888	0.934	0.956



ROC on different features



ROC on two fusion strategies

- 'Pool3', 'Fc1' and 'Fc2' denote the features extracted from layer pool3, fc1 and fc2; '+' and 'Vote-' denotes feature selection and voting strategies.
- Features generated by CNNs overwhelm low-level features, thus it is feasible to transfer CNN features to ultrasound images domain.
- Two feature fusion methods can generally improve the accuracy by 7%-8% and 10%-14%, and increase sensitivity by11-20%, compared with the deep features and the low-level features.
- Feature-selected method aims to improve the TP rate when FP rate stays low, while voting makes effort on ensuring the classification accuracy of positive samples.

CONCLUSION

In this paper, a feature extraction method for ultrasound images is presented to classify the thyroid nodules into benign and malignant.

- The pre-trained CNN model is transferred to ultrasound domain to gain semantic features.
- We considered both low-level and high-level features, and proposed two feature fusion strategies for classification.
- Hybrid methods outperformed both the pre-trained CNN model and the traditional single-type feature method.