

## 针对病变区域的多视图集成甲状腺分类网络

徐蕾<sup>1,2,4</sup>, 郭锐锋<sup>1,2,4</sup>, 王广义<sup>3</sup>, 郭佳慧<sup>5</sup>, 王鸿亮<sup>1,2,4</sup>

<sup>1</sup>(中国科学院沈阳计算技术研究所, 沈阳 110168)

<sup>2</sup>(中国科学院大学, 北京 100049)

<sup>3</sup>(中国人民解放军31411部队, 沈阳 110000)

<sup>4</sup>(辽宁省基于数字孪生的交互系统工程研究中心, 沈阳 110168)

<sup>5</sup>(中国医科大学附属第一医院 内分泌与代谢科, 沈阳 110001)

E-mail: wanghl@sict.ac.cn

**摘要:** 甲状腺结节是一种常见的内分泌疾病。近年来,随着深度学习的发展,提出了大量的甲状腺良恶性辅助诊断(CAD)方法。尽管CAD在甲状腺分类任务上表现良好,但目前大多数研究都是采用黑盒模式,并没有从医生的诊疗过程出发,忽视了医生是从肿瘤的多个角度及其周围组织信息进行诊断。因此本文提出一种多视图专家集成分类网络。首先,利用Curvelet变换得到更多图像特征;其次,将不同视图输入到多个专家网络模型中,以分而治之的策略进行诊断;最后,结合主治专家网络与多个辅助专家网络,实现对甲状腺结节良恶性的综合判断。实验结果表明,基于多视图集成网络的分类性能显著优于单视图方法,并在多个评价指标上表现出明显的提升。

**关键词:** 集成学习;深度学习;甲状腺结节;多视图

中图分类号: TP391

文献标识码: A

文章编号: 1000-1220(2026)02-0421-07

## Multi-view Ensemble Thyroid Classification Network Based on Lesion Regions

XU Lei<sup>1,2,4</sup>, GUO Ruifeng<sup>1,2,4</sup>, WANG Guangyi<sup>3</sup>, GUO Jiahui<sup>5</sup>, WANG Hongliang<sup>1,2,4</sup>

<sup>1</sup>(Shenyang Institute of Computing Technology, University of Chinese Academy of Sciences, Shenyang 110168, China)

<sup>2</sup>(University of Chinese Academy of Sciences, Beijing 100049, China)

<sup>3</sup>(The Chinese People's Liberation Army 31411 Army, Shenyang 110000, China)

<sup>4</sup>(Liaoning Provincial Engineering Research Center for Interactive Systems Based on Digital Twins, Shenyang 110168, China)

<sup>5</sup>(Department of Endocrinology and Metabolism, The First Hospital of China Medical University, Shenyang 110001, China)

**Abstract:** Thyroid nodules are a common endocrine disorder. In recent years, with the advancement of deep learning, numerous auxiliary diagnostic methods for thyroid malignancy (CAD) have been proposed. Although CAD systems have demonstrated good performance in thyroid classification tasks, most of the existing studies adopt black-box models, which do not align with the clinical diagnostic process. These methods overlook the fact that physicians assess the tumor from multiple perspectives while considering the surrounding tissue information. To address this issue, this paper proposes a multi-view expert ensemble classification network. First, Curvelet transform is employed to extract additional image features, enhancing the detail and complexity of the image representation. Second, different views are fed into multiple expert network models, where a divide-and-conquer strategy is applied for diagnosis. Finally, the main expert network is combined with several auxiliary expert networks to provide a comprehensive evaluation of thyroid nodule malignancy. Experimental results demonstrate that the multi-view ensemble network significantly outperforms single-view methods, showing notable improvements across several evaluation metrics.

**Keywords:** ensemble learning; deep learning; thyroid nodule; multi-view

## 0 引言

在过去的30年间,甲状腺癌的发病率在世界范围内的不同人群中都有所增加<sup>[1]</sup>。成人中有多达3/4的人会出现甲状腺结节<sup>[2]</sup>。自20世纪70年代末以来,FNB一直是甲状腺结节诊断和治疗中的常用方法<sup>[3]</sup>。尽管FNB作为一种诊断手段被广泛应用,但它在临床实践中可能带来一定的伤害和不适,

尤其在操作不当或患者特定情况时。随着技术的进步,医学影像(如超声、CT、MRI等)在临床诊断中发挥着越来越重要的作用,减少不必要的侵入性操作、降低患者的痛苦和风险。

与此同时,计算机辅助诊断(CAD)作为一种新兴的技术手段,正日益成为辅助临床医生进行疾病诊断的重要工具。随着深度学习技术的不断发展,CAD结合了各类算法模型,使其能够自动分析影像数据,提升诊断效率和准确性。超声影像

收稿日期:2025-03-04 收修稿日期:2025-03-04 基金项目:辽宁省基础研究计划项目(2023JH26/10300017)资助。作者简介:徐蕾,女,1998年生,博士研究生,研究方向为计算机视觉、医学图像处理;郭锐锋,男,1968年生,博士,研究员,CCF会员,研究方向为计算机视觉、医学图像处理;王广义,男,1986年生,助理工程师,研究方向为计算机网络通信;郭佳慧,女,1994年生,博士,讲师,研究方向为甲状腺相关疾病;王鸿亮(通信作者),男,1982年生,博士,研究员,研究方向为计算机视觉、医学图像处理。

是识别甲状腺主流方法,例如 Avola 等人提出一种基于知识驱动的端到端的甲状腺结节分类框架,利用专家集合指导的 DenseNet 对甲状腺超声检查数据进行结节分类,有效提高了计算机辅助诊断系统的性能<sup>[4]</sup>. Zhao 等人受领域知识启发,提出一种结合局部和全局特征的双路径网络,并且使用一种基于自注意力机制的组织解耦块(TAD)连接双路径,模仿人类视觉的双路径结构实现对甲状腺结节的自动诊断<sup>[5]</sup>. Sun 等人提出了一种基于视觉变换器(ViT)的甲状腺结节分类模型 TC-ViT,利用 ViT 提取出甲状腺结节的全局特征,并结合 ROI 图像增强局部特征,最终模型分类准确率达到 86.9%<sup>[6]</sup>.

由于超声图像低分辨率、斑点噪声和高可变性特点<sup>[7]</sup>,使得图像处理技术难以有效应用.此外,超声诊断高度依赖于操作者的经验和技能,具有较强的主观性,增加了诊断结果的不稳定性<sup>[8]</sup>.因此研究者们还提出了基于灰度超声图像(US)、SWE 和 CDU 3 种模态甲状腺结节的自动诊断方法<sup>[9]</sup>. Zhang 等人采用 US 和红外热成像 IRT 图对甲状腺结节分类实现了很好的效果<sup>[10]</sup>.与超声成像相比,CT 影像具有更高的分辨率和更稳定的图像质量,能够提供更加清晰的解剖结构和病灶细节,尤其是在甲状腺结节形态复杂或边缘模糊的情况下,CT 成像的优势更加明显<sup>[11]</sup>.

然而,现有的基于深度学习的甲状腺分类模型仍旧存在一些问题,一方面,传统的甲状腺病变分类模型都是基于单视图方法<sup>[12,13]</sup>,该类模型忽略了结节及其周围组织可以从多个角度进行观察的事实,仅考虑单一视角的图像作为网络输入,难以全面提取结节的尺度信息和多分辨率信息.此外在甲状腺的检查与诊断中,需要特别关注甲状腺两侧的病变情况.然而,整体图像的输入会带来更多的噪声信息.选择 ROI(Region of Interest) 病变区域是经验性的,不同的医生视角观察的视角不同,并且没有统一的标准.最终的 ROI 的划分对最后的分类结果可能有显著的影响<sup>[14]</sup>.

另一方面,在临床诊断中,良性和恶性甲状腺结节的区别往往较为微妙如图 1 所示,需要医生捕捉一些细微的特征.如结节的形态、钙化、边界等<sup>[15]</sup>,基于深度学习的分类模型在处理这类问题时困难重重.

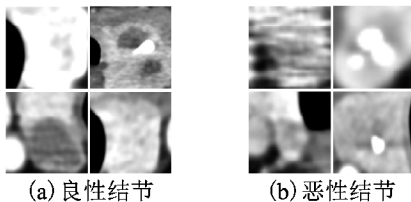


图1 甲状腺良恶性结节 CT 图像

Fig. 1 CT images of benign and malignant thyroid nodules

所以,本文提出了一种基于 CT 影像病变区域的多视图集成甲状腺分类模型.设计了病变区域 ROI 的提取方法,实现多尺度多视图的病变区域扩展分析.构建多个专家网络,采用模型集成方法,以学习不同角度的病变区域特征信息,模拟不同医生在不同视角下的诊断结果.融合学习各个专家网络的输出特征,对甲状腺 CT 影像病变区域进行分类.本文的贡献如下:

1) 提出一种基于多视图的甲状腺结节自动诊断方法.以

病变区域为中心,从边缘角度出发,获取结节不同视角特征输入到网络中,有效提高了模型的表达能力.

2) 提出 C-DSC (Curvelet-Depthwise Separable Convolution) 模块,丰富 CT 图像在不同尺度和角度上的细节特征. C-DSC 模块将 Curvelet 变换与深度可分离卷积相结合,充分融合病变区域在频域与空间域上的局部特征,增强模型泛化能力.

3) 提出一种基于多视图的甲状腺病变区域集成分类模型,采用分而治之的诊断策略,构建多个针对不同视图的专家模型,分别对各自视图进行深入诊断.采用复合模式,结合主治医师专家网络与多个辅助专家网络,通过集成多个专家模型的诊断结果,实现对甲状腺结节良恶性的综合判断.

4) 所提出的方法在甲状腺结节 CT 数据集上进行实验,模型在多个指标上均优于传统的单视图分类模型.

## 1 相关工作

### 1.1 集成学习

集成学习最早是在机器学习领域中提出的,其核心思想是通过组合多个学习器的预测结果,以提高模型的整体性能.集成学习的目标是将多个基学习器的预测结果进行融合.

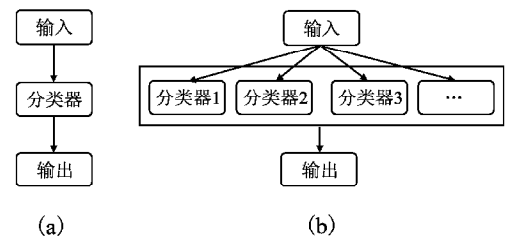


图2 (a)为单分类器,(b)为集成分类器

Fig. 2 Single classifier(a) and ensemble classifier(b)

如图 2 所示,图 2(a)为单一的学习器,特征输入一个单一分类器后直接输出分类结果.图 2(b)为集成学习,特征输入多个分类器中,各分类器独立进行预测,随后通过融合机制整合各分类器的结果,从而获得比单一学习器更准确且更具鲁棒性的决策.常见的集成学习方法有 Bagging<sup>[16]</sup>、Boosting<sup>[17]</sup>等,随着深度学习的快速发展,集成学习逐渐被应用于神经网络中. Chent 等人提出采用多个卷积神经网络 CNN 进行集成,以实现高光谱图像分类<sup>[18]</sup>. Krishna<sup>[19]</sup>, Guo<sup>[20]</sup>, 孙<sup>[21]</sup>等人的研究也证明了集成模型相比单一模型能够显著提升性能.

### 1.2 曲波变化 (Curvelet Transform)

随着 CNN 在视觉方面的突破,有很多研究工作将频域应用于视觉任务,Rao 等人提出 GTNet 全局滤波网络,通过在频域中以对数线性复杂度学习长期空间依赖关系实现图像分类任务<sup>[22]</sup>. Lin 等人使用快速傅立叶变换(FFT)、参数可学习滤波器和逆 FFT 来捕获全局上下文信息,实现遥感图像语义分割性能<sup>[23]</sup>. Finder 等人利用小波变换实现大感受野而不过度参数化,有效提升卷积神经网络在图像分类和下游任务中的性能<sup>[24]</sup>. 相比于傅里叶和小波变换,曲波变换需要更少的系数并具有高方向选择性和高各向异性特性<sup>[25]</sup>,能够很好的

关注到图像纹理和轮廓. Nencini 等人实验证明曲波变换能够有效提升空间分辨率<sup>[26]</sup>. 由于医学影像的特异性, 相比于其他变换, 曲波 Curvelet 可从基函数能够沿着图像的曲线结构进行方向性分解, 使其能更精准地提取病变和组织的边缘和轮廓. Srivastava 等人利用 Cuverlet 曲波变换的多尺度、多方向特性, 提出多模态医学图像融合方法. 证明比基于单个像素的融合规则更有效.

### 1.3 深度可分离卷积

传统卷积操作是通过一个大小为  $K \times K$  为的卷积核如图 3(a) 所示, 以全局的角度融合学习不同通道中的特征, 该种

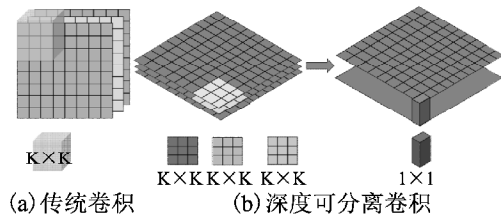


图 3 传统卷积和深度可分离卷积对比图

Fig. 3 Traditional convolution and depthwise separable convolution comparison

方法容易遗漏单通道中的局部特征, 导致无法有效学习部分细节信息.

对于具有  $C_{in}$  个通道的输入, 传统的卷积需要  $K \times K \times C_{in} \times C_{out}$  个参数. 为了增强对单通道局部特征的学习能力, 以及降低模型的参数量降低参数量和计算复杂度, Sifre 等人 2014 年首次提出深度可分离卷积概念<sup>[27]</sup>, 随后在 MobileNet<sup>[28]</sup> 和 Xception<sup>[29]</sup> 网络广泛应用. 深度可分离网络如图 3(b) 所示通过将传统的卷积操作分解为两个独立部分: 深度卷积核逐点卷积. 在深度卷积中, 每个输入的通道都单独与  $K \times K$  的滤波器进行卷积计算, 每个通道与卷积核的卷积产生一个特征图, 逐点卷积是一个  $1 \times 1$  的卷积操作, 将深度卷积的输出进行通道间的线性组合. 经过深度可分离卷积后, 参数量变为  $K \times K \times C_{in} + C_{in} \times C_{out}$ , 在保持高效特征提取能力的同时, 大幅度降低了计算复杂度和参数量.

Shang 等人使用深度可分离卷积代替标准卷积, 独立处理极化合成孔径雷达图像中每个通道的特征<sup>[30]</sup>. KC 等人提出一种基于深度可分离卷积的植物病害检测模型, 显著减少了参数量的同时取得了 98.34% 的分类准确率<sup>[31]</sup>. Khan 等人通过引入深度可分离网络设计了并行网络 CNN-DSCK, 显著提高了模型的效率<sup>[32]</sup>.

## 2 方法

本章将详细介绍针对病变区域的多视图集成甲状腺分类网络, 整体框架如图 4 所示, 分为病变区域提取模块, C-DSC

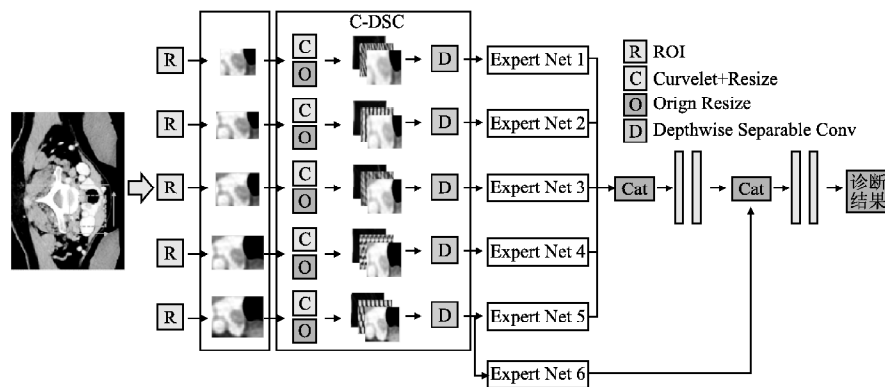


图 4 模型整体框架图

Fig. 4 Model framework diagram

模块和集成专家网络诊断模块.

病变区域提取模块用于提取图像的病变位置信息, 并进行不同程度的扩张, 模拟医生从多个角度观察病变细节过程, 以这种方式捕获病变区域在不同尺度和视角下的不同特征. C-DSC 模块通过频域和空间域之间转换, 分别从频域和空间域提取关键信息, 进行特征融合, 丰富特征提取的过程, 提升模型的表达能力. 最后经过专家网络诊断模块, 不同尺度和视角的病变区域输入到不同的专家网络中, 针对不同病变区域的特点进行诊断, 采用复合模式, 阶梯型分析病变结果. 最终实现对甲状腺结节良恶性的综合判断.

### 2.1 病变区域提取

肿瘤边缘浸润在判断肿瘤良恶性中起着关键作用, 因此研究人员常通过肿瘤区域分割, 提取图像中不同肿瘤的特征及其扩展范围. 如图 5 所示, 甲状腺呈蝴蝶状, 位于颈部中央, 左右两侧通过峡部相连.

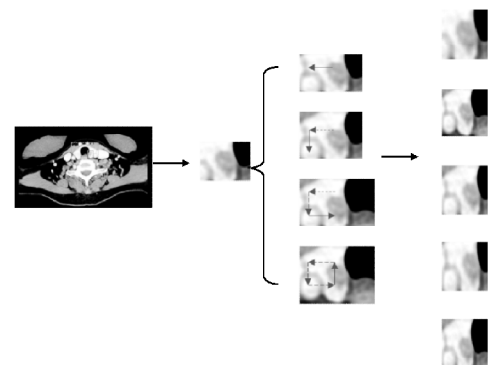


图 5 病变区域提取过程

Fig. 5 Lesion region extraction process

不同的恶性肿瘤边缘侵害程度不同, 所得信息不同, 基于此本文设计了 4 种病变区域扩展方法. 以 ROI 病变区域为中

心,按照逆时针方向进行扩充,每次扩充比例按照原有 ROI 的长宽的 1/4,生成 4 张不同扩充率的病变区域。

## 2.2 C-DSC 模块

为了提取出 CT 影像的丰富特征,本文设计了 C-DSC 模块如图 6 所示。对 CT 影像进行 Curvelet 变换,首先通过快速傅里叶变换 (FFT) 将其从空间域转换为频域表示。其次通过设定随机方向和角度对频域数据进行旋转和重新采样,以捕获更加多样化的边缘和方向特征。最后通过逆快速傅里叶变换 (IFFT),将频域信息转换为空间域。最终得到两个曲波分量,高频分量  $H(x,y)$  捕获图像的局部纹理信息,低频分量  $L(x,y)$  提取图像的全局结构信息。

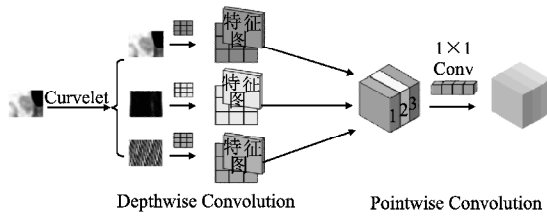


图 6 深度可分离卷积

Fig. 6 Depthwise separable convolution

将高频分量、低频分量和原始灰度图形成三通道的输入。其中  $G(x,y)$  为原始灰度图,  $I$  为堆叠后的三通道图像,  $x,y$  是图像水平和垂直方向的坐标,  $c$  为通道数。如公式 (1) 所示:

$$I(x,y,c) = [G(x,y), H(x,y), L(x,y)] \quad (1)$$

将构造的三通道图像输入到深度可分离卷积中进行特征提取。首先进行 Depthwise 卷积,对输入在三通道分别进行空间卷积,提取每个通道的空间特征,如公式 (2) 所示。保留各个通道的信息独立性同时避免了通道间的特征干扰。然后通过  $1 \times 1$  卷积,在通道维度上将独立的特征进行跨通道融合, Pointwise 卷积能够有效整合灰度、高频和低频特征。如公式 (3) 所示:

$$Z_d(x,y,c) = \sum_{i,j} K_d(i,j,c) \cdot I(x+i,y+j,c) \quad (2)$$

$$Z_p(x,y) = \sum_c K_p(c) \cdot Z_d(x,y,c) \quad (3)$$

其中  $Z_d(x,y,c)$  为 Depthwise 卷积输出的特征图,  $Z_p(x,y)$  为 Pointwise 卷积输出的特征图,  $K$  为卷积核,  $i,j$  为卷积核在垂直核水平方向的偏移量。

## 2.3 集成专家网络诊断模块

集成专家网络诊断模块由多个微调网络堆叠而成,包括: AlexNet<sup>[33]</sup>、DenseNet<sup>[34]</sup>、GoogleNet<sup>[35]</sup>、ResNet<sup>[36]</sup>、ShuffleNet<sup>[37]</sup>、Xception<sup>[29]</sup>、MobileNet<sup>[28]</sup>、ConvNeXt<sup>[38]</sup>、VGG<sup>[39]</sup>、Efficientnet<sup>[40]</sup>。所有专家网络都 ImageNet 上进行预训练,使其具有丰富的知识学习能力。鉴于医学图像中病变信息密度低,特征跳跃性大特点,考虑到模型堆叠产生的计算复杂度,参数量等问题。

采用复合专家模式对甲状腺结节进行诊断。Efficientnet 在多个医学图像任务上表现出色,能够有效捕捉医学图像的全局信息,在处理高维特征中可避免过拟合现象<sup>[41-43]</sup>,因此选择 Efficientnet 作为主治专家网络,能够处理图像中病变区域密度较低的情况。对于主治专家网络采用非冻结模型,输入扩张率最大的病变 ROI 区域,负责提取全局关键信息从而进

行全面诊断。对于辅助专家网络,采用冻结部分模型参数,对特定特征或局部信息进行精细化诊断。冻结参数主要集中在辅助网络的早期层次,以保持这些网络的基础特征提取能力,而让主治专家网络能够专注于全面的诊断任务。

经过集成专家网络后,输入到两个全连接网络中,使用 leakyReLU 激活函数。在训练过程中采用交叉熵损失函数 (CrossEntropyLoss) 如公式 (4) 所示。最后根据初步诊断和全面诊断的结果,会诊后给出最终的结果。

$$L(y,p) = -[y \log(p) + (1-y) \log(1-p)] \quad (4)$$

## 3 实验方案

### 3.1 数据集

论文数据集由中国医科大学附属第一医院提供,共收集 1160 张 CT 影像。由专业医师进行标注,并以标注区域为基础裁剪后进行多视图扩充,统一调整为  $224 \times 224$  的图像大小。

表 1 实验数据集

Table 1 Experimental dataset

	训练集	测试集	总计
恶性	820	214	1034
良性	453	123	576
总计	1273	337	1610

表 1 显示了实验所使用数据集的良恶性的分布信息,其中包括 1034 张恶性结节,576 张的良性结节。由于良恶性数据分布不均衡,为避免模型在训练中过度关注数量较多的恶性结节,分别对良性和恶性样本按照 8:2 的比例独立划分训练集和测试集,从而确保每类样本在训练和测试阶段的分布一致性。

### 3.2 实验相关设置

本文的实验环境为 Ubuntu 18.04 LTS, GPU 选用 NVIDIA RTX 3090,软件环境为 CUDA 12.2,使用 Intel Xeon 处理器及

表 2 实验参数

Table 2 Experimental parameters

Parameters	Values
Epoch	500
Dropout	0.5
Batch Size	16
Learning rate	1e-4
Betas	(0.9, 0.999)
Weight_decay	1e-4

Pytorch 2.2.0 框架。所有实验都应用 Adam 优化器,训练 500 epoch,使用早停止技术以避免过拟合实验中使用的关键参数如表 2 所示。

### 3.3 评价指标

作为分类任务,本模型采用精准率 (Precision)、准确率 (Accuracy)、灵敏度 (Sensitivity) 和 F1-Score 对实验结果进行评估。各指标信息介绍如下:

#### 3.3.1 精确率

精确率指标是评估模型对正样本的预测精度水平,是在衡量模型在预测为正类的样本中,实际为正类的比例的重要

评价指标. 计算公式如式(5)所示:

$$\text{Precision} = \frac{T_{TP}}{T_{TP} + F_{FP}} \quad (5)$$

### 3.3.2 准确率

准确率评价指标可以有效地衡量模型正确预测的样本数目, 占总样本数目的比例. 计算公式如式(6)所示:

$$\text{Accuracy} = \frac{T_{TP} + T_{TN}}{T_{TP} + T_{TN} + F_{FP} + F_{FN}} \quad (6)$$

### 3.3.3 灵敏度

该评价指标可以有效地衡量在所有实际为正类的样本中, 模型正确预测为正类的比例. 高灵敏度的模型可以有效地减少漏诊情况. 计算公式如式(7)所示:

$$\text{Sensitivity} = \frac{T_{TP}}{T_{TP} + F_{FN}} \quad (7)$$

### 3.3.4 F1-Score

F1 值是一种综合指标, 是精确率和灵敏度的调和平均值, 能够有效避免模型单独依赖某个指标, 用于综合评估模型的诊断能力. 计算公式如式(8)所示:

$$F1 = 2 \times \frac{\text{Precision} \times \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}} \quad (8)$$

其中, 真正例  $T_{TP}$  为正确预测的正例, 假正例  $F_{FP}$  为错误预测的正例, 真负例  $T_{TN}$  为正确预测的负例, 假负例  $F_{FN}$  为错误预测的负例.

## 4 结 果

### 4.1 单分类器结果分析

为证明本文所提出多视图集成方法的有效性, 首先在单视图下对单个分类器中进行训练, 实验结果如表 3 所示.

表 3 单视图-单分类器训练结果

Models	Result (%)			
	Precision	Accuracy	Sensitivity	F1
AlexNet	<b>75.37</b>	<b>75.89</b>	<b>71.67</b>	<b>72.57</b>
VGG	67.84	69.94	65.63	66.11
GoogleNet	64.59	67.26	63.02	63.33
ResNet	63.56	66.37	62.3	62.57
Xception	64.67	67.26	63.65	63.94
DenseNet	66.55	68.75	66.27	66.39
MobileNet	63.81	66.07	63.81	63.81
Efficientnet	67.44	69.64	65.56	66.00
ShuffleNet	60.71	64.29	58.57	58.44
ConvNeXt	70.32	72.02	70.79	70.51

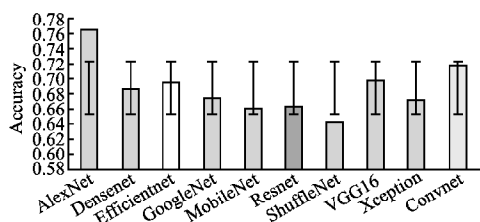


图 7 单视图单模型 ACC 对比图

Fig. 7 Single view single model ACC comparison chart

根据表中数据, 多数分类器准确率均在 60% 左右, 其中

AlexNet 分类器在测试集上的准确率为 75.37%, 这显示出良性与恶性结节的特征可能存在显著重叠, 导致单个分类器难以准确区分. 由于数据集中良恶性样本分布不均衡, 各个分类器的灵敏度 (Sensitivity) 指标均整体表现一般. 此外为了更加直观地对比单个分类器的性能, 图 7 中展示了各个分类器的 ACC 指标信息.

### 4.2 集成模型实验结果分析

本文提出的多视图集成分类模型在甲状腺结节良恶性分类任务上取得了显著的性能提升. 在表 4 的消融实验分析,

表 4 消融实验  
Table 4 Ablation study

Models	Result (%)			
	Precision	Accuracy	Sensitivity	F1
single (w/o)	75.37	75.89	71.67	72.57
C-DSC(w/o)	80.05	80.16	80.16	77.51
all	<b>85.62</b>	<b>86.22</b>	<b>82.83</b>	<b>84.00</b>

“w/o”表示移除模型中某个模块, “w/”表示使用模型中的某个模块. 首先在单分类下进行实验, 各项指标均达到 70% 以上. 在表第 2 行仅使用多专家诊断模块, 相比于单个分类器下, 模型准确率均得到有效提升, 验证了多专家协作机制对复杂医学影像分类任务的有效性. 加入所提出的 C-DSC 模块后, 模型显著增强了对甲状腺良性结节的区分能力, 分类准确率达到 85.62%, 相比于单分类器提升 10.25 个百分点.

表 5 对比实验

Table 5 Comparison experiment

Models	Result (%)				
	Precision	Accuracy	Sensitivity	F1	
文献[4]	80.70	81.54	79.36	79.90	
其他	文献[44]	72.28	73.51	69.28	69.99
	文献[45]	71.91	73.82	71.53	71.70
	文献[46]	73.35	73.12	68.93	69.52
	3-辅助集成	79.10	79.46	76.11	77.04
本文	4-辅助集成	84.75	83.92	80.63	81.91
	5-辅助集成	<b>85.62</b>	<b>86.22</b>	<b>82.83</b>	<b>84.00</b>

此外在表 5 中展示了本文方法与其他分类方法的对比结果. 尽管 3-辅助集成的表现略低于文献[4], 但实验结果表明, 随着辅助专家网络数量的增加, 本文所提出的针对病变区域的多视图集成分类网络在提升模型的精度和鲁棒性方面具有显著效果. 与文献[46]采用单一模型、单视图方法相比, 本文方法在精度上提高了 12.27%, 多视图专家集成方法能够有效避免单一模型的误差和不稳定性.

## 5 总结与展望

本文提出了一种基于多视图集成甲状腺分类模型. 针对传统 CAD 方法的黑盒模式及忽视医生多视角诊断特征的局限性, 设计了 C-DSC 模块和多专家会诊策略的诊断框架. 通过采用 Curvelet 曲波变换结合深度可分离卷积, 提升对 CT 影像中的细节特征的提取能力. 受病理学家诊断过程的启发, 采用多专家进行会诊的思路, 进一步提升了模型的稳定性和鲁棒性.

尽管本文所提出的方法在甲状腺分类任务上表现出色,但仍存在一些局限性.首先,对于病变区域的提取依赖于医生的手动标注,具有一定的主观性,未来的工作将着力于实现自动化标注技术,减少模型对人工的依赖.其次,在实际的诊断过程中,医生不仅依靠于影像数据,同时还会结合多种临床信息进行综合考虑,后续的研究将引入更多类型的临床数据,贴近医生的诊断流程,提高模型诊断的全面性.此外,实验所使用的数据集存在样本分布不均的问题,未来工作希望使用和整合更多的样本,进一步提高模型的泛化能力.模型有望能进一步提升其性能,为甲状腺结节的智能辅助诊断提供更高效、更可靠的解决方案.

### References:

- [ 1 ] MirandaFilho A, LortetTieulent J, Bray F, et al. Thyroid cancer incidence trends by histology in 25 countries; a population-based study [J]. *The Lancet Diabetes & Endocrinology*, 2021, 9(4) :225-234.
- [ 2 ] Staibano P, Forner D, Noel C W, et al. Ultrasonography and fine-needle aspiration in indeterminate thyroid nodules; asystematic review of diagnostic test accuracy [J]. *The Laryngoscope*, 2022, 132(1) :242-251.
- [ 3 ] Sakorafas G H. Thyroid nodules; interpretation and importance of fine-needle aspiration (FNA) for the clinician-Practical considerations [J]. *Surgical Oncology*, 2010, 19(4) :e130-e139.
- [ 4 ] Avola D, Cinque L, Fagioli A, et al. Multimodal feature fusion and knowledge-driven learning via experts consult for thyroid nodule classification [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2022, 32(5) :2527-2534.
- [ 5 ] Zhao S X, Chen Y, Yang K F, et al. A local and global feature disentangled network; toward classification of benign-malignant thyroid nodules from ultrasound image [J]. *IEEE Transactions on Medical Imaging*, 2022, 41(6) :1497-1509.
- [ 6 ] Sun J, Wu B, Zhao T, et al. Classification for thyroid nodule using ViT with contrastive learning in ultrasound images [J]. *Computers in Biology and Medicine*, 2023, 152: 106444, doi: 10. 1016/j. compbiomed. 2022. 106444.
- [ 7 ] Das D, Iyengar M S, Majdi M S, et al. Deep learning for thyroid nodule examination; a technical review [J]. *Artificial Intelligence Review*, 2024, 57(3) :47, doi:10. 1006/s10462-023-10635-9.
- [ 8 ] Choi S H, Kim E K, Kwak J Y, et al. Interobserver and intraobserver variations in ultrasound assessment of thyroid nodules [J]. *Thyroid*, 2010, 20(2) :167-172.
- [ 9 ] Xiang Z, Zhuo Q, Zhao C, et al. Self-supervised multi-modal fusion network for multi-modal thyroid ultrasound image diagnosis [J]. *Computers in Biology and Medicine*, 2022, 150: 106164, doi: 10. 1016/j. compbiomed. 2022. 106164.
- [ 10 ] Zhang N, Liu J, Jin Y, et al. An adaptive multi-modal hybrid model for classifying thyroid nodules by combining ultrasound and infrared thermal images [J]. *BMC Bioinformatics*, 2023, 24(1) :315, doi:10. 1186/s12859-023-05446-2.
- [ 11 ] Lalonde M N, Correia R D, Sykietis G P, et al. Parathyroid imaging [J]. *Seminars in Nuclear Medicine*, 2023, 53(4) :490-502.
- [ 12 ] Moussa O, Khachnaoui H, Guetari R, et al. Thyroid nodules classification and diagnosis in ultrasound images using fine-tuning deep convolutional neural network [J]. *International Journal of Imaging Systems and Technology*, 2020, 30(1) :185-195.
- [ 13 ] Wang Y, Yue W, Li X, et al. Comparison study of radiomics and deep learning-based methods for thyroid nodules classification using ultrasound images [J]. *IEEE Access*, 2020, 8: 52010-52017, doi: 10. 1109/ACCESS. 2020. 2980290.
- [ 14 ] Xu H, Wang X, Guan C, et al. Value of whole-thyroid CT-based radiomics in predicting benign and malignant thyroid nodules [J]. *Frontiers in Oncology*, 2022, 12:828259, doi:10. 3389/fonc. 2022. 828259.
- [ 15 ] Kwak J Y, Han K H, Yoon J H, et al. Thyroid imaging reporting and data system for US features of nodules; astep in establishing better stratification of cancer risk [J]. *Radiology*, 2011, 260(3) :892-899.
- [ 16 ] Breiman L. Bagging predictors [J]. *Machine Learning*, 1996, 24(2) :123-140.
- [ 17 ] Chen T, Guestrin C. XGBoost: ascalable tree boosting system [C]//*Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016:785-794.
- [ 18 ] Chen Y, Wang Y, Gu Y, et al. Deep learning ensemble for hyperspectral image classification [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2019, 12(6) :1882-1897.
- [ 19 ] Krishna T B, Kokil P. Standard fetal ultrasound plane classification based on stacked ensemble of deep learning models [J]. *Expert Systems with Applications*, 2024, 238: 122153, doi: 10. 1016/j. eswa. 2023. 122153.
- [ 20 ] Guo Y, Wang X, Xiao P, et al. An ensemble learning framework for convolutional neural network based on multiple classifiers [J]. *Soft Computing*, 2020, 24(5) :3727-3735.
- [ 21 ] SUN F Q, CONG C L, ZHANG K, et al. Benign and malignant diagnosis of breast cancer pathological medical images based on multi-model convolutional neural networks [J]. *Journal of Chinese Computer Systems*, 2020, 41(4) :732-735.
- [ 22 ] Rao Y, Zhao W, Zhu Z, et al. GFNet; global filter networks for visual recognition [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, 45(9) :10960-10973.
- [ 23 ] Bai L, Lin X, Ye Z, et al. MsanIfNet; semantic segmentation network with multiscale attention and nonlocal filters for high-resolution remote sensing images [J]. *IEEE Geoscience and Remote Sensing Letters*, 2022, 19:1-5, doi:10. 1109/LGRS. 2022. 3185641.
- [ 24 ] FINDER S E, AMOYAL R, TREISTER E, et al. Wavelet convolutions for large receptive fields [C]//*European Conference on Computer Vision*, Cham; Springer Nature Switzerland, 2024:363-380.
- [ 25 ] Kanagaraj K, Priya G G L. Curvelet transform based feature extraction and selection for multimedia event classification [J]. *Journal of King Saud University-Computer and Information Sciences*, 2022, 34(2) :375-383.
- [ 26 ] Nencini F, Garzelli A, Baronti S, et al. Remote sensing image fusion using the curvelet transform [J]. *Information Fusion*, 2007, 8(2) :143-156.
- [ 27 ] Sifre L, Mallat S. Rigid-motion scattering for texture classification

- [J]. arXiv preprint arXiv:1403.1687,2014.
- [28] Howard A G, Zhu M, Chen B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications[J]. arXiv preprint arXiv:1704.04861,2017.
- [29] Chollet F. Xception: deep learning with depthwise separable convolutions[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR),2017:1800-1807.
- [30] Shang R, He J, Wang J, et al. Dense connection and depthwise separable convolution based CNN for polarimetric SAR image classification[J]. Knowledge-Based Systems,2020,194:105542, doi:10.1016/j.knsys.2020.105542.
- [31] Kc K, Yin Z, Wu M, et al. Depthwise separable convolution architectures for plant disease classification[J]. Computers and Electronics in Agriculture,2019,165:104948, doi:10.1016/j.compag.2019.104948.
- [32] Khan Z Y, Niu Z. CNN with depthwise separable convolutions and combined kernels for rating prediction[J]. Expert Systems with Applications,2021,170:114528, doi:10.1016/j.eswa.2020.114528.
- [33] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM,2017,60(6):84-90.
- [34] Huang G, Liu Z, Van Der Maaten L, et al. Densely connected convolutional networks[C]//IEEE Conference on Computer Vision and Pattern Recognition(CVPR),2017:2261-2269.
- [35] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[C]//IEEE Conference on Computer Vision and Pattern Recognition(CVPR),2015:1-9.
- [36] Wang F, Jiang M, Qian C, et al. Residual attention network for image classification[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR),2017:6450-6458.
- [37] Zhang X, Zhou X, Lin M, et al. ShuffleNet: an extremely efficient convolutional neural network for mobile devices[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition,2018:6848-6856.
- [38] Liu Z, Mao H, Wu C Y, et al. A ConvNet for the 2020s[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR),2022:11966-11976.
- [39] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J]. CoRR,abs/1409.1556,2014.
- [40] Tan M, Le Q. EfficientNet: rethinking model scaling for convolutional neural networks[C]//Proceedings of the 36th International Conference on Machine Learning; VOL 97, PMLR,2019:6105-6114.
- [41] Raza R, Zulfiqar F, Khan M O, et al. Lung-EffNet: lung cancer classification using EfficientNet from CT-scan images[J]. Engineering Applications of Artificial Intelligence,2023,126:106902, doi:10.1016/j.engappai.2023.106902.
- [42] Siddique N, Paheding S, Angulo A A R, et al. Fractal, recurrent, and dense U-Net architectures with EfficientNet encoder for medical image segmentation[J]. Journal of Medical Imaging,2022,9(6):064004, doi:10.1117/1.JMI.9.6.064004.
- [43] Islam M M, Talukder Md A, Uddin M A, et al. BrainNet: precision brain tumor classification with optimized EfficientNet architecture[J]. International Journal of Intelligent Systems,2024,2024(1):3583612, doi:10.1155/2024/3583612.
- [44] Nguyen D T, Kang J K, Pham T D, et al. Ultrasound image-based diagnosis of malignant thyroid nodule using artificial intelligence[J]. Sensors,2020,20(7):1822, doi:10.3390/s20071822.
- [45] Zhao H bo, Liu C, Ye J, et al. A comparison between deep learning convolutional neural networks and radiologists in the differentiation of benign and malignant thyroid nodules on CT images[J]. Endokrynologia Polska,2021,72(3):217-225.
- [46] Lee J H, Ha E J, Kim J H. Application of deep learning to the diagnosis of cervical lymph node metastasis from thyroid cancer with CT[J]. European Radiology,2019,29(10):5452-5457.

#### 附中文参考文献:

- [21] 孙福权, 丛成龙, 张 琨, 等. 基于多模型卷积神经网络的乳腺癌病理医疗图像良恶性诊断[J]. 小型微型计算机系统,2020,41(4):732-735.