arXiv:2602.01000v1 [cs.CV] 1 Feb 2026

CortiNet：一种用于超声胆囊疾病诊断的物理-感知混合皮层启发式双流网络

Vagish Kumar
印度理工学院德里分校应用力学系 vagish.kumar@am.iitd.ac.in

Souvik Chakraborty
印度理工学院德里分校应用力学系
亚迪人工智能学院 (ScAI)
souvik@am.iitd.ac.in

摘要

胆囊疾病日益普遍，若未能在早期确诊，可能导致严重并发症。超声成像具有无创、成本低且易于普及等特点，因此是主要的诊断方式。然而，超声图像固有的低分辨率和散斑噪声严重影响诊断可靠性，常常促使人们采用深度且计算量庞大的卷积神经网络，而这类网络难以部署于常规临床环境中。在本研究中，我们提出了 CortiNet，这是一种轻量级、受皮层启发的双流神经架构，用于胆囊疾病自动诊断，并将具有物理可解释性的多尺度信号分解与感知驱动特征学习相结合。受人类视觉皮层中的并行处理通路启发，CortiNet 明确区分低频结构信息与高频感知细节，并通过专门的编码流分别进行处理。该架构直接处理结构化的频率选择性表征，而非原始像素强度，因此融入了强有力的基于物理的归纳偏置，能够以大幅减少的参数量实现高效特征学习。后期皮层式 融合机制在保持计算效率的同时，整合互补的结构与纹理线索。此外，我们提出了一种结构感知可解释性框架，其中梯度加权类激活映射仅应用于所提出 CortiNet 架构的结构分支。这一设计使模型仅关注结构特征，从而增强其抵御散斑噪声的鲁棒性。我们在 10,692 张专家标注图像上评估 CortiNet，这些图像涵盖九种具有临床意义的胆囊疾病类别。实验结果表明，CortiNet 仅需传统深度卷积模型所需参数的一小部分，即可实现较高的诊断准确率（98.74%）。准确率、鲁棒性和计算效率的结合，使 CortiNet 尤其适合部署于资源受限环境及即时诊疗临床环境。

关键词 胆囊疾病诊断 · 轻量级深度学习 · 超声检查 · 皮层启发神经网络 · 计算机辅助诊断

1 引言

胆囊疾病包括胆石症、胆囊炎、胆囊息肉和胆囊癌；若未能及早发现，可能造成严重的临床后果，已成为日益严峻的全球健康问题 [1]。在这些疾病中，胆囊癌尤其令人担忧，因为其进展迅速且早期缺乏特异性症状，往往导致诊断延迟、预后不良。流行病学证据表明，在生活方式改变、饮食习惯、肥胖和人口老龄化的推动下，胆囊疾病患病率持续上升；而胆结石等既存疾病则会显著增加恶性转化的风险 [2, 3]。根据 GLOBOCAN 2022，仅胆囊癌每年就在全球造成约 89,055 人死亡；若在晚期确诊，五年生存率仅约为 5% [4, 5]。超声检查（USG）

是评估胆囊异常的主要成像方式，具有无创、经济、可实时成像且易于普及等优点 [6]。它在检测胆结石、评估胆囊壁增厚、识别息肉以及提供恶性肿瘤的早期征象方面发挥着关键作用。然而，超声成像存在固有局限，包括散斑噪声、低对比度、空间分辨率有限以及依赖操作者，这些因素使细微病理特征难以识别，并导致观察者间差异 [7, 8]。影像检查量快速增长，更使这些挑战雪上加霜；近期分析表明，为满足当前临床需求，放射科医生平均每 3 至 4 秒就需判读一幅图像，这引发了对诊断一致性、疲劳和延迟发现的担忧 [9, 10]。

面对日益加重的诊断负担以及超声判读固有的挑战，深度神经网络（DNNs），尤其是卷积神经网络（CNNs），已成为自动化医学图像分析的强大工具 [11, 12]。CNNs 在多种应用中展现出卓越性能，包括皮肤癌检测 [13, 14, 15], 、肺结节恶性程度预测 [16, 17], 、结核病诊断 [18, 19],、脑肿瘤识别 [20, 21, 22], 、乳腺癌检测 [23, 24, 25], 以及新冠肺炎诊断 [26, 27, 28]。它们在医学成像中的广泛应用，源于其能够自动学习层次化且具有判别性的特征，从而减少对耗费人力且可能带有偏见的特征工程的需求。这一能力自然促使研究者尝试将 CNNs 应用于基于超声图像的胆囊疾病检测。然而，CNNs 在临床决策中的应用日益广泛，也引发了人们对模型可解释性的担忧。这些模型被视为黑箱系统。为应对这一挑战，梯度加权类激活映射（Grad-CAM） [29] 等事后可解释性技术已得到广泛采用，通过突出显示对模型预测贡献最大的图像区域，提供视觉解释 [30]。

早期深度学习（DL）方法在基于超声图像的胆囊（GB）疾病自动诊断中，主要侧重于在受限诊断场景下验证可行性。Jeong 等人 [31] 提出了一种基于CNN的框架，利用人工选取的感兴趣区域对良性与恶性胆囊息肉进行二分类，凸显了深度模型在超声分析中的潜力，但由于人工干预和诊断范围受限，可扩展性受到限制。后续研究探索了自动区域强调以及更深的卷积骨干网络（包括VGG、Inception、ResNet和MobileNet架构）的使用，以提升分类性能[32]。尽管这些方法提高了准确率，但它们在很大程度上依赖像素域学习以及深层、过度参数化的模型，从而引发了对计算效率、对超声伪影的鲁棒性以及在临床环境中实际部署的担忧。近期的医学人工智能（AI）研究则探索了结构化表征与先进架构设计，以更好地刻画胆囊病变的异质视觉特征。GBCNet [33] 结合了多尺度特征聚合与高阶统计建模，并采用课程学习策略以减轻纹理偏差，展示了更丰富表征的优势。RadFormer [34] 进一步利用注意力机制与基于变换器的推理，将全局解剖上下文与局部病理线索相融合。尽管这些架构具有出色的诊断性能，但它们引入了显著的模型复杂度与计算开销。相比之下，本研究提出了CortiNet，一种轻量级皮层启发式双流架构，将物理感知多尺度信号分解与感知驱动特征学习作为端到端框架的一等组件加以嵌入。通过显式编码领域感知的归纳偏置并分离结构与感知信息通路，CortiNet应对了超声医疗人工智能中的关键挑战，即鲁棒性、效率与可部署性，同时保持了较高的诊断准确率。

尽管取得了这些成功，将卷积神经网络应用于基于超声的胆囊疾病诊断仍面临诸多挑战。医学图像通常质量相对较低，分辨率有限且固有地存在散斑噪声，用于胆囊异常的标注数据集也明显小于一般计算机视觉任务中的数据集。这些因素可能导致模型性能下降，因为网络可能会关注无关纹理，而非具有临床意义的特征，例如病变区域的形状、边界或形态 [35]。因此，模型可解释性在这一场景中至关重要。可视化技术有助于验证网络注意力是否集中于具有临床意义的胆囊区域，而不是背景模式或成像伪影 [36]。此外，传统的深度卷积神经网络计算开销较大，需要大量内存和处理资源，这限制了其在资源受限的临床环境中的实时部署可行性。因此，尽管深度神经网络在提高诊断一致性和效率方面展现出巨大潜力，但其有效应用于胆囊超声分析仍然具有挑战性，探索这一特定领域的研究相对较少。

传统深度学习方法在胆囊超声分析中的局限性，促使人们开发兼具计算效率以及对图像噪声和变化的鲁棒性的架构。具体而言，设计融入领域特定知识的网络，例如超声信号形成中的波动相关物理学，以及人类专家采用的感知策略，能够在降低对伪特征过拟合风险的同时，提高诊断可靠性。此外，在资源受限的临床环境中部署模型的需求还要求

模型在参数量极少的情况下仍能保持较高的准确率。受人类视觉皮层并行处理流的启发，双流架构为分离互补信息（例如粗粒度结构模式和细粒度细节）并以连贯的方式将其融合提供了自然的框架。通过结合物理感知多尺度表示与感知驱动特征学习，此类架构有望实现鲁棒、可解释且轻量化的诊断性能，从而应对基于超声图像进行胆囊疾病检测所固有的计算和临床挑战。在这一框架中，Grad-CAM 等可解释性技术通过提供定性视觉证据，补充架构设计选择，从而在量化性能之外支持具有临床意义的决策支持。这一思路促使我们设计了 CortiNet——一种皮层启发式双流网络，将基于物理的归纳偏置与感知特征融合融为一体，以实现准确、高效的胆囊疾病诊断。主要科学贡献如下：

- **皮层启发式双流架构**：提出 CortiNet，这是一种受人类视觉皮层并行处理通路启发的双流网络，通过明确分离低频结构信息与高频纹理细节，增强特征表示。
- **物理-感知混合设计**：将基于小波的多尺度分解与感知驱动编码相结合，将超声图像形成的物理机制融入网络，同时利用互补的感知线索。
- **轻量高效**：以极少的参数实现出色的诊断性能，可部署于资源受限的即时诊疗临床环境，同时不牺牲准确率。
- **稳健的特征学习**：通过强调胆囊形态、边界和细粒度病理模式等临床相关特征，提高模型在噪声较多、对比度较低的超声图像中的可靠性。
- **模型可解释性**：提出一种结构感知可解释性框架，可抵御超声数据中的噪声，从而有助于建立对底层模型的信任。

本文其余各节的结构如下：第 2 节介绍所提出方法的详细内容；第 3 节展示开展的各项实验及模型性能；第 4 节阐述所提方法的关键见解及所得结果；最后，第 5 节总结全文，并探讨拓展本研究范围的潜在未来方向。

2 CortiNet：物理-感知混合、皮层启发式双流网络

本节介绍所提出的神经架构。我们将该架构命名为 CortiNet，是“皮层启发式双流网络”的简称，以强调其设计理念既植根于人类视觉感知，也植根于具有物理基础的信号表征。CortiNet 将原理明确的多尺度信号分解与模拟视觉皮层并行处理流的神经架构相结合，从而能够针对超声等易受噪声影响的医学成像模态，实现鲁棒且可解释的特征学习。CortiNet 采用物理-感知混合架构，将具有物理可解释性的信号分解与受生物学启发的视觉处理紧密结合。从物理角度看，医学图像，尤其是超声图像，是复杂波-组织相互作用的结果，其信息天然分布于多个空间频率中。从感知角度看，人类视觉系统能够高效地解读此类信号，将其分解为具有不同尺度和方向选择性的成分，并通过并行的皮层通路进行处理。通过融合这两项原理，CortiNet 将低频结构信息与高频感知细节明确分离，并通过专门的编码流对其进行处理，从而增强抗噪性并提升诊断表示学习能力。

2.1 感知基础、物理可解释的分解

CortiNet 的第一阶段对输入图像进行具有物理意义的多分辨率分解，这与早期视觉皮层中观察到的频率选择性接受野类似。我们不单纯依赖学习滤波器，而是采用一种具有坚实数学基础的小波变换，将图像分解为相互补充、对应不同空间尺度和方向的成分。对于输入图像$\boldsymbol{x} \in \mathbb{R}^{(C\times H\times W)}$，该分解可得到

$$P(x) = \{A_1, D_1^{HL}, D_1^{LH}, D_1^{HH}\}, \tag{1}$$

其中 A_1 编码粗粒度结构内容，而 D_1^{HL}、 D_1^{LH}、 D_1^{HH}捕捉方向感知高频响应。该分解采用 Daubechies基函数实现，具有紧致空间支撑，

正交性，并且能够抵御超声特异性噪声伪影，从而将物理归纳偏置直接嵌入学习流程。

2.2 用于结构与细节的双皮层流

受视觉皮层中并行处理通路的启发，CortiNet 采用两个参数互不共享的独立编码器：(a)**结构流（近似通路）**，用于学习全局解剖上下文、空间连续性和器官级形态；(b) **细节流（感知通路）**，用于学习对临床判别至关重要的局部纹理、边缘和方向变化。令 $\mathcal{S}(\cdot;\,\boldsymbol{\theta}_S)$ 和 $\mathcal{T}(\cdot;\,\boldsymbol{\theta}_T)$ 分别表示结构编码器和细节编码器。相应的特征表示为

$$F_S = \mathcal{S}(A_1;\boldsymbol{\theta}_S), \tag{2}$$

$$F_T = \mathcal{T}([D_1^{HL}, D_1^{LH}, D_1^{HH}];\boldsymbol{\theta}_T). \tag{3}$$

这种明确的分离体现了皮层特化，并防止高频噪声污染全局结构表征。

2.3 通过后期融合实现皮层整合

在人类视觉中，感知整合仅在早期皮层区域完成特化处理后才会发生。CortiNet 通过后期特征融合体现了这一原则。自适应全局平均池化可生成紧凑且与尺度无关的描述符，

$$\mathbf{f}_S = GAP(F_S), \quad \mathbf{f}_T = GAP(F_T), \tag{4}$$

将其拼接，形成统一的感知嵌入，

$$\mathbf{f} = [\mathbf{f}_S, \mathbf{f}_T]. \tag{5}$$

这种融合机制能够联合利用互补的结构线索和感知线索进行推理，同时保留二者在语义上的区别。融合表示通过带有 softmax激活的全连接层映射为类别概率，

$$Z \;=\; \mathbf{f}\,W^\top + b, \tag{6}$$

$$\hat{\mathbf{y}} = \mathrm{softmax}(Z), \tag{7}$$

其中，W 和 b 分别表示全连接层的权重和偏置。完整的参数集

$$\boldsymbol{\theta}_{\text{NN}} = \boldsymbol{\theta}_S \cup \boldsymbol{\theta}_T \cup \{\boldsymbol{W}, \boldsymbol{b}\} \tag{8}$$

通过多类交叉熵损失进行优化，从而得到一个端到端可训练的系统，该系统兼顾物理信号特性与感知处理原理。整体架构示意图如图 1 所示。训练阶段的完整算法见算法1。

2.4 CortiNet 中面向噪声感知推理的自适应分支选择

超声图像天然会受到散斑噪声和采集相关伪影的影响，而这些因素对高频信号分量的污染尤为严重。为明确考虑这一特性，CortiNet采用噪声感知自适应推理策略，在推理过程中动态决定是否启用感知（细节）路径。该决策依据结构（近似）路径的经验可靠性；该路径主要编码低频解剖信息，且本身对噪声具有更强的鲁棒性。

设通向预测的结构路径记为

$$f_a : \mathcal{X} \to \Delta^C, \tag{9}$$

其中 $\mathcal{X}$ 表示超声图像空间，且 Δ^C 表示 C类概率单纯形。给定一个小型校准数据集

$$\mathcal{D}_{\text{cal}} = \{(x_i, y_i)\}_{i=1}^{N_C}, \tag{10}$$

从与测试数据相同的分布中抽取，我们将结构路径的经验分类准确率评估为

$$\mathcal{A}_a = \frac{1}{N}\sum_{i=1}^{N_C} \mathbb{I}\left[\arg\max f_a(x_i) = y_i\right], \tag{11}$$

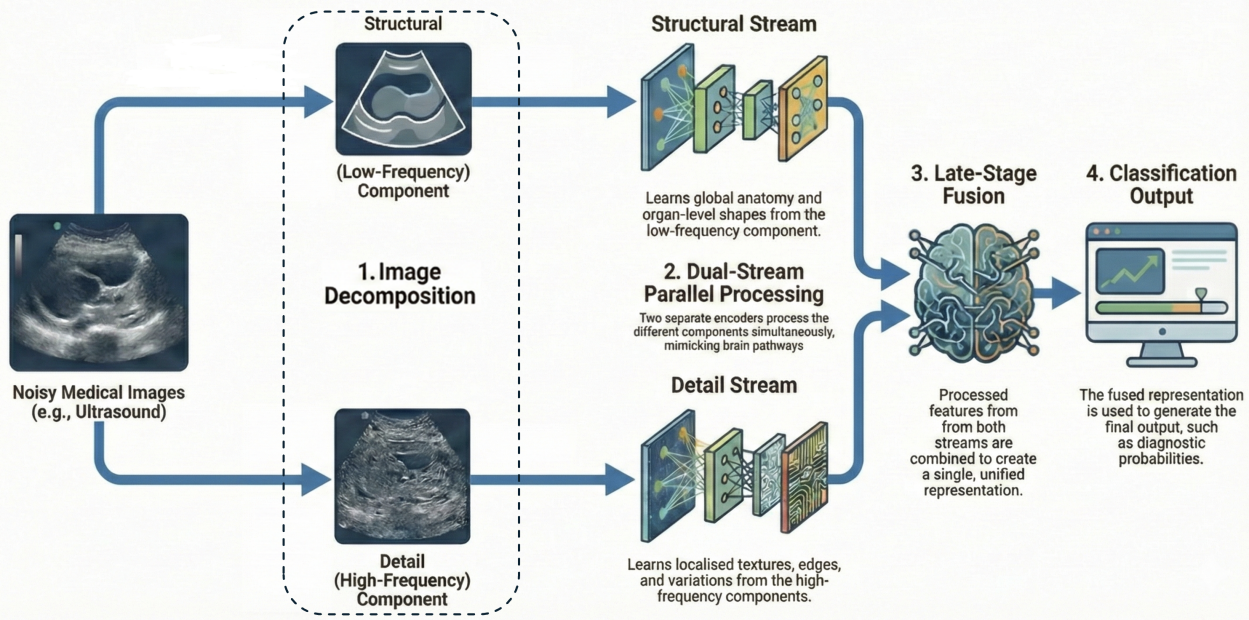


图 1：所提出的基于超声的胆囊疾病诊断 CortiNet 架构概览。该框架采用皮层启发式双流设计，明确分离结构与感知信息通路。首先，通过多尺度信号分解将输入超声图像转换为结构化的频率选择性表征，得到互补的低频分量和高频分量：前者编码全局解剖结构，后者捕捉精细纹理与边缘信息。随后，两个专用编码流使用轻量级卷积块并行处理这些分量，从而在超声特异性噪声和分辨率限制下实现高效且有针对性的特征提取。为整合互补线索，后期的皮层风格融合模块结合学习到的结构与感知表征，同时保持计算效率并尽量减少参数增长。接着，融合表示被传递至紧凑型分类头，以生成多类别胆囊疾病预测。CortiNet 通过处理多分辨率表征并融入强大的领域感知归纳偏置，以显著缩减的参数量实现稳健的诊断性能，因此适用于资源受限和床旁临床场景。

其中 $\mathbb{I}(\cdot)$ 表示示性函数。该校准准确率可作为结构可靠性和有效噪声严重程度的分布级代理。引入阈值 $\gamma \in (0,1)$，以控制感知通路的激活。具体而言，推理策略定义为

$$\text{Activate perceptual pathway} \iff \mathcal{A}_a < \gamma. \tag{12}$$

当 $\mathcal{A}_a \geq \gamma$ 时，仅使用结构路径进行预测，从而抑制噪声主导的高频信息并降低计算开销。

从统计学习的角度来看，高频图像表征往往具有较低的偏差但较高的方差，尤其是在超声采集等存在噪声的测量过程中。相比之下，低频结构表征会引入更大的偏差，但能显著降低方差，从而在分布偏移下获得更强的鲁棒性。因此，无条件融合两种表征可能会放大噪声驱动特征，导致泛化能力下降。所提出的准确率引导自适应推理策略明确利用了这种偏差—方差权衡。结构路径充当抗噪估计器，而感知通路则作为条件细化机制，仅在分布层面可靠性证据支持时才被激活。校准准确率为有效噪声严重程度提供了一种合理的数据驱动估计，使模型能够在诊断模糊病例中保持判别能力的同时，尽量减少不必要的方差。

所提出的自适应推理策略符合既有的放射学实践，即通常先进行粗略的解剖评估，再开展细致检查。在高质量的超声采集中，结构线索往往足以支持可靠诊断，过度依赖细微纹理细节则可能引入伪影模式。相反，诊断模糊病例则受益于更深入的感知审查。CortiNet 根据校准准确率动态调整推理路径，从而适应特定地点的采集条件和扫描仪变异

算法1： 所提方法的算法

输入： 一组图像 $X = \{x_1, x_2, \ldots, x_N\}$，其中 $x_n \in \mathbb{R}^{C\times H\times W}$ 及相应标签 $C = \{c_1, c_2, \ldots, c_N\}$，其中 $c_n \in \mathbb{Z}^+$，以及超参数

输出： 优化后的网络参数 $\theta^{\star}_{\mathrm{NN}}$

对于 $t \leftarrow 1$ **到** T **执行**
  遍历 小批量 $\{(x_i, y_i)\}_{i\in\mathcal{B}}$, $|\mathcal{B}| = B$ **执行**
    // 小波分解
    计算 x_i 的离散小波变换： $\mathcal{W}^{(1)}_{\psi}(x_i) = \{A_1, D_1^{\mathrm{HL}}, D_1^{\mathrm{LH}}, D_1^{\mathrm{HH}}\}$; ▷ 公式 (1)
    // 双编码器
    将 A_1 输入近似系数编码器： $F_S \leftarrow \mathcal{S}(A_1; \theta_{\mathrm{A}})$; ▷ Eq. (2)
    将细节系数传入细节编码器
    $F_T \leftarrow \mathcal{T}([D_1^{\mathrm{HL}}, D_1^{\mathrm{LH}}, D_1^{\mathrm{HH}}]; \theta_{\mathrm{D}})$; ▷ Eq. (3)
    // 晚期融合
    应用自适应全局平均池化
    $\mathbf{f}_S \leftarrow \mathrm{GAP}(F_S)$; $\mathbf{f}_T \leftarrow \mathrm{GAP}(F_T)$; ▷式（4）
    拼接以形成融合表示： $\mathbf{f} \leftarrow [\mathbf{f}_S;\ \mathbf{f}_T]$; ▷ 式（5）
    // 分类头
    将融合特征映射为 logits： $Z \leftarrow \mathbf{f}W^{\top} + b$; ▷ Eq. (6)
    应用 softmax 以获得类别概率： $\hat{\mathbf{y}} \leftarrow \mathrm{softmax}(Z)$; ▷ 公式（7）
    // 损失与参数更新
    计算交叉熵损失： $\ell(\boldsymbol{\theta}_{\mathrm{NN}}) = -\frac{1}{B}\sum_{i\in\mathcal{B}} -\sum_{k=1}^{K} \mathbf{1}[y_i = k]\log \hat{y}_{i,\,k}$;
    计算梯度 $\nabla\ell(\boldsymbol{\theta}_{\mathrm{NN}})$;
    更新参数： $\boldsymbol{\theta}_{\mathrm{NN}} \leftarrow \mathrm{AdamUpdate}(\boldsymbol{\theta}_{\mathrm{NN}},\ \nabla\ell(\boldsymbol{\theta}_{\mathrm{NN}}))$;
  end
结束

算法 2： CortiNet 中准确率引导的噪声感知自适应推理

输入： 测试超声图像 x、校准数据集 $\mathcal{D}_{\mathrm{cal}}$、准确率阈值 γ

输出： 预测的胆囊疾病标签 $\hat{y}$

校准阶段：

初始化计数器 $c \leftarrow 0$

对（$x_i,\ y_i$）$\in \mathcal{D}_{\mathrm{cal}}$ **执行**
  计算 $p_a(y \mid x_i) \leftarrow f_a(x_i)$
  若$\arg\max p_a(y \mid x_i) = y_i$ **则**
    $c \leftarrow c + 1$
  end
end

计算校准准确率 $\mathcal{A}_a \leftarrow \frac{c}{|\mathcal{D}_{\mathrm{cal}}|}$

推理阶段：

若 $\mathcal{A}_a \geq \gamma$ **则**
  // 结构路径被认为可靠；感知细节可能受噪声主导
  计算 $p_a(y \mid x)$
    $\hat{y} \leftarrow \arg\max p_a(y \mid x)$
else
  // 结构路径准确率降低表明存在噪声或诊断歧义
  激活感知通路和融合模块
  计算融合预测 $p_f(y \mid x)$
    $\hat{y} \leftarrow \arg\max\ p_f(y \mid x)$
end
return $\hat{y}$

且无需重新训练。这种具备噪声感知能力、由准确率驱动的适应性，增强了计算机辅助诊断系统在真实世界胆囊成像中的鲁棒性、可解释性和可部署性。

3 结果

本节全面评估所提出的 CortiNet 架构在基于超声图像的多类别胆囊疾病分类任务中的表现。我们首先介绍数据集特征和实验设置，随后分析定量性能、与现有方法的比较、计算效率及鲁棒性。此外，我们还开展了补充分析，以评估各个架构组件的贡献，并深入探讨模型的可解释性和特征敏感度。

3.1 数据集描述与实验设置

数据集描述。 实验采用一个公开的胆囊超声图像数据集，其中包含 10,692 张图像，涵盖九种具有临床意义的疾病类别 [37]。该数据集涵盖广泛的胆囊病变，包括良性和恶性疾病，反映了常规临床实践中遇到的真实诊断差异。各类别的图像数量分别为：胆结石（1,326）、腹部及腹膜后异常（1,170）、胆囊炎（1,146）、膜性及坏疽性胆囊炎（1,224）、胆囊穿孔（1,062）、息肉和胆固醇结晶（1,020）、胆囊腺肌病（1,164）、癌（1,590），以及各种原因导致的胆囊壁增厚（990）。尽管类别分布略有不平衡，但没有任何单一类别占据主导地位，因此能够开展有意义的多类别评估。图 2 展示了总体类别分布。该数据集包含以多种空间分辨率采集的超声图像，图像尺寸包括 (2400×1800)、(1170×876)、(900×1200) 和 (1200×900) 像素。这种差异反映了采集方案和超声系统的不同，为自动化分析带来了额外挑战。 **数据划分策略**。 采用分层训练集—验证集—测试集划分，比例为 60:20:20，确保各子集的类别分布与完整数据集中的分布基本一致。这种分层划分可减轻抽样偏差，并支持对所有疾病类别进行可靠的性能评估。 **预处理与数据增强**。 训练期间，所有图像均调整为320×320 像素，以确保样本具有一致的空间尺寸。为提升泛化能力和鲁棒性，采用了数据增强，包括随机水平翻转、随机旋转，以及亮度和对比度变化均设为 0.2 的颜色抖动。随后，将图像转换为张量，并将像素强度归一化至 $[0, 1]$ 范围。验证和测试期间，仅执行图像尺寸调整至 320×320 及张量转换，不进行数据增强，以确保评估结果不受偏差影响。 **评估方案**。 所有实验均遵循固定的留出法评估方案，使用预先定义的训练集、验证集和测试集划分。模型选择和超参数调优基于验证集进行，所有报告结果均对应留出测试集上的性能。**实现细节**。 我们使用 PyTorch 深度学习框架实现模型，并在 NVIDIA RTX A5000 GPU 上进行训练。该计算设置能够高效地完成训练和推理，同时确保实验结果可复现。论文录用后，源代码将在 http://www.github.com/csccm/CortiNet 上公开。

3.2 胆囊疾病谱中的诊断性能

我们首先评估所提出的框架在数据集中涵盖的整个胆囊病变谱中的诊断能力。我们采用互补的整体指标和分类指标来评估性能，以同时反映总体诊断准确率和针对特定疾病的可靠性。除定量指标外，我们还通过混淆分析考察错误模式，以识别临床相关病症之间的系统性混淆。**评估指标：** 为全面评估九种胆囊疾病类别的诊断性能，我们采用了多种与临床相关的评估指标。除总体准确率外，我们还报告每个类别的精确率、召回率（敏感度）、特异度、阴性预测值（NPV）、F1分数以及受试者工作特征曲线下面积（AUC）。准确率虽能衡量整体正确性，但在多分类医学诊断中可能掩盖临床上重要的失效模式。因此，我们着重考察召回率和特异度，分别量化模型正确识别患病病例和排除未患病病例的能力。精确率和阴性预测值进一步反映阳性与阴性预测的可靠性，而F1分数则体现敏感度与精确率之间的平衡。我们报告AUC，以反映不受决策阈值影响的分类判别能力；这在临床风险分层场景中尤为重要。

混淆矩阵分析与分类别性能： 图 3(a) 中的混淆矩阵呈现出明显的对角线优势，表明模型对全部九种胆囊疾病类别均能做出较高比例的正确预测。非对角线元素很少，说明类别间混淆有限，即便是超声表现相似的临床相关病症也是如此。具体而言，胆结石（类别 1）、膜性及坏疽性胆囊炎（类别 4）、息肉和胆固醇结晶（类别 6）以及胆囊腺肌病（类别 7）几乎都能被完美分类，体现了模型从超声图像中捕捉疾病特异性结构与纹理线索的能力。对于诊断难度较大的类别，例如穿孔（类别 5）以及由不同病因导致的胆囊壁增厚

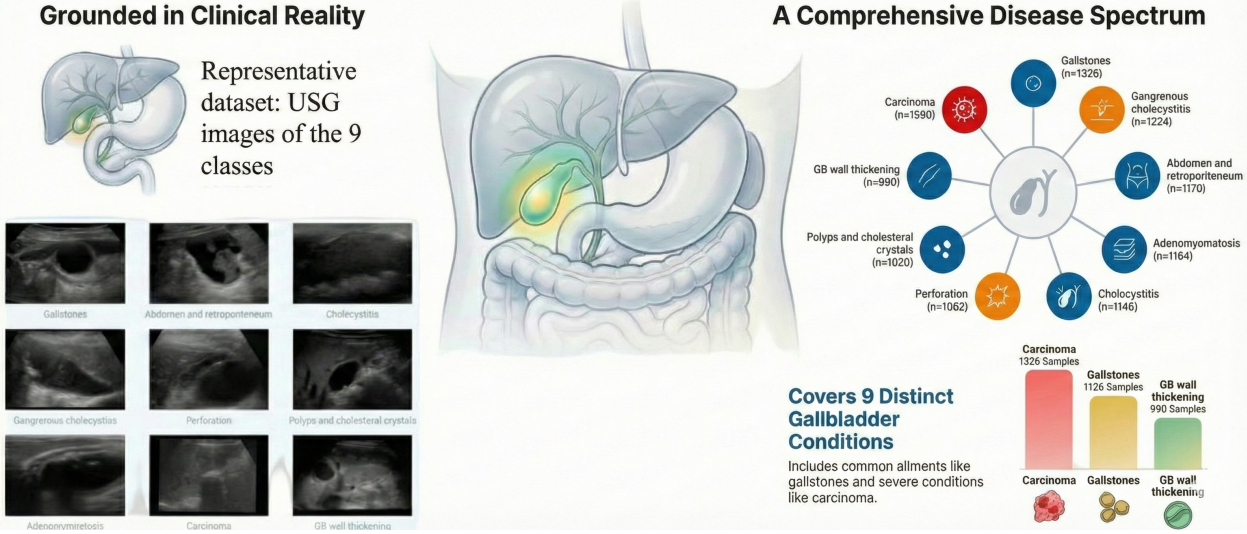
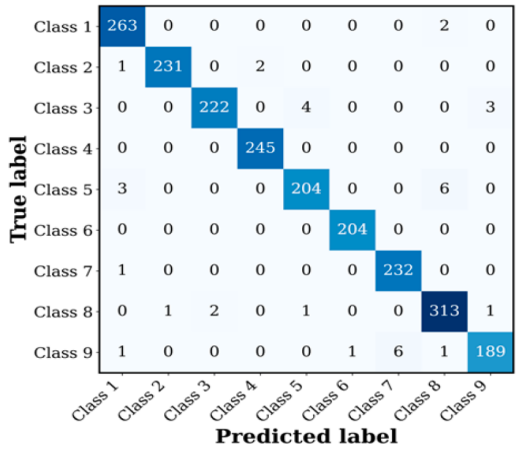


图2：**胆囊疾病诊断的临床背景与数据集概览**。该图综合展示了本研究使用的数据集所体现的解剖学相关性、成像方式及疾病多样性。左：代表性超声图像展示了各胆囊疾病类别之间广泛的视觉异质性，包括良性、炎症性和恶性疾病，突出了散斑噪声、低对比度以及细微形态学差异所带来的挑战。中：人体胆道系统示意图，其中胆囊被突出显示，将计算任务置于其生理与解剖上下文之中。右（上）：所考虑的胆囊病变谱示意图，涵盖九种不同的疾病类别，从胆结石和胆囊炎等常见疾病到癌等较为罕见但临床侵袭性强的病变。右（下）：各疾病类别的数据集构成摘要，表明同时覆盖高患病率和低患病率疾病，类别不平衡程度中等，且无单一类别占主导。这些图板共同强调了支撑所提出计算框架的超声队列的临床真实性、诊断多样性与代表性。

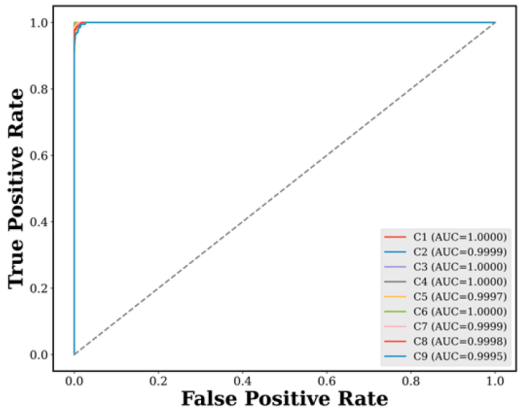
（类别 9），误分类率略高，但仍处于较低水平（见图 3(c)）。值得注意的是，这些错误并非系统性错误，也未显示模型偏向某个主导类别，表明其决策边界稳定，而非偏向患病率较高的病症。癌（类别 8）是临床上至关重要的类别，其精确率和召回率始终较高，凸显了模型识别恶性模式且不过度产生假阳性的潜在价值。受试者工作特征分析（见图 3(b)）显示，模型在整个胆囊疾病谱上均具有近乎顶尖的判别性能，所有类别的曲线下面积值均为 ≥ 0.999，表明其可分性强且不受分类阈值影响。**定量性能总结：** 图 3(d) 总结了各类别的定量性能。所有疾病类别的准确率均保持较高水平，范围为 0.9545 至 1.0000。包括膜性及坏疽性胆囊炎（类别 4）以及息肉和胆固醇结晶（类别 6）在内的若干类别，其召回率和准确率均达到 100%，同时保持较高的特异度和阴性预测值。所有类别的 F1分数均超过 0.96，表明敏感度与精确率之间保持了良好平衡。曲线下面积值始终较高（≥ 0.999），进一步证实模型对所有疾病类别均具有出色的判别能力，也证明了其学习到的表征具有鲁棒性。总体而言，这些结果表明，所提出的模型不仅具有较高的整体准确率，还能在各种胆囊病变中实现具有临床意义的可靠性，涵盖常见病症以及严重或恶性病例。在敏感度、特异度和预测值方面均衡的性能对于后续临床决策支持至关重要，因为漏诊和误报都会造成重大影响。

与最先进方法的比较： 为更好地评估所提出框架的诊断性能，我们将其与一系列常用于医学图像分类的代表性最先进深度学习模型进行基准比较。这些模型既包括传统卷积架构，也包括较新的注意力机制和基于 Transformer 的设计，它们在超声和腹部成像任务中均展现出较强的性能。为确保比较公平且可复现，所有参比模型均在完全相同的实验条件下进行训练和评估，采用相同的分层训练集—验证集—测试集划分、图像分辨率、预处理流程和评估方案。评估采用临床相关指标，包括总体准确率、各类别 F1分数以及受试者工作特征曲线下面积（曲线下面积），并特别关注癌和严重炎症状态等高风险疾病的敏感度。在所有评估指标上，所提出的模型均持续表现出优越的

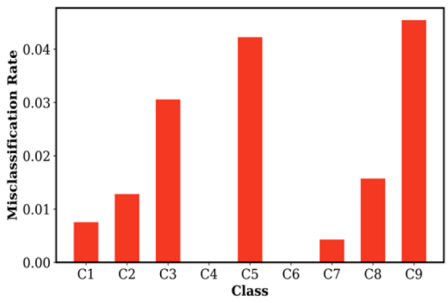
(a) Inter-class Confusion Analysis



(b) ROC Analysis



(c) Class-wise misclassification rate



(d) Class-wise Quantitative Performance Metrics

Class	Metrics						
	Acc.	Prec.	Recall	Spec.	NPV	F1-score	AUC
C1	0.9924	0.9776	0.9924	0.9967	0.9989	0.9850	0.9999
C2	0.9871	0.9956	0.9871	0.9994	0.9984	0.9914	0.9998
C3	0.9694	0.9910	0.9694	0.9989	0.9963	0.9801	0.9999
C4	1.000	0.9919	1.000	0.9989	1.000	0.9959	1.000
C5	0.9577	0.9760	0.9577	0.9974	0.9953	0.9668	0.9997
C6	1.000	0.9951	1.000	0.9994	1.000	0.9975	1.000
C7	0.9957	0.9747	0.9957	0.9968	0.9994	0.9851	0.9999
C8	0.9842	0.9720	0.9842	0.9950	0.9972	0.9781	0.9997
C9	0.9545	0.9792	0.9545	0.9979	0.9953	0.9667	0.9995

图3：**所提出模型在胆囊疾病谱中的诊断性能**。(a) 类间混淆分析。九种胆囊疾病类别的混淆矩阵显示出明显的对角线优势，表明分类准确率较高，类间混淆极少，即使在超声影像表现相似的病况中也是如此。(b) ROC分析。所有类别的一对其余ROC曲线均呈现近乎理想的形态，曲线下面积接近1，证实模型在不同决策阈值下具有出色的判别能力。(c) 分类别误分类率。根据混淆矩阵计算的误分类率在所有类别中均较低；形态学上具有挑战性的病况误差略高，但不存在系统性偏倚，表明模型行为稳定可靠。(d) 分类别定量性能指标。各类别的准确率、精确率、召回率、特异度和阴性预测值均表现出持续较高的水平，其中若干类别的敏感度和准确率接近完美或达到完美。

性能，优于现有方法。值得注意的是，对于具有细微形态学差异和较高类内变异、在临床上具有挑战性的类别，性能提升最为显著；在这些类别上，基线模型更容易产生混淆。相比之下，所提出的方法在常见和罕见疾病类别上均保持稳定的敏感度和精确率，体现出更强的特征判别能力和鲁棒性。重要的是，这些性能提升并未伴随模型复杂度的大幅增加，说明改进源于更有效的表征学习，而非过参数化。这些结果表明，所提出的框架在诊断准确率与实际可部署性之间实现了良好平衡，而这对于在真实临床环境中应用至关重要。

3.3 性能、效率与临床可部署性的比较

接下来，我们将 CortiNet 与成熟的深度学习架构及近期提出的胆囊专用模型进行基准比较，以评估其诊断性能和计算效率（图 4）。在所有评估指标上，CortiNet 均持续展现出更优的性能，同时计算开销显著更低。在分类准确率方面［图 4(a)］，CortiNet 的准确率达到 0.9874，高于 VGG16（0.9360）、InceptionV3（0.9804）、ResNet152（0.9794）和 MobileNet（0.9809），也优于 GBCNet（0.9275）和 RadFormer（0.9682）等胆囊聚焦模型。在召回率（0.9869）、精确率（0.9873）、F1分数（0.9870）和曲线下面积（0.9999）方面也呈现出相似趋势，表明其对胆囊疾病谱具有持续稳健的鉴别能力。值得注意的是，CortiNet 还保持了极高的特异度（0.9984）和阴性预测值（0.9984），凸显了其排除非病理病例的可靠性，而这对于临床决策支持系统至关重要。更重要的是，这些性能提升伴随着模型复杂度的大幅降低［图 4(b)］。CortiNet 仅需 132 万个可训练参数，而 VGG16、InceptionV3 和 ResNet152 分别需要 1.38 亿、2.18 亿和 5816 万个参数。其计算需求也同样大幅降低：CortiNet 的计算量仅为 1.52 十亿次浮点运算，比其他架构低数个数量级，同时实现了最快的推理速度，每张图像仅需 2.93 毫秒［图 4(c)］。相比之下，更深、更庞大的模型推理延迟显著更高，这可能限制其在对时效性要求较高的临床环境中的应用。即使与 GBCNet 和 RadFormer 等胆囊专用架构相比，CortiNet 仍以显著更小的计算开销实现了更高的诊断准确率，体现出有利于实际部署的准确率–效率权衡。

3.4 Ablation与鲁棒性分析：自适应噪声感知处理的实证证据

为阐明CortiNet双流架构的功能作用及其在临床真实扰动下的行为，我们联合分析了架构消融与抗噪性。结果如图5所示。图5(a)中的消融研究表明，仅感知（高频）分支会出现明显的性能下降（准确率0.8728，召回率0.8668），表明在缺乏全局结构约束的情况下处理细尺度纹理线索时，易受噪声放大影响。相比之下，仅结构（低频）分支取得了显著更强且更稳定的性能（准确率0.9832），凸显了粗粒度解剖与形态学信息在胆囊成像中的诊断重要性。缺乏显式频率分离的原始像素基线，其表现始终低于基于小波的配置，说明嵌入频率感知归纳偏置具有益处。完整的CortiNet模型在所有指标上均取得最佳性能（准确率0.9874，F1分数0.9870，曲线下面积0.9999），表明对感知细节进行选择性整合而非无条件纳入，能够产生最可靠的预测。

为评估模型在真实图像退化条件下的鲁棒性，我们考察了模型对加性高斯噪声的敏感度，以反映超声成像中常见的采集变异性。如图 5(c) 所示，除一个噪声水平外，CortiNet 在所有噪声水平下均取得最高分类准确率；在最高噪声水平（$\sigma = 0.3$）下，各算法的性能差异尤为明显。这表明，所提出的架构不仅能提升不利条件下的鲁棒性，也能在理想成像质量下实现更优的判别能力。观察到的稳定性可归因于噪声感知自适应设计（图 5(b)）：随着噪声增加，结构连贯的低频表征在决策中占据主导，从而避免虚假的高频激活干扰预测。在乘性散斑噪声下也观察到一致的趋势（结果未展示）；即使在 $\sigma = 0.3$ 下，CortiNet 的性能几乎没有下降。在其他胆囊专用模型中，GBCNet 对噪声表现出明显的敏感性，而 RadFormer 则展现出相对更高的韧性，但在各种噪声条件下，其表现始终不及 CortiNet。从临床角度看，这种双重优势——更高的基线准确率以及在图像损坏情况下持续保持鲁棒性——对于在常规超声工作流中可靠部署至关重要，因为实际应用中经常会遇到理想和退化的成像条件。

3.5 基于梯度归因的结构感知可解释性

除了预测性能之外，超声成像人工智能系统的临床实用性，在很大程度上取决于其可解释性和解剖学合理性。在胆囊超声中，诊断推理本质上由结构驱动，依赖壁增厚、管腔变形、声影和肿块诱导的形态畸变等线索，而非仅凭细粒度纹理。因此，对噪声或高频伪影敏感的可解释性机制可能会削弱临床信任，尤其是在存在散斑和采集变异性的情况下。为此，我们采用梯度加权类激活映射（Grad-CAM） [29], , 仅针对 CortiNet 的结构（低频）分支研究模型的可解释性。这一设计选择是有意为之：结构分支编码粗尺度的形态学表征，其本身较不易受散斑噪声影响，也更符合放射学推理；此前呈现的消融和鲁棒性

(a) CortiNet vs. Benchmarks

Metrics	VGG16	InceptionV3	ResNet	MobileNet	GBCNet	RadFormer	CortiNet
Acc.	0.9360	0.9804	0.9794	0.9809	0.9275	0.9682	0.9874
Recall	0.9358	0.9768	0.9783	0.9816	0.9281	0.9684	0.9869
Spec.	0.9919	0.9987	0.9974	0.9934	0.9909	0.9960	0.9984
Prec.	0.9371	0.9898	0.9797	0.9864	0.9283	0.9685	0.9873
NPV	0.9920	0.9987	0.9974	0.9901	0.9909	0.9960	0.9984
F1 Score	0.9361	0.9892	0.9788	0.9853	0.9274	0.9683	0.9870
AUC	0.9977	0.9997	0.9985	0.9998	0.9927	0.9991	0.9999

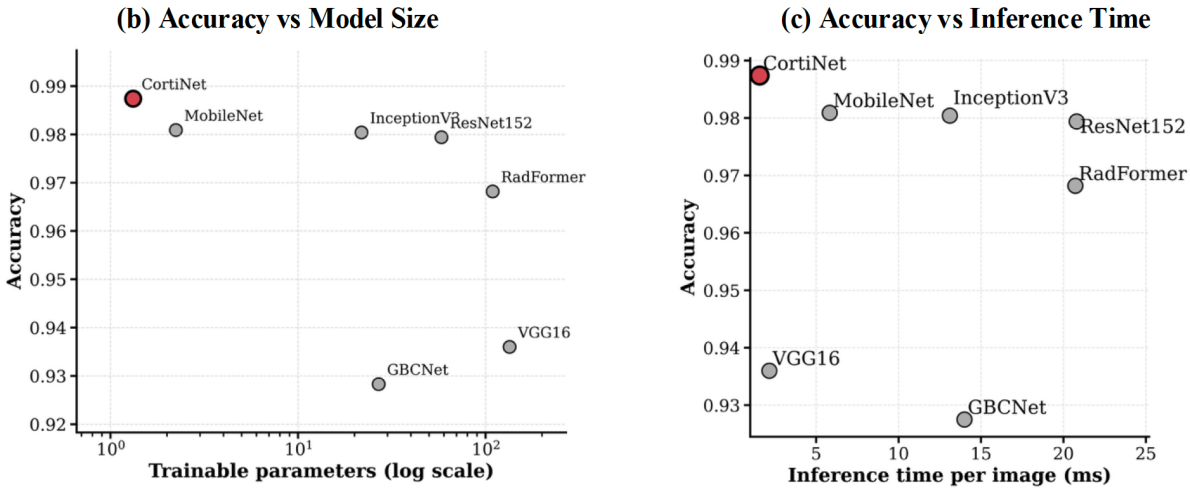


图4：**CortiNet 与基线模型的诊断性能和计算效率比较。**（a）代表性卷积、轻量级和胆囊专用深度学习架构的总体分类性能，采用准确率、召回率、精确率、F1分数和曲线下面积进行评估。CortiNet 在所有指标上均持续取得最高性能。（b）以可训练参数数量为指标的模型复杂度比较，结果表明，与通用模型和特定任务模型相比，CortiNet 在大幅减少参数的同时，取得了最准确的结果。（c）每张图像的推理时间，表明 CortiNet 以最低延迟实现了更高的诊断准确率，为实时且资源受限的临床部署提供支持。

分析结果也印证了这一点。图6对比了传统 Grad-CAM 与所提出的结构感知归因策略。如图6(b)所示，传统 Grad-CAM 生成的激活区域较为弥散且空间分散，延伸至胆囊边界之外，并突出显示受散斑主导的背景区域。相比之下，源自 CortiNet 结构分支的结构感知 Grad-CAM 生成了定位清晰且符合解剖结构的激活图，与临床相关的感兴趣区域相吻合，包括胆囊壁–腔界面及发生病理性变形的区域（图6(c)）。重要的是，在散斑噪声扰动下，这些归因模式仍保持空间稳定（图6(d)），表明解释本身继承了底层结构表征的鲁棒性。从临床角度看，这一特性与专家的放射学解读高度一致：诊断依据是整体形态学线索，而非像素级纹理。通过在解释过程中将感知细节与结构推理解耦，CortiNet 生成的归因图不仅更加稳定，也更便于临床解读。更广泛而言，这些结果支持了以下假设：皮层启发式的结构与感知处理分离，为构建可解释、可信且适于部署的医疗人工智能系统提供了一条有原则的方法路径。

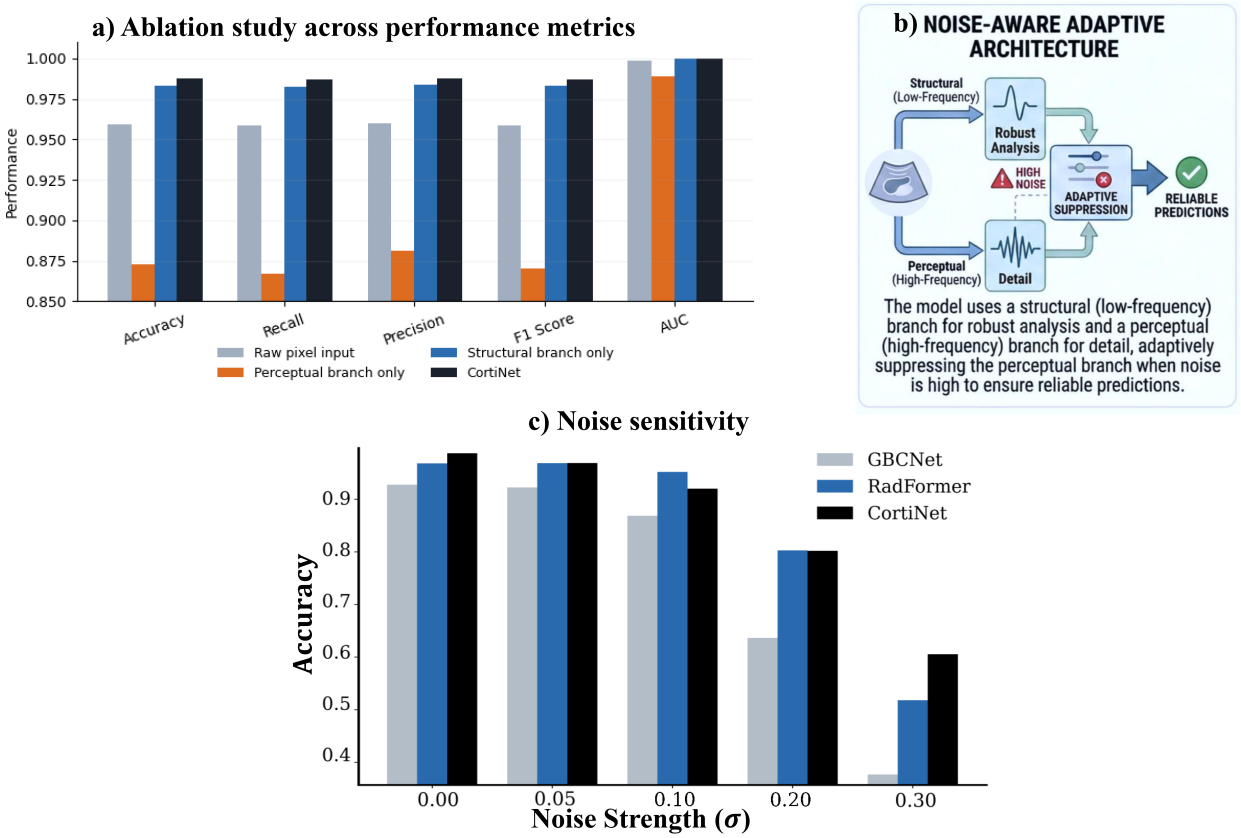


图 5：**所提出的自适应双分支架构的消融与噪声鲁棒性分析**。(a) 对关键性能指标进行定量消融研究，比较原始像素输入、仅感知（高频）分支、仅结构（低频）分支以及完整的 CortiNet 模型。尽管两个分支都有助于提升性能，但二者结合后，在准确率、F1分数和曲线下面积方面始终表现更优，凸显了结构表征与感知表征的互补作用。(b) 所提出的噪声感知自适应架构示意图，其中结构分支执行稳健的低频分析，感知分支则捕捉细粒度细节；在高噪声条件下，自适应抑制会降低感知分支的贡献，从而确保预测稳定。(c) CortiNet 与代表性胆囊聚焦模型的分类准确率对噪声强度增加的敏感度。噪声水平较高时，CortiNet 的准确率下降幅度最小，表明自适应分支选择和结构主导推理提升了模型的鲁棒性。

4 讨论

用于超声诊断的深度学习模型面临着一个根本性的矛盾：既要对细微视觉线索保持敏感，又要抵御噪声、操作者变异性和采集伪影的影响。在胆囊成像中，解决这一矛盾尤具挑战性，因为具有诊断意义的结构与散斑主导纹理及低信噪比并存。本文结果表明，CortiNet 通过明确分离并自适应地整合感知表征与结构表征来应对这一挑战，而非将二者纠缠在单一的特征层次中。

性能提升的机制性解读： 消融研究直接证明，CortiNet 的性能优势源于自适应分支选择，而非简单的特征聚合。尽管结构（低频）分支单独使用时已经表现出较强性能——这反映了解剖连续性在超声诊断中的重要性——完整的 CortiNet 模型还会在感知纹理信息具有信息价值时选择性地将其整合，从而进一步提升准确率、召回率和校准表现。相比之下，仅使用感知分支会导致性能显著下降，凸显了在超声判读中单独依赖高频线索的不可靠性。从数学角度看，这种行为可解释为一种由输入条件决定的互补表征自适应加权，使 CortiNet 能够在噪声较大或低对比度条件下突出结构主导特征，并在信号质量允许时利用纹理线索。这与来自人类视觉皮层的证据相一致，

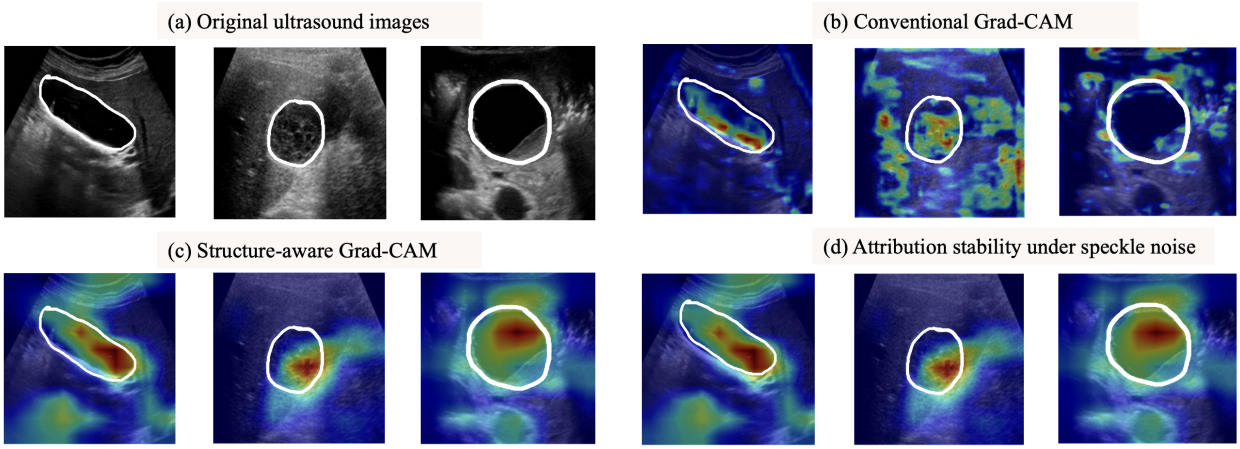


图6：**基于梯度归因的结构感知可解释性**。（a）具有专家标注解剖边界的代表性胆囊超声图像。（b）应用于标准端到端模型的传统Grad-CAM会产生弥散且对噪声敏感的归因模式，并延伸至胆囊区域之外。（c）由CortiNet结构（低频）分支得到的结构感知Grad-CAM可产生解剖上连贯的激活，集中于临床相关区域，包括胆囊壁和管腔界面。（d）散斑噪声扰动下的归因稳定性表明，尽管图像质量下降，结构感知解释在空间上仍保持一致。总体而言，所提出的可解释性策略使模型注意力与放射学推理相一致，并抑制虚假的纹理驱动激活，从而支持可信的临床解读。

其中分层的an　　其中并行通路会根据任务动态调节对空间频率成分的敏感度　　k
需求与信号re　　责任。

抗噪性与临床可靠性：在散斑噪声水平不断升高的条件下进行的鲁棒性分析揭示了CortiNet与基线模型之间的关键差异。随着噪声强度增加，仅感知模型和原始像素模型的准确率出现明显下降，而CortiNet在较宽的噪声范围内保持近乎恒定的性能，总体准确率下降最小。这一行为可由噪声敏感性指数简洁概括，该指数直接反映最坏情况下的性能下降——一项具有临床意义的指标。从部署角度来看，这种鲁棒性至关重要。超声图像质量因操作者、设备和患者解剖结构而差异很大。在此类变异下预测仍保持稳定的模型，对于真实世界的临床工作流程本质上更为可靠。结构分支在噪声下的稳定性，结合CortiNet的自适应融合策略，确保诊断决策不会受到瞬时散斑伪影的过度影响。

效率与临床可部署性：除了准确率之外，与常用的深度卷积神经网络相比，CortiNet 在实现这些结果的同时，可训练参数和浮点运算次数少了几个数量级。这种高效性直接转化为更低的推理延迟和更少的硬件需求，使 CortiNet 非常适合床旁部署、便携式超声系统以及资源受限的临床环境。从这个意义上说，CortiNet 挑战了长期以来的普遍假设：更高的诊断性能必然需要规模更大、计算量更高的模型。

结构感知可解释性：尽管详细的可解释性分析另行呈现，但值得指出的是，将基于梯度的归因限制在结构分支内，可产生与已知解剖标志相符、符合临床认知的显著性模式。这进一步表明，CortiNet 的预测依据的是具有诊断意义的结构，而非虚假的纹理相关性；这是赢得临床信任和获得监管认可的必要条件。

5 结论

本研究提出了 CortiNet，这是一种皮层启发式、结构感知深度学习框架，从准确率、鲁棒性和可部署性三个关键维度推动了超声图像分类的发展。通过明确解耦并自适应融合感知信息与结构信息，CortiNet 在保持计算开销较低且能够抵御噪声引起的性能下降的同时，实现了卓越的诊断性能。本研究的关键贡献不在于渐进式的架构改进，而在于证明了生物启发式归纳偏置，尤其是自适应的结构主导处理，能够为安全关键型医学成像任务带来切实收益。CortiNet 展示了以人类视觉认知为基础的原则性架构设计，能够超越更大型且更

复杂的模型，同时提供更出色的鲁棒性和可解释性。展望未来，该框架为超声领域及更广泛场景中的可泛化、结构感知人工智能系统开辟了新的发展方向，尤其适用于具有异质图像质量和有限计算资源特征的场景。通过将皮层启发与临床实用主义相结合，CortiNet 朝着可信赖且可部署的医疗人工智能迈出了有意义的一步。

尽管 CortiNet 展现出了出色的诊断性能和鲁棒性，但仍有若干局限值得考虑。首先，本研究基于回顾性收集的超声图像开展。尽管该数据集涵盖了广泛的胆囊病变，但仍需在多机构队列和异质成像方案中进行前瞻性验证，才能充分确立其临床可推广性。其次，目前的可解释性分析侧重于结构分支中的基于梯度的归因；尽管这一选择提高了解剖学可解释性，但尚未涵盖在实际放射学决策中常见的时间推理或多视图推理。最后，皮层启发式自适应分支选择机制通过经验评估，而非直接进行神经生理学验证；虽然观察到的提升支持结构感知自适应整合有益这一假设，但进一步的理论依据和神经科学基础有助于强化这一联系。通过前瞻性研究、更完善的伪影建模以及扩展的可解释性分析来解决这些局限，是未来研究的重要方向。

致谢

V. Kumar 感谢印度教育部（MoE）以总理研究奖学金（PMRF）的形式提供的经费支持。

代码可用性

论文被接收后，通讯作者将在 GitHub 上公开本研究中用于复现结果的所有源代码。

利益冲突

作者声明不存在任何利益冲突。

参考文献

[1] Piero Portincasa, A Di Ciaula, Ornella de Bari, Gabriella Garruti, Vincenzo Ostilio Palmieri 和 DQ-H Wang。胆结石及其相关并发症的管理。胃肠病学与肝病学专家评论，10(1):93–112，2016。[2] Laura M Stinton 和 Eldon A Shaffer。胆囊疾病的流行病学：胆石症与癌症。肠道与肝脏，6(2):172，2012。[3] Dan Huang、Hyundeok Joo、Nan Song、Sooyoung Cho、Woosung Kim 和 Aesun Shin。胆结石与胆道癌风险之间的关联：系统综述与荟萃分析。*Epidemiology and Health*，43:e2021011，2021。[4] Freddie Bray、Mathieu Laversanne、Hyuna Sung、Jacques Ferlay、Rebecca L Siegel、Isabelle Soerjomataram 和 Ahmedin Jemal。2022年全球癌症统计数据：Globocan对185个国家36种癌症的全球发病率和死亡率估算。*CA*：临床医生癌症杂志，74(3):229–263，2024。[5] Pankaj Gupta、Kesha Meghashyam、Yashi Marodia、Vikas Gupta、Rajender Basher、Chandan Krushna Das、Thakur Deen Yadav、Santhosh Irrinki、Ritambhra Nada 和 Usha Dutta。局部晚期胆囊癌：影像学标准及作用综述。*Abdominal Radiology*，46:998–1007，2021。[6] Danilo Avola、Luigi Cinque、Alessio Fagioli、Gianluca Foresti 和 Alessio Mecca。超声医学成像技术：综述。*ACM Computing Surveys (CSUR)*，54(3):1–38，2021。[7] Thomas L Szabo。诊断性超声成像：由内而外。Academic Press，2013。[8] A Lodhi、K Waite 和 I Alam。超声检查诊断胆囊息肉的准确率。*Radiography*，26(2):e52–e55，2020。[9] Robert J McDonald、Kara M Schwartz、Laurence J Eckel、Felix E Diehn、Christopher H Hunt、Brian J Bartholmai、Bradley J Erickson 和 David F Kallmes。断层成像使用情况变化和技术进步对放射科医师工作负荷的影响。*Academic Radiology*，22(9):1191–1198，2015。

[10] Medicus Healthcare Solutions。放射科医师短缺：弥合供需缺口，2025。访问日期：2025-05-05。[11] Jan Egger、Christina Gsaxner、Antonio Pepe、Kelsey L Pomykala、Frederic Jonske、Manuel Kurz、Jianning Li 和 Jens Kleesiek。医学深度学习——系统性元综述。生物医学中的计算机方法与程序，221:106874，2022。[12] Maryam Fallahpoor、Subrata Chakraborty、Biswajeet Pradhan、Oliver Faust、Prabal Datta Barua、Hossein Chegeni 和 Rajendra Acharya。PET/CT成像中的深度学习技术：从正弦图到图像空间的全面综述。生物医学中的计算机方法与程序，243:107880，2024。[13] Hao Wang、Euijoon Ahn、Lei Bi 和 Jinman Kim。用于多标签皮肤病变分类的自监督多模态学习。生物医学中的计算机方法与程序，265:108729，2025。[14] Mario Fernando Jojoa Acosta、Liesle Yail Caballero Tovar、Maria Begonya Garcia-Zapirain 和 Winston Spencer Percybrooks。利用深度学习技术诊断皮肤镜图像中的黑色素瘤。*BMCMedicalImaging*，21:1–11，2021。[15] Ivan A Bratchenko、Lyudmila A Bratchenko、Yulia A Khristoforova、Alexander A Moryatov、Sergey V Kozlov 和 Valery P Zakharov。使用卷积神经网络分析拉曼光谱进行皮肤癌分类。生物医学中的计算机方法与程序，219:106755，2022。[16] Rahul Paul、Matthew Schabath、Robert Gillies、Lawrence Hall 和 Dmitry Goldgof。卷积神经网络集成可准确预测两年后的肺结节恶性程度。生物学与医学中的计算机，122:103882，2020。[17] Tan-Cong Nguyen、Tien-Phat Nguyen、Tri Cao、Thao Thi Phuong Dao、Thi-Ngoc Ho、Tam V Nguyen 和 Minh-Triet Tran。MANet：用于肺结节检测与分割的多分支注意力辅助学习。生物医学中的计算机方法与程序，241:107748，2023。[18] Paras Lakhani 和 Baskaran Sundaram。胸部X线摄影中的深度学习：使用卷积神经网络自动分类肺结核。放射学，284(2):574–582，2017。[19] Maoneng Hu、Zichen Wang、Xinxin Hu、Yi Wang、Guoliang Wang、Huanhuan Ding 和 Mingmin Bian。基于级联深度监督U-Net的尘肺病合并肺结核高分辨率计算机断层扫描诊断。生物医学中的计算机方法与程序，226:107151，2022。[20] Ahmad Saleh、Rozana Sukaik 和 Samy S Abu-Naser。使用深度学习进行脑肿瘤分类。载于*2020*年辅助与康复技术国际会议（*iCareTech*），第131–136页。IEEE，2020。[21] Yoonseok Choi、Mohammed A Al-Masni、Kyu-Jin Jung、Roh-Eul Yoo、Seong-Yeong Lee 和 Dong-Hyun Kim。用于有限磁共振图像模态下脑肿瘤分割的单阶段知识蒸馏网络。生物医学中的计算机方法与程序，240:107644，2023。[22] Rita Appiah、Venkatesh Pulletikurthi、Helber Antonio Esquivel-Puentes、Cristiano Cabrera、Nahian I Hasan、Suranga Dharmarathne、Luis J Gomez 和 Luciano Castillo。将本征正交分解与深度学习网络相结合进行脑肿瘤检测。生物医学中的计算机方法与程序，250:108167，2024。[23] Caroline Barcelos Gonçalves、Jefferson R Souza 和 Henrique Fernandes。利用仿生算法优化CNN架构，以检测红外图像中的乳腺癌。生物学与医学中的计算机，142:105205，2022。[24] Shuo Dai、Xueyan Liu、Wei Wei、Xiaoping Yin、Lishan Qiao、Jianing Wang、Yu Zhang 和 Yan Hou。用于准确乳腺肿瘤分割的多尺度、多任务融合UNet模型。生物医学中的计算机方法与程序，258:108484，2025。[25] Carlos Aumente-Maestro、Jorge Díez 和 Beatriz Remeseiro。用于超声成像中乳腺癌分割与分类的多任务框架。生物医学中的计算机方法与程序，260:108540，2025。[26] Moutaz Alazab、Albara Awajan、Abdelwadood Mesleh 和 Salah Alhyari。使用深度学习预测和检测COVID-19。计算机信息系统与工业管理应用国际期刊，12:14–14，2020。[27] Rutwik Gulakala、Bernd Markert 和 Marcus Stoffel。利用渐进式增长生成对抗网络和CNN优化快速诊断COVID-19感染。生物医学中的计算机方法与程序，229:107262，2023。[28] Hong Ju、Yanyan Cui、Qiaosen Su、Liran Juan 和 Balachandran Manavalan。CodeNet：一种用于COVID-19检测的深度学习模型。生物学与医学中的计算机，171:108229，2024。

[29] Ramprasaath R Selvaraju、Michael Cogswell、Abhishek Das、Ramakrishna Vedantam、Devi Parikh 和 Dhruv Batra。Grad-CAM：通过基于梯度的定位，从深度网络中生成视觉解释。载于*IEEE*计算机视觉国际会议论文集，第 618–626 页，2017。[30] Erico Tjoa 和 Cuntai Guan。可解释人工智能（XAI）综述：迈向医学可解释人工智能。*IEEE*神经网络与学习系统汇刊，32(11):4793–4813，2020。[31] Younbeom Jeong、Jung Hoon Kim、Hee-Dong Chae、Sae-Jin Park、Jae Seok Bae、Ijin Joo 和 Joon Koo Han。基于深度学习的超声检查肿瘤性胆囊息肉诊断决策支持系统：初步结果。*Scientific Reports*，10(1):7700，2020。[32] Ahmed Mahdi Obaid、Amina Turki、Hatem Bellaaj、Mohamed Ksantini、Abdulla AlTaee 和 Alaa Alaerjan。利用深度学习检测胆囊疾病类型：一种有效的医学方法。*Diagnostics*，13(10):1744，2023。[33] Soumen Basu、Mayank Gupta、Pratyaksha Rana、Pankaj Gupta 和 Chetan Arora。超越人类准确率：采用课程学习从超声图像中检测胆囊癌。载于*IEEE/CVF*计算机视觉与模式识别会议（*CVPR*）论文集，第 20886–20896 页，2022。[34] Soumen Basu、Mayank Gupta、Pratyaksha Rana、Pankaj Gupta 和 Chetan Arora。RadFormer：采用全局-局部注意力的 Transformer，实现可解释且准确的胆囊癌检测。医学图像分析，83:102676，2023。[35] Shagun Sharma 和 Kalpna Guleria。用于图像分类的深度学习模型：比较与应用。载于*2022*年第*2*届先进计算与工程创新技术国际会议（*ICACITE*），第 1733–1738 页。IEEE，2022。[36] Amitojdeep Singh、Sourya Sengupta 和 Vasudevan Lakshminarayanan。医学图像分析中的可解释深度学习模型。*Journal of Imaging*，6(6):52，2020。[37] Amina Turki、Ahmed Mahdi Obaid、Hatem Bellaaj、Mohamed Ksantini 和 Abdulla AlTaee。Uidatagb：用于胆囊疾病检测的多类别超声图像数据集。*Data in Brief*，54:110426，2024。