DOI:10.20033/j.1003-7241.(2026)07-0114-05

一种基于深度学习的散斑噪声抑制方法

丁礼磊

(安康学院 电子与信息工程学院,陕西 安康 725000)

摘　要:散斑噪声是由一个分辨单元内众多散射体的反射波叠加形成的,广泛存在于光学检测和医学成像等领域的图像中。以散斑噪声为主的噪声干扰严重影响视网膜光学相干层析(optical coherence tomography,OCT)图像质量,降低了图像的灰度分辨率,模糊了图像相关结构信息。传统抑制散斑噪声方法在保留图像结构信息方面效果不是很好,有监督学习方法存在训练数据匮乏和难以获得作为标签的无噪声图像的难题。针对此,提出一种无监督学习抑制散斑噪声的方法,仅以单张有噪声的图像为训练样本,利用卷积神经网络强大的拟合能力来学习从图像到其相应的高质量图像的复杂映射。实验表明:所提方法的评价具有良好的降噪效果,验证了方法在数据不充分时,仍能较好地抑制图像噪声,有效地保留OCT图像中结构信息,且有很强的泛化能力,具有很强的实用性。

关键词:散斑噪声;深度学习;无监督学习;光学相干层析成像;图像降噪;卷积神经网络

中图分类号:TP391.41;TP18;TP273　　**文献标志码**:A　　**文章编号**:1003-7241(2026)07-0114-05

A speckle noise suppression method based on deep learning

Ding Lilei

(School of Electronics and Information Engineering, Ankang University, Ankang 725000, Shaanxi, China)

Abstract: Speckle noise is formed by the superposition of reflected waves from numerous scatterers within a resolution unit, and is widely present in images in fields such as optical detection and medical imaging. The noise interference mainly caused by speckle noise seriously affects the quality of retinal optical coherence tomography (OCT) images, reduces the grayscale resolution of the images, and blurs the relevant structural information of the images. Traditional methods for suppressing speckle noise are not very effective in preserving image structural information, and supervised learning methods face challenges such as insufficient training data and difficulty in obtaining noise free images as labels. A method for unsupervised learning to suppress speckle noise is proposed, which only uses a single noisy image as a training sample and utilizes the powerful fitting ability of convolutional neural networks to learn complex mappings from the image to its corresponding high-quality image. The experiment shows that the evaluation of the proposed method has good noise reduction effect, which verifies that the method can effectively suppress image noise and preserve structural information in OCT images even when the data is insufficient. It also has strong generalization ability and strong practicality.

Keywords: speckle noise; deep learning; unsupervised learning; optical coherence tomography; image denoising; convolutional neural network

散斑噪声是OCT成像所固有的、无法直接去除的噪声[1-2]。散斑噪声对图像分辨率和图像对比度的破坏作用严重影响OCT技术在临床,特别是高分辨率成像方面的应用[3]。OCT图像散斑噪声抑制的主要目标是在抑制噪声的同时,能够保持图像结构细节的完整性[4],目前主要有硬件和软件两种方式。硬件方式需要对OCT系统的整体架构进行设计和改进,价格相对昂贵[5]。软件方式是将生成的OCT图像利用数字图像处理技术进行后期处理,该方式大致可以分为传统的数字图像处理方法和基于深度学习的方法。传统的数字图像处理方法有利用滤波器进行降噪,比较经典的有高斯滤波器、均值滤波器和中值滤波器等[6-7],这些方法通常会模糊图像中的特征和边缘信息。为了解决这一问题,出现了许多改进算法,如非局部均值(non-local means, NLM)[8]、三维块匹配滤波(block matching 3D, BM3D)[9]、字典学习[10-11]和多帧Bscan图像平均等方法,这些算法一定程度上解决了降噪的同时丢失了图像结构特征信息的问题,但是也出现了参数调整复杂和图像采样时间增加等问题。

近几年来,随着计算机软硬件的飞速发展和大数据、云计算时代的到来,基于深度学习的方法,已经被广泛应用于图像处理中,如图像分类、图像识别和图像分割等,将传统的降噪算法融入深度学习中也是近几年的研究热点。刘建宾等[12]将深度学习技术和非局部均值算法结合起来探索新的降噪方法,将含噪图像所有相邻点的像素值、相

收稿日期:2025-01-09;**录用日期**:2025-03-13

基金项目:陕西省教育厅科学研究计划项目(25JU007);陕西省教育厅科学研究计划项目(20JK0475)

作者简介:丁礼磊(1985—),男,硕士研究生,研究方向为深度学习、嵌入式系统设计。

引用本文:丁礼磊. 一种基于深度学习的散斑噪声抑制方法[J]. 自动化技术与应用, 2026,45(7):114-117,122. (Ding Lilei. A speckle noise suppression method based on deep learning[J]. Techniques of Automation and Applications, 2026,45(7):114-117,122.)

似性值作为深度学习网络的输入,将原图像作为输出进行训练,使训练好的模型可直接用于滤波。郝富贵[13]使用基于 FFDNet 网络,将有噪声的原始图像和相应的无噪声标签图像构成的训练数据集对网络进行训练,然后利用网络学习原始图像中的噪声,用原始图像减去学习到的噪声便可实现图像的降噪。基于深度学习的降噪方法在降噪效果和保留图像结构特征方面取得很好的表现,但是模型训练需要大量无噪声的图片作为训练集,获取无噪声的图片又比较困难。近年来,无监督深度学习给图像降噪提供了新思路。Zhu 等[14]受 Noise2Noise(N2N)策略启发,提出了一种针对低剂量 CT(low-dose computed tomography, LDCT)的自监督去噪方法。经典 N2N 需要两张独立含噪图像构成 Noisy-Noisy 训练对,而该方法通过向原始 LDCT 图像中人工叠加二次噪声来生成第二张含噪图像,从而绕开了临床难以获取双扫描的障碍。吴广义等[15]提出了基于 N2N 训练策略的无监督深度残差稀疏注意力网络(DRSA Net)用于视网膜 OCT 图像降噪,并取得了很好效果,但此种方法需要保证成对的图像拍摄时是相同场景,实际使用的时候,两次不同的观测很难是相同的场景。比如医学图像,两次拍摄的时间不一样,也许器官在位置上发生了细微的变化,所以很难得的相同场景下成对的图像。针对 N2N 策略中需要在相同场景下成对照片这一缺点,Krull 等[16]在 2019 年提出了 Noise2Void(N2V)策略,N2V 完全采用了和 N2N 不一样的思路,仅用一张或多张有噪声的图像作为训练集,即可完成降噪模型的训练和图像的降噪任务。本文将 N2V 策略应用到视网膜 OCT 图像降噪中,取得了较好的降噪效果,该方法最接近实际使用场景,因此能得到很好的利用。

1 方法

1.1 N2V 无监督训练原理

有监督学习是利用一组带有标签的数据,学习从输入到输出的映射关系,然后将这种映射关系应用到未知数据上,达到分类或回归的目的,即给网络一个输入 x, 获得预测输出 y, 不断调节网络参数 θ, 通过让 y 接近实际目标,经过不断的训练,最终 θ 保存了使网络取得最小损失值时网络参数,通过学习可完成降噪的任务。

$$f(x;\theta) \rightarrow y \tag{1}$$

式中, x 是网络输入, y 是网络输出, θ 是网络参数。

Noise2Void(N2V)无监督训练学习目标是利用一张或几张实际拍摄的含噪声图像完成训练,不需要无噪声图像作为标签数据。该策略原理是给定含噪声的训练图像,在图像上确定一个感受野(如 5 × 5 像素块),将感受野的中心部分挖掉一个像素,然后剩余的部分作为输入 x, 把网络看不到的这个中心像素点(被挖掉的像素点)称为盲点,将该中心像素作为目标 T, 网络的输出预测值 y 和目标 T 之间做 Loss,就可以促进网络学会降噪,原理示意图如图 1 所示。

该原理核心是通过中心像素周围的信息预测真实信号(不含噪声的信号),所以该策略有三个假设:第一,不同像素位置,信号是相互关联的,某个像素的值会受到其周围像素值的影响,这也符合一般的成像规律;第二,不同像素位置,噪声是相互独立的;第三,噪声的均值为 0。由于生物组织的高散射性,照射到生物组织的入射光被生物组织内的颗粒所散射,形成无规则分布的颗粒状结构的衍射图样即为散斑,散斑的形成是无规律、随机的[17],散斑噪声之间是相互独立的,从统计学角度看,大量散斑噪声的均值为 0,第二、三个假设也满足。

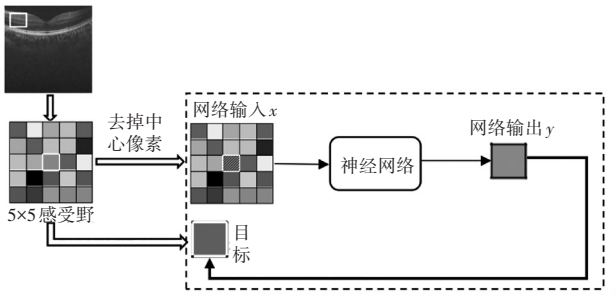


图 1 N2V 策略原理示意图

Fig. 1 Schematic diagram of N2V strategy principle

从以上基本策略和假设可以看出,N2V 无监督训练学习利用图像像素之间的关联性,通过周围像素的信号推测出中心像素信号,但是噪声无法通过周围像素的噪声推测出中心像素的噪声,解决这一问题分析如下:将网络输出预测的中心像素 X_{pre} 和输入像素块中被屏蔽的中心像素 X_{in} 做 Loss。

$$X_{pre} - X_{in} = s_{pre} + n_{pre} - (s_{in} + n_{in}) = s_{pre} - s_{in} + (n_{pre} - n_{in}) \tag{2}$$

式中, s 和 n 分别表示图片的信号和噪声,网络的输出如(3)式。

$$X_{pre} = s_{pre} + n_{pre} \rightarrow E(s_{in} + \text{随机噪声}) = s_{in} + E(\text{随机噪声}) \tag{3}$$

根据噪声零均值假设 $E(\text{随机噪声}) = 0$, 有

$$X_{pre} \rightarrow s_{in} + E(\text{随机噪声}) = s_{in} \tag{4}$$

即通过周围像素的预测值 X_{pre} 可作为真实信号 S_{in}。让网络学习一个点周围所有点到该点的映射,当网络有大量点到点学习的时候,网络会优先输出目标点的均值,由于噪声均值假设为 0,所以输出结果就是图像中不含噪声的真实信号。

1.2 训练方式

实验神经网络选用 U-Net, U-Net 网络结构最早由 Ronneberger 等于 2015 年提出一种基于卷积神经网络的图像分割框架,U-Net 凭借其出色的性能在众多的分割网络中脱颖而出,研究者以 U-Net 为基础相继提出了多种改进变体[18]。整体网络结构成 U 型结构,包括编码部分和解码部分,两部分对应阶段之间使用跳跃连接结构,复用低层次特征信息,以更好还原图像细节信息。编码部分通过卷积层+RELU 激活函数+最大池化层来提取图像特征,

解码部分使用反卷积进行上采样,上采样之后同样使用卷积层,逐步回复图像的细节信息,并最终恢复特征图至输入图像的尺寸[19]。

输入一张(多张)噪声图像得到干净图像基本过程如下:以一个像素为中心将噪声图像进行分割若干块,将块作为神经网络的输入(input),并以这个中心像素作为目标(target),该网络将会学习直接输入块中心的像素映射到网络的输出位置上。为了提高训练效率,将一张给定用于训练的噪声图像,随机裁剪出若干小块(如 64×64,32×32),任取一个小块,在该块上随机选取 N 个像素点,对于每一个点裁剪出以其为中心、以感受野为大小的块,在这个块中用附近的一个像素的值替换中心位置的像素值,并以这个中心像素作为目标,不断优化网络参数,达到去噪的目的。

1.3 实验环境和参数设置

BM3D 算法是最具代表的优秀降噪算法,实验使用 N2V 训练策略和 U-Net 类型的 CNN,并与 BM3D 降噪进行对比试验,实验流程如图 2 所示。深度学习框架基于 Tensorflow 构建、训练和测试,批大小设置为 64,学习率为 0.001,Adam 为优化器,最大迭代次数设置为 35,在 35 次迭代中具有最小损失值的模型被保存以供进一步评估,选取 MSE 作为损失函数。

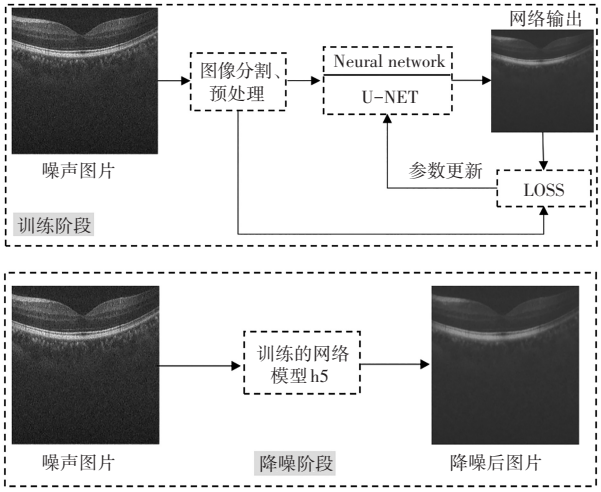


图 2 OCT 无监督训练及降噪流程

Fig. 2 Flowchart of OCT unsupervised training and denoising

1.4 评价指标

评价指标选取 4 种常用的定量指标对图像的降噪效果进行评价,包括峰值信噪比(peak signal-to-noise ratio,PSNR)、结构相似性(structural similarity index,SSIM)、等效视数(equivalent number of looks,ENL)、边缘保留指数(edge preservation index,EPI)。

PSNR 是全局性评价指标,为降噪图像和真值图像对应像素之间的差异,是最经典的降噪评价指标,单位为 dB,具体定义如下。

$$\mathrm{PSNR} = 20 \times \lg\left(\frac{\mathrm{MAX}_I}{\mathrm{MSE}}\right) \tag{5}$$

式中,MSE 表示降噪图像和真值图像之间的均方差,MAX_I 代表图像 I 中最大的强度值。

SSIM 是比较降噪图像和真值图像之间结构相似性指标,具体定义如下。

$$\mathrm{SSIM} = [l(I_\mathrm{d},I_\mathrm{c})^{\alpha} \cdot c(I_\mathrm{d},I_\mathrm{c})^{\beta} \cdot s(I_\mathrm{d},I_\mathrm{c})^{\gamma}] \tag{6}$$

式中,$(I_\mathrm{d},I_\mathrm{c})$、$c(I_\mathrm{d},I_\mathrm{c})$、$s(I_\mathrm{d},I_\mathrm{c})$ 分别可表示为

$$l(I_\mathrm{d},I_\mathrm{c}) = \frac{2\mu_\mathrm{d}\mu_\mathrm{c} + c_1}{\mu_\mathrm{d}^2 + \mu_\mathrm{c}^2 + c_1} \tag{7}$$

$$c(I_\mathrm{d},I_\mathrm{c}) = \frac{2\sigma_\mathrm{d}\sigma_\mathrm{c} + c_2}{\sigma_d^2 + \sigma_\mathrm{c}^2 + c_2} \tag{8}$$

$$s(I_\mathrm{d},I_\mathrm{c}) = \frac{\sigma_\mathrm{dc} + c_3}{\sigma_\mathrm{d}\sigma_\mathrm{c} + c_3} \tag{9}$$

式中,I_d、I_c 分别为原始噪声图像和降噪图像;μ_d、μ_c 分别为 I_d、I_c 的平均值;σ_d、σ_c 分别为 I_d、I_c 的标准差;σ_dc 为 I_d、I_c 的协方差;c_1、c_2、c_3 为常数。α、β、γ 通常设置为 1。

EPI 是反映降噪图像边缘细节保持程度的性能指标,感兴趣区域(ROI)的 EPI 可定义为

$$\mathrm{EPI} = \frac{\sum_W \sum_H |I_\mathrm{n}(x+1,y) - I_\mathrm{n}(x,y)|}{\sum_W \sum_H |I_\mathrm{d}(x+1,y) - I_\mathrm{d}(x,y)|} \tag{10}$$

式中,W、H 分别为图像的宽度和高度。

ENL 用于测量降噪图像中均匀区域的平滑度,ENL 可定义为

$$\mathrm{ENL} = \frac{\mu^2}{\sigma^2} \tag{11}$$

式中,μ 为感兴趣区域(region of interest,ROI)的平均值,σ 为标准差。

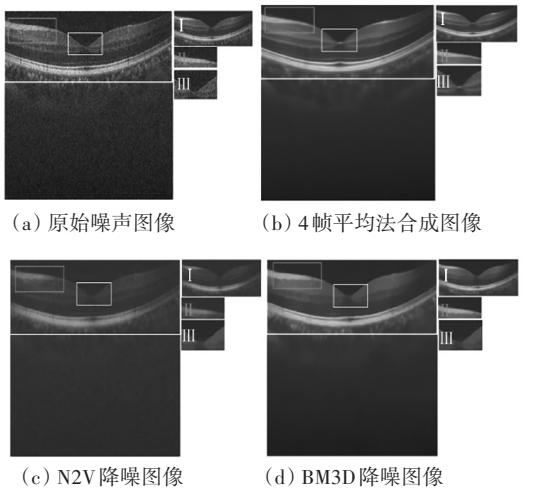
2 结果与分析

为了验证去噪网络和训练策略的有效性,选用在视网膜 OCT 图像降噪任务中被广泛使用的 Duke 大学公开数据集[20]来训练和评估网络的去噪效果。该数据集采用 Bioptigen 公司的 840 nm 光谱域光学相干层析成像(spectral domain OCT,SD-OCT)系统获取,采自患有或未患有年龄相关性黄斑病变的受试者。数据集共含有 18 组图像,每组由连续位置拍摄的 4 张相邻噪声图像和由这 4 张图像采用多帧平均法合成的无噪声图像 1 张,像素大小为 450 × 900(高×宽),本次实验选用第一组 4 张中的 1 张图像作为网络训练输入的噪声图像,使用多帧平均法合成的图像作为无噪声图像进行相关指标的计算。

图 3 为原始噪声图像、4 帧平均法合成图像、N2V 降噪图像和 BM3D 降噪图像。OCT 图像的前景视网膜区域是需要重点关注的,这个区域包含了较多医学信号,为了摒弃背景区域的干扰,实验验证时选择 3 个感兴趣区域(ROI),一个 ROI 是图像上半部分(Ⅰ矩形框)计算 SSIM 和 EPI,其他两个 ROI(Ⅱ和Ⅲ矩形框)计算 ENL。

表 1 为 BM3D 算法和 N2V 无监督学习降噪图像的客观数值评价指标,由表 1 可知,与原始噪声图像相比,两种

降噪方法在图像的各项评价指标方面都实现了较大的提升,相比块匹配算法 BM3D,N2V 无监督学习降噪时间大幅缩短,可在 2 s 内完成对 OCT 图像的高水平降噪,SSIM、EPI 和 ENL 指标与 BM3D 相当或高于 BM3D,但 PSNR 指标略低于 BM3D,从图上也可以直观看出 BM3D 的背景去噪效果略高于 N2V。



(a) 原始噪声图像　(b) 4帧平均法合成图像

(c) N2V降噪图像　(d) BM3D降噪图像

图 3　视网膜 OCT 图像降噪结果

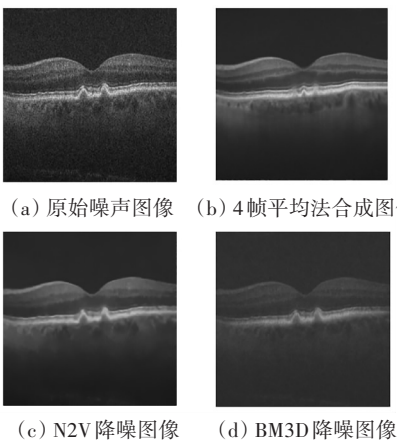
Fig. 3　Denoising results of retinal OCT images

表 1　降噪数值评价表

Tab. 1　Quantitative evaluation of denoising

Method	PSNR	SSIM	EPI	ENL	Time/s
Noisy	17.303	0.126	-	Ⅰ区域:4.060 Ⅱ区域:6.685	-
BM3D	25.520	0.563	0.220	Ⅰ区域:6.730 Ⅱ区域:25.120	126
N2V	24.018	0.564	0.405	Ⅰ区域:11.379 Ⅱ区域:26.710	<2

为了验证模型的泛化能力,从数据集里随机抽取一张噪声图片,使用已训练好的 h5 模型进行降噪,降噪结果如图 4 所示,各项指标如表 2 所示。从图可以看出 BM3D 和 N2V 都有很好的降噪效果,BM3D 降噪图像有些块状的模糊结构,但背景的降噪效果比 N2V 较好,从表 2 指标也可以看出,BM3D 降噪图像的 PSNR 值略高于 N2V,SSIM、EPI 和 ENL 3 项指标 N2V 较高,这与上述验证结果相同,该方法具有很好的泛化能力。



(a) 原始噪声图像　(b) 4帧平均法合成图像

(c) N2V降噪图像　(d) BM3D降噪图像

图 4　模型泛化能力验证

Fig. 4　Validation of model generalization ability

表 2　验证模型泛化能力去噪数值评价结果

Tab. 2　Quantitative evaluation of denoising for model generalization validation

Method	PSNR	SSIM	EPI	ENL	Time/s
Noisy	17.154	0.075	—	3.408	—
BM3D	26.206	0.407	0.198	7.828	128
N2V	24.845	0.518	0.422	25.298	<2

3　结论

根据 N2V 无监督学习策略,选用 U-Net 网络结构作为神经网络,利用单张噪声图片完成高质量的训练,在实现较高水平降噪的同时,保留了视网膜 OCT 图像的一些特征信息。分别从 PSNR、SSIM、EPI、ENL 4 种数值评价对 BM3D 降噪算法和 N2V 无监督学习进行对比分析,在数据匮乏的情况下,除 PSNR 指标外,N2V 无监督学习在各项评价指标上均取得较好的结果,且具有较强的泛化性,降噪时间也大幅缩短。

参考文献

[1]周旭东, 陈明惠, 马文飞, 等. 多尺度条件卷积的 OCT 视网膜图像降噪研究[J]. 光学技术, 2022, 48(1):102-108.

[2]蔡怀宇, 韩晓艳, 娄世良, 等. 扫频光学相干层析视网膜图像配准去噪算法[J]. 中国光学, 2021, 14(2):289-297.

[3]Vb Silva, Dad Jesus, S Klein, et al. Signal-carrying speckle in optical coherence tomography:a methodological review on biomedical applications[J]. Journal of Biomedical Optics, 2022, 27(3):66-75.

[4]谢巧雪, 马宗庆, 祝连庆,等. 基于结构相似约束生成对抗网络的视网膜 OCT 图像去噪算法[J]. 电子测量与仪器学报, 2023, 37 (3):11-20.

[5]潘玲佼, 范伟伟, 吴全玉. 基于双边随机投影算法的光学相干层析成像降噪[J]. 数据采集与处理, 2021, 36(4):713-721.

[6] Saini A, Dogra A. Unleashing the power of image denoising: a comprehensive review of classical to deep learning methods[J]. Journal of Computer Science, 2025, 21(8):1872-1888.

[7]Sharma A, Dogra A, Goyal B , et al. From spatial domain to patch-based models: a comprehensive review and comparison of multimodal medical image denoising algorithms[J]. Computers Materials& Materials. 2025,85(1):367-481.

[8]Mao J , Sun L , Chen J , et al. Overview of research on digital image denoising methods[J]. Sensors, 2025, 25(8):2615.

[9]崔程程, 周先春, 昝明远, 等. 基于自适应滤波的 BM3D 降噪算法[J]. 电子测量技术, 2021, 44(12):97-101.

[10]毛静. 改进的多尺度学习字典图像去噪算法[J]. 计算机与数字工程, 2023, 51(4):899-905.

[11]汤力, 梁国迪, 王菁, 等. 基于机器视觉技术的表单图像缺陷矫正及自动识别方法[J]. 自动化技术与应用, 2024, 43(10):52-55, 138.

[12]刘建宾, 刘保中. 一种基于深度学习的非局部均值图像降噪方法[J]. 计算机仿真, 2020, 37(8):228-234.

[13]郝富贵. 基于 FFDNet 的散斑去噪方法研究与应用[D]. 天津:天津大学, 2019.

[14]Zhu Y, He Q, Yao Y, et al. Self-supervised noise2noise method utilizing corrupted images with a modular network for LDCT denoising [J]. Pattern Recognition, 2025, 161:111285.

(下转第 122 页)