

## 基于变分自编码器的眼科光学相干断层成像图像生成

赵蒙蒙, 鲁贞贞, 朱书缘, 冯继宏\*

北京工业大学环境与生命学部, 智能化生理测量与临床转化北京市国际科研合作基地, 北京 100124

**摘要** 提出了一种基于变分自编码器的眼科光学相干断层成像(OCT)图像生成方法,缓解了深度学习任务中数据不足的问题,有助于提高眼科疾病辅助诊断算法的性能。搭建了基于变分自编码器的 OCT 图像生成网络,并基于目前两个公开的视网膜 OCT 图像数据集,构建了老年性黄斑变性、糖尿病性黄斑水肿和正常三种类别的视网膜 OCT 图像样本数据集,并分别训练网络得到各自的图像生成模型。采用主观视觉评价和客观实验验证两种方式验证了所提图像生成方法是有效的。主观视觉评价和客观实验结果均表明所提方法可以有效生成三种类别的视网膜 OCT 图像。

**关键词** 医用光学; 图像生成; 视光学; 光学相干断层成像; 深度学习; 变分自编码器

中图分类号 TP391.7

文献标志码 A

doi: 10.3788/AOS202141.1417001

## Generation of Optical Coherence Tomography Images in Ophthalmology Based on Variational Auto-Encoder

Zhao Mengmeng, Lu Zhenzhen, Zhu Shuyuan, Feng Jihong\*

Faculty of Environment and Life, Beijing University of Technology, Beijing International Science and Technology Cooperation Base for Intelligent Physiological Measurement and Clinical Transformation, Beijing 100124, China

**Abstract** We proposed an image generation method of ophthalmic optical coherence tomography (OCT) based on the variational auto-encoders to alleviate the problem of insufficient images in deep learning and improve the performance of computer-aided diagnosis algorithms in ophthalmology. We created a generation network of OCT images based on the variational auto-encoders. In addition, we constructed three kinds of retinal OCT image datasets of age-related macular degeneration, diabetic macular edema, and normal situation based on the two public retinal OCT image datasets to train the network and obtain the generation models, respectively. The effectiveness of the image generation method was verified by subjective visual observation and objective experiments. Both the subjective visual observation and the objective experiments show that our method can effectively generate three kinds of retinal OCT images.

**Key words** medical optics; image generation; optometry; optical coherence tomography; deep learning; variational auto-encoder

**OCIS codes** 170.4580; 170.4470; 170.4500; 100.4996

## 1 引言

光学相干断层成像(OCT)技术<sup>[1]</sup>是应用于眼科的重要检查技术,具有无创、非接触、成像结果可以清晰展示病变部位和层次等优点<sup>[2-4]</sup>。随着我国

人口老龄化的加重,眼部疾病患者逐年增多,特别是老年性黄斑变性(AMD)患者<sup>[5-6]</sup>和糖尿病性黄斑水肿(DME)患者<sup>[7-8]</sup>,若不及时有效诊断可能会导致失明<sup>[9]</sup>。目前对这两种疾病的诊断主要依靠眼科医生观察视网膜 OCT 图像,而患者人数较多,给眼科

收稿日期: 2021-01-04; 修回日期: 2021-01-31; 录用日期: 2021-03-08

基金项目: 国家自然科学基金(81971683)、北京市自然科学基金(2202002)

通信作者: jhfeng@bjut.edu.cn

医生带来了巨大的工作量。此外,人工诊断依赖于医生的临床经验,费时费力,还容易造成误诊和漏诊。随着深度学习技术的不断发展,很多研究者提出将计算机辅助诊断系统应用于眼部疾病的筛查和诊断<sup>[10-15]</sup>。为了使辅助诊断系统获得较好的性能,需要使用大量的数据进行训练。但目前公开的视网膜 OCT 图像较少,并且大量视网膜 OCT 图像的获取和标记为眼科医生带来了巨大的工作量。因此,开发一种有效的视网膜 OCT 图像生成方法显得尤为重要。

变分自编码器(VAE)是一种特殊形式的自编码器模型<sup>[16]</sup>,由 Kingma 等于 2014 年提出,在图像生成方面得到了十分广泛的应用<sup>[17-19]</sup>。Wan 等<sup>[20]</sup>采用一种基于变分自编码的手写数字图像生成方法来解决样本数据不均衡的问题,使用 MNIST 数据集中相对较少类别的图像训练网络并生成仿真图像,该方法相比其他方法在预测率、召回率和特异性等方面都表现更优。Cristovao 等<sup>[21]</sup>提出了一种基于变分自编码器生成非连续图像的中间图像的方法,使用非连续的面部图像和茶壶图像训练网络,并结合图像插值法来生成具有不同角度的中间图像,相比于标准图像和其他研究中的图像,该方法虽然精确地保持了脸部和茶壶中间图像的结构,但生成的中间图像质量不稳定。Hou 等<sup>[22]</sup>提出了一种基于变分自编码器的人脸图像和自然图像生成模型,该模型将图像生成网络、预训练网络和生成对抗网络相结合,提升了生成图像的质量;Hou 等使用 VGGNet 来提取真实图像和生成的仿真图像的特征,并将其输入到基于生成对抗网络(generative adversarial networks)的判别器中重构图像的损失,然后通过反向传播算法来优化模型的参数;该方法生成的图像保留了图像的更多细节,同时背景清晰。Karatsiolis 等<sup>[23]</sup>提出了基于条件去噪自编码器的图像生成模型,通过在图像生成网络中嵌入基于卷积神经网络的分类器,生成了高质量的图像样本;Karatsiolis 等分别使用 MNIST、SVHN 和 CelebA 三个数据集训练网络,并生成相应的仿真图像;与其他模型相比,该模型生成的图像与真实图像的相似度更高。Yun 等<sup>[24]</sup>提出了一种基于条件卷积变分自编码器(CCVAE)的金属表面不同缺陷类型图像的生成方法,用以解决样本不均衡的问题;他们通过在网络中加入标签,实现了不同类别图像的生成;结果表明,使用 CCVAE 生成的图像训练的分类模型的性能优于使用真实图像训练得到的分类模型。

Zheng 等<sup>[25]</sup>提出一种基于条件变分自编码的手写图像生成模型,并将图像标签引入到编码器中进行特定类别的数据生成;他们使用 MNIST 和 Fashion-MNIST 数据集进行图像生成实验,与其他方法相比,该方法生成的图像更具多样性,生成的模糊图像较少,但是仅适用于手写图像的生成。Yan 等<sup>[26]</sup>提出了一种基于变分自编码器的分层图像生成模型,先分别生成前景和背景,再将它们进行结合,实现图像的生成;他们使用人脸图像和鸟类自然图像对模型进行测试,测试结果表明,与真实图像相比,该模型生成的图像较为清晰,但网络结构较大且参数较多。

目前,国内外尚未有公开的文献将变分自编码器技术应用于视网膜 OCT 图像生成的研究中。本文提出了一种基于变分自编码器的视网膜 OCT 图像生成方法,其中:编码器部分基于卷积神经网络构建,输入的是真实的视网膜 OCT 图像,编码器将其映射为服从正态分布的均值和方差;解码器部分基于反卷积神经网络构建,其输入是随机采样的服从正态分布的两个数,输出为生成的仿真图像。然后,基于两个公开的视网膜 OCT 图像数据集,分别构建了 AMD、DME 和正常(Normal, NOR)三种类别的样本数据集,并分别训练生成网络,在此过程中调整网络的结构和参数,得到各自的图像生成模型。最后,使用视觉评价和实验验证两种方式验证生成的仿真图像的有效性。当生成的图像满足视觉评价标准之后,本课题组设计了两组对比实验,验证了生成的仿真图像的有效性,同时证明了有效的图像生成可以缓解深度学习任务中数据不足的问题。主观视觉观察和客观实验验证结果均表明,本文提出的基于变分自编码器的视网膜 OCT 图像生成方法可以有效生成 AMD、DME 和 NOR 三种视网膜 OCT 图像。

## 2 基于变分自编码器的视网膜 OCT 图像生成方法

### 2.1 变分自编码器的原理

变分自编码器是一种深度隐空间生成模型,在数据生成方面具有巨大的应用价值<sup>[18]</sup>。变分自编码器分为编码器和解码器两部分,编码器用于从输入图像中学习图像的特征分布并将其映射为服从正态分布的均值和方差,而解码器则学习如何将这两个参数重构为输入图像。在训练过程中,计算输入图像和生成图像之间的损失函数,采用梯度下降法

结合反向传播算法更新网络中的参数。训练完成后,随机采样服从正态分布的两个数,并将其输入解

码器生成新图像。变分自编码器的结构图如图 1 所示。

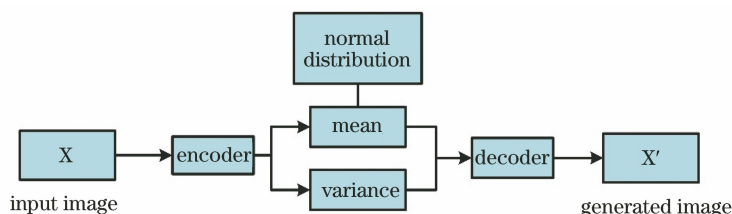


图 1 变分自编码器结构图

Fig. 1 Structure diagram of variational auto-encoder

### 2.2 基于变分自编码器的图像生成

首先构建基于变分自编码器的图像生成网络及 AMD、DME 和 NOR 三个视网膜 OCT 图像样本数据集,然后分别使用这三个数据集训练图像生成网络,并将解码器部分作为图像生成模型,分别命名为 VAE-AMD、VAE-DME 和 VAE-NOR。随机采样

一组服从正态分布的两个数,并将其输入模型中生成仿真图像,然后采用视觉评价和实验验证两种方式评价生成的仿真图像的有效性。在实验过程中,当生成的仿真图像达到视觉评价标准后,设计两组对比实验来验证生成的图像的有效性。基于变分自编码器的仿真图像生成流程图如图 2 所示。

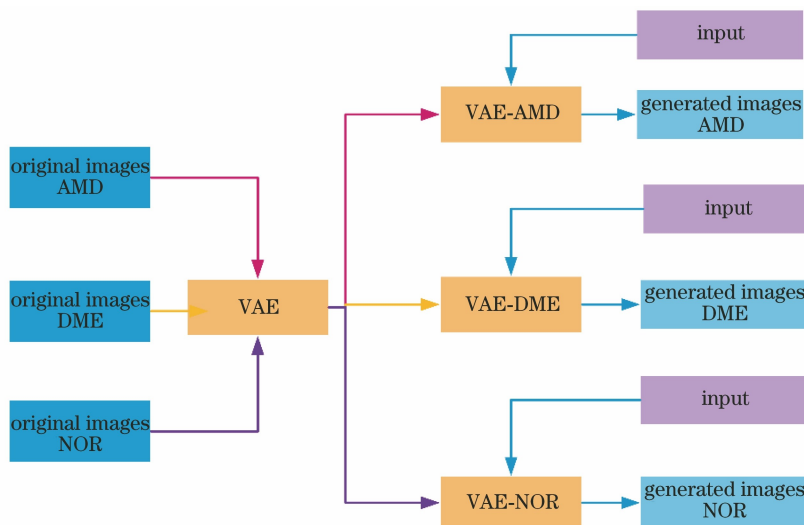


图 2 基于变分自编码器的仿真图像生成方法流程图

Fig. 2 Flow chart of simulation image generation method based on variational auto-encoder

## 3 实验

### 3.1 样本数据集

变分自编码器的工作原理是学习图像的特征并进行编码和解码重构,因此用于训练网络的样本仅具有某一类特征时效果最好。但公开的 2014\_BOE\_Srinivasan 数据集<sup>[27]</sup>中的 AMD 患者和 DME 患者的视网膜 OCT 图像并不完全都是病变图像;因此,本课题组在眼科医生的帮助下从该数据集中的 15 个 AMD 患者视网膜 OCT 图像中选取了 157 张图像,从 15 个 DME 患者视网膜 OCT 图像中选取了 261 张图像,从 15 个正常受试者视网膜 OCT 图像中选取了 1350 张图像。同时,为了增加数据的多样

性,从另一个公开数据集 OCT2017<sup>[28]</sup>中随机选取了 623 张 AMD 图像和 309 张 DME 图像。这些图像共同组成了本次实验的样本数据集,用来训练基于变分自编码器的图像生成网络。样本数据集中的图像分布如表 1 所示(为了方便表示,使用 SD\_AMD 表示 AMD 患者视网膜 OCT 图像的样本数据集,SD\_DME 表示 DME 患者视网膜 OCT 图像的样本数据集,SD\_NOR 表示正常受试者视网膜 OCT 图像的样本数据集)。

如表 1 所示,样本数据集 SD\_AMD 中共 780 张图像,SD\_DME 中共 570 张图像,SD\_NOR 中共 1350 张图像。在进行网络训练时,为了减少冗余特征、节省计算机资源和缩短运算时间,将图像统一裁

剪为 400 pixel×200 pixel。

表 1 样本数据集中的图像分布

Table 1 Image distribution in our data sets

Data set	SD_AMD	SD_DME	SD_NOR
2014_BOE_Srinivasan	157	261	1350
OCT2017	623	309	0
Sum	780	570	1350

### 3.2 最优网络结构和实验参数设置

由于卷积神经网络(CNN)在图像特征提取方面表现较好<sup>[29-30]</sup>,因此本文基于 CNN 搭建编码器,基于反卷积神经网络搭建解码器。对于本文构建的样本数据集,最优网络结构设计如图 3 所示。其中: Conv 层为卷积层;flatten 层为“压平层”,用来将多维输入一维化;dense 层为全连接层;Deconv 为反卷积层;relu,selu 和 tanh 为激活函数。编码器部分由一个输入层、三个 Conv 层、一个 flatten 层和一个 dense 层构成,卷积层的激活函数均为 relu,卷积核的大小为 3×3,个数分别为 16、32 和 32,dense 层的激活函数为 selu。编码器的输入为 400 pixel×200 pixel 大小的图像,输出为服从正态分布的均值和方差这两个参数。解码器由一个输入层、一个 dense 层、一个 reshape 层、一个 Deconv 层、一个 Conv 层和一个输出层构成,dense 层和反卷积层的激活函数均为 selu,卷积层的激活函数为 tanh,反卷积层和卷积层的卷积核大小均设置为 3×3,个数分别为 16 和 32。解码器的输入为服从正态分布的两个随机数,输出为 400 pixel×200 pixel 的仿真图像。

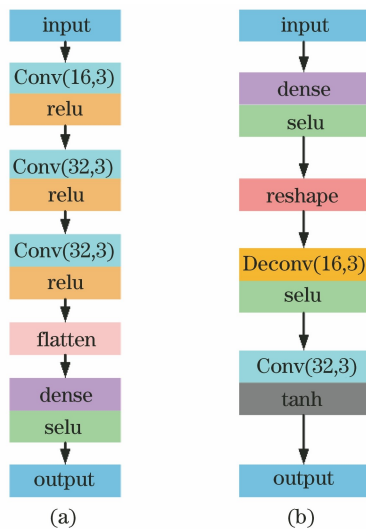


图 3 编码器与解码器的结构。(a)编码器;(b)解码器

Fig. 3 Structures of encoder and decoder.

(a) Encoder; (b) decoder

图像生成模型基于 TensorFlow-gpu V1.7.0 和 Keras V2.1.6,搭配 CUDA V9.0.176 和 cuDNN V7.3,使用 NVIDIA GTX 1060 6 GB GPU 进行加速计算。网络训练采用自适应优化器,损失函数为 binary\_crossentropy。

### 3.3 网络的训练

层次性数据格式(HDF5)文件是一种用于存储科学数据的文件格式,特别适合进行大量数据的存储和操作。因此在训练生成网络之前,将图像及其标签同时存储在 HDF5 文件中,并通过编码器的 input 层将其输入网络。在实际训练过程中,分别读入三个样本数据集 SD\_AMD、SD\_DME 和 SD\_NOR 对应的 HDF5 文件。Batch Size 设置为 80,即每 80 张图像作为一组输入网络进行训练,所有图像均输入网络训练后完成第一轮次的训练。Epoch 设置为 3000,即网络训练 3000 轮次后保存生成模型。

训练时,将图像按照 9:1 的比例划分为训练集和验证集,分别用于网络的训练和验证,并根据验证集损失函数的变化情况和人工观察仿真图像的效果来调整网络的结构和参数,直到生成质量较好的图像;此时保存解码器部分作为图像生成模型,用于后续的图像生成。损失函数的值越低,说明仿真图像与真实图像的差异越小。

### 3.4 仿真图像生成与质量评价

目前大多数研究使用视觉观察的方式来验证生成的仿真图像的质量,该方法虽然简单便捷,但是取决于观察者的主观感受,仅使用该方法说服力不强。因此,本文采用主观视觉观察和客观实验验证两种方式来验证所生成的仿真图像的有效性。

#### 3.4.1 主观视觉评价

这里通过人工视觉观察所生成的仿真图像中是否包含了原始图像的特征信息来评价生成图像的效果。图 4(a)、(b)为 AMD 患者视网膜的 OCT 图像,图 4(c)~(e)为生成的仿真图像。该类患者视网膜病变表现是视网膜色素上皮层出现大小不一的隆起,如图 4(a)、(b)所示。从图 4(c)~(e)可以看出,在生成的仿真图像中,视网膜色素上皮层也出现了大小不一的隆起,具有真实图像的特征。图 4(f)、(g)为 DME 患者视网膜的 OCT 图像,图 4(h)~(j)为生成的仿真图像。该类患者视网膜病变的表现是视网膜色素上皮层出现低反射空洞区域,仿真图像也出现了与真实图像相似的空洞区域。图 4(k)、(l)为正常受试者视网膜的 OCT 图像,图 4(m)~(o)为生成的仿真图像。该类 OCT 图像的特

点是视网膜各层光滑,而且生成的仿真图像中视网膜各层也光滑。

综上所述,使用本文中的图像生成方法生成的

三种类别的仿真图像具有真实图像的特征,同时仿真图像中视网膜色素上皮层隆起和水肿的位置和大小是随机的,增加了数据的多样性。

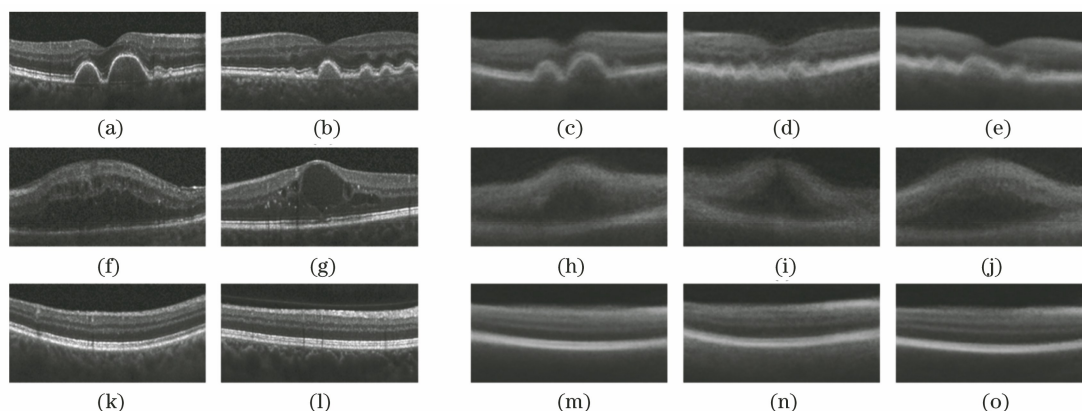


图 4 生成的仿真图像。(a)(b) AMD 患者视网膜的 OCT 图像;(c)~(e)生成的仿真图像;(f)(g)为 DME 患者视网膜的 OCT 图像;(h)~(j)生成的仿真图像;(k)(l)为正常受试者视网膜的 OCT 图像;(m)~(o)生成的仿真图像

Fig. 4 Generated images. (a)(b) Real OCT images of AMD patients' retina; (c)~(e) generated images; (f)(g) real OCT images of DME patients' retina; (h)~(j) generated images; (k)(l) real OCT images of normal subjects; (m)~(o) generated images

### 3.4.2 客观实验验证

为了客观证明生成的仿真图像的有效性,本课题组设计了两组实验,分别为实验一和实验二。实验一是从生成的 AMD、DME 和 NOR 三种仿真图像中各选取 1000 张图像训练一个此前本课题组

建的基于 CNN 的三分类网络<sup>[31]</sup>,得到模型一。训练过程中模型的准确率和损失值的变化如图 5(a)、(b)所示。从图中可以看出,随着 Epoch 增加,准确度不断上升,损失值不断下降,当这两个值趋于稳定时模型收敛。

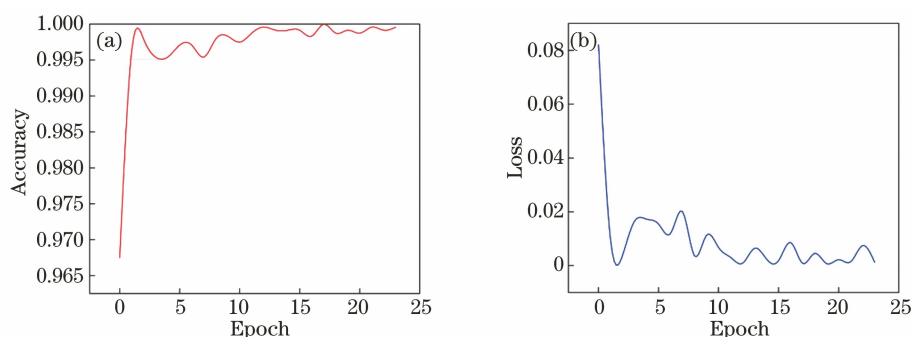


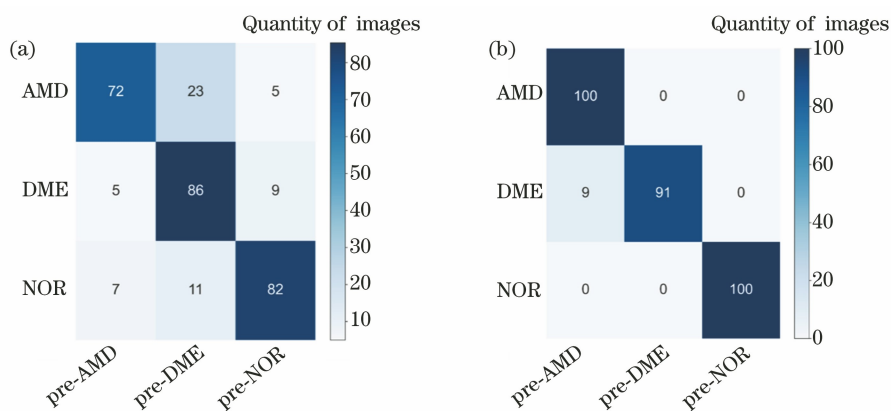
图 5 准确度和损失值的变化图。(a)准确度;(b)损失值

Fig. 5 Variations of accuracy and loss. (a) Accuracy; (b) loss

当模型收敛后,从每种类别的真实图像中随机选出 100 张图像组成测试集来测试分类模型。通过输出混淆矩阵来可视化模型对真实图像的分类结果,如图 6(a)图所示。由混淆矩阵计算得出分类模型的准确率为 80%,灵敏度为 79%,特异度为 82%。

实验二是使用包含三种类别的 2700 张真实图像训练基于 CNN 的三分类网络<sup>[31]</sup>,得到模型二。然后从生成的仿真图像中随机取出 100 张图像对模

型进行测试,并输出混淆矩阵,如图 6(b)所示。由混淆矩阵可计算得出分类模型的准确率为 97%,灵敏度为 95.5%,特异度为 100%。实验一和实验二的客观实验验证结果表明,使用仿真图像训练的分类模型可以有效识别真实图像,使用真实图像训练的分类模型也可以有效地对仿真图像进行分类,证明本文提出的基于变分自编码器的视网膜 OCT 图像生成方法是有效的。



Notes: the horizontal ordinate represents the predicted results, and the ordinate represents the true results.

图 6 混淆矩阵。(a)实验一;(b)实验二

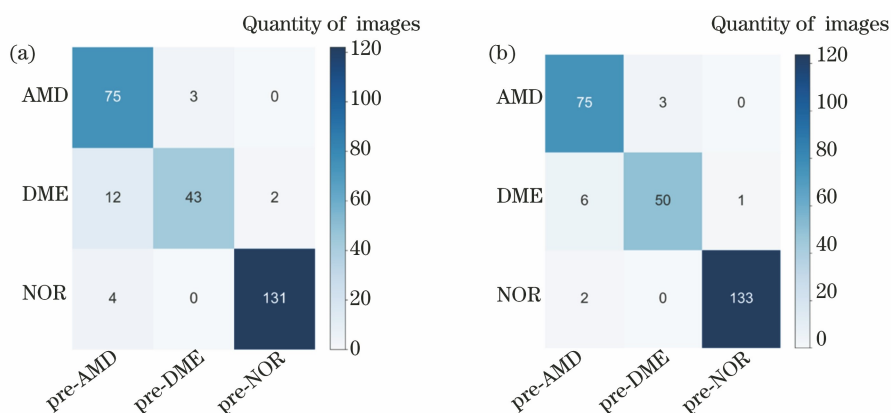
Fig. 6 Confusion matrices. (a) Experiment 1; (b) experiment 2

此外,本课题组设计了实验三和实验四来证明对图像进行有效扩充可以提升分类模型的性能。实验三仅使用有限的真实图像训练分类网络,得到分类模型后对真实图像进行识别。实验四用真实图像与生成的仿真图像组成的混合图像训练分类网络,然后对真实图像进行识别。首先从三类真实图像中各取出 10%(共 270 张图像)作为实验三和实验四的测试集,不参与网络训练。实验三用取出测试集后剩余的 2430 张真实图像训练一个基于 CNN 的三分类网络<sup>[31]</sup>,得到模型三。实验四用 2430 张真实图像与生成的 3000 张仿真图像组成的混合图像来训练分类网络<sup>[31]</sup>,得

到模型四。最后分别使用测试集来测试模型三和模型四,并分别输出混淆矩阵查看两个模型的分

类效果。图 7(a)为模型三的混淆矩阵,图 7(b)为模型四的混淆矩阵。通过混淆矩阵计算可得出模型三的准确率为 92.22%,灵敏度为 87.41%,特异度为 97.04%,模型四的准确率为 95.56%,灵敏度为 92.59%,特异度为 98.52%。

从实验三和实验四的结果可以看出:对于相同的测试集,模型四的分类性能优于模型三。这表明,有效的数据生成可以提升分类模型的性能,缓解深度学习中数据不足的问题,同时也证明了本文图像生成方法是有效的。



Notes: the horizontal ordinate represents the predicted results, and the ordinate represents the true results.

图 7 混淆矩阵。(a)实验三;(b)实验四

Fig. 7 Confusion matrices. (a) Experiment 3; (b) experiment 4

## 4 结 论

本课题组将变分自编码器技术应用于视网膜 OCT 图像的生成中,用训练生成的图像生成模型生成了三种类别的视网膜 OCT 图像。从视觉上看,所生

成的仿真图像均具备了真实图像的特征信息,并且特征的大小和位置更具随机性,增加了图像的多样性。设计了两组对比实验,验证了所提图像生成方法是有效的,同时证明了有效的图像生成可以缓解深度学习任务中数据不足的问题。所提图像生成方

法可以有效生成三种类别的视网膜 OCT 图像,有助于提高眼科疾病辅助诊断系统的准确度和泛化能力,也可以用于生成样本数据集中少数类别的图像,解决了样本分布不均衡的问题。本文所提方法的不足之处是只能用于单一类别数据的生成,不能使用一个模型生成多种类别的图像,后续研究拟尝试在训练过程中引入标签信息来实现多类别数据的生成。

### 参 考 文 献

- [1] Huang D, Swanson E A, Lin C P, et al. Optical coherence tomography [J]. *Science*, 1991, 254 (5035): 1178-1181.
- [2] Farsiu S, Chiu S J, O'Connell R V, et al. Quantitative classification of eyes with and without intermediate age-related macular degeneration using optical coherence tomography [J]. *Ophthalmology*, 2014, 121(1): 162-172.
- [3] Markomichelakis N N, Halkiadakis I, Pantelia E, et al. Patterns of macular edema in patients with uveitis: qualitative and quantitative assessment using optical coherence tomography [J]. *Ophthalmology*, 2004, 111(5): 946-953.
- [4] Medeiros F A, Zangwill L M, Bowd C, et al. Evaluation of retinal nerve fiber layer, optic nerve head, and macular thickness measurements for glaucoma detection using optical coherence tomography [J]. *American Journal of Ophthalmology*, 2005, 139(1): 44-55.
- [5] Coleman H R, Chan C C, Ferris F L, et al. Age-related macular degeneration [J]. *The Lancet*, 2008, 372(9652): 1835-1845.
- [6] Jager R D, Mieler W F, Miller J W. Age-related macular degeneration [J]. *New England Journal of Medicine*, 2008, 358(24): 2606-2617.
- [7] Ciulla T A, Amador A G, Zinman B. Diabetic retinopathy and diabetic macular edema: pathophysiology, screening, and novel therapies [J]. *Diabetes Care*, 2003, 26(9): 2653-2664.
- [8] Lang G E. Diabetic macular edema [J]. *Ophthalmologica*, 2012, 227(s1): 21-29.
- [9] Bressler N M. Age-related macular degeneration is the leading cause of blindness [J]. *JAMA*, 2004, 291 (15): 1900-1901.
- [10] Fang L Y, Cunefare D, Wang C, et al. Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative AMD patients using deep learning and graph search [J]. *Biomedical Optics Express*, 2017, 8(5): 2732-2744.
- [11] Garvin M K, Abramoff M D, Kardon R, et al. Intraretinal layer segmentation of macular optical coherence tomography images using optimal 3-D graph search [J]. *IEEE Transactions on Medical Imaging*, 2008, 27(10): 1495-1505.
- [12] Shah A, Zhou L X, Abrámoff M D, et al. Multiple surface segmentation using convolution neural nets: application to retinal layer segmentation in OCT images [J]. *Biomedical Optics Express*, 2018, 9(9): 4509-4526.
- [13] Sun Y S, Liu Y H, Zhang D F, et al. Diagnosis method of diabetic retinopathy based on deep learning [J]. *Laser & Optoelectronics Progress*, 2020, 57 (24): 241701.  
孙雨琛, 刘宇红, 张达峰, 等. 基于深度学习的糖尿病视网膜病变诊断方法 [J]. *激光与光电子学进展*, 2020, 57(24): 241701.
- [14] Zhang T F, Zhong S C, Lian C M, et al. Deep learning feature fusion-based retina image classification [J]. *Laser & Optoelectronics Progress*, 2020, 57(24): 241025.  
张添福, 钟舜聪, 连超铭, 等. 基于深度学习特征联合的视网膜图像分类 [J]. *激光与光电子学进展*, 2020, 57(24): 241025.
- [15] Lian C M, Zhong S Z, Zhang T F, et al. Transfer learning-based classification of optical coherence tomography retinal images [J]. *Laser & Optoelectronics Progress*, 2021, 58(1): 117002.  
连超铭, 钟舜聪, 张添福, 等. 光学相干断层扫描视网膜图像的迁移学习分类 [J]. *激光与光电子学进展*, 2021, 58(1): 117002.
- [16] Kingma D P, Welling M. Auto-encoding variational bayes [EB/OL]. (2013-12-20) [2021-01-02]. <https://arxiv.org/abs/1312.6114v4>.
- [17] Jiang Z T, He Y T. Infrared and visible image fusion method based on convolutional auto-encoder and residual block [J]. *Acta Optica Sinica*, 2019, 39 (10): 1015001.  
江泽涛, 何玉婷. 基于卷积自编码器和残差块的红外与可见光图像融合方法 [J]. *光学学报*, 2019, 39 (10): 1015001.
- [18] Zhai Z L, Liang Z M, Zhou W, et al. Research overview of variational auto-encoders models [J]. *Computer Engineering and Applications*, 2019, 55 (3): 1-9.  
翟正利, 梁振明, 周炜, 等. 变分自编码器模型综述 [J]. *计算机工程与应用*, 2019, 55(3): 1-9.
- [19] Lu D, Wei X, Cao X, et al. Fast reconstruction method for fluorescence molecular tomography based on autoencoder [J]. *Acta Optica Sinica*, 2019, 39 (6): 0617001.

- 卢笛, 卫潇, 曹欣, 等. 基于自编码器的荧光分子断层成像快速重建 [J]. 光学学报, 2019, 39(6): 0617001.
- [20] Wan Z Q, Zhang Y Z, He H B. Variational autoencoder based synthetic data generation for imbalanced learning [C] // 2017 IEEE Symposium Series on Computational Intelligence (SSCI), November 27-December 1, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 1-7.
- [21] Cristovao P, Nakada H, Tanimura Y, et al. Generating in-between images through learned latent space representation using variational autoencoders [J]. IEEE Access, 2020, 8: 149456-149467.
- [22] Hou X X, Sun K, Shen L L, et al. Improving variational autoencoder with deep feature consistent and generative adversarial training [J]. Neurocomputing, 2019, 341: 183-194.
- [23] Karatsiolis S, Schizas C N. Conditional generative denoising autoencoder [J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 31(10): 4117-4129.
- [24] Yun J P, Shin W C, Koo G, et al. Automated defect inspection system for metal surfaces based on deep learning and data augmentation [J]. Journal of Manufacturing Systems, 2020, 55: 317-324.
- [25] Zheng K, Cheng Y F, Kang X J, et al. Conditional introspective variational autoencoder for image synthesis [J]. IEEE Access, 2020, 8: 153905-153913.
- [26] Yan X C, Yang J M, Sohn K, et al. Attribute2Image: conditional image generation from visual attributes[M]//Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Cham: Springer, 2016, 9908: 776-791.
- [27] Srinivasan P P, Kim L A, Mettu P S, et al. Fully automated detection of diabetic macular edema and dry age-related macular degeneration from optical coherence tomography images [J]. Biomedical Optics Express, 2014, 5(10): 3568-3577.
- [28] Kermany D S, Goldbaum M, Cai W J, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning [J]. Cell, 2018, 172(5): 1122-1131.
- [29] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [J]. Communications of the ACM, 2017, 60(6): 84-90.
- [30] Litjens G, Kooi T, Bejnordi B E, et al. A survey on deep learning in medical image analysis [J]. Medical Image Analysis, 2017, 42: 60-88.
- [31] Zhao M M, Zhu S Y, Shan H, et al. Automatic classification of retinal OCT images based on convolutional neural network [J]. Proceedings of SPIE, 2020, 11511: 115110Z.