

FC层的目的是通过计算类的分数来执行较高级的推理。在该层每个神经元连接所有神经元在前面层产生全球语义信息。

CNN的最后一层是输出层 (O)，这里的SOFTMAX运营商通常用于CLASSI网络阳离子任务。的最佳参数 (θ ，两种常见的符号 w 和 b) 用于特定任务可以通过最小化损失函数 (来确定 L)

神经网络连接定义的任务。在数学上，对于 \tilde{n} 输入 - 输出关系 ($\{ \text{一世} N, \emptyset N \}; \tilde{n} \in [1, N]$ 和相应的标签 G_n 的损失

被推断为：

$$L = \frac{1}{N} \sum_{n=1}^N \text{LN}(\theta; G_n, \emptyset N) \quad (2)$$

哪里 \tilde{n} 表示训练图像的数量， $\text{一世} N, \emptyset N$ 和 G_n 相当于 \tilde{n} 训练图像。在这里，在训练CNN的一个重大的挑战，从训练样本的数量有限，相比于需要为手头的任务进行优化，可学参数的数量就出现了。最近的研究已经开发的一些关键技术，以更好地培养和优化深等车型数据增强，重量初始化，随机梯度下降 (SGD)，批标准化，快捷的连接，正规化。为了更多的了解与在CNN的进步，读者建议指纸 Gu等。(2018)。

越来越多地使用CNN的作为的许多视觉上的任务，准备用于不同的目的 (如分割，CLASSI网络阳离子或本地化) 和可用的数据，骨干取得了架构解决问题搜索一个主要渠道。

在这种挑战，主要表现为病情严重程度分级的问题，参与者可以直接利用现有的有线电视新闻网的变体或合奏他们将输入图像划分到中提到的类别之一 表4。一些配置简单和CNN的变体是在文献中可用；一些最流行的是AlexNet (Krizhevsky等人, 2012)，VGG (西蒙尼扬和Zisserman 2014)，GoogLeNet (Szegedy等人, 2015年) 和RESNET (他等人。年, 2016年) 由于其对物体识别任务不同的基准性能优越。这些架构的演进典型的趋势是，网络已经变得更深，

例如，RESNET比AlexNet，VGGNet和分别GoogLeNet更深约19，图8和7倍。而深度的增加提高了特征表示和预测性能，这也增加了复杂性，使之成为二FPI崇拜优化和甚至变得易于过度拟合。此外，越来越多的层 (即，网络深度) 导致消失梯度问题由于大量的乘法运算的结果。因此，很多球队选择了DenseNet (Iandola等人, 2014)，其在一个前馈方式的各层连接到每一个其他层，减少的训练参数和减轻消失梯度问题的数量。DenseNet展品 (+1) 在/2个连接

代替的层网络中，只有，如在网络男人 - 上述tioned。这使得整个网络，导致更紧凑的内部表示，反过来，提高其预测精度功能重用。另一种选择方法，深层聚合 (DLA) 的结构 (Yu等人, 2017年)，扩展了“浅”跳过DenseNet连接将更多的深度和所述特征共享。DLA采用两种结构 - 迭代深聚合 (IDA) 和深层次聚集 (HDA)，其迭代和分层保险丝功能层次 (即语义空间)，使网络工作，更好的精度和更少的参数。最近的全卷积网络 (FCN) (龙等人, 2015年) 调整和扩展深CLASSI网络阳离子架构 (VGG和GoogLeNet) 为完全卷积网络，并通过网络NE-调整转让自己所学的交涉分割任务。它德科幻定义了一个跳过的架构，从深，外观IN-粗糙层联合机语义信息

从浅，科幻NE层形成，以产生准确和详细的分割。

对于病变划分任务，大多数参赛团队利用掌中宽带架构 (Ronneberger等人, 2015年)。在U形网络系统结构的主要思想是通过添加连续的层，在那里执行，而不是池操作上采样 (通过去卷积)，以补充通过对称膨胀路径通常订购网络。上采样部分由大量特征的通道，其允许网络上传播的上下文信息到更高分辨率的层。从承包路径高分辨率功能合并上取样的输出和输入到软最大CLASSI网络呃是基于像素CLASSI网络阳离子。该网络可与极少数的训练图像和重叠瓦战略的手段使高分辨率图像的无缝分割。其它类似的架构SegNet (Badrinarayanan等人, 2015) 是由一队选择; 它由编码器和解码器的网络，其中，所述编码器网络是由一个使用SoftMax层代替拓扑等同于VGG16并且其中FC层CONV层。而，解码器网络包括解码器的分层结构中，一个对应于每个编码器。解码器使用MAX-池指数上采样的编码器输入产生稀疏的特征图。后来，它与卷积的可训练滤波器银行稀疏特征映射到致密他们。最后，解码器输出被馈送到一个软最大CLASSI音响ER为分割图的产生。一队选择面膜RCNN (他等人, 2017年)，主要基于的区域建议网络 (RPN的技术)，其股与所述检测网络的整个图像，从而使区域的建议定位和进一步段正常和异常结构在视网膜的卷积功能。RPN是一个完全卷积网络有助于在每个位置同时预测对象边界和“对象性”的分数。

以下部分呈现由相对于参赛队伍到三个子挑战而设计的解决方案。表6汇总数据增强，规范化，按照各队进行预处理任务。






















5.1。次挑战 - 1: 病变划分

对于给定的图像，这个任务试图得到一个像素是病变 (无论是MA，HE，EX或SE) 的概率。虽然不同的眼底病变具有不同的局部特征，例如，MA，HE，EX，SE具有不同的形状，颜色和分布特征，这些病变共享类似全局特征。因此，大多数参赛团队建立了一个总体框架，但适用于病变部位不同，总结如下分割：

5.1.1。VRT (Jaemin Son等人) 中

儿子等人。MODI音响编U形网 (Ronneberger等人, 2015年) 在这样的方式上采样层具有相同数量的功能映射与级联层。它是基于，具有初始层和上取样层到分割同样重要的动机。另外，它们调节maxpooling使最大损伤的半径跨越在粗糙层中的像素的数量。在EX和HE的情况下，MAX-池做六次，而对于SE和MA是做四次，两次。此外，为应对MA的，他们用逆像素舒荷兰国际集团ff i第一个1280转换 $\times 1280 \times 3$ 个像素的图像 $640 \times 640 \times 12$ ，用于网络输入和像素舒而第ING (施等人。年, 2016年) 转换成 $640 \times 640 \times 4$ 分割地图分成 $1280 \times 1280 \times 1$ 个像素。后来，归一化的眼底图像和参考按地真理的对进料至网络，以产生在范围[0, 1]的分割结果。他们用二进制加权交叉熵 (墨菲, 2012) 的损失

表6
数据增强，归一化和概述预处理在竞争解决方案。其中，RF，RR，RS，RT，RC代表随机FLIP，旋转，缩放，平移分别和作物。

任务	队名	数据增强						数据标准化	数据预处理
		RF	RR	RS	RT	RC	其他		
次挑战 - 1	 VRT						剪断		FOV裁剪，通过255，那么平均减法除法
	 科大讯飞						×		病变补丁提取
	 PATech			×		×	颜色→ 		RGB到LUV，对比度调节
	 SDNU			×	×	×	×	-	-
	 越早			×	×		×		平均减法，病变补丁提取
	 LzyUNCC		×	×	×		随机和光度量 	-	FOV裁剪，图像增强
	 SAIHST			×	×	×	×		CLAHE，高斯平滑
次挑战 - 2	 LzyUNCC		×	×	×		颜色→  , 随机和光度量 	-	FOV裁剪，图像增强
	 VRT	×	×	×	×	×	×		平均减
	 长毛象					×	颜色	×	形态学打开和关闭
	 AVASAVA		×	×	×		×		强度缩放
	 HarangIM1	×	×	×	×	×	×		FOV种植
	 HarangIM2	×	×	×	×	×	×		-
次挑战 - 3	 DeepDR	×	×	×	×		OD，黄斑中心凹区		FOV种植，平均减
	 VRT						剪切和裁剪OD		FOV裁剪，对比度调节
	 ZJU-BII-SGEX	×	×	×	×	×	×		FOV种植
	 SDNU		×		×	×	×	-	-
	 IITkgpKLIV			×	×	×	×		-
	 CBER	×	×	×	×	×	×	-	-

→参考：Krizhevsky等。(2012)
Ⓢ参考：霍华德(2013年)

函数给出

$$L = 1 - \frac{\sum_{N=1}^{\bar{n}} [-\alpha G_n \log(1 - G_n) - (1 - G_n) \log(1 - \varnothing_n)]}{\bar{n}} \quad (3)$$

哪里 \bar{n} 表示在间歇式的对的数目， G_n 和 \varnothing_n 表示为真分割和预测分割 \bar{n} 图片。的价值 α 测定如下：

$$\alpha = \frac{0}{Y F_{-1}^{\#}} \quad (4)$$

表7
γ 价值观式。(4)。

EXS	社企	的HE均线
64	512	8 32

哪里 Z_{n0} 和 F_n 表示背景和前景的数量
以像素为单位 \bar{n} 图片。因为背景在淹没病变分割任务前景中，这种损失函数被设计以惩罚假阴性到升压灵敏度，在检测病变的重要因素。也， γ 被留下作为超参数和选择出来{0.25，0.5，1，2，4，8，16，32，64，256，512}，以产生上的验证集的最高AUPR。选择的网络连接最终 γ 针对不同的病变，汇总在表7。

他们在用亚当优化300个时代的培训网络(金玛和Ba 2014)，用超参数 $\beta_1=0.5$ ， $\beta_2=0.999$ 和2学习率 $E=4$ 直到250个历元和2 $E=5$ 直到最后。所有执行是由Keras完成2.0.8

使用与8 TITAN X (帕斯卡) 的服务器张量流后端1.4.0。源代码可在 https://bitbucket.org/woalsdnd/isbi_2018_fundus_challenge。

5.1.2. 科大讯飞，MIG (峰岩Wang等人。) 王等人。提出了视网膜病变分割带U-Net的一种新颖的级联CNN为基础的方法(Ronneberger等人，2015年)作为基本模型。它包括三个阶段，在第一个阶段是粗分割模型得到初始分割的面具，那么第二个阶段是设计用于假阳性减少，最后一个级联CLASSI网络呢，一个科幻NE分段模型被用来重新连接的NE结果从以前的阶段。第一阶段模型是用大小为256的补丁受训 \times 256个像素集中在一个特定的病灶之间MA，HE或EX和320 \times 320个像素为SE，导致粗分割结果。前一级的结果是这样一个事实，由于粗非聚焦区域

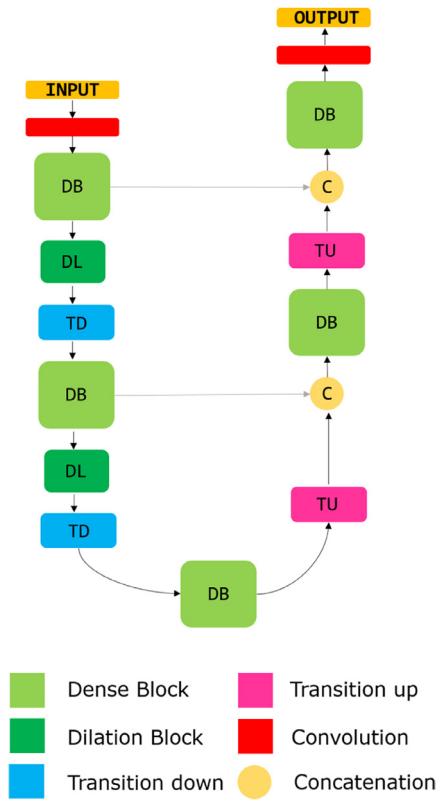


图4。建议架构病划分。

(非靶病灶)在学习过程中导致高的假阳性计数未使用。在第二阶段中,不同于使用中的损伤的第一个分割模型为中心从输入数据集池样品,使用概率来自前级的映射中提取候选区域。在此,供给到型号为SE输入大小为 $320 \times 320 \times 3$ 个像素,对于HE和EX是 $256 \times 256 \times 3$ 个像素,而对于MA是MODI音响ED $80 \times 80 \times 3$ 个像素考虑到其小的外观。在该步骤中,一个候选区域被视为阳性样品,如果它与地面实况相交-过联合比给定的阈值(即0.5)更大。通过这种方式,最微不足道的非重点区域进行有效的拒绝。但是,这是在测试误报的一小部分仍然存在identi网络版,所以额外的模型引入到重新连接的NE分割结果。在最后阶段,从第二级幸存候选区域被用作导致更准确的分割结果输入补丁。对于第一个和第三个阶段,他们用二进制交叉熵或骰子损耗函数(多模型训练),反之,对于第二阶段,它们仅使用二进制交叉熵作为损失函数。在第一个,第二个和第三个阶段模型训练100,300个100分别时期具有0.9的势头。其中,对于网络连接的初始学习速率RST和第三级被设置在0.1和是由10次,每次30个历元减少,并且用于所述第二阶段,设定为0.001由每80个信号出现时间的10倍降低。MXNET平台被用于训练模型。

5.1.3. PATech (刘力红等人。)

丽红等。开发了一种新颖的基于补丁CNN模型(如图4),其中他们创新组合DenseNets (Iandola等人,2014)中并用UNET扩张块(Ronneberger等人,2015),以获得更多的背景信息和多尺度特征。

该模型由下采样路径的与4向下转变(TD),4块扩张(DL)和上采样路径

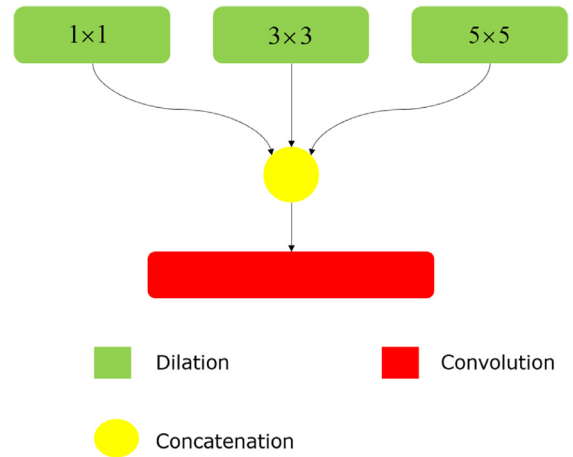


图5。架构扩张块。

用4向上转变(TU)。为了捕捉多尺度特征,DL(见

图5)被用于以1,3和5被连接用于卷积扩张速率。致密块(DB)是由四层构成。扩张背后的卷积的新组合的想法是更好地处理出现在不同的尺度,病灶小地方扩张速度密切注意微小病变的特点,较大的扩张速度专注于大病灶。在另一方面,使用DB的启用更深的电子FFI client网络。

最初,它们提取的区域内FOV从图像,然后归一化他们消除局部对比度的差异和不均匀的照明。后来,他们用小块 256×256 个像素在64(128为MA)步骤来生成训练样本(仅修补与病变地面实况重叠的),随后数据扩张馈送到模型之前。为了处理数据的高度不平衡蔓延,他们设计的是骰子的功能组合的损失函数(Sudre等人,2017)和二维交叉熵如下:

$$L = -\text{平均}(W_{10} * G * \text{商标}) + w^{A_{11}} * (1 - G) * \text{日志}(1 - O) + w^{A_2} * \text{骰子}(G) \quad (5)$$

哪里 $w^{A_{10}}$ 和 $w^{A_{11}}$ 是用来保持正和负像素之间的平衡的因素,并 w^{A_2} 是用来控制骰子和交叉熵损失之间显著性的因素。值 $w^{A_{10}}$, $w^{A_{11}}$ 和 w^{A_2} 被凭经验设定为0.7,0.3和

0.4分别。该机型采用默认参数适当优化的训练, $\beta_1=0.9$ 和 $\beta_2=0.999$ 。初始学习速率设置为 2×10^{-4} ,然后在20每20个历元划分。这个模型是使用PyTorch1.12和特斯拉M60平台,用于CentOS的7.2操作系统培训来实现。

5.1.4. 越早(云芝Wang等人。)

王等人。采用了U型网(Ronneberger等人,2015)architecture解决视网膜病变划分问题。该网络需要一个 380×380 个像素的眼底图像补丁作为输入和预测内196视网膜病变的二进制掩模 $\times 196$ 个像素的输入跳线的中央区域中。通过减去每个颜色通道的地方意味着他们预处理眼底图像和进行随机flipping用于数据扩张。批标准化被利用来提高训练efficiency并通过“有效”补白的所有卷积运算。对于训练,他们随后为每种类型的病变(即MA,HE,EX和SE)的一个三阶段过程。对于第一个阶段,它们根据给定的地面实况掩模提取从训练集中正图像块,和随机抽取从与眼底图像负图像块

并且没有明显的视网膜病变。目标函数是用于MA, HE, EX和SE交叉熵损失函数的总和。亚当算法用于优化的参数。在第二阶段中, 它们连接使用提取的补丁的每个病变类型NE-调谐U形网。随后, 他们在训练集中施加优化U形网上的眼底图像和由提取的U形网产生的假阳性的补丁。他们进一步音响使用正图像块与假阳性斑(硬负贴剂)作为第三阶段一起NE-调谐U形网。在测试阶段, 它们提取使用滑动窗口重叠的图像块并馈送这些补丁到网络中以得到相应的概率图。最初的学习速度设为 $E-4$ 并使用的步骤的固定成本数用作停止标准。他们实现了基于TensorFlow库的Nvidia GeForce GTX GPU 1080Ti U型网络架构。

5.1.5. LzyUNCC (中宇Li等人。)

Li等人。通过嵌入DLA结构(开发了基于FCN的方法 Yu等人, 2017年)为EX的和SE的分割。由于病变位于分散, 不定期, DLA结构与FCN嵌入能够从本地和全局层面的语义和空间信息更好的聚集提供了承认其存在提振。他们使用来自IDRID和E-Ophtha数据库pixellevel地面实况注解视网膜图像。他们第一个采用数据预处理和增强了一系列的方法。随后, 考虑到EX的和SE的之间的关系, 他们第一个训练EX的分割的初始模型。他们选择了更小的模式, 即, DLA-34与二进制交叉熵作为损失函数训练分割网络。最后, 受过训练的深层模型是科幻NE-调整为SE的分割。虽然EX细分的模型训练, 折衷参数(罚分) 中的损失函数被分配到控制前景像素的权重, 并试图从1不同的惩罚值至16。最后, 这些分割结果进行融合以自适应地计算最佳性能。他们采用了原DLA城市景观分割实验设置和训练模型100个时代与批量大小4, 其中乘学习率为($1 - 10^{-4}$)

学习率设为0.01。

5.1.6. SAIHST (尹何Choi等人) 中

Choi等人。提出了基于掌中EX的分割模型(Ronneberger等人, 2015年), 其中编码器路径的CON V层被替换为DB的。然而, 他们的模型的解码器路径保持等同于一般的U型网。他们建立DB 12和3的生成因子 $\times 3$ CONV层, 批量归一化, 和RELU活化。最后一层产生用于通过乙状结肠激活函数EXS像素级预计地图。对于训练, 他们仅使用眼底图像的绿色通道和使用对比度受限自适应直方图均衡(CLAHE) 增强它。后来, 每个图像被填充到尺寸 4352×3072 个像素和裁剪成 512×512 个补丁。这些补丁被进一步增强, 并用于训练。该损失是由二进制交叉熵计算。该模型被训练为20个时期为10的小批量大小和它们使用的亚当优化为2的初始学习速率 $E-4$, β_1 0.9和

β_2 0.999。该模型在Keras 2.1.4与1.3.0 TensorFlow后端服务进行编程。

5.1.7. SDNU (小旦Sui等。)

Sui等。提出了基于掩码R-CNN结构分割病变从眼底图像的方法。他们从采用面膜R-CNN的实施 阿卜杜拉(2017年) 为解决这个问题。该方法能检测不同的物体, 而同时生成例如分割掩码。网络训练之前的数据增加处理和二进制交叉熵

用作损失函数。最初的学习速度设为

0.02与0.9的势头。他们选择RESNET-101作为主干。它们实现的算法在Keras与张量流作为后端和8分NVIDIA TITAN XP中的GPU处理。实验环境中的Ubuntu 16.06下建造的。

5.2. 次挑战 - 2 : 疾病分级

对于给定的图像, 这个任务试图得到一个解决疾病的产生严重性等级即DR (5类的问题) 和DME (3类的问题)。参与方案的概要如下 :

5.2.1. LzyUNCC (中宇Li等人。)

Li等人。开发通过嵌入DLA结构用于自动分级DR和DME的基于RESNET的方法。对于这项工作, 他们使用IDRID和Kaggle数据集。最初, 对于给定的训练图像, 他们进行数据预处理和数据增强。随后, 基于设计RESNET与DLA结构, 初始模型是使用来自Kaggle数据集35000个的视网膜图像训练。以后, 它们通过网络连接5倍交叉验证技术NE-调谐使用IDRID数据集的模型。最后, 网络连接已经输出合奏在一起作为音响NAL牌号为输入图像。需要注意的是网络的博士和DME分别分级的培训是非常重要的。通过SGD与64的小批量大小进行训练, 而从0.001学习率开始, 并随后通过10个每20个时期划分, 用于总共30个历元。Yu等人, 2017年)。

5.2.2. VRT (Jaemin Son等人) 中

儿子等人。使用的网络(儿子等人, 2018) 为DR分级。Kaggle数据集最初是用于预训练网络, 然后将模型是网络连接的NE-调谐使用IDRID数据集。倒数第二层是全球平均汇集(GAP) 中, 用FC层连接。整个输出是从被相对于真标签计算L2损失的单个值。SGD用的湿斯捷罗夫势头使用

0.9作为优化器。学习速率设置为 10^{-3} 。该模型被训练为100个时代。眼底图像是在[0, 1]的范围内归一化并且平均中减去信道, 明智的。对于DME的分级, 分段EXS(使用子挑战提出的分割网络 - 1), 局部中央凹和分段OD(使用在子挑战提出的分割网络 - 3) 用于使音响最终决定。有了这些信息, 分段OD的半长轴(R) 估计。另外, 眼底图像被划分为三个区域为黄斑区 :

$\|X - C\| \leq R$, 近黄斑区 : $R < \|X - C\| \leq 2R$ 和剩余区域 : $2R < \|X - C\|$ 。 哪里 X 表示图像中的点。此外, 一些功能, 如强度为分段EX, 高于阈值的像素的数量的总和 (在[0,255]刻度178), 为最小的像素的数量和最大的斑点, 平均像素对斑点的数目是提取用于每个指示OD是否被分段区域和二进制FL AG。现在, 在初始训练由于梯度升压众多特征中选择具有高重要性的特征 (例如, XGBoost) 很可能超过网络吨当与过度冗余特征提供。获月数据集添加到给定的数据和外面的图像的10%的被留下作为验证集。组Hyper-参数进行网格搜索方法搜索。其产生在验证组精度最高的超参数的组合为最低子重量 : 2, 子样本 : $A : 9.0$, $\alpha : 1.0$, 深度 : 6, 其他超参数被设置为默认值。所有的实现是由PyTorch v0.4.1使用带有8 TITAN X (帕) 的服务器来完成。源代码可在 https://bitbucket.org/woalsdnd/isbi_2018_fundus_challenge。