



可在目录列表 ScienceDirect

## 医学图像分析

期刊主页 : [www.elsevier.com/locate/media](http://www.elsevier.com/locate/media)

## IDRiD : 糖尿病视网膜病变 - 分割和分级挑战

人员 Prasanna Porwal <sup>a, b, 1, \*</sup>, Samiksha Pachade <sup>-#</sup>, Manesh Kokare <sup>-#</sup>, 吉里什德希穆克 <sup>c, 1</sup>,Jaemin 儿子 <sup>d</sup>, 雄妻 <sup>d</sup>, 刘利红 <sup>E</sup>, 剑宗王 <sup>E</sup>, 刘新会 <sup>E</sup>, 孙良高 <sup>E</sup>,天博武 <sup>E</sup>, 晶科 <sup>E</sup>, 王峰岩 <sup>F</sup>, 尹宝才 <sup>F</sup>, 云芝王 <sup>G</sup>, Gopichandh Danala <sup>G</sup>,林盛他 <sup>G</sup>, 尹浩彩 <sup>H</sup>, 永陈利 <sup>H</sup>, 桑赫荣 <sup>H</sup>, 中宇力 <sup>一世</sup>,李晓丹 <sup>J</sup>, 吴张俊彦 <sup>J</sup>, 李泉龙 <sup>K</sup>, 周婷 <sup>K</sup>, 亚诺什·托特 <sup>Ø</sup>, 艾格尼丝·巴兰 <sup>Ø</sup>,古里阿维纳什 <sup>p</sup>, 西 Saketh Chennamsetty <sup>p</sup>, 穆罕默德萨夫万 <sup>p</sup>, Varghese 表示亚历克斯 <sup>p</sup>,Xingzheng <sup>昌</sup> <sup>昌</sup>, 李成 <sup>R</sup>, d, 桑浩楚 <sup>小号</sup>, 李彭城 <sup>小号</sup>, 新吉 <sup>T</sup>, 张三元 <sup>昌</sup>,沉雅欣 <sup>ü</sup>, v, 凌带 <sup>ü</sup>, v, Oindrila 萨哈 <sup>X</sup>, Rachana Sathish 所在 <sup>X</sup>, 塔尼娅梅洛 <sup>y</sup>, 邓丽君阿劳霍 <sup>y, z</sup>,鲍拉日 Harangi <sup>Ø</sup>, 盛彬 <sup>ü</sup>, v, 若谷方 <sup>w^A</sup>, Debdoott 表 <sup>X</sup>, 安德拉斯·豪伊杜 <sup>Ø</sup>,郑渊洁郑 <sup>J</sup>, 安娜·玛丽亚·门多萨 <sup>y, z</sup>, 少汀张 <sup>一世</sup>, 奥雷里奥 Campilho <sup>y, z</sup>, 郑斌 <sup>G</sup>,顶岗沉 <sup>k</sup>, 卢卡 Giancardo <sup>b, 1</sup>, Gwenolé Quellec <sup>一种</sup>, 1, 法布里斯 Mériadeau <sup>Z, C, 1</sup><sup>-#</sup> 工程技术, 桑德拉, 印度的至尊宗师 Gobind Singhji 研究所<sup>b</sup> 德克萨斯大学健康科学中心的生物医学信息学院, 大学在休斯敦, 美国<sup>c</sup> 眼科诊所, 医院 Sushrusha, 桑德拉, 印度马哈拉施特拉邦<sup>d</sup> VUNO 公司, 首尔, 韩国<sup>E</sup> 平安科技(深圳)有限公司, 中国<sup>F</sup> 科大讯飞的研究, 中国合肥<sup>G</sup> 和计算机电气工程学院, 美国俄克拉荷马州大学<sup>H</sup> 韩国成均馆大学三星综合技术院健康科学与技术 (SAIHST), 首尔, 韩国<sup>-#</sup> 北卡罗来纳大学夏洛特, 美国计算机科学系<sup>J</sup> 山东师范大学信息科学与工程学院, 中国学院<sup>K</sup> 放射学和金砖四国, 北卡罗莱纳大学的系在教堂山, 美国<sup>L</sup> Cleerly 公司, 纽约, 美国<sup>\*</sup> 弗吉尼亚理工大学, 弗吉尼亚州, 美国<sup>n</sup> 布法罗大学, 纽约州, 美国<sup>o</sup> 德布勒森, 信息学 4002 德布勒森学院, POB 400, 匈牙利大学<sup>p</sup> 个别研究员, 印度<sup>q</sup> 浙江大学计算机科学与技术学院, 杭州, 中国的大学<sup>r</sup> 机器学习的生物影响分析小组, 生物信息学研究所, A-STAR, 新加坡<sup>s</sup> 新加坡, 新加坡国立大学计算机学院,<sup>t</sup> 北京上工医疗科技有限公司, 中国<sup>u</sup> 上海交通大学计算机科学与工程, 中国系<sup>v</sup> ARTI 科幻官方智能, 人工智能研究所, 上海交通大学, 中国教育部重点实验室<sup>w^A</sup> 生物医学工程佛罗里达大学, 美国 J. 克雷顿普鲁伊特家庭部,<sup>x</sup> 技术克勒格布尔斯, 印度的印度理工学院<sup>y</sup> INESC TEC - 研究所系统与计算机工程, 技术和科学, 波尔图, 葡萄牙<sup>z</sup> FEUP - 波尔图大学的工程, 波尔图, 葡萄牙学院<sup>-#</sup> INSERM, UMR 1101, 布雷斯特, 法国<sup>z</sup> 和电气电子工程系, 国油大学, 马来西亚<sup>c</sup> ImVIA / IFTIM, Université 电动机员兼法国第戎<sup>d</sup> 和计算机电气工程系, 加拿大阿尔伯塔大学<sup>\*</sup> 在卓越中心的信号与图像处理, 工程和技术, 桑德拉 (MS), 印度的至尊宗师 Gobind Singhji 研究所通讯作者。<sup>#</sup> 电子邮件地址 : [\(P. Porwal\)](mailto:porwal.prasanna@sggs.ac.in)。<sup>1</sup> 这些作者共同举办的挑战。所有其他的贡献文件中提出的算法 (S) 的结果

## 文章 信息

## 抽象

## 文章历史：

获得11月2019修订2019年9月9日接受2019年9月16日可在线2019年10月3日

## 关键词：

糖尿病视网膜病变视网膜图像分析  
深度学习挑战

糖尿病视网膜病变 (DR) 是可以避免的视力丧失的最常见的原因，主要影响在全球范围内的劳动年龄人口。筛选DR，再加上及时的咨询和治疗，是全球值得信赖的政策，以避免视力减退。然而，实施DR筛查项目是具有挑战性，由于能够在为DR的风险筛选不断增长的全球糖尿病人群的医疗专业人员的稀缺性。在视网膜图像分析计算机辅助疾病诊断可以提供这样的大规模筛查工作可持续发展的途径。在计算能力和机器学习方法最近科学的进步C提供生物医学科学家们实现这一目标的途径。针对推进状态的最先进的自动诊断DR，关于“糖尿病性视网膜病变一个重大挑战

- 分割与分级”与IEEE国际研讨会生物医学成像 (- 2018 ISBI) 共同举办。在本文中，我们报道了设置和这一挑战的结果主要是基于印度的糖尿病视网膜病变的图像数据集 (IDRID)。有三个主要subchallenges：病变划分，疾病的严重程度分级和视网膜地标和分割的定位。在这个挑战这些多任务允许测试算法的普遍性，这是什么使得它从现有的不同。它接收到来自与其在这一挑战有效的进入495个148登记提交的材料科学的C社区的积极响应。本文概述了挑战，它的组织，使用的数据集，评价方法和表现最出色的解决方案参与的结果。的顶级表现的方法利用的临床信息，数据扩张，以及模型的合奏的共混物。这些发现必须启用视网膜图像分析和基于图像-DR特别是筛选新的发展潜力。

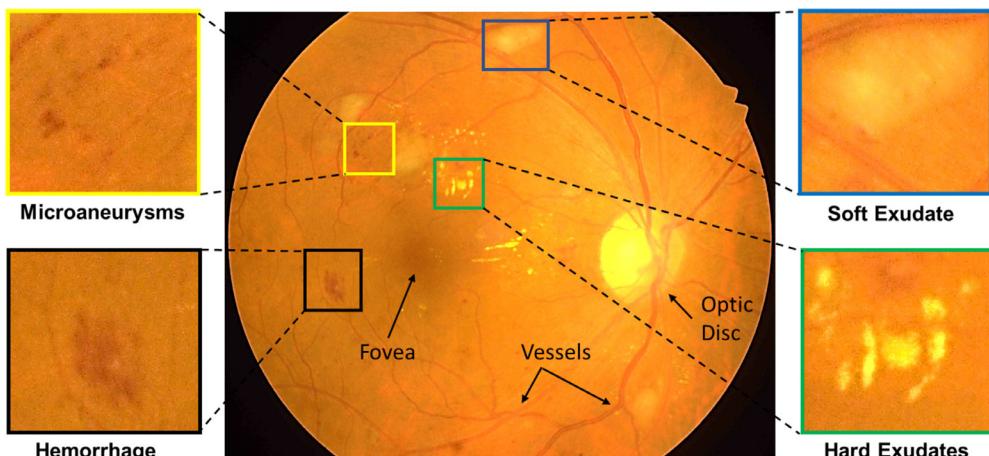
©2019爱思唯尔保留所有权利。

## 1.简介

糖尿病视网膜病变 (DR) 和糖尿病性黄斑水肿 (DME) 是由于引起视网膜微血管最常见的威胁视力的医疗状况的变化糖尿病 (触发 [Lambert和SALZ 2015](#))，主要影响的劳动年龄人口在世界上 ([阿特拉斯，2017年](#))。DR导致在血管结构的逐渐变化 (包括血管迂曲，分支角度和口径) 和所得的异常 (微动脉瘤，出血和渗出物)，而，DME的特征在于FL流体的保留或黄斑的肿胀在DR的任何阶段中可能出现的 ([Bandello等人，2010; Ciulla等人，2003](#))。据国际糖尿病联盟 ([阿特拉斯，2017年](#)) 估计，目前，受糖尿病影响的个体的全球数量为425万美元，并可能上升到6.93亿由2045他们当中，三分之一的人的估计有某种形式的DR，并十分之一是容易视力威胁DR ([ICO，2017年; Bourne等人，2013](#))。DR是通过目测检查视网膜眼底图像的一个或多个视网膜病变的存在等微动脉瘤 (MAS)，出血 (HES)，软诊断

渗出物 (SES) 和硬性渗出 (EXS) ([Wong等人，2016](#))，如图 [图 1](#)。

早期诊断和治疗DR可以防止视力下降。因此，糖尿病患者通常称为视网膜筛查一次或每年两次 ([摩天，1993; Kollias和Ulbig，2010; 汤等人，2016](#))。糖尿病眼部护理，主要是依赖于眼科医生和必要的卫生保健基础设施 (数量 [琼斯和爱德华兹，2010; Lin等人2016](#))。在印度，眼科医生与人口的比例为1:1 07000，然而，在城市地区这一比例为1:9000，而在农村地区存在对6080000居民只有一个眼科医生 ([拉曼等人，2016](#))。到2045年，仅印度预计将有大约1.51亿人患有糖尿病和三分之一的人预计将有DR ([阿特拉斯，2017年](#))。计划到屏幕这样一个人口众多的与实施，管理，人力平地机的可用性和长期财政可持续性DR对垒的问题。因此，需要计算机辅助诊断工具用于筛查这样一个人口众多，需要持续跟进的DR，并减少对眼科医生的负担有效地促进 ([Jelinek和克里，2009年; Walter等人，2002](#))。这样的工具可以帮助医生在identi科幻阳离子，



**图 1。** 扩大区域 (在左侧) 的MA，和的HE和 (右) 的SE，并EXS：视网膜图像 (在中心) 通过突出显示与DR相关的正常结构 (血管，视盘和中央凹中心) 和异常的图示。

解释和视网膜异常的测量，并最终筛选和监测疾病。在计算能力和机器学习方法最近科学的进步<sup>1</sup>提供途径，生物医学科学家的临床实践中迫切要求满足（Shortliffe和布卢瓦，2006；巴顿等人，2006年）。为了满足这种需要，用精确的像素或图像级别专家注释沿原始图像（<sup>2</sup>又名基础事实）播放，以促进开发，验证研究团体的重要作用，以及DR病变分割技术的比较（特鲁科等人，2013）。与DR如MA，请的HE，SES和EXS相关的损伤的精确像素级注释是用于评估的个体病变分割技术的精度非常宝贵的资源。这些精确的分段病变有助于确定疾病的严重程度，并进一步充当可以在后续程序协助疾病的自来水进展的路线图。同样，在另一方面，DR和DME的疾病严重程度映像级专家的标签是像分析和检索算法的开发和评估很有帮助。这必然导致几个研究小组开发并共享视网膜图像数据集，即获月（Decencière等人，2014），E-Ophtha（Decencière等人，2013<sup>3</sup>），DiaretDB（考皮等人，2012），驾驶（面包车Ginneken等人，2004年），凝视（胡佛，1975年），ARIA（派普电子等人，2008年）和HEI-MED（Giancardo等人，2012）。此外，两个挑战在DR，即视网膜病变在线挑战赛（ROC）的背景下举办<sup>2</sup>和Kaggle DR检测挑战<sup>3</sup>。ROC与检测MA的目标组织。然而，Kaggle挑战的目的是获得解决方案，用于确定DR的严重级别。这些挑战促进科学的调查一下社区的来自世界各地的上有竞争力的参与在对科学的进步<sup>4</sup>同时建设性设置启用的网络场的进步。以前的努力已经取得使用图像CLASSI科幻阳离子，模式识别和机器学习很好的进展。经过近二十年的进展，已有若干研究小组进行了系统回顾（巴顿等人，2006年。络纱机等人，2009；阿布拉莫夫等人，2010；Mookiah等人，2013a；Jordan等人，2017；Nørgaard和Grauslund，2018）。

DR和DME的严重性级别为数据库中的每个图像。这种挑战汇集了计算机视觉和生物医学研究人员的最终目标，以进一步刺激和促进研究，以及为一个实用的软件工具的开发，将支持电子FFI古老而精确的测量和视网膜图像的分析提供了一个独特的平台，可以在DR管理有用。最初，地面实况一起训练数据集提供给与会者的其算法的开发。后来，结果被判定在对测试数据集，这些算法的性能。成功是由算法的结果如何密切配合地面实况测量。有三个主要子挑战：病变划分，疾病的严重程度分级，并定位和视网膜地标分割。在IDRID挑战这些多任务允许测试的算法的普遍性，这是什么使得它从现有的不同。此外，这一挑战寻求的自动化解决方案来同时预测DR和DME的严重性。据预测作为一个单独的任务来增加这一挑战相比数得分的任务。

本文的其余部分的结构如下：第2节给出了自动化DR筛查的发展做前期工作的简短评论，第3节提  
供参考数据集的细节，

第4节描述了竞争的组织通过各种阶段和第5节详细介绍了顶级表演竞争对手的解决方案。第6节礼物性能在这个挑战中使用的评估措施。然后，第7节呈现结果，分析和相应的参赛队所有子挑战的排名。

第8节提供的结果，限制，并从这一挑战，并在最后结论的经验教训进行简短的讨论。除了本文中，附录A包括提供与所述数据集IDRID不同国家的最先进的可公开获得的数据库的比较。

## 2. 审查用于检测DR的视网膜图像分析的

尽管许多人的努力已经在网络场已取得一自动化DR筛选过程，病变检测仍然是一个具有挑战性的任务，由于以下几个方面：（1）病变的结构复杂（形状，尺寸，强度）病变，（B）检测在镶嵌图像和噪声（亮边界重新FL ections，脉冲噪声，光重新FL ections），（c）中的高级间相似度的存在（即，之间MA-HE和EX-SE），和不那么罕见（d）外观-lesion结构（神经网络BER重新FL ections，血管再FL ections，玻璃疣）使得二FFI崇拜以建立病变分割一个灵活的和稳健的模型。据我们所知，这个挑战之前，所有病变（MA，HE，SE和EX）可以同时在同一框架下发展没有任何报道。也有，有缺乏通用平台来测试的决定在同一组图像中的正常和异常的视网膜结构的方法的稳健性。此外，没有的像素级注释和DR和DME的同时分级可用性有限（参见表附录A）。为了解决这些问题，我们引入了一种叫做印度的糖尿病视网膜病变的图像数据集（IDRID）新的数据集（Porwal等人，2018A）。此外，它被用来作为一个基础数据集盛大挑战的“糖尿病视网膜病变-分割与分级”组织连同ISBI - 2018年IDRID数据集提供了典型的DR病变和正常视网膜结构的专家标记。它还提供疾病

自动图像处理已被证明是对视网膜眼底图像的分析，并将其应用到未来的眼部护理一个有前途的选择。引进的DR筛查项目的自动化技术和快速增长的深学习技术实现有趣的结果是成功的故事和未来的潜力成就的例子。特别是，研究人员在追（Krizhevsky等人，2012）深学习过的国家的最先进的ImageNet挑战基于模型显示出显着的改善，出现了深度学习的基于潮模式在医学图像分析。因此，我们决定基于他们是否在DR的情况下使用深学习CLASSI科幻阳离子展示最新的相关工作。

### 2.1. 非深学习方法

通过传统手工特征为视网膜图像分析的总体框架为基础的方法涉及若干阶段，通常为：对比度增强或非均匀性均衡，图像分割，特征提取，并且CLASSI音响阳离子的预处理阶段。特征提取策略根据目标参与其中，即视网膜病变检测，疾病筛查或地标定位而变化。2006年，一个研究小组（巴顿等人，2006年）在其视网膜图像分析是基于并讨论用于检测与DR相关的视网膜标和病变的初始技术中概述的原则。后来，复卷机等。（2009年）报道DR的自动化分析工作，分析1998年-2008年，他们归类文学为一系列的操作或步骤的预处理，

<sup>2</sup> <http://webeye.ophth.uiowa.edu/ROC/>

<sup>3</sup> <https://www.kaggle.com/c/diabetic-retinopathy-detection>

脉管分割，本地化和视神经盘（OD）的分割，黄斑中心凹和，检测和病变的分割的定位。一些评论文章（[阿布拉莫夫等人，2010; 乔丹等人，2017年](#)）提供的简要介绍为眼底图像的重点是视网膜病变和自动化技术的identi音响阳离子对于大规模筛选视网膜疾病的分析的定量方法。

在文献中的尝试大部分是针对独家检测和/或一种类型的从图像病变（无论是MA，请的HE，EXS或SES）的分割。一些参与了肿瘤分割常见的方法是数学形态学（[乔希和Karule, 2019; Hatanaka等人, 2008; 张某等人, 2014年](#)），区域生长（[Fleming等人, 2006; 李和Chutatape, 2004年](#)），和监督的方法（[Wu等人, 2017年; 周某等人, 2017年; Garcia等人, 2009; 唐等人, 2013](#)）。模板匹配的除了这些方法，在均线的情况下，最初的研究表明效力（[Quellec等人, 2008年](#)），篇幅值（[Das等人, 2015年](#)），气空间（[Giancardo等人, 2011](#)），稀疏表示（[张某等人, 2012; Javid等人, 2017年](#)），黑森基于区域描述符（[阿达勒等人, 2014](#)）和字典学习（[罗沙等人, 2012](#)）。在另一方面，为的HE的独占分割，超象素的基于特征（[Tang等, 2013; 罗梅罗 - Oraá等人, 2019](#)）被发现是有效的。这些红色病变（二者MAS和HES）也经常使用动态形状特征（一起检测 [Seoud等人, 2016](#)），滤波器响应和多个内核学习（[Srivastava等人, 2017年](#)）和混合特征提取的方法（[Niemeijer等人, 2005年](#)）。同样，对于EXS，研究人员喜欢的聚类方法依赖（[Osareh等人, 2009](#)），基于模型的（[Sanchez等人, 2009; Harangi和豪伊杜2014](#)），蚁群算法（ACO）（[Pereira等人, 2015](#)）和上下文信息（[Sanchez等人, 2012](#)），而对于社会企业的研究人员利用尺度不变特征变换（SIFT）（[纳克维等人, 2018](#)），自适应阈值和ACO（[SRENG等人, 2019](#)）。另外，有几种方法被设计成用于多种病变检测诸如多尺度幅度调制频率调制（[阿古等人, 2010](#)），机器学习（[Roychowdhury等人, 2014](#)），黑森多尺度分析，变分割和纹理特征的组合（[菲格雷等人, 2015年](#)）。这些技术被示出为通常涉及在检测到的解剖结构（即OD和中央凹）与病变检测的相互依存关系，而这又确定自动化DR筛选结果。

定位和OD的分割和中央凹视网膜促进病变程度的检测以及评估（基于这些病变的几何位置）和DR和DME的监测进展。因此，有几种方法已被提出用于OD的定位，且大多使用像强度，形状，颜色，纹理等的OD性质和许多其他显示数学形态学的有效性（[莱斯等人, 2013; 马林等人, 2015年](#)），模板匹配（[Giachetti等人, 2014](#)），可变形模型（[Yu等, 2012; 吴等人, 2016年](#)）和强度廓线分析（[Kamble等人, 2017; 乌里韦 - 瓦伦西亚和马丁内斯Carballido, 2019](#)）。此外，方法用于OD分割是基于水平集（[Yu等人, 2012](#)），阈值（[马林等人, 2015年](#)），活动轮廓（[玛丽等人, 2015年](#)）和形状建模（[Cheng等人, 2015年](#)），聚类（[塔库尔和Juneja, 2017年](#)），以及混合（[Bai等, 2014](#)）方法。类似地，中央凹大多使用具有OD和容器通过形态（几何关系检测 [Welfer等人, 2011](#)），阈值（[Gegundez-Arias等人, 2013](#)），模板匹配（[Kao等人, 2014](#)）和强度廓线分析（[Kamble等人, 2017年](#)）技术。在检测正常解剖结构的性能不佳可能病变的检测和筛查准确度产生不利影响。例如，考虑在2002年提出的基于数学形态学技术（[Walter等人, 2002](#)），2008年（[Sopharak等人, 2008年](#)）和2014（[张某等人, 2014年](#)）。这些作品展示了如何从形态

通过包括用于网络连接的最终目标渗出物检测的多个步骤演进逻辑基于处理的办法。在最初的努力，

[Walter等。（2002年）](#) 设计用于OD和EXS分割，之后除去OD，以获得候选EX的技术。同样的，

[Sopharak等。（2008年）](#) 与OD和血管的检测和清除来实现相同的目标。最近，一种方法被提出 [张某等人。（2014）](#) 实现更好的结果，但它涉及到（a）中的空间校准，暗和亮的解剖结构诸如血管和OD的（b）中的检测还分别（c）中明亮的边界区的检测候选项的实际提取之前。此外，还有基于纹理其他技术（[莱斯等人, 2017; Porwal等人, 2018C](#)）中级（[皮雷斯等人, 2017年](#)），其为放弃DR筛选病变分步骤视网膜图像的功能。然而，大多数这些技术依赖于上面提到的中间步骤。在一方式中基于机器学习（[Roychowdhury等人, 2014](#)），作者检测出亮部和暗病变作为第一步和以后进行分层病变CLASSI音响阳离子以产生用于DR一个严重性等级。同样的，

[安塔尔和豪伊杜（2014）](#) 建议涉及到图像的质量水平评估的策略，预先筛选之后病变的解剖学特征提取，以科幻应受决定有关DR的使用合奏CLASSI网络ERS的存在。此外，对于DR的不同阶段，形态区域属性的identi音响阳离子（[Yun等人, 2008年](#)），纹理参数（[阿查等人, 2012; Mookiah等人, 2013b](#)），高阶谱的非线性特性（[阿查里雅等人, 2008年](#)），混合（[达拉等人, 2015](#)）和信息融合（[Niemeijer等人, 2009年](#)）方法被发现是有用的。由于DME是基于黄斑EXS，许多研究人员的位置分级（[Giancardo等人, 2012; Medhi和Dandapat 2014; 佩尔多莫等人, 2016; 马林等人, 2018](#)）提出了基于EXS功能，以确定DME的严重程度。虽然几个人（[迪帕克和Sivaswamy, 2012; Mookiah等人, 2015; 阿查等人, 2017年](#)）已经提出了各种特征提取技术等级DME阶段没有分割EXS，主要用于本节中的方法，特征是基于颜色，亮度，尺寸，形状，边缘强度，纹理，并且在空间和/或变换域的像素簇的上下文信息。而CLASSI科幻阳离子通过CLASSI网络实现ERS如K近邻（KNN），朴素贝叶斯，支持向量机（SVM），ARTI网络官方神经网络（ANN），决策树等。

这些病变检测或筛选技术被示出为通常涉及与检测其他地标相互依存关系。然而，也缺乏一个单一的平台，以测试他们对每个目标的性能。对于这样的基于手工功能的方法，这一挑战提供了一个独特的平台，比较和对比算法的性能进行检测的解剖结构，病变以及DR和DME的筛选。

## 2.2. 深学习方法

深度学习是一个总称，以能够同时直接从数据学习低级别的代表性和更高级别的参数德科幻NE多层次神经网络。这种表示的学习能力大大减少了工程即席功能的需要，但是，learningbased满端至端培训的深一般方法需要样品的显着的数量。它在最近一个时期的快速发展主要是由于大规模在得知启用多层（两个以上）网络的建设算法计算能力的数据，进步和发展的通量（[欣顿, 2018; Voulodimos等人, 2018](#)）。这一进展引起的利益在创建基于机器视觉的分析，数据驱动模型的健康信息学（[Ching等人, 2018; Ravi等人, 2017年](#)）。因此，它正在成为机器学习的有效工具，有前途的重塑自动化医学图像分析的未来（[斯潘等人, 2016; Litjens等人, 2017; 铃木, 2017年; Shen等人, 2017; Kim等人,](#)

2018; KER等人, 2018)。在深度学习的各种方法的变体, 卷积神经网络(细胞神经网络或ConvNets)是最流行的医学图像分析的科幻场(HOO-Chang等人, 2016; 凯琳和Pencina, 2018)。一些配置简单和CNN的变体可在文献中, 一些最流行的是AlexNet(Krizhevsky等人, 2012), VGG(西蒙尼扬和Ziserman 2014), GoogLeNet(Szegedy等人, 2015年)和RESNET(他等人, 年, 2016年)。

深度学习也被广泛应用在视网膜图像分析, 因为其独特的保护地方形象的关系特征。在文献中采用利用“关闭的, 现成的CNN”深学视网膜图像的方法多数设有作为补充信息信道到其它手工特征或局部显着性映射用于检测与DR相关的异常(胡齐克等人, 2018; 奥兰多等人, 2018; Dai等人, 2018), OD的分割(Zilly等人, 2017; Fu等人, 2018), 和DR的检测(Rangrej和Sivaswamy, 2017年)。作者(Fu等人, 2016)雇用完全连接条件随音响视场与CNN沿着集成像素之间的区别容器概率图和远距离相互作用, 以获得网络连接最终二进制脉管系统。而一些方法初始化参数与预先训练的模型(非医学图像), 然后在网络连接NE-调整(塔杰巴赫什等人, 2016)为DR筛选网络参数(高尔衫等人, 2016; 卡森Lam等人, 2018)。在另一种方法研究人员使用二维(2D)图像块作为输入, 而不是对病变检测fullsized图像(Tan等人, 2017b; 面包车Grinsven等人, 2016; Lam等, 2018; 胡齐克等人, 2018; Khojasteh等人, 2018), 以及OD和中央凹检测(Tan等人, 2017A)。

加西亚等人。(2017)训练“从零开始CNN”, 并与基于其他两个现有架构网络NE-调整结果进行了对比。最近, Shah等。(2018)表明, 在学习视觉内核的字典发现异常的自动编码器可以刺激多样性的整体训练。而

Giancardo等。(2017)提出了一种新的方法来计算血管嵌入, 充分利用新的编码器增强CNN的内部表示, 在DR CLASSI网络阳离子和检索任务演示的改善。有一个在DR采用CNN模型在最近的时间自动ident i科幻阳离子的显着的发展。定制的CNN(Gargeya和冷, 2017年)提出了DR筛选和使用从EyePACS系统获得的75137幅图像(训练夸德罗斯和Bresnick, 2009年), 其中一个附加CLASSI音响ER上CNN衍生特征进一步用于确定如果图像是具有或不具有视网膜病变。同样, 谷歌公司(高爾衫等人, 2016)开发了用于图像CLASSI音响阳离子网络优化(FI netuning), 其中一个CNN通过利用由128175个图像的带标签的回顾性开发数据库训练。也有一些混合算法, 其中多个, 半依赖性CNN的基于视网膜损伤的外观训练有素(阿布拉莫夫等人, 2016; Quellec等人, 2016)。更进一步, 研究人员(Quellec等人, 2017年)展示了基于CNN病变分割的能力训练像级别CLASSI音响阳离子。然而, Lynch等。(2017)证明了基于多个半依赖的细胞神经网络的混合算法可能会提供DR转诊筛查更稳健的选择, 强调病变细分的重要性。对于进一步的细节, 读者建议遵循检测分泌物的最新评论(FRAZ等人, 2018), 红色病变(Biyan和Patre, 2018)和重点是DR的计算机辅助诊断系统综述(Mookiah等人, 2013a; Nørgaard和Grauslund, 2018)。在阿尔季科幻官方情报这目前的进展提供了一个机会, 研究人员为提高DR转诊系统的性能提升到一个更强大的诊断系统, 可以提供针对多种疾病匹配的临床相关的国际标准的定量信息。因此, 所提出的挑战设计提供一个途径, 以评估精确DR严重性状态和OP-

portunity提供病灶准确的措施, 这可能甚至在后续的研究帮助, 观察视网膜图谱的变化。

### 3.印度糖尿病性视网膜病变的图像数据集

#### 该IDRID数据集(Porwal等人, 2018A)从实际创建

在位于Nanded的, (MS), 印度眼科诊所获得的临床检查的受糖尿病影响视网膜的人使用的照片兴和重点黄斑与被抓获VX-10α眼底照相机。之前的图像采集, 所有受者的瞳孔与浓度为0.5%托吡卡胺的一滴扩张。所捕获的图像具有50·网络视场来看, 4288号决议×2848个像素, 并存储在JPG格式。对FI最终数据集由516个图像分为连接的五个DR(0-4)和三个DME(0-2)类与根据临床相关的国际标准以及德音响定义特性。它提供了典型的DR病变和正常视网膜结构的专家标记。它还提供了DR和DME的数据库中的每个图像的疾病严重程度。三种类型的地面真理的数据集中可用:

**1.像素级译注:** 在技术在图像内和分割出的从背景感兴趣区域定位各个病灶这种类型的注释的是有用的。与DR的迹象八十一眼底彩色照片在pixellevel发展的均线, 社企, EXS和的HE地面实况进行了注释。该二进制掩码(如在所示图2)对于每种类型的病变在TIF文件音响格式提供。此外, OD也被注解在像素级和二进制掩模以相同的格式被提供所有81个图像。所有这些注解发挥研究至关重要的作用图像内的分割病变的计算分析。

**2.图像等级分级:** 它由意在描述与整个图像相关整体风险因素的信息。两名医学专家提供的裁决一致等级的全套516个的图像与各种DR和DME的病理状态。分级所有图像是在CSV文件连接可用。糖尿病视网膜图像根据国际临床糖尿病性视网膜病变表CLASSI音响编成单独的基团(Wu等人, 2013), CON组f奈德到下观察图像, 如图表格1。基于EXS的出现接近黄斑中心区(DME严重程度决定Decencière等人, 2014), 如图表2。

3. OD和中央窝中心坐标的OD和中央凹中心位置被标记为所有516个图像和标记可以作为一个单独的CSV文件中。

表格1

DR严重程度分级。NPDR: 非增殖性DR和PDR: 增殖DR

DR等级	发现
0: 异常1的没有明显的视网膜病变无明显迹象: 轻度NPDR	均线的存在只
2: 中度NPDR	不仅仅是均线, 但低于严重NPDR
3: 严重NPDR	以下任何方式: <ul style="list-style-type: none"> <li>· &gt; 20点视网膜内的HE</li> <li>· 静脉串珠</li> <li>· 视网膜微血管异常</li> <li>· 没有PDR的迹象</li> </ul>
4: PDR	任一或两个以下的: 新生血管形成玻璃体/预视网膜HE

表2

风险二甲醚。

DME级	发现
0	没有明显的EX(S)
1	EX(多个)的一个盘直径的半径外的存在从黄斑中心
2	一个盘直径为黄斑中心的半径内EX(一个或多个)的存在

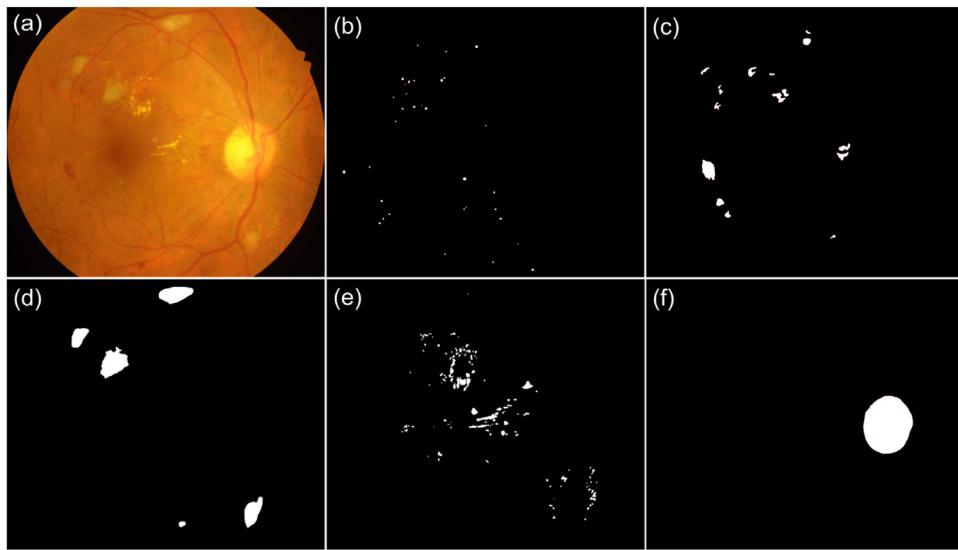


图2中。视网膜照片和不同的像素级别的注释：选自 (a ) 样品的眼底图像的IDRID数据集；样品基础事实为 ( BF ) 的MA，的HE，社会企业，分别EXS和OD。

该IDRID数据集可从IEEE数据端口库<sup>4</sup>

下一个知识共享署名4.0许可。有关数据更详细的信息在数据描述符是可用的 ( Porwal等人, 2018B )。

**A1** 和 **A2** 强调这个数据集的相对优势相对于现有的数据集。IDRID是提供三种类型的上述注释的唯一数据集。这种精简的注解集将允许其在研究和带来更好的普及机型的发展，为图像分析利用，实现自动化DR诊断方面取得进一步进展。

表3

在像素级注释为不同类型的视网膜病变的视网膜图像的分层连接的阳离子。

病变类型	设置 - 一个图片	设置 - B图片
嘛	54	27
他	53	27
SE	26	14
EX	54	27

表4

视网膜图像的分层网络阳离子分级的DR和DME。

DR等级	集-A	设置-B DM级	集-A	集-B
0	134	34	0	177
1	20	五	1	41
2	136	32	2	195
3	74	19		48
4	49	13		

#### 4. 挑战组织

在“糖尿病视网膜病变 - 分割和分级挑战赛”是由成不同阶段，给人一种井井有条的工作过程中对大赛的成功增强。图3 描述了整个挑战组织的工作 - 溢流。面临的挑战是OFFI在ISBI officially公布 - 2018网站<sup>5</sup>

15日十月

2017年。

我们面临的挑战又分为三个子挑战如下：

1. 病变分割：分割与DR如MA，请的HE，EXS和SE相关的视网膜病变。

2. 疾病评分标准：根据DR和DME的严重性级别眼底图像CLASSI音响阳离子。

3. OD检测与分割，以及中央窝检测：OD的自动定位和中央凹中心坐标，OD的分割。

我们面临的挑战涉及4个阶段，具体如下：

**阶段1。数据准备和分发：**该IDRID数据集对这一挑战，专家VERI网络版，所有的图像是足够的质量，临床相关的方式，即无图像被复制和疾病分层网络阳离子代表DR和DME的合理混合物的存在。与基础事实沿数据集分成训练集和测试集。用于与像素级注释的图像，数据被分离为2/3训练（设置-A）和1/3用于测试（设置-B）（见 表3 ）。

类似地，对于OD分割（分挑战的一部分 - 3）数据以相同的比例成集-A（54个图像）和SETB（27个图像）被划分。由于算法的输出是代表了解到感知模式。对于病变和OD分割任务的数据进行了仔细划分，因为它提供了学习足够代表性的数据，并且以后可以用来衡量算法的性能保持性比例这样的方式。这应该是每个子集中的病变和OD分割任务图像的百分比（子挑战 - 1和子挑战的一部分 - 3）是由研究结果的支持（多宾和Simon, 2011）这表明分离数据转换成2/3（训练）：1/3（测试）是用于从50的样本量为200。对于其他子挑战（疾病分级和OD和中央凹中心位置）的最佳选择，数据分离在80（训练集：设置-A）：20（测试集：设置-B）的比例。数据分裂的百分比，在这种情况下，是为了提供分为不同严重性级别的数据的足够量。请注意，该数据集是分层网络根据拆分前的DR和DME成绩编。的数据集的细节的细目示出

<sup>4</sup> <https://ieee-dataport.org/open-access/indian-diabetic-retinopathy-imagedataset-idrid>

<sup>5</sup> <https://biomedicallimaging.org/2018/challenges/>

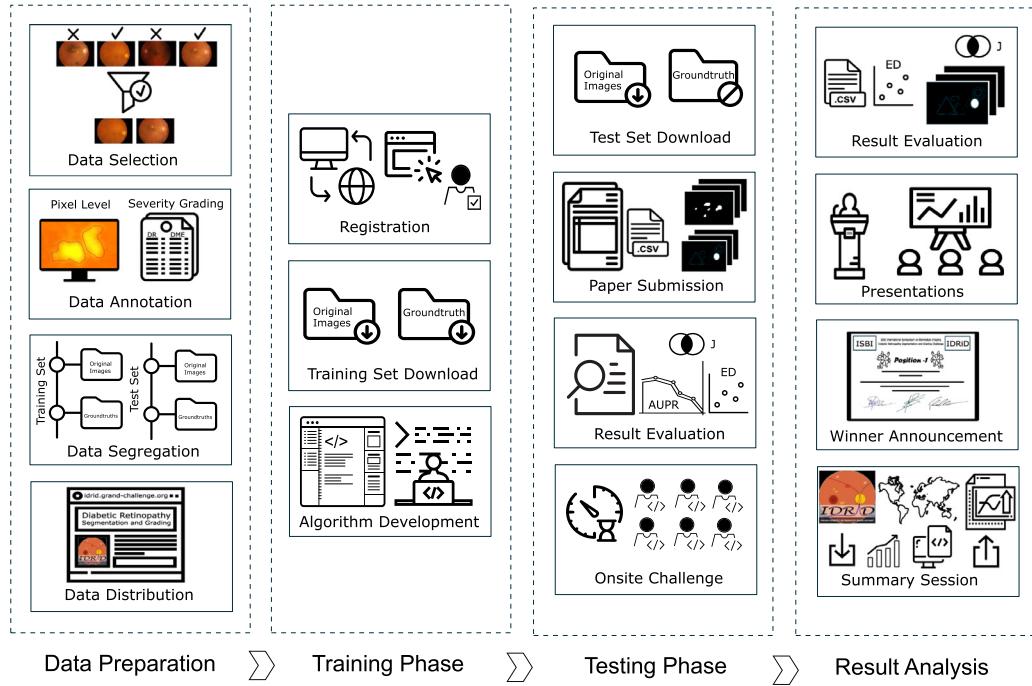


图3。工作ISBI的溢流年 - 2018年 : 糖尿病视网膜病变 - 分割和分级的挑战。

我们面临的挑战托管于在生物医学成像平台大挑战<sup>6</sup>，对于生物医学成像相关的比赛中流行的平台之一。一个挑战网站成立，并于25推出于2017年10月以传播挑战相关的信息。它也被用于注册，数据分发，提交结果和纸张，以及通信组织者和参加者之间。

#### 任务 - 5 : DR和DME分级

3次挑战 - 3 : OD和中央窝检测任务 - 6 : OD中心本地化

#### 任务 - 7 : 黄斑中心凹中心本地化

#### 任务 - 8 : OD分割

挑战网站公开面向提交从12月二月和参与者可以提交他们的成果和文件，说明他们的方法主办方直到11月游行。参赛者可以提交最多三个方法每队被评估为每一个任务，只要有技术之间的显着差异，超越了简单的改变或参数改变。对于任务1~4 (即子挑战 - 1) 和任务-8，团队被要求提交输出概率映射为灰度图像和所有其他的任务，以CSV格式被接受。通过挑战组织者提交的结果进行评估，并显示在挑战网站的排行榜它们的性能。对于子挑战 - 1，团队是基于对提交的关于测试的结果集的性能进行评估，而对于其它两个子挑战使用通过留一交叉验证方法得到的训练结果集进行。在这个阶段，它收到了来自148个提交了37支不同的球队，这出16支队伍入围参与到现场的挑战研究界一个很好的反响。邀请当中，13支球队CON网络Rmed指其参与的现场挑战，而两队拒绝参加由于其他承诺，一个团队是不是能够安排在有限的时间财政支持。

**阶段2。注册和发行训练数据：** 挑战考虑ISBI现场比赛的报名是从发射大挑战网站的打开 (即25月)

2017年)，直到最后期限提交的结果 (即11月游行

2018)。有兴趣的研究团队可以通过挑战网站为一个或所有子挑战注册。第一个数据，即组A (图像和地面真理) 的一部分，该网络已经提供给挑战的参与者20月2018年一月参加者可以下载数据集，并开始开发或他们的方法作案科幻阳离子。此外，他们也被允许使用其他数据集为他们的方法的发展，符合条件外部数据集也应该公开。

**阶段3。测试数据的发布：** SET-B (仅适用于图像) 的subchallenge - 1发布于20月2018年对于其他两个子挑战2月，集-B发布了4月四月一日是“现场”挑战的一部分。从子挑战的现场评估忍住了主办方 - 在结果图像分割任务的评估1个考虑时序约束。

提出的材料都寻求任一的对应于三个子挑战 (1 - 病变分割，2 - 疾病分级，3 - OD和中央窝检测) )以下8个不同的任务如下：

1分挑战 - 1 : 病变划分

任务1 : MA分割

任务 - 2 : 何分割

任务 - 3 : SE分割

任务 - 4 : EX分割

2.分挑战 - 2 : 疾病分级

**阶段4。ISBI挑战活动：**面临的主要挑战活动是在同期举行ISBI年 - 2018年四月，4月2018年使用SET-B (仅适用于图像) 的子挑战 - 2和3提供给通过挑战网站 (联机模式)，以及在挑战网站 (离线模式) 的便携式设备的参与者。参与者被要求出示一小时内的各个挑战任务结果。参赛队伍可以把他们自己的系统或远程系统运行测试。此外，没有关于可能被用来产生结果的机器数量没有限制。但是，考虑到用于处理的定时约束，

<sup>6</sup> <https://grand-challenge.org/>

表5

所有参赛球队名单入围并参加了“现场”的挑战。所有团队颜色代码在所有上市进一步参考更容易。该DL表示提交的算法是否是基于深度学习。其中，子挑战 - 1 ( SC1 ) 对应于病变分割诸如微动脉瘤 ( MA )，出血 ( HE )，软性渗出 ( SE ) 和硬性渗出 ( EX )。而，亚挑战 - 2 ( SC2 ) 表示疾病的严重程度分级对应于DR和ODM E。类似地，子挑战 - 3 ( SC-3 ) 与视盘检测 ( ODD )，视网膜中央凹的检测 ( FD ) 和视盘分割 ( ODS ) 的交易。Harangi等。参加两种方法HarangiM1和Harangi M2，为简单起见，共同具有单一颜色代码表示为HarangiM1-M2。同样，李等人。参加两种方法LzyUNCC和LzyUNCC\_Fusion ( 文中为LzyUNCC-II重命名 ) ( 文中为LzyUNCC-I改名 )，它们共同具有相同的颜色代码表示为LzyUNCC。然而，这些不同的方法在文中单独提及无论它是必要的。·团队“现场”的挑战不能参加，但后来收到的结果，组织者。

队名	作者	DL	SC1			SC2		SC3	
			MA	HE	SE	EX	奇	FD	ODS
VRT	Jaemin儿子等人。								
科大讯飞，MIG	峰岩Wang等。					x	x	xx	
PATech	刘力红等人。				x	x	x	xx	
越早	云芝Wang等。					x	x	xx	
SAIHST	尹浩Choi等人。	x	x	x	x	x	x	xx	
LzyUNCC	中宇李等人。	x	x			x		xx	
SDNU	李晓丹Sui等。					x			
长毛象	王俊彦Wu等人。	x	x		xx		x	xx	
HarangiM1-M2	的Balazs Harangi等人，	x	x		xx		x	xx	
AVSASVA	Varghese表示亚历克斯等人。	x	x		xx		x	xx	
DeepDR	凌带等。	x	x		xxx			x	
ZJU-BLL-SGEX	Xingzheng吕等人。	x	x		xxx				
IITkgpKLIV	Oindrila Saha等。	x	x		xxx		x	x	
· CBER	安娜门多萨等人。	xx		x	xxx				

这之前曾与一个以上的溶液进入一些球队决定只用自己的表现最佳的解决方案。

此外，从子挑战前三名的球队 - 1予陈述他们的工作机会。在此期间，一些组织团队成员编为subchallenge结果 - 2个3队分别获得为7分钟，他们的方法的介绍和3分钟，保留questionanswers。在第一个演示文稿会议持续了大约30分钟，在子挑战的演讲结束 - 1的结果子的挑战 - 2和3被宣布。同样，这些子挑战前三名表演队给他们的工作，简短的介绍。现场挑战活动结束时，6之后四月，挑战和分析结果的总结提出，其中包括了网络最终排名竞争的解决方案。此信息挑战网站上还可以访问。很多团队曾参与多个子挑战中列出的是要注意重要 表5 而本文的其余部分与被选择的挑战，这些方法只处理。

resentation - 这使得它计算EFFI client ( Shen等人，2017年 )。CNN的体系结构通常由一个输入层，后面是卷积 ( CONV ) 的序列，子采样 ( POOL )，完全连接 ( FC ) 层和F1应受一个使用SoftMax或回归层的，以产生所需的输出。所有层的功能作详细说明如下：

一组被用来与输入层执行2D卷积独立滤池 ( 或内核 ) 的CONV层包含 ( 一世 ) 以产生特征 ( 或激活 ) 映射 ( 一种 ) 这给在每个空间位置内核的响应。在数学上，用于输入接插 ( 一世  $x, y$  ) 在位置为中心 (  $X, Y$  ) 的  $\alpha$  层，在特征值一世  $\beta$  特征图，一种  $x, y, i, t$ ，被获得为：

$$\text{一种 } x, y, i, t = F((W_i)_t \cdot \text{一世 } x, y + b_i) = F(Q_{\alpha, x, y, i, t}) \quad (1)$$

其中，参数  $w^{\alpha, i}$  和  $b_i$  是权重向量和的偏差项一世  $\beta$  的滤波器  $\alpha$  层，和  $F(\cdot)$  是一个非线性激活函数如乙状结肠，整流音频编线性单元 ( RELU ) 或双曲正切 ( 双曲正切 )。值得注意的是，内核是非常重要的  $w^{\alpha, i}$  产生特征地图  $C$ ：，：，一世共享，降低了模型的复杂性，使网络更容易培养。

## 5.同类竞争解决方案

参赛队的多数建议对该挑战解决任务的CNN基础的方法。本节详细介绍有关CNN和参赛队伍利用及其变种的基本术语和缩写。此外，总结了解决方案及相关技术SPECI网络阳离子。对于一个特定的方法的详细说明，请参阅ISBI大挑战研讨会在诉讼 [HTTPS : //idrid.grand-challenge.org/](https://idrid.grand-challenge.org/) 组织/ Challenge\_Proceedings /。

POOL层旨在通过减少特征地图的分辨率，实现平移不变性。在POOL层的特征图中的每个单元使用来自相应的卷积特征映射稀疏连接内的单元的子集的。最常见的池操作是平均池和最大池。它执行下采样操作，并通常置于两个CONV层之间实现一组分层的图像特征。在初始CONV层内核检测低级别的功能诸如边缘和曲线，而在较高的层中的内核被学习以编码更多的抽象的功能。的几个CONV和POOL层的顺序逐渐提取更高级别的特征表示。

对于输入图像，CNN变换在一端上的原始图像的像素来产生在另一端的单个可微得分函数。它利用三个机制 - 稀疏连接 ( 又名当地接受网络ELD )，重量共享和不变 ( 或等变 ) REP-

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局语义信息。

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 ( $\theta$ ，两种常见的符号  $w^A$  和  $b$ ) 用于特定任务可以通过最小化损失函数 (来确定  $L$ )

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

被推断为：

$$L = 1 - \frac{\sum \bar{n}}{\bar{n}} \ln(\theta; G_n, \emptyset N) \quad (2)$$

哪里  $\bar{n}$  表示训练图像的数量，一世  $N$ ， $\emptyset N$  和

$G_n$  相当于  $\bar{n}$  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，

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形网体系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约网络。上采样部分由大量特征的通道，其允许网络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。

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FFI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 et (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

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：

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

#### 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ffING (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接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FL IP，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		
	VRT						剪断	FOV裁剪，通过255，那么平均减法除法
次挑战 - 1	科大讯飞					x		病变补丁提取
	PATech		x		x		颜色 <sup>a</sup>	RGB到LUV，对比度调节
	SDNU	x		x	x		-	-
	越早	x		x		x		平均减法，病变补丁提取
	LzyUNCC	x		x	x		随机和光度量 <sup>b</sup>	FOV裁剪，图像增强
	SAIHST		x	x	x	x	x	CLAHE，高斯平滑
	LzyUNCC	x	x	x	x		颜色 <sup>a</sup> ， 随机和光度量 <sup>b</sup>	FOV裁剪，图像增强
次挑战 - 2	VRT	x	x	x	x	x	x	平均减
	长毛象					x	颜色	形态学打开和关闭
	AVASAVA	x	x	x	x		x	intensity scaling
	HarangiM1	x	x	x	x	x	x	FOV cropping
	HarangiM2	x	x	x	x	x	x	-
Sub-challenge - 3	DeepDR	x	x	x	x		OD, fovea region	FOV cropping, mean subtraction
	VRT						shear and cropped OD	FOV cropping, contrast adjustment
	ZJU-BII-SGEX	x	x	x	x	x	x	FOV cropping
	SDNU	x		x	x	x	-	-
	IITkgpKLIV		x	x	x	x	-	-
	CBER	x	x	x	x	x	x	-

<sup>a</sup> Reference: Krizhevsky et al. (2012)

<sup>b</sup> Reference: Howard (2013)

function given by

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

where  $N$  denotes the number of pairs in a batch,  $G_n$  and  $O_n$

represent true segmentation and predicted segmentation for  $n$ th image. The value of  $\alpha$  was determined as follows:

$$\alpha = \frac{B_1}{F_1} \quad (4)$$

where  $B_1$  and  $F_1$  denote the number of background and foreground pixels in  $n$ th image. Since background overwhelms foreground in lesion segmentation task, this loss function was designed to penalize false negatives in order to boost sensitivity, an important factor in detecting lesions. Also,  $\gamma$  was left as a hyper-parameter and chosen out of {0.25, 0.5, 1, 2, 4, 8, 16, 32, 64, 256, 512} to yield the highest AUPR on validation set. The final selected  $\gamma$  values for different lesions are summarized in Table 7.

They trained the network over 300 epochs using Adam optimizer (Kingma and Ba, 2014) with hyper-parameters of  $\beta_1 =$

$0.5$ ,  $\beta_2 = 0.999$  and learning rate of  $2 \times 10^{-4}$  until 250 epochs and  $2 \times 10^{-5}$  until the end. All implementation was done by Keras 2.0.8

Table 7

$\gamma$  values in Eq. (4).

EXs	SEs	HEs	MAs
64	512	8	32

with tensorflow backend 1.4.0 using a server with 8 TITAN X (pascal). The source code is available at [https://bitbucket.org/woalsnd/ isbi\\_2018\\_fundus\\_challenge](https://bitbucket.org/woalsnd/ isbi_2018_fundus_challenge).

### 5.1.2. IFLYTEK-MIG (Fengyan Wang et al.)

Wang et al. proposed a novel cascaded CNN based approach for retinal lesion segmentation with U-Net (Ronneberger et al., 2015) as a base model. It consists of three stages, the first stage is a coarse segmentation model to get initial segmentation masks, then the second stage is a cascade classifier which was designed for false-positive reduction, at last, a fine segmentation model was used to refine results from previous stages. First stage model was trained using the patches of size  $256 \times 256$  pixels centered on a particular lesion amongst MA, HE or EX and  $320 \times 320$  pixels for SE, resulting in the coarse segmentation outcome. Results of the previous stage are coarse due to the fact that non-focus regions

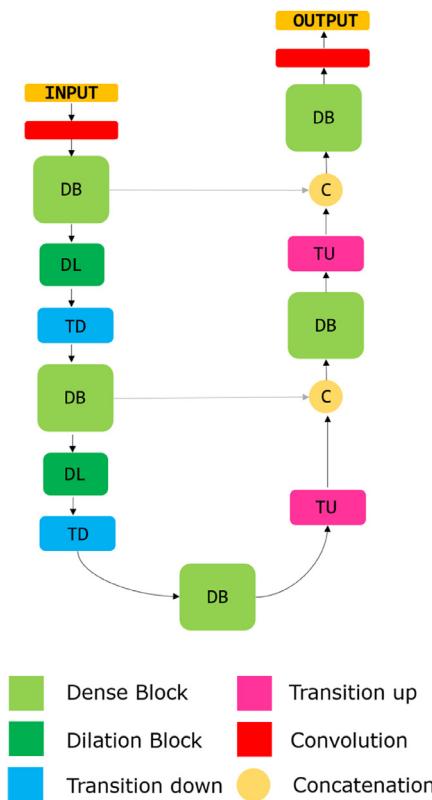


Fig. 4. Proposed architecture for lesion segmentation.

(non-target lesions) were not utilized in the learning process leading to high false-positive count. In the second stage, unlike the first segmentation model which used a lesion centered sample from input dataset pool, candidate regions were extracted using probability maps from the previous stage. Here, the input size fed to model for SE was  $320 \times 320 \times 3$  pixels, for HE and EX it was  $256 \times 256 \times 3$  pixels, and for MA it was modified to  $80 \times 80 \times 3$  pixels considering its small appearance. In this step, a candidate region was regarded as a positive sample if its intersection-over-union with the ground truth was greater than the given threshold (i.e. 0.5). In this way, most trivial non-focus regions were effectively rejected. However, it was identified in the test that a small proportion of false positives still exist, so an additional model was introduced to refine the segmentation results. In the last stage, candidate regions survived from the second stage were utilized as the input patches resulting in more accurate segmentation results. For the first and third stage, they used binary cross-entropy or dice loss function (multi-model training), whereas, for the second stage, they used only binary cross-entropy as a loss function. The first, second and third stage models were trained for 100, 300 and 100 epochs respectively with the momentum of 0.9. In which, the initial learning rate for the first and third stage was set 0.1 and is reduced by 10 times every 30 epochs, and for the second stage it was set to 0.001 reduced by 10 times every 80 epochs. MXNET platform was used for training the models.

### 5.1.3. PATEch (Liu lihong et al.)

Lihong et al. developed a novel patch-based CNN model (as shown in Fig. 4) in which they innovatively combined the DenseNets (Iandola et al., 2014) and dilation block with UNet (Ronneberger et al., 2015) to capture more context information and multi-scale features.

The model is composed of a down-sampling path with 4 Transitions Down (TD), 4 Dilatation Block (DL) and an up-sampling path

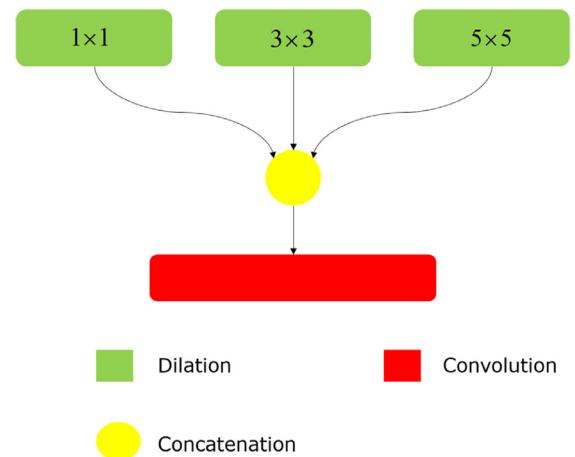


Fig. 5. Architecture for dilation block.

with 4 Transitions Up (TU). To capture multi-scale features, DL (see Fig. 5) is used with dilation rate of 1, 3 and 5 are concatenated for the convolution. The dense block (DB) is constructed by four layers. The idea behind novel combination of dilation convolution is to better deal with the lesions appearing at different scales, where small dilation rate pay closer attention to the characteristics of tiny lesions, larger dilation rate focus on large lesions. On the other hand, use of DB's enabled a deeper and more efficient network.

Initially, they extracted regions within FOV from the images and then normalized them to eliminate local contrast differences and uneven illumination. Later, they used small patches  $256 \times 256$  pixels at a stride of 64 (128 for MA) to generate the training samples (only patches that overlap with the lesion ground truth) followed by data augmentation before feeding to the model. To deal with highly imbalanced spread of data, they designed a loss function that is a combination of dice function (Sudre et al., 2017) and 2D cross Entropy as follows:

$$\begin{aligned} L = & -\text{mean}(w_{10} * G * \log(O) \\ & + w_{11} * (1 - G) * \log(1 - O) \\ & + w_2 * \text{dice}(G)) \end{aligned} \quad (5)$$

where  $w_{10}$  and  $w_{11}$  are the factors utilized to keep a balance between the positive and negative pixels, and  $w_2$  is the factor utilized to control significance between dice and cross entropy loss. The values of  $w_{10}$ ,  $w_{11}$  and  $w_2$  were empirically set to 0.7, 0.3 and

0.4 respectively. The models were trained using Adam optimizer with default parameters,  $\beta_1 = 0.9$  and  $\beta_2 = 0.999$ . The initial learning rate was set to  $2 \times 10^{-4}$ , and then divided by 20 in every 20 epochs. This model was implemented with PyTorch1.12 and Tesla M60 platform was utilized for training on the CentOS 7.2 operating system.

### 5.1.4. SOONER (Yunzhi Wang et al.)

Wang et al. adopted U-Net (Ronneberger et al., 2015) architecture for solving retinal lesion segmentation problem. The network takes a  $380 \times 380$  pixels fundus image patch as input and predicts the binary mask of the retinal lesion within  $196 \times 196$  pixels central region of an input patch. They pre-processed fundus images by subtracting the local mean of each color channel and performed random flipping for data augmentation. Batch normalization was utilized to improve training efficiency and all convolution operations adopted 'valid' paddings. For training, they followed a three-stage process for each type of lesions (i.e. MA, HE, EX and SE). For the first stage, they extracted positive image patches from the training set according to given ground truth mask, and randomly extracted negative image patches from fundus images with

and without apparent retinopathy. The objective function was a summation of cross-entropy loss functions for MA, HE, EX and SE. Adam algorithm was employed to optimize the parameters. In the second stage, they fine-tuned U-Net using the extracted patches for each lesion type. Subsequently, they applied optimized U-Net on fundus images in the training set and extracted false-positive patches generated by U-Net. They further fine-tuned U-Net using positive image patches together with false-positive patches (hard negative patches) as a third stage. In the testing phase, they extracted overlapped image patches using a sliding window and fed these patches into the network to get corresponding probability maps. The initial learning rate was set to  $e-4$  and the fixed number of steps was used as a stopping criterion. They implemented U-Net architecture based on TensorFlow library with Nvidia GeForce GTX 1080Ti GPU.

#### 5.1.5. LzyUNCC (Zhongyu Li et al.)

Li et al. developed a method based on FCN by embedding DLA structure ([Yu et al., 2017](#)) for segmentation of EX's and SE's. As the lesions are located dispersively and irregularly, the embedding of DLA structure with FCN enables better aggregation of semantic and spatial information from local and global level provides a boost in recognizing their presence. They used retinal images with pixellevel ground truth annotations from both IDRiD and E-Ophtha database. They first adopted a series of methods for data preprocessing and augmentation. Subsequently, considering the correlation between EX's and SE's, they first trained an initial model for segmentation of EX. They chose a smaller model, i.e., DLA-34 to train the segmentation network with binary cross-entropy as a loss function. At last, the trained deep model was fine-tuned for segmentation of SE. While the model training of EX segmentation, a trade-off parameter (penalty) was assigned in loss function to control the weights of foreground pixels, and tried different penalty value from 1 to 16. At last, these segmentation results were fused to adaptively compute the best performance. They adopted original DLA cityscapes segmentation experimental settings and trained the model for 100 epochs with batch size 4, where the poly learning rate was  $(1 - \frac{epoch}{total epoch})^{0.9}$  with the momentum of 0.9. The initial learning rate was set to 0.01.

#### 5.1.6. SAIHST (Yoon Ho Choi et al.)

Choi et al. proposed a model for segmentation of EX based on U-Net ([Ronneberger et al., 2015](#)), in which CONV layers of encoder path are replaced with DB's. Whereas, the decoder path of their model was kept identical to that of general U-Net. They built DB with a growth factor of 12 and  $3 \times 3$  CONV layers, batch normalization, and ReLU activation. The last layer generates a pixel level prediction map for EXs through the sigmoid activation function. For training, they utilized only the green channel of fundus image and enhanced it using Contrast Limited Adaptive Histogram Equalization (CLAHE). Later, each image was padded to a size of  $4352 \times 3072$  pixels and cropped into 204 patches of  $512 \times 512$  pixels. These patches are further augmented and used for training. The losses were calculated by binary cross-entropy. The model was trained for 20 epochs with a mini-batch size of 10 and they used Adam optimizer with an initial learning rate of  $2 \cdot e-4$ ,  $\beta_1$  of 0.9 and  $\beta_2$  of 0.999. The model was programmed in Keras 2.1.4 served with TensorFlow 1.3.0 backend.

#### 5.1.7. SDNU (Xiaodan Sui et al.)

Sui et al. proposed a method based on Mask R-CNN structure to segment lesions from the fundus image. They adopted the implementation of Mask R-CNN from [Abdulla \(2017\)](#) for solving the problem. This method could detect different objects while simultaneously generating instance segmentation mask. Network training precedes the data augmentation process and binary cross-entropy

was used as a loss function. The initial learning rate was set to 0.02 with a momentum of 0.9. They chose ResNet-101 as a backbone. They implemented an algorithm in Keras with Tensorflow as backend and processed on 8 NVIDIA TITAN Xp GPUs. The experimental environment was built under Ubuntu 16.06.

#### 5.2. Sub-challenge – 2: Disease grading

For a given image, this task seeks to get a solution to produce severity grade of the diseases i.e. DR (5 class problem) and DME (3 class problem). The summary of participating solutions is as follows:

##### 5.2.1. LzyUNCC (Zhongyu Li et al.)

Li et al. developed a method based on ResNet by embedding DLA structure for automated grading of DR and DME. For this work, they used IDRiD and Kaggle dataset. Initially, for the given training images, they perform data preprocessing and data augmentation. Subsequently, based on the designed ResNet with DLA structure, initial models are trained using 35,000 retinal images from Kaggle dataset. Later, they fine-tuned the model using IDRiD dataset through 5 fold cross-validation technique. Finally, the five outputs are ensembled together as final grades for input images. It is important to note that networks for grading of DR and DME were trained separately. The training was performed by SGD with a mini-batch size of 64, while the learning rate starts from 0.001 and it is then divided by 10 every 20 epochs, for 30 epochs in total. The other hyper-parameters are fixed to settings of original DLA ImageNet classification ([Yu et al., 2017](#)).

##### 5.2.2. VRT (Jaemin Son et al.)

Son et al. used network ([Son et al., 2018](#)) for DR grading. Kaggle dataset was initially used to pre-train the network and then the model was fine-tuned using IDRiD dataset. The penultimate layer was Global Average Pooled (GAP) and connected with FC layer. The entire output is a single value from which L2 loss was calculated against the true label. SGD was used with Nesterov momentum of

**0.9 as an optimizer. Learning rate was set to  $10^{-3}$ . The model was trained for 100 epochs.** Fundus image was normalized in the range [0, 1] and the mean was subtracted channel-wise. For grading of DME, segmented EXs (using the segmentation network proposed in sub-challenge – 1), localized fovea and segmented OD (using the segmentation network proposed in sub-challenge – 3) were utilized for making the final decision. With this information, the semi-major axis of segmented OD ( $r$ ) was estimated. Further, the fundus image was divided into three regions as macular region:

**$\|x - c\| < r$ , near macular region;  $r < \|x - c\| < 2r$  and remaining region;  $2r < \|x - c\|$ .** where  $x$  denotes a point in the image. Furthermore, several features such as sum of intensity for segmented EX, the number of pixels above threshold (178 in the [0, 255] scale), the number of pixels for smallest and largest blob, mean of the number of pixels for blobs are extracted for each area, and binary flag that indicates whether the OD is segmented. Now, features with high importance were selected among numerous features in the initial training due to gradient boosting (for instance, XGBoost) was likely to overfit when provided with overly redundant features. Messidor dataset was added to the given data and out of which 10% of images were left as the validation set. Set of hyper-parameters were searched by grid-search approach. The combination of hyper-parameters that yielded the highest accuracy in the validation set was min child weight: 2, subsample: 0.2, colsample by tree: 0.2,  $\lambda$ : 9.0,  $\alpha$ : 1.0, and depth: 6. Other hyperparameters are set to default values. All implementations were done by PyTorch v0.4.1 using a server with 8 TITAN X (pascal). The source code is available at [https://bitbucket.org/woalsnd/isbi\\_2018\\_fundus\\_challenge](https://bitbucket.org/woalsnd/isbi_2018_fundus_challenge).

### 5.2.3. Mammoth (Junyan Wu et al.)

Wu et al. proposed a unified framework that combines deep feature extractor and statistical feature blending to automatically predict the DR and DME severity scores. For DME, they used DenseNet ( [Iandola et al., 2014](#) ) to directly predict the severity score. Whereas for DR, Kaggle training dataset was used to pretrain the DenseNet model through a dynamic sampling mechanism to balance the training instances and later fine-tuned using IDRiD dataset. Initially, the background of all images was cropped and resized to  $512 \times 512$  pixels. Later, morphological opening and closing are utilized to preserve bright and dark regions. For instance, the morphological opening can erase the EXs and highlight the MAs. Whereas, the closing operation can remove MAs and preserve EXs. These operations can be used to denoise specific levels of classifications, for example, the risk of DME only depends on the location of the EXs. Further, several standard data augmentation methods (as shown in [Table 6](#)) are also employed. Mean Squared Error (MSE) and cross-entropy with five classes were the loss functions employed to train the network and SGD for optimization. The initial learning rate was set to 0.0005 with a decrement of 0.1 after every 30 epochs. The initial training was done by 200 epochs and fine-tuning by 50 epochs. Afterward, the last layer was removed before the final prediction, and its statistical features were aggregated together into a boosting tree. Specifically, 50 pseudo-random augmentations were performed to get 50 outputs from last second FC layer (size of 4096), then the mean and standard deviation of 50 feature vectors for each image was computed, and both vectors were then concatenated together for training in LightGBM. The output from the second last layer of fine-tuning experiments was used to train a blending model, strategy adopted from team o\_O's solution of Kaggle DR challenge. Finally, for the disease grading prediction, gradient boosting tree model was built on a combined second last layer from the pre-trained network and fine-tuned network.

### 5.2.4. Harangim1 (Balazs Harangi et al.)

Harangi et al. proposed an approach for the classification of retinal images via the fusion of two AlexNet ( [Krizhevsky et al., 2012](#) ) and GoogLeNet ( [Szegedy et al., 2015](#) ). For this aim, they removed FC and classification layers and interconnect them by inserting a joint FC layer followed by the classic softmax/ classification layers for the final prediction. In this way, single network architecture was created which allows to train the member CNN's simultaneously. For each  $O^{(n)}$ , let us denote the outputs of the final FC layers of the member CNN's by  $O_1^{(n)}, O_2^{(n)}$ .

$O_1^{(n)}, O_2^{(n)}$  The FC layer of

their ensemble aggregates them via

$$O^{(n)} = A_1^{(n)} + A_2^{(n)} \quad (6)$$

where weight matrices  $A_1, A_2$  were of size  $5 \times 5$  and initialized as

$$\begin{array}{cccccc} \cdots & 1/5 & & & & \cdots \\ & 0 & 0 & 0 & 0 & \\ A_1 = A_2 = & 0 & 1/5 & 0 & 0 & \\ & 0 & 0 & 1/5 & 0 & \\ & 0 & 0 & 0 & 1/5 & \\ & 0 & 0 & 0 & 0 & 1/5 \end{array} \quad (7)$$

The last two layers of the ensemble were a Softmax and a classification one. Let  $O^{(n)}$

be an output of a former layer, the MSE was used for optimization as a loss function:

$$MSE = \frac{1}{2N} \sum_{n=1}^N (O^{(n)} - O^{(n)})^2 \quad (8)$$

During the training phase, back-propagation is applied to minimize the loss via adjusting all parameters of member CNNs and weight matrices  $A_1, A_2$ .

For the grading of DME, the final layers of member CNNs consist of 3 neurons, and weight matrices  $A_1, A_2$  were  $3 \times 3$ , initialized as

$$\begin{array}{ccc} [ & 1/3 & ] \\ & 0 & 0 \\ A_1 = A_2 = & 0 & 1/3 \\ & 0 & 0 & 1/3 \end{array} \quad (9)$$

For training, they merged IDRiD and Kaggle training set. The parameters of architectures were found by SGD algorithm in 189 and 50 epochs respectively for DR and DME classification tasks. Learning rate was set to 0.0001. Training times required on the datasets for DR and DME were 96.6 (189 epochs) and 23.4 (50 epochs) hours respectively. Implementation of this work was done in MATLAB 2017b. The training was performed using an NVIDIA TITAN X GPU card with 7 TFlops of single-precision performance,

336.5 GB/s of memory bandwidth, 3072 CUDA cores, and 12 GB memory.

### 5.2.5. AVSASVA (Varghese Alex et al.)

Alex et al. used ensembles of pre-trained CNNs (on ImageNet dataset), namely, ResNets ( [He et al., 2016](#) ) and DenseNets ( [Iandola et al., 2014](#) ) for the task of disease grading. For DR grading, two ensembles of CNN's namely "primary" and "expert" classifiers were used. The primary classifier was trained to classify a fundus image as one of the 4 classes viz; Normal, Mild NPDR, Moderate NPDR or S-(N)-PDR, a class formed by clubbing Severe NPDR and PDR. The expert classifier was trained exclusively on Severe NPDR or PDR images and was utilized to demarcate the input image as one of the aforementioned classes. During inference, each fundus image was resized to a dimension of  $256 \times 256$  pixels. For the task of grading of DR in fundus images, they used test time augmentation through the "Ten Crop" function defined in PyTorch. The images were first passed through the primary classifier and then through the expert classifier, only if the image was classified as S-(N)-PDR by the primary classifier. The final prediction was achieved by using a majority voting scheme.

For DME grading, two ensembles were trained in a one versus rest approach. Ensemble 1 was trained to classify the input as either "image with no apparent EXs" (Grade 0) or "presence of EXs in image" (Grade 1 & Grade 2), while the Ensemble 2 was trained to classify an image as "Grade 2" DME or not (Grade 0 & Grade 1). During inference, the resized images were fed to both ensembles, and the final prediction was obtained by combining the two predictions by utilizing a set of user-defined rules. Briefly, the user-defined rules were: an image was classified as Grade 0 DME if ensemble 1 and ensemble 2 predict the absence of EXs and the absence of grade 2 DME respectively. A scenario wherein ensemble 2 predicts the presence of grade 2 DME, images were classified under category "Grade 2 DME" irrespective of the prediction from ensemble 1. Lastly, images were classified as Grade 1 DME if none of the above conditions were satisfied.

Both models for DR and DME were initialized with the pretrained weights and the parameters of networks were optimized by reducing cross-entropy loss with ADAM as an optimizer. The learning rate was initialized to 10<sup>-3</sup> for DR and 10<sup>-4</sup> for DME. For DR, the learning rate was reduced by a factor of 10% every instance when the validation loss failed to drop. Each network was trained for 30 epochs and the model parameters that yielded the lowest validation loss were used for inference. For DME, the learning rate was annealed step-wise with a step size of 10 and the multiplicative factor of learning rate decay value of 0.9.

### 5.2.6. Harangim2 (Balazs Harangi et al.)

Harangi et al. combined self-extracted, CNN-based features with traditional, handcrafted ones for disease classification. They modified AlexNet ( [Krizhevsky et al., 2012](#) ) to allow the embedding of

handcrafted features via the FC layer. In this way, they created a network architecture that could be trained in the usual way and additionally uses domain knowledge. They extended the FC layer, to get *FC<sub>fuse</sub>*, originally containing 4096 neurons of AlexNet by adding 68-dimensional vector containing handcrafted features. Then, the  $4164 \times 5$  (or  $4164 \times 3$  for DME) layer *FC<sub>class</sub>* was considered for DR (or DME) classification task. In this way, both final weighings *FC<sub>class</sub>* of handcrafted features were obtained and the 4096 AlexNet features were trained by backpropagation.

To obtain 68 handcrafted features used by CNN, they employed one image level and two lesion-specific methods. The amplitude-frequency modulation (AM-FM) method extracts information from an image by decomposing its green channel at different scales into AM-FM components (Havlicek, 1996). As a result, a 30-element feature vector was obtained, which reflects the intensity, geometry, and texture of structures contained in the image (Agurto et al., 2010). Whereas to extract features related to the lesions MA and EX, they employed two detector ensembles (Antal and Hajdu, 2012; Nagy et al., 2011), which consist of a set of <preprocessing method (PP), candidate extractor (CE)> pairs organized into a voting system. Such a <PP, CE> pair was formed by applying PP to the retinal image and CE to its output. This way, a <PP, CE> pair extracts a set of lesion candidates from the input image, acting as a single detector algorithm. They used the output of these ensembles to obtain 38 features related to the number and size of MA's and EX's. Parameters of the architectures were optimized by SGD algorithm in 85 and 50 epochs for DR and DME respectively. Training times were 83.1 (85 epochs) and 46.2 (50 epochs) hours on the datasets for DR and DME. Implementation of this work was done in MATLAB 2017b. Training has been performed using an NVIDIA TITAN X GPU card with 7 TFlops of single-precision, 336.5 GB/s of memory bandwidth, 3072 CUDA cores, and 12 GB memory.

### 5.3. Sub-challenge – 3: Optic disc and fovea detection

For a given image, this task seeks to get a solution to localize the OD and Fovea. Further, it seeks to get the probability of a pixel being OD (OD segmentation). Summary of approaches is detailed as follows:

#### 5.3.1. Deepdr (Ling Dai et al.)

Dai et al. proposed a novel deep localization method, which allows coarse-to-fine feature encoding strategy for capturing the global and local structures in fundus images, to simultaneously model two-task learning problem of the OD and fovea localization. They took advantage of prior knowledge such as the number of landmarks and their geometric relationship to reliably detect the OD and fovea. Specifically, they first designed a global CNN encoder (with a backbone network of ResNet-50 (He et al., 2016)) to localize the OD and fovea centers as a whole by solving a regression task. All max-pooling layers were replaced with average pooling layers as compared to original ResNet architecture, due to the fact that max-pooling could lose some useful pixel-level information for regression to predict the coordinates. This step was used to simultaneously perform the two detection tasks, because of the geometric relationship between OD and fovea, the performance of multi-task learning is better than a single task. The predicted output coordinates of this global CNN encoder component were used for detecting the bounding boxes of the target OD and fovea. Then the current center coordinates are refined through a local encoder (with a backbone network of VGG-16 (Simonyan and Zisserman, 2014)) which only localizes the OD center or fovea center of their related bounding boxes. During the training stage, they designed an effective data augmentation scheme to solve the problem of insufficient training data. In particular, to build the training set of a local encoder, bounding boxes were randomly selected

based on the ground truth, for each object several bounding boxes of different positions and scales were cropped. The local encoder can be reused multiple times to approximate the target coordinates. The local encoder was iterated twice for refining centers comprehensively. All three models were initialized from the pretrained ImageNet network and replaced the network's last FC layer and Softmax layer by the center coordinates regressor. The regression loss for the central location was the Euclidean loss. The modified loss function for global and local encoders was  $0.045(L_{od} + L_{fovea})$  and  $0.045(L_{od}/L_{fovea})$  respectively. Where  $L_{od}$  and  $L_{fovea}$  are losses for OD and fovea, and scaling factor was introduced since the original Euclidean distance is too large in practice to converge. The proposed learning model was implemented in Caffe framework and trained using SGD with momentum. The FC layers for center regression were initialized from zero-mean Gaussian distributions with standard deviations 0.01 and 0.001. Biases were initialized to

0. The global encoder was trained for 200 epochs, local encoders (OD and fovea both) for 30 epochs respectively. The batch size for global encoder was 16, and 64 for the other two local encoders. The learning rate was set as 0.01 and was divided by 10 when the error plateaus.

#### 5.3.2. VRT (Jaemin Son et al.)

Son et al. proposed an OD segmentation model consisting of U-Net (Ronneberger et al., 2015) and CNN that takes a vessel image and outputs  $20 \times 20$  activation map whose penultimate layer is concatenated to the bottleneck layer of U-Net. Initially, original images were cropped ( $3500 \times 2848$  pixels), padded ( $3500 \times 3500$  pixels) and then resized ( $640 \times 640$  pixels). Each image was standardized with its mean and standard deviation (SD). When calculating the mean and SD, values less than 10 (usual artifacts in the black background) are ignored. Vessel images were prepared with an external network (Son et al., 2017). Pixel values in a vessel image range from 0 to 1. It uses external datasets DRIONSDB (Carmona et al., 2008) and DRIVE (van Ginneken et al., 2004) available with OD and vessel ground truths respectively. For augmentation, the fundus images were affine-transformed and additionally OD was cropped and randomly placed on the image for a random number of times (0 to 5). This augmentation was done to prevent the network from segmenting OD solely by brightness. Pairs of a fundus image and vessel segmentation were provided as input and OD segmentations in the resolution of  $640 \times 640$  and  $20 \times 20$  pixels are given as the ground truth. Binary crossentropy is used as a loss function for both U-Net and vessel network with the loss of  $L_{total} = L_{U-Net} + 0.1 \cdot L_{vessel}$ . Total 800 epochs were trained via Adam optimizer and decreasing learning rate with hyper-parameters of  $\beta_1 = 0.5$ ,  $\beta_2 = 0.999$ . The learning rate was  $2 \times 10^{-4}$  until 400 epochs and  $2 \times 10^{-5}$  until the end. Weights and biases were initialized with Glorot initialization method (Glorot and Bengio, 2010).

They also proposed a four branch model in which two branches were dedicated to the prediction of locations for OD and fovea from vessels (vessel branches) and the other two branches aim to predict the locations from both fundus and vessels (main branches). Similar to OD segmentation, penultimate layers of vessel branches were depth-concatenated to the main branches. After deriving an activation map that represents the probability of containing an anatomical landmark, a hard-coded matrix was multiplied to yield co-ordinates. Original images were cropped as in the segmentation task and standardized with an identical method and later augmented by flip and rotation to ease implementation efforts. Mean absolute error was used as loss function for both outputs with the loss of  $L_{total} = L_{main} + 0.3 \cdot L_{vessel}$ . SGD was used with Nesterov momentum of 0.9 as an optimizer. Learning rate was set to  $10^{-3}$  from 1 to 500 epochs and  $10^{-4}$  from 501 to 1000 epochs. All implementation was done in Keras 2.0.8

with TensorFlow backend 1.4.0 using a server with 8 TITAN X (pascal). Source code is available at [https://bitbucket.org/porwal/pascal-isbi\\_2018\\_fundus\\_challenge](https://bitbucket.org/porwal/pascal-isbi_2018_fundus_challenge).

### 5.3.3. ZJU-BII-SGEX (Xingzheng Lyu et al.)

Lyu et al. utilized Mask R-CNN ([He et al., 2017](#)) to localize and segment OD and fovea simultaneously. It scans the image and generates region proposals by 2D bounding boxes. Then the proposals were classified into different classes and compute a binary mask for each object. They firstly preprocessed the original retinal image into fixed dimensions as network input. A feature extractor (ResNet-50) with feature pyramid networks (FPN) generates feature maps at different scales, which could be used for regions of interest (ROI) extraction. Then a region proposal network (RPN) scans over the feature maps and locates regions that contain objects. Finally, a ROI head network (RHN) is employed to obtain the label, mask, and refined bounding box for each ROI. They also incorporated prior knowledge of retinal image as a post-processing step to improve the model performance. They used IDRiD dataset and two subsets in RIGA dataset ([Almazroa et al., 2018](#)) (Messidor and BinRushed, 605 images) with OD mask provided. They applied the transfer learning technique to train the model. They firstly trained the RHN network by freezing all the layers of FPN and RPN networks and then fine-tuned all layers. The model was implemented in TensorFlow 1.3 and Python 3.4 (source code was modified from [Abdulla \(2017\)](#)). The learning rate started from 0.001 and a momentum of 0.9 was used. The network was trained on one GPU (Tesla K80) with 20 epochs.

### 5.3.4. IITkgpKliv (Oindrila Saha et al.)

Saha et al. used SegNet ([Badrinarayanan et al., 2015](#)) for segmentation of lesions and OD. OD was added as an additional class in the same problem as lesion segmentation so that the model could better differentiate EXs and OD which have similar brightness levels. However, in contrast to original SegNet, the final decoder output is fed to a sigmoid layer to produce class probabilities for each pixel independently in 7 channels. Each channel has the same size as input image: 536 × 356 pixels and consists of activations in the range [0, 1] where 0 corresponds to background and 1 to the presence of a corresponding class. Apart from 5 classes

i.e. MA, HE, SE, EX and OD, two additional classes: (i) retinal disk excluding the lesions and OD, and (ii) black background form the 7 channels. Images were downsampled to 536 × 356 pixels, preserving the aspect ratio. Additionally, Drishti-GS ([Sivaswamy et al., 2014](#)) dataset was used for data augmentation to account for the case of absence of lesions. Further, horizontal, vertical and 180°

flipped versions of the original images were taken. The network was trained using binary cross-entropy loss function and Adam optimizer with learning rate 10<sup>-3</sup> and  $\beta = 0.9$ . Early stopping of the training based on the validation loss is adopted to prevent overfitting. It was observed that the validation loss started to increase after 200 epochs. One more softmax layer is introduced after the Sigmoid layer for normalizing the value of a pixel for each class across channels. The segmented output is finally upsampled for each class to 4288 × 2848 pixels. All implementations were done in PyTorch using 2x Intel Xeon E5 2620 v3 processor with GTX TITAN X GPU 12 GB RAM and 64 GB System RAM.

### 5.3.5. SDNU (Xiaodan Sui et al.)

Sui et al. used Mask R-CNN ([He et al., 2017](#)) for solving all tasks in this sub-challenge. Mask R-CNN could realize accurate target detection based on proposed candidate object bounding boxes of RPN to achieve the objective of OD and Fovea localization. At the same time, it could also get the OD segment at the mask predicting branch. The head architecture of Mask R-CNN (ResNet101 as a backbone) consists of three parallel branches for clas-

sification, bounding box regression, and predicting mask. By this method, the localization of OD and fovea, and segmentation of OD could be achieved directly. They retrained the network to get the new weight parameter of the framework. During the training phase, the dataset of this challenge was augmented by flipping, resizing and trained by 10-fold cross-validation. After training 2000 epochs, the last trained model is obtained. They implemented this algorithm in TensorFlow and it is processed on 8 NVIDIA TITAN Xp GPUs. The experiment environment is built under Ubuntu

16.06.

### 5.3.6. CBER (Ana Mendonça et al.)

Mendonça et al. proposed hand-crafted features based approach for the localization and segmentation tasks in this sub-challenge. Distinct methodologies have been developed for detecting and segmenting these structures, mainly based on color and vascular information. The methodology proposed in the context of this challenge includes three inter-dependent modules. Each module performs a single task: OD localization, OD segmentation or fovea localization. While the modules responsible for the OD localization and segmentation were an improved version of two methods previously published ([Mendonça et al., 2013; Dashtbozorg et al., 2015](#)), the method proposed for fovea localization was completely new. Initially, the module associated with the OD localization receives a fundus image and segments the retinal vasculature. Afterward, the entropy of the vessel directions is computed and combined with the image intensities in order to find the OD center coordinates. For OD segmentation, the module responsible for this task uses the position of the OD center for defining the region where the sliding band filter ([Pereira et al., 2007; Esteves et al., 2012](#)) is applied. The positions of the support points which give rise to the maximum filter response were found and used for delineating the OD boundary. Since a relation between the fovea-OD distance and the OD diameter was known ([Jonas et al., 2015](#)), the module responsible for the fovea localization begins by defining a search region from the OD position and diameter. The fovea center is then assigned to the darkest point inside that region.

## 6. Evaluation measures

The performance of each sub-challenge was evaluated based on different evaluation metrics. Following evaluation measures were used for different sub-challenges:

### 6.1. Sub-challenge – 1

In this sub-challenge, the performance of algorithms for lesion segmentation tasks was evaluated using submitted grayscale images and available binary masks. As in the lesion segmentation task(s) background overwhelms foreground, a highly imbalanced scenario, the performance of this task was measured using area under precision (a.k.a. Positive Predictive Value (PPV)) recall (a.k.a.

Sensitivity (SN)) curve (AUPR) ([Saito and Rehmsmeier, 2015](#)).

$$SN = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (10)$$

$$PPV = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \quad (11)$$

The curve was obtained by thresholding the results at 33 equally spaced instances i.e. [0, 8, 16, ..., 256] in gray levels or [0, 0.03125, 0.0625, ..., 1] in probabilities. The AUPR provides a single-figure measure (a.k.a. mean average precision (mAP)), computed over the Set-B, was used to rank the participating methods. This performance metric was used for object detection in The PASCAL Visual Object Classes (VOC) Challenge ([Everingham et al., 2010](#)).

**The AUPR measure is more realistic ( Boyd et al., 2013; Saito and Rehmsmeier, 2015 ) for the lesion segmentation performance over the area under Receiver Operating Characteristic (ROC) curve.**

### 6.2. Sub-challenge - 2

Let the expert labels for DR and DME be represented by  $DR_{gt}(n)$  and  $DME_{gt}(n)$ . Whereas  $DR_{gt}(n)$  and  $DME_{gt}(n)$  are the predicted results, then correct instance is the case when the expert label for DR and DME matches with the predicted outcomes for both DR and DME. This was done since, even with the presence of some exudation that may be categorized as mild DR, its location on the retina is also an important governing factor (to check DME) to decide the overall grade of disease. For instance, EXs presence in the macular region can affect the vision of patient to a greater extent and hence, it should be dealt with priority for referral (that may otherwise be missed or cause a delay in treatment with the present convention of only DR grading) in the automated screening systems. Hence, disease grading performance accuracy for this sub-challenge, from the results submitted in CSV format for test images (i.e.  $N = 103$ ), is obtained by **algorithm 1** as follows:

---

#### Algorithm 1: Computation of disease grading accuracy.

---

```

Data: Method Results and Labels with DR and DME Grading
Result: Average disease grading accuracy for DR and DME

1 for  $n = 1, 2, \dots, N$  do
2   | Correct = 0;
3   | If ( $DR_{gt}(n) == DR_{pred}(n)$ ) and ( $DME_{gt}(n) == DME_{pred}(n)$ ) then
4     |   | Correct = Correct + 1;
5   | end
6 end

7 Average Accuracy =  $\frac{\text{Correct}}{N}$ 

```

---

### 6.3. Sub-challenge - 3

For the given retinal image, the objective of sub-challenge – 3 (task - 6 and 7) was to predict the OD and fovea center coordinates. The performance of results submitted in CSV format was evaluated by computing the Euclidean distance (in pixels) between manual (ground truth) and automatically predicted center location. Lower Euclidean distance indicates better localization. After determining these distances for each image in the Set-B, i.e. for 103 images, the average distance representing the whole dataset was computed and used to rank the participating methods.

The optic disc segmentation (task - 8) performance is evaluated using Jaccard index (  $J$  ) ( Jaccard, 1908 ). It represents the proportion of overlapping area between the segmented OD (  $O$  ) and the ground truth (  $G$  ).

$$J = \frac{|O \cap G|}{|O \cup G|} \quad (12)$$

Higher  $J$  indicates better segmentation. For the segmented results, images in range [0, 255], it was computed at 10 different equally spaced thresholds [0, 0.1,

, 0.9] and averaged to obtain final

score.

### 7. Results

This section reports and discusses the results of all subchallenges. Performance of all competing solutions on the Set-B for all eight subtasks are divided into three sub-challenge categories and discussed including their leaderboard rank.

### 7.1. Sub-challenge – 1

In this section, we present the performance of all competing solutions for the lesion segmentation task. All results received from the participating teams were analyzed using the validation measure given in **Section 6.1**. This measure generated a set of precision-recall curves for each of the different techniques. Out of the total 37 teams that participated in the challenge, 22 teams participated (a complete list is available on the challenge website) in the sub-challenge-1 whose results were evaluated and ranked using the AUPR values. Amongst them, 7 teams (see **Table 5**) having performance within top 4 positions in either of lesion segmentation task were invited for the challenge workshop and 3 teams having overall better performance, i.e. solutions developed by the teams that ranked amongst the top three for at least three different lesion segmentation tasks, presented their work at ISBI.

**Table 8** summarizes the individual performance (Off-site evaluation) of each solution listed in order of their final placement for each subtask. It also contains various approaches followed and external dataset (if any) used for training the models. A higher rank indicates more favorable performance for the individual task(s). The top-3 entries according to the individual lesion segmentation task are VRT, iFLYTEK-MIG and PATech. Some sample lesion segmentation results illustrated in **Fig. 6** and their corresponding overall evaluation score from **Table 8** gives a better idea of how the evaluation scores correlate with the quality of segmentation. **Fig. 7**

summarizes the performance of top-4 teams per lesion segmentation task. The different curves represent the performance of the participating methods for various lesions (MAs, HEs, SEs and EXs). Team VRT achieved highest AUPR score for HE and SE segmentation task. Whereas, team PATech and iFLYTEK-MIG obtained best score for EX and MA segmentation task respectively.

### 7.2. Sub-challenge – 2

This section presents the results achieved (On-site evaluation) by participating teams for the DR and DME grading task. It is important to note that this task was evaluated for simultaneous grading of DR and DME using the validation algorithm outlined in

**Section 6.2** on the Set-B. This algorithm produced an average grading accuracy of joint DR and DME on all images. **Table 9** summarizes the result of teams for the on-site challenge along with the approach followed and external dataset used for training the model by respective team.

The top-performing solution at the “on-site” challenge was proposed by team LzyUNCC followed by team VRT and team Mammoth. **Fig. 8** shows the average accuracy of competing solutions for the individual as well as simultaneous grading of DR and DME. Teams are observed to perform poorly in the DR grading task that reduced the overall accuracy for simultaneous grading of DR and DME. Major reason seems to be the difficult test set, difficulty in accurately discriminating the DR severity grades.

### 7.3. Sub-challenge – 3

This section presents an evaluation of “On-site” results for participating teams in the sub-challenge – 3, for all three subtasks. The results for subtasks of OD and Fovea center localization were evaluated by computing Euclidean distance, whereas OD segmentation results were evaluated and ranked using Jaccard similarity score as outlined in **Section 6.3**. Results from the on-site evaluations are reported in **Table 10** and **Table 11** that summarises the performance of all participating algorithms for all three subtasks.

The winning methods for localization tasks were developed by team DeepDR and team VRT, with DeepDR performing best in both

**Table 8**

Sub-challenge – 1 “Off-site” leaderboard highlighting top 4 teams from each lesion (MAs, HEs, SEs and EXs) segmentation task on the testing dataset. It details the approach followed by respective team and external dataset used for training their model (if any).

Lesion	Team name	AUPR	Approach	Ensemble	Input Size (Pixels)	External dataset
Microaneurys	iFLYTEK	0.5017	Cascaded CNN		<b>320 × 320</b>	x
	VRT	0.4951	U-Net	x	<b>1280 × 1280</b>	x
	PATech	0.4740	DenseNet+U-Net		<b>256 × 256</b>	x
	SDNU	0.4111	Mask R-CNN	x	<b>3584 × 2380</b>	x
Hemorrhages	VRT	0.6804	U-Net	x	<b>640 × 640</b>	x
	PATech	0.6490	DenseNet+U-Net		<b>256 × 256</b>	x
	iFLYTEK	0.5588	Cascaded CNN		<b>320 × 320</b>	x
	SOONER	0.5395	U-Net	x	<b>380 × 380</b>	x
Soft Exudates	VRT	0.6995	U-Net	x	<b>640 × 640</b>	x
	LzyUNCC-I	0.6607	FCN+DLA	x	<b>1024 × 1024</b>	E-ophtha
	iFLYTEK	0.6588	Cascaded CNN		<b>320 × 320</b>	x
	LzyUNCC-II	0.6259	FCN+DLA	x	<b>1024 × 1024</b>	E-ophtha
Hard Exudates	PATech	0.8850	DenseNet+U-Net		<b>256 × 256</b>	x
	iFLYTEK	0.8741	Cascaded CNN		<b>320 × 320</b>	x
	SAIHST	0.8582	U-Net	x	<b>512 × 512</b>	x
	LzyUNCC-I	0.8202	FCN+DLA	x	<b>1024 × 1024</b>	E-ophtha

**Table 9**

Sub-challenge – 2 “On-site” leaderboard highlighting performance of top 6 teams for DR and DME grading on the test dataset. It details the approach followed by respective team and external dataset used for training their model.

Team Name	Accuracy	Approach	Ensemble	Input Size (Pixels)	External Dataset
LzyUNCC	0.6311	Resnet +DLA	5	<b>896 × 896</b>	Kaggle
VRT	0.5534	CNN	10	<b>640 × 640</b>	Kaggle, Messidor
Mammoth	0.5146	DenseNet		<b>512 × 512</b>	Kaggle
HarangiM1	0.4757	AlexNet +GoogLeNet	2	<b>224 × 224</b>	Kaggle
AVSASVA	0.4757	ResNet +DenseNet	DR-8, DME-5	<b>224 × 224</b>	DiaretDB1
HarangiM2	0.4078	AlexNet +Handcrafted features	2	<b>224 × 224</b>	Kaggle

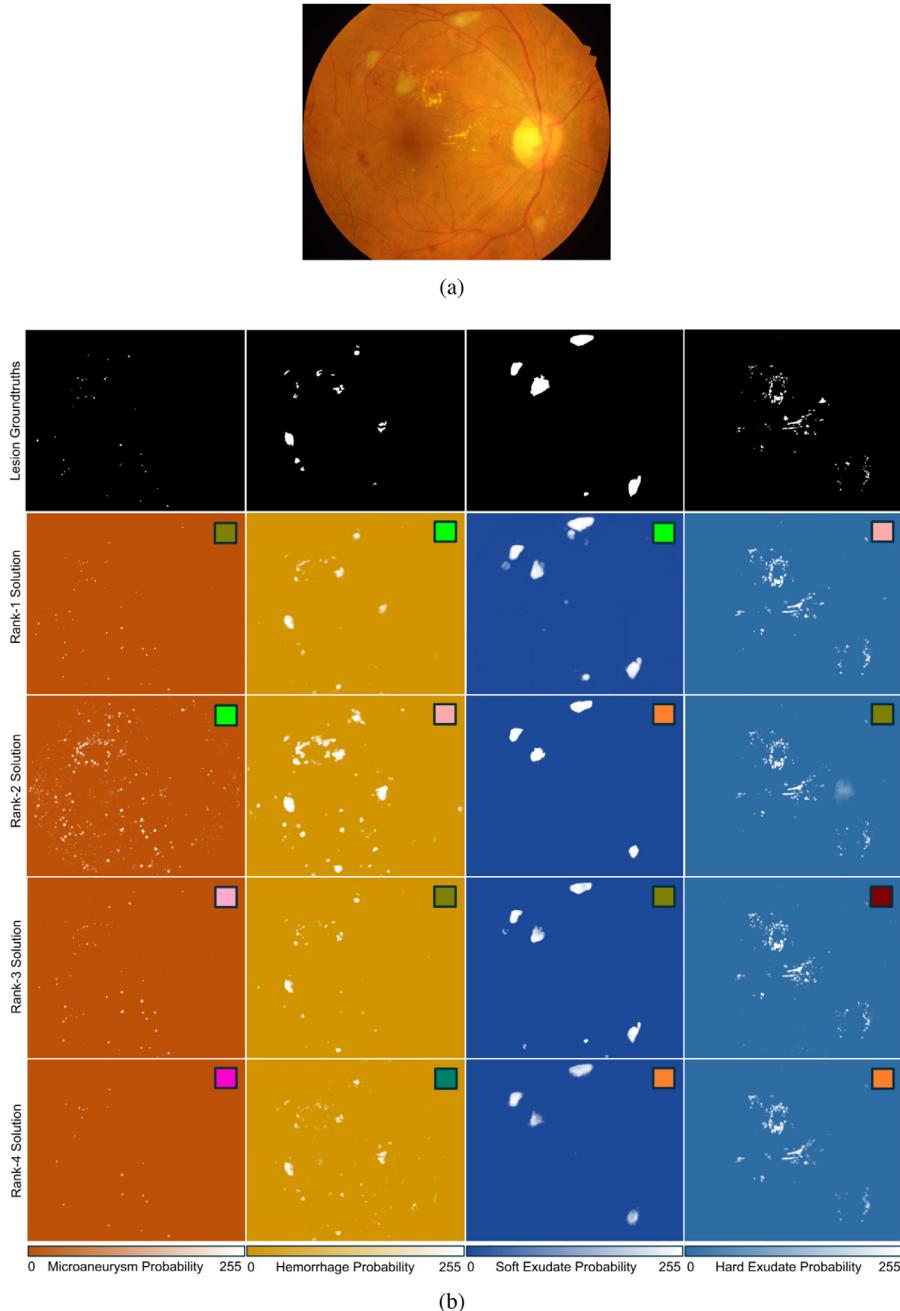
**Table 10**

“On-site” leaderboard highlighting performance of top 5 teams in OD and fovea localization task on the test dataset. It highlights the approach followed by respective team and external dataset used for training their model (if any). ED: Euclidean distance.

Localize	Team Name	ED (Pixels)	Rank	Approach	Input Size (Pixels)	External Dataset
Optic Disc	DeepDR	21.072	1	ResNet +VGG	<b>224 × 224, 950 × 950</b>	–
	VRT	33.538	2	U-Net	<b>640 × 640</b>	DRIVE
	ZJU-BII-SGEX	33.875	3	Mask R-CNN	<b>1024 × 1024</b>	RIGA
	SDNU	36.220	4	Mask R-CNN	<b>1984 × 1318</b>	–
	CBER	29.183	–	Handcrafted Features	<b>536 × 356</b>	–
Fovea	DeepDR	64.492	1	ResNet +VGG	<b>224 × 224, 950 × 950</b>	–
	VRT	68.466	2	U-Net	<b>640 × 640</b>	DRIVE
	SDNU	85.400	3	Mask R-CNN	<b>1984 × 1318</b>	–
	ZJU-BII-SGEX	570.133	4	Mask R-CNN	<b>1024 × 1024</b>	RIGA
	CBER	59.751	–	Handcrafted Features	<b>536 × 356</b>	–

OD and Fovea detection tasks. But the winning entries for OD segmentation task were from teams ZJU-BII-SGEX, VRT and IITKgpKLIV. Some sample OD segmentation results from these teams are illustrated in Fig. 9.

Fig. 10 shows box-plots illustrating the range of Euclidean distances from the center of (a) OD and (b) fovea as well as (c) spread of Jaccard index for OD segmentation.

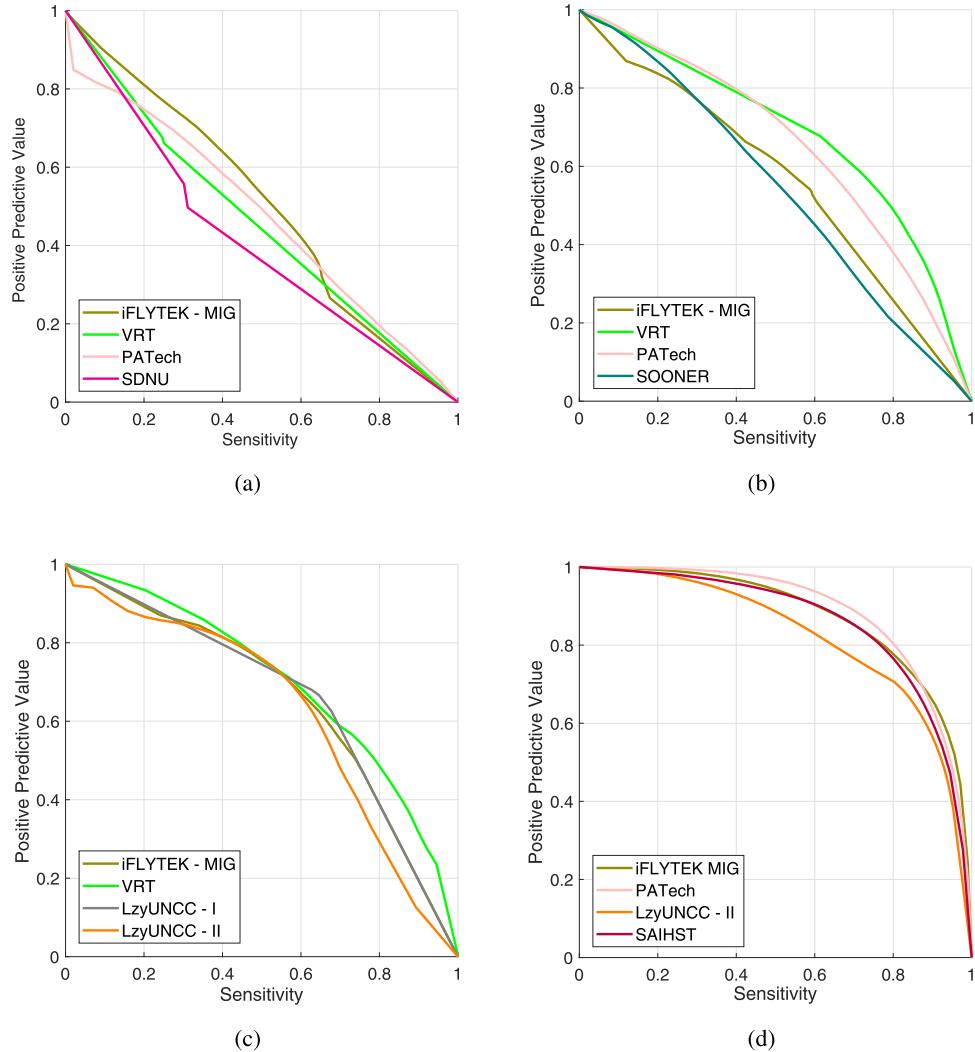


**Fig. 6.** Illustration of lesion segmentation results: (a) sample image and (b) segmentation outcome of top-4 teams (from left to right) (i) MAs, (ii) HEs, (iii) SEs, and (iv) EXs in retinal fundus images. Top row corresponds to ground truths, second row to entry from top performing team, similarly, third, fourth and fifth rows correspond to entries from other three teams respectively. The lesion segmentation entries are colored for better illustration and separation from each type of lesion.

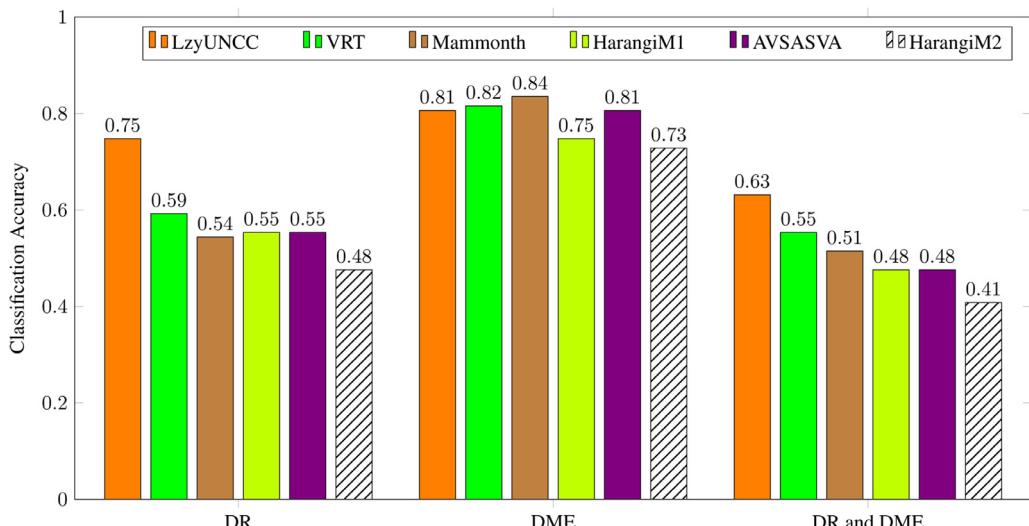
**Table 11**

“On-site” leaderboard highlighting performance of top 5 teams in OD segmentation task on the test dataset. It details the approach followed by respective team and external dataset used for training their model (if any). J: Jaccard index.

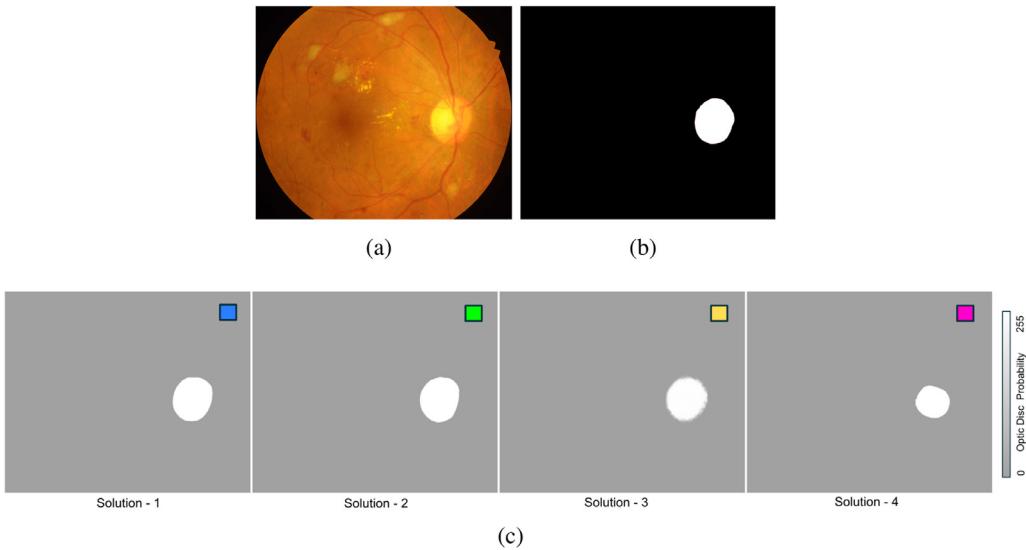
Team name	J	Rank	Approach	Input size (Pixels)	External dataset
ZJU-BII-SGEX	0.9338	1	Mask R-CNN	1024 × 1024	RIGA
VRT	0.9305	2	U-Net	640 × 640	DRIVE, DRIONS-DB
IITKgpKLIV	0.8572	3	SegNet	536 × 356	Drishti-GS
SDNU	0.7892	4	Mask R-CNN	1984 × 1318	–
CBER	0.8912	–	Handcrafted Features	536 × 356	–



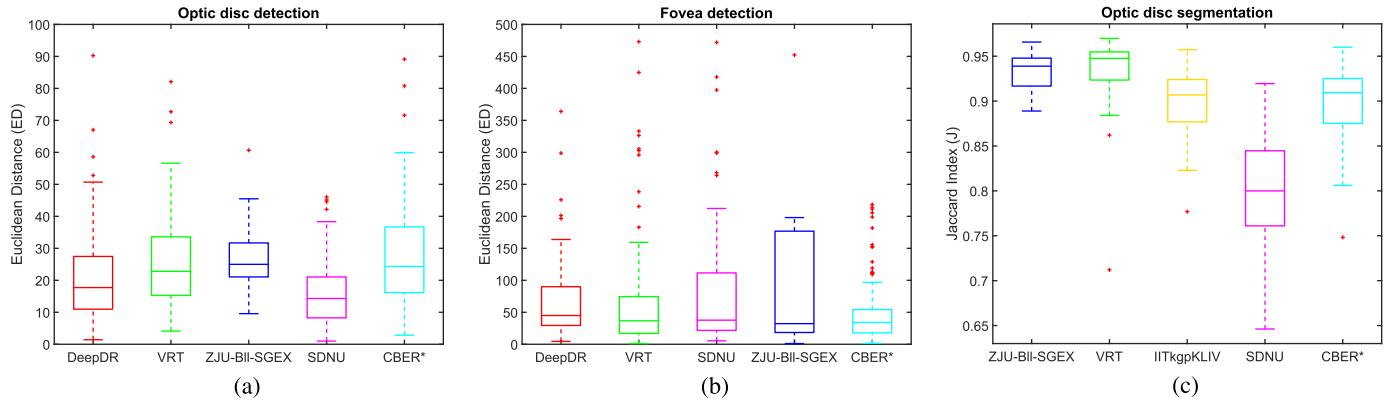
**Fig. 7.** The AUPR curves for the four top performing individual methods on the test dataset. These curves plot the sensitivity versus the positive predictive values for the different lesions, namely, (a) MAs, (b) HEs, (c) SEs, and (d) EXs.



**Fig. 8.** Barplots showing separate and simultaneous classification accuracy of solutions developed by top - 6 teams for grading of DR and DME.



**Fig. 9.** Illustration of OD segmentation results: (a) sample image. (b) OD ground truth and (c) segmentation outcome of top-4 teams (from left to right).



**Fig. 10.** Boxplots (a,b) showing dispersion of Euclidean distance for individual methods for OD and fovea and (c) showing the dispersion of Jaccard index for OD segmentation task. Boxplots show quartile ranges of the scores on the test dataset; plus sign indicate outliers (full range of data is not shown).

## 8. Discussion and conclusion

In this paper, we have presented the details of IDRiD challenge including information about the data, evaluation metrics, an organization of the challenge, competing solutions and final results for all sub-tasks, i.e., lesion segmentation, disease grading and localization and segmentation of other normal retinal structures. Given the significant number of participating teams (37) and results obtained, we believe this challenge was a success. To the organizational end, efforts have been made in creating a relevant, stimulating and fair competition, capable of advancing collective knowledge in the research community. This section presents a discussion, limitations, and lessons learned from this challenge.

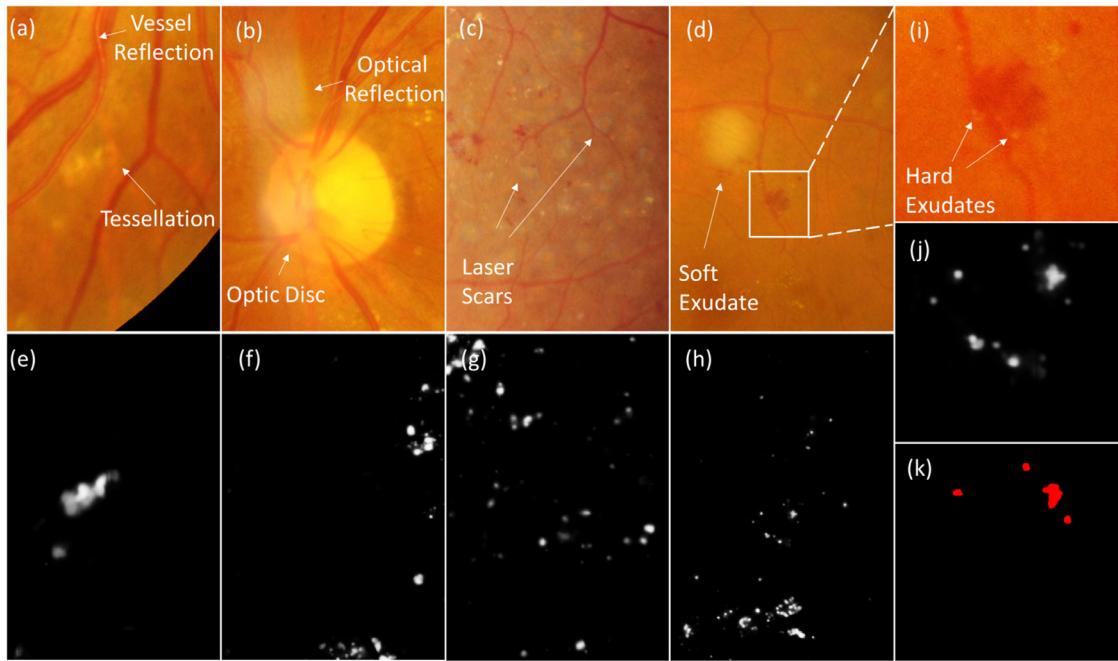
dition of dense block to extract features efficiently, boosting results significantly. Fig. 11 highlights the performance of top solution for EX that performs significantly well in the presence of normal retinal structures and different challenging circumstances.

From the top-performing approaches, it is evident that solving the data imbalance problem improves the model performance significantly. Since background overwhelms foreground i.e. there are more background pixels than lesion pixels (see Fig. 6), the loss during training is more effectively back-propagated than that of the foreground that penalizes false negatives, boosting the sensitivity of lesion segmentation. In general, the architectural modifications to U-Net-based networks provided widely varying results for the different types of lesion.

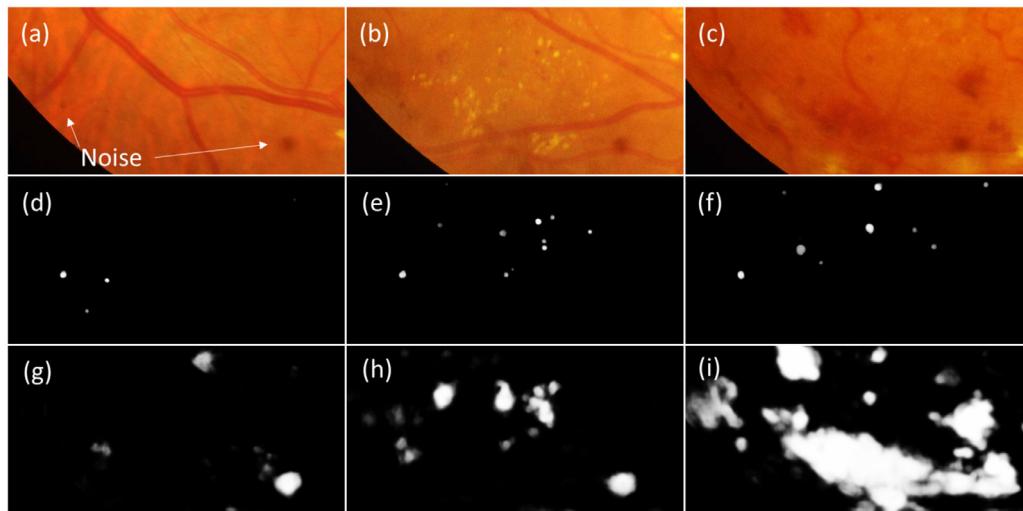
The first sub-challenge was conducted in an off-site mode in which 22 teams participated with their lesion segmentation methods. The results of these methods on the Set-B were evaluated by the organizers and amongst them, top-4 performing methods per lesion segmentation task are included in this paper. The computed AUPR values ranged between 0.4111 (for MAs) and 0.885 (for EXs). When the performance of top solutions was analyzed by computing the area under ROC curve (AUC) at the pixel level, in threshold range [0:0.01:1], it resulted in AUC of 0.8263, 0.9716, 0.9540 and

0.9883 for MA, HE, SE and EX respectively. The best approach for lesion segmentation used U-Net, with data augmentation and ad-

In the on-site disease-grading task, six methods were compared and contrasted. When assessed using the test data set hidden from the participants, the grading accuracy ranged between 0.4078 and



**Fig. 11.** Illustration of (a-d) different challenging circumstances for segmentation of EXs, (e-h) segmentation results (probability map) of the top-performing team for EXs, (i) enlarged part of Fig. (d), and (j) depicts its performance to be better than (k) the human annotator (The annotator tool had a limitation of the markup capability when there is an overlap of multiple types of lesion. In this case, EXs and HE).



**Fig. 12.** Illustration of results by top performing solutions for (a-c) different images with noise causing most common false positives in the segmentation of (d-f) MAs, and (g-i) HEs respectively.

0.6311 as shown in Table 9. Notably, all teams except AVASAVA used the external Kaggle DR dataset for pre-training their models. This dataset contains a large number of retina images annotated with the disease level, in contrast, team AVASAVA pre-trained their model on ImageNet, a dataset containing natural images and object annotations, effectively showing the network a much smaller number of retina images at the training stage, approximately 1% compared to the other teams. This indicates that in the presence of a limited number of labeled data, transfer learning approaches along with the good model pruning could yield comparable and competitive results. However, while the models do determine the variability of performance, the number, type, and quality of training data is a crucial factor for a fair comparison of competing solutions. There is still work needed on simultaneous grading of DR

and DME as the reported results do not yet reach the performance needed for a clinically viable automatic screening. Considering the misclassified instances in confusion matrices shown in Table 12, along with the lesion information, it is essential to give attention towards characterization of intra-retinal microvascular abnormalities (IRMA's) and venous beading for improvement in the overall grading results.

In the sub-challenge – 3, another on-site challenge, four teams were evaluated for the task of OD/fovea localization and OD segmentation. For the task of OD localization, the Euclidean distance varied between 21.072 and 36.22 (lower values indicate better performance). However, for Fovea localization task the same performance metric ranged between 64.492 and 570.133. This massive variation is due to outliers, e.g. team ZJU-BII-SGEX had 23 outliers

**Table 12**

Confusion matrix of retinal images predicted by top performing solution for DR (5 class) and DME (3 class).

		Predicted							Predicted		
		0	1	2	3	4	Actual	0	1	2	3
Actual	0	30	0	2	1	1	Actual	0	40	2	3
	1	3	1	1	0	0		1	5	2	3
	2	3	2	22	4	1		2	5	2	41
	3	2	0	1	13	3					
	4	1	0	1	0	11					

whose Euclidean distance exceeded 700. In the OD segmentation task, the average Jaccard similarity index score amongst the participants ranged between 0.7892 and 0.9338. The top-performing solutions developed by DeepDR and VRT leveraged prior clinical knowledge, such as the number of landmarks and their geometric relationship to detect another retinal landmark. It is also observed that data augmentation and ensemble of models yield substantial improvements in terms of accuracy. Considering the clinical significance of OD diameter while DME severity grading, we further compute the average OD diameter (in pixels) for each image of the test set. The average diameter of OD ground truth is 516.61 pixels, whereas, corresponding values for the results of solutions developed by the teams ZJU-BII-SGEX, VRT, IITKgpKLIV, CBER and SDNU are 514.25, 519.21, 513.48, 508.04 and 460.19 pixels respectively. Team CBER submitted their results after the competition and they were not included in the leaderboard.

judicated consensus grades yield better performance when compared to datasets consisting of poor-quality (non-gradable) images and images captured in varied acquisition settings. Therefore, this challenge provides data collected in the routine clinical practice using an acquisition protocol consistent for all images. The data was acquired after pupil dilation with the same camera at the same resolution, ensuring consistent quality. This dataset did not include non-gradable images and images with substantial disagreement amongst the expert annotators. Even after these efforts to provide the best possible data, the annotation process is still inherently subjective, and the annotator judgment is a limiting factor for the method performance which is mostly trained and evaluated in a supervised manner. We also note that images captured with different retinal cameras or with different diseases would have allowed for a better estimation of the generalization ability of the proposed methods since they might be more representative from clinical settings. Further, while we believe that data challenges like ours foster "methodology diversity", the majority of competing solutions used deep convolutional networks. These approaches are comparably easier to implement than approaches based on feature engineering and do generalize well to multiple medical imaging domains, which in turn, dramatically reduces the need for specialized task knowledge. Notably, amongst the competing solutions in this challenge that utilized the deep learning approach along with the task-relevant subject knowledge have demonstrated superior performance. However, it seems there might be some impact of challenge duration, apart from the number of submissions, on the quality of developed solutions. Considering the time span from data availability to deadline of results submission, about one and a half month, was considerably tight for managing all tasks at the same time. For the team VRT who had been working on analyzing fundus images for more than a year when participated in the competition that attempting all tasks were possible, still, it was challenging for them to commit all the tasks. However, it would be highly challenging for a newcomer to succeed in multiple tasks. In that sense, the competition period was not sufficient for perfecting all tasks. However, it would be enough for a competent participant,

As expected, we found that image resolution is a vital factor for the model performance, especially for the task of segmentation of small objects such as MAs or EXs. In fact, the top-performing approaches processed the images patch-wise, which allow models to have a local high-resolution image view or directly with the high-resolution image as a whole. This is essential as MAs or small EXs lesions span very few pixels in some cases, and reducing the original image size would prevent an accurate segmentation. Similarly, image resolution plays a very important role for disease classification task (see Table 9), the most likely reason is that presence of the disease is determined by the presence of lesions in the image, including the small ones that might be invisible at low resolution. This is corroborated by the confusion matrices in Table 12 which show misclassified instances in DR (particularly, grade 1 and 2) as well as DME (5 images each belonging to grade 1 and 2 are predicted as grade 0). For the localization tasks, all participants were asked to identify retinal structures with coordinates at full image resolution. Most of them performed these tasks by scaling image to the smaller size and then converted their predictions in the original image space. Comparative analysis indicates that the input image resolution has limited effect on the results of the localization problem. For instance, in the case of OD localization, the top-performing team utilized two image resolutions, one (224 × 224 pixels) for approximate location prediction and other (cropped ROIs 950 × 950 pixels) for refining that estimate. Similarly, teams CBER and VRT resized the image to 536 × 356 pixels and 640 × 640 pixels respectively to get an approximate center location whereas the team SDNU utilized the input size of 1984 × 1318 pixels. Considering the OD average diameter of approximately 516 pixels, limited performance variation (10 to 15 pixels) is observed as compared to the top-performing solution for huge variation (multiple times) in input resolutions (see

e.g. new entrants in the field as team SAIHST, to finish one task if the participant can focus on the competition completely. Also, in this challenge, the results were evaluated all at once after the result submission deadline. However, a continuous on-line assessment of participating solutions would have facilitated the submission procedure by providing real-time feedback to the teams performance. This would have enabled a maximum number of submissions during the challenge period, probably boosting the final count of submissions. However, this would have introduced a risk of overfitting the test data by continuous submissions based on the system's performance on the test set.

This challenge led to the development of a variety of new robust solutions for lesion segmentation, detection, and segmentation of retinal landmarks and disease severity grading. Despite the complexity of the tasks, less than one-and-a-half month time for development, it received a very positive response, and the topperforming solutions were able to achieve results close to the human annotators. Still, there is room for improvement, especially in the lesion segmentation and disease-grading tasks. Though the competition is now completed, the dataset has been made publicly available for research purposes to attract newcomers to the problem and to encourage the development of novel solutions to meet current and future clinical standards.

Table 10). This is because the retinal structures to be identified, OD and fovea, are very unlikely to disappear due to a reduction of image resolution and they have clear geometrical constraints.

As confirmed by recent studies (Krause et al., 2018; Son et al., 2019), we hypothesized that algorithms developed using images with fine visibility and images having high resolution with ad-

#### Declaration of Competing Interest

The authors have no conflicts of interest to declare.

**Table A.1**

Summary of technical specifications and hardware used in different databases.

Name of Database	Number of Images	Technical Details					NMY Format
		Image Size(s)	FOV	Camera			
ARIA	212	768 × 576	50	Zeiss FF450+			TIFF
DIARETDB	130+89	1500 × 1152	50	Zeiss FF450+			PNG
DRIVE	40	768 × 584	45	Canon CR5			JPEG
E-Ophtha	47EX+35H 148MA+233H 1440 × 960 - 2048 × 1360 (4)		45	Canon CR - DGI & Topcon TRC - NW6			JPEG
HEIMED	169	2196 × 1958	45	Zeiss Visucam PRO			JPEG
Kaggle	88,702	433 × 289 - 3888 × 2592	Varying	Any camera (EyePACS Platform)		-	TIFF
MESSIDOR	800 MY+ 400 NMY+ 1756 1440 × 960, 2240 × 1488, 2304 × 1536 45			3CCD/ Topcon TRC NW6			Both TIFF
ROC	100	768 × 576, 1058 × 1061, 1389 × 1383	45	Topcon NW100 & NW200 Canon CR5 - 45 NM			JPEG
STARE	397	605 × 700	35	Topcon TRV-50		x	PPM
IDRID	516 (81 with LA)	4288 × 2848	50	Kowa VX-10 α			JPG

EX - Hard Exudate, MA - Microaneurysms, H - Healthy, MY - Mydriatic, NMY - Non-Mydriatic, FOV - Field of View, LA - Lesion Annotation.

**Table A.2**

Comparison of different databases with the IDRID database.

Name of database	Normal fundus structures		Abnormalities			Multiple experts	DR grading	DME grading
	OD	VS	FA	MA	HE			
ARIA				x	x	x x		x
DIARETDB1	x	x x					4	x
DRIVE	x		x	x	x x		3	x
E-Ophtha	x	x x		x	x	x	2	x
HEIMED	x	x x		x	x	x	1	x
Kaggle	x	x x		x	x	x x	2	
MESSIDOR	x	x x		x	x	x x x	1	
ROC	x	x x		x	x x		4	x
STARE		x		x	x	x x	2	x
IDRID		x					2	

OD - Optic Disc, VS - Vessels, FA - Fovea, MA - Microaneurysms, HE - Hemorrhage, EX - Hard Exudate, SE - Soft Exudate,

# - Number of Experts

## Acknowledgments

This work is sponsored by the Shri Guru Gobind Singhji Institute of Engineering and Technology, Nanded (M.S.), INDIA. The authors would like to thank the following people for their help in various aspects of organizing the ISBI-2018 Diabetic Retinopathy Segmentation and Grading Challenge: Prof. Emanuele Trucco (University of Dundee, Scotland), Tom MacGillivray (University of Edinburgh, Scotland), Ravi Kamble (SGGS Institute of Engineering and Technology, Nanded), Prof. Vivek Sahasrabuddhe (Government Medical College, Nanded) and Désiré Sidibé (Université de Bourgogne, France). We would also like to thank Prof. Jorge Cuadros, University of California, Berkeley (Organizer of Kaggle Diabetic Retinopathy challenge) for his kind permission for reporting the results of the models trained on their dataset. **VRT:** This study was supported by the Research Grant for Intelligence Information Service Expansion Project, which is funded by. [National IT Industry Promotion Agency \(NIPA-C0202-17-1045\)](#) in South Korea. **DeepDR:**

This work was supported in part by the [National Natural Science Foundation of China](#) under Grant [Grant 61872241](#), [Grant 61572316](#), in part by the [National Key Research and Development Program of China](#) under Grant [2016YFC1300302](#) and [Grant 2017YFE0104000](#), in part by the [Science and Technology Commission of Shanghai Municipality](#) under Grant [16DZ0501100](#) and Grant [17411952600](#).

**HarangiM1-M2:** Research was supported in part by the Janos Bolyai

Research Scholarship of the Hungarian Academy of Sciences and the project

EFOP-3.6.2-16-2017-00015 supported by the European Union and the State of Hungary, co-financed by the European Social Fund. **ZJU-BII-SGEY:** This work is supported by Beijing Shanggong Medical Technology Co., Ltd., which provided ocular healthcare solutions in China. This research is partially supported by the A - STAR A1818g0022 grant of Singapore. Many thanks to the labeled images from Image Annotation Group of Beijing Shanggong Medical Technology. Team **CBER** (A.M. Mendonça, T. Melo, T. Araújo and A. Campilho) is financed by the ERDF European Regional Development Fund through the Operational Programme for Competitiveness and Internationalisation - COMPETE 2020 Programme, and by National Funds through the FCT Fundação para a Ciência e a Tecnologia ([Portuguese Foundation for Science and Technology](#)) within project [CMUP-ERI/TIC/0028/2014](#). **Teresa Araújo** is funded by the FCT grant SFRH/BD/122365/2016. **SDNU:** This study was supported by the National Natural Science Foundation of China (Grant No. 61572300).

## Appendix A. Comparison of Publicly Available Retinal Image Databases

**Table A.1** and **Table A.2** provides the summary of technical specifications and available ground truths in several existing datasets and the IDRID dataset.

## Supplementary material

- Supplementary material associated with this article can be found, in the online version, at doi: 10.1016/j.media.2019.101561.
- ### References
- Abdulla, W., 2017. Mask r-CNN for object detection and instance segmentation on keras and tensorflow. [https://github.com/matterport/Mask\\_RCNN](https://github.com/matterport/Mask_RCNN).
- Abràmoff, M.D., Garvin, M.K., Sonka, M., 2010. Retinal imaging and image analysis. *IEEE Rev. Biomed. Eng.* 3, 169–208.
- Abràmoff, M.D., Lou, Y., Erginay, A., Clarida, W., Amelon, R., Folk, J.C., Niemeijer, M., 2016. Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. *Investigat. Ophthalmol. Vis. Sci.* 57 (13), 5200–5206.
- Acharya, R., Chua, C.K., Ng, E., Yu, W., Chee, C., 2008. Application of higher order spectra for the identification of diabetes retinopathy stages. *J. Med. Syst.* 32 (6), 481–488.
- Acharya, U.R., Mookiah, M.R.K., Koh, J.E.W., Tan, J.H., Bhandary, S.V., Rao, A.K., Hagiwara, Y., Chua, C.K., Laude, A., 2017. Automated diabetic macular edema (DME) grading system using DWT-DCT features and maculopathy index. *Comput. Biol. Med.* 84, 59–68.
- Acharya, U.R., Ng, E.Y.-K., Tan, J.-H., Sree, S.V., Ng, K.-H., 2012. An integrated index for the identification of diabetic retinopathy stages using texture parameters. *J. Med. Syst.* 36 (3), 2011–2020.
- Adal, K.M., Sidibé, D., Ali, S., Chaum, E., Karnowski, T.P., Mériau-deau, F., 2014. Automated detection of microaneurysms using scale-adapted blob analysis and semi-supervised learning. *Comput. Method. Progr. Biomed.* 114 (1), 1–10.
- Agurto, C., Murray, V., Barriga, E., Murillo, S., Pattichis, M., Davis, H., Russell, S., Abràmoff, M., Soliz, P., 2010. Multiscale AM-FM methods for diabetic retinopathy lesion detection. *IEEE Trans. Med. Imag.* 29 (2), 502–512.
- Almazroa, A., Alodhayb, S., Osman, E., Ramadan, E., Hummadi, M., Dlaim, M., Alkatee, M., Raahemifar, K., Lakshminarayanan, V., 2018. Retinal fundus images for glaucoma analysis: the RIGA dataset. In: *Medical Imaging 2018: Imaging Informatics for Healthcare, Research, and Applications*, Vol. 10579. International Society for Optics and Photonics, p. 10579B.
- Antal, B., Hajdu, A., 2012. Improving microaneurysm detection using an optimally selected subset of candidate extractors and preprocessing methods. *Pattern Recognit.* 45 (1), 264–270.
- Antal, B., Hajdu, A., 2014. An ensemble-based system for automatic screening of diabetic retinopathy. *Knowl.-Based Syst.* 60, 20–27.
- Atlas, I.D.F.D., 2017. Brussels, belgium: international diabetes federation. *Int. Diabet. Federat. (IDF)*. <http://diabetesatlas.org/resources/2017-atlas.html>
- Badrinarayanan, V., Kendall, A., Cipolla, R., 2015. Segnet: a deep convolutional encoder-decoder architecture for image segmentation. *arXiv preprint arXiv:1511.00561*.
- Bai, J., Miri, M.S., Liu, Y., Saha, P., Garvin, M., Wu, X., 2014. Graph-based optimal multi-surface segmentation with a star-shaped prior: application to the segmentation of the optic disc and cup. In: *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*. IEEE, pp. 525–528.
- Bandello, F., Parodi, M.B., Lanzetta, P., Loewenstein, A., Massin, P., Menchini, F., Veritti, D., 2010. Diabetic macular edema. In: *Macular Edema*, Vol. 47. Karger Publishers, pp. 73–110.
- Biyani, R.S., Patre, B.M., 2018. Algorithms for red lesion detection in diabetic retinopathy: a review. *Biomed. Pharmacother.* 107, 681–688.
- Bourne, R.R.A., Stevens, G.A., White, R.A., Smith, J.L., Flaxman, S.R., Price, H., Jonas, J.B., Keeffe, J., Leasher, J., Naidoo, K., et al., 2013. Causes of vision loss worldwide, 1990–2010: a systematic analysis. *Lancet Global Health* 1 (6), e339–e349.
- Boyd, K., Eng, K.H., Page, C.D., 2013. Area under the precision-recall curve: point estimates and confidence intervals. In: *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer, pp. 451–466.
- Carmona, E.J., Rincón, M., García-Feijoo, J., Martínez-de-la Casa, J.M., 2008. Identification of the optic nerve head with genetic algorithms. *Artif. Intell. Med.* 43 (3), 243–259.
- Carson Lam, D.Y., Guo, M., Lindsey, T., 2018. Automated detection of diabetic retinopathy using deep learning. *AMIA Summit. Translat. Sci. Proc.* 2017, 147.
- Cheng, J., Yin, F., Wong, D.W.K., Tao, D., Liu, J., 2015. Sparse dissimilarity-constrained coding for glaucoma screening. *IEEE Trans. Biomed. Eng.* 62 (5), 1395–1403.
- Ching, T., Himmelstein, D.S., Beaulieu-Jones, B.K., Kalinin, A.A., Do, B.T., Way, G.P., Ferrero, E., Agapow, P.-M., Zietz, M., Hoffman, M.M., et al., 2018. Opportunities and obstacles for deep learning in biology and medicine. *J. R. Soc. Interface* 15 (141), 20170387.
- Chudzik, P., Majumdar, S., Calivá, F., Al-Diri, B., Hunter, A., 2018. Microaneurysm detection using fully convolutional neural networks. *Comput. Method. Progr. Biomed.* 158, 185–192.
- Cuilla, T.A., Amador, A.G., Zinman, B., 2003. Diabetic retinopathy and diabetic macular edema: pathophysiology, screening, and novel therapies. *Diabetes Care* 26 (9), 2653–2664.
- Cuadros, J., Bresnick, G., 2009. EyePACS: an adaptable telemedicine system for diabetic retinopathy screening. *J. Diabetes Sci. Technol.* 3 (3), 509–516.
- Dai, L., Fang, R., Li, H., Hou, X., Sheng, B., Wu, Q., Jia, W., 2018. Clinical report guided retinal microaneurysm detection with multi-sieving deep learning. *IEEE Trans. Med. Imag.* 37 (5), 1149–1161.
- Das, V., Puhan, N.B., Panda, R., 2015. Entropy thresholding based microaneurysm detection in fundus images. In: *2015 Fifth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCPVRIPG)*. IEEE, pp. 1–4.
- Dashtbozorg, B., Mendonça, A.M., Campilho, A., 2015. Optic disc segmentation using the sliding band filter. *Comput. Biol. Med.* 56, 1–12.
- Decencière, E., Cazuguel, G., Lay, B., Cochener, B., Trone, C., Gain, P., Ordonnéz Varela, J.-R., Massin, P., Erginay, A., et al., 2014. Feedback on a publicly distributed image database: the messidor database. *Image Anal. Stereol.* 33 (3), 231–234.
- Deepak, K.S., Sivaswamy, J., 2012. Automatic assessment of macular edema from color retinal images. *IEEE Trans. Med. Imag.* 31 (3), 766–776.
- Dhara, A.K., Mukhopadhyay, S., Bency, M.J., Rangayyan, R.M., Bansal, R., Gupta, A., 2015. Development of a screening tool for diabetic retinopathy in fundus images. In: *Medical Imaging 2015: Computer-Aided Diagnosis*, Vol. 9414. International Society for Optics and Photonics, p. 94140H.
- Dobbin, K.K., Simon, R.M., 2011. Optimally splitting cases for training and testing high dimensional classifiers. *BMC Med. Genom.* 4 (1), 31.
- Esteves, T., Quelhas, P., Mendonça, A.M., Campilho, A., 2012. Gradient convergence filters and a phase congruency approach for in vivo cell nuclei detection. *Mach. Vis. Appl.* 23 (4), 623–638.
- Everingham, M., Van Gool, L., Williams, C.K.I., Winn, J., Zisserman, A., 2010. The pascal visual object classes (voc) challenge. *Int. J. Comput. vis.* 88 (2), 303–338.
- Farnell, D.J.J., Hatfield, F.N., Knox, P., Reakes, M., Spencer, S., Parry, D., Harding, P., 2008. Enhancement of blood vessels in digital fundus photographs via the application of multiscale line operators. *J. Franklin Inst.* 345 (7), 748–765.
- Ferris, F.L., 1993. How effective are treatments for diabetic retinopathy? *JAMA* 269 (10), 1290–1291.
- Figueiredo, I.N., Kumar, S., Oliveira, C.M., Ramos, J.A.D., Engquist, B., 2015. Automated lesion detectors in retinal fundus images. *Comput. Biol. Med.* 66, 47–65.
- Fleming, A.D., Philip, S., Goatman, K.A., Olson, J.A., Sharp, P.F., 2006. Automated microaneurysm detection using local contrast normalization and local vessel detection. *IEEE Trans. Med. Imag.* 25 (9), 1223–1232.
- Fraz, M.M., Badar, M., Malik, A.W., Barman, S.A., 2018. Computational methods for exudates detection and macular edema estimation in retinal images: a survey. *Arch. Comput. Method. Eng.* 1–28.
- Fu, H., Cheng, J., Xu, Y., Wong, D.W.K., Liu, J., Cao, X., 2018. Joint optic disc and cup segmentation based on multi-label deep network and polar transformation. *arXiv preprint arXiv:1801.00926*.
- Fu, H., Xu, Y., Wong, D.W.K., Liu, J., 2016. Retinal vessel segmentation via deep learning network and fully-connected conditional random fields. In: *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on*. IEEE, pp. 698–701.
- García, G., Gallardo, J., Mauricio, A., López, J., Del Carpio, C., 2017. Detection of diabetic retinopathy based on a convolutional neural network using retinal fundus images. In: *International Conference on Artificial Neural Networks*. Springer, pp. 635–642.
- García, M., Sanchez, C.I., Poza, J., López, M.I., Hornero, R., 2009. Detection of hard exudates in retinal images using a radial basis function classifier. *Annal. Biomed. Eng.* 37 (7), 1448–1463.
- Gargyea, R., Leng, T., 2017. Automated identification of diabetic retinopathy using deep learning. *Ophthalmology* 124 (7), 962–969.
- Gegundez-Arias, M.E., Marin, D., Bravo, J.M., Suero, A., 2013. Locating the fovea center position in digital fundus images using thresholding and feature extraction techniques. *Comput. Med. Imag. Graph.* 37 (5–6), 386–393.
- Giacchetti, A., Ballerini, L., Trucco, E., 2014. Accurate and reliable segmentation of the optic disc in digital fundus images. *J. Med. Imag.* 1 (2), 024001.
- Giancardo, L., Meriaudeau, F., Karnowski, T.P., Li, Y., Garg, S., Tobin, K.W., Chaum, E., 2012. Exudate-based diabetic macular edema detection in fundus images using publicly available datasets. *Med. Image Anal.* 16 (1), 216–226.
- Giancardo, L., Meriaudeau, F., Karnowski, T.P., Li, Y., Tobin, K.W., Chaum, E., 2011. Microaneurysm detection with radon transform-based classification on retina images. In: *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, pp. 5939–5942.
- Giancardo, L., Roberts, K., Zhao, Z., 2017. Representation learning for retinal vasculature embeddings. In: *Fetal, Infant and Ophthalmic Medical Image Analysis*. Springer, pp. 243–250.
- Glorot, X., Bengio, Y., 2010. Understanding the difficulty of training feed-forward neural networks. In: *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, pp. 249–256.
- GreenSPAN, H., Van Ginneken, B., Summers, R.M., 2016. Guest editorial deep learning in medical imaging: overview and future promise of an exciting new technique. *IEEE Trans. Med. Imag.* 35 (5), 1153–1159.
- van Grinsven, M.J., van Ginneken, B., Hooyng, C.B., Theelen, T., Sánchez, C.I., 2016. Fast convolutional neural network training using selective data sampling: application to hemorrhage detection in color fundus images. *IEEE Trans. Med. Imag.* 35 (5), 1273–1284.
- Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J., et al., 2018. Recent advances in convolutional neural networks. *Pattern Recognit.* 77, 354–377.

- Gulshan, V., Peng, L., Coram, M., Stumpe, M.C., Wu, D., Narayanaswamy, A., Venugopalan, S., Widner, K., Madams, T., Cuadros, J., et al., 2016. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* 316 (22), 2402–2410.
- Harangi, B., Hajdu, A., 2014. Detection of exudates in fundus images using a markovian segmentation model. In: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, pp. 130–133. Hatanaka, Y., Nakagawa, T., Hayashi, Y., Kakogawa, M., Sawada, A., Kawase, K.,
- Hara, T., Fujita, H., 2008. Improvement of automatic hemorrhage detection methods using brightness correction on fundus images. In: Medical Imaging 2008: Computer-Aided Diagnosis, Vol. 6915. International Society for Optics and Photonics, p. 6915E.
- Havlicek, J.P., 1996. Am-Fm Image Models. University of Texas at Austin. He, K., Gkioxari, G., Dollár, P., Girshick, R., 2017. Mask r-cnn. In: Computer Vision (ICCV), 2017 IEEE International Conference on. IEEE, pp. 2980–2988. He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778.
- Hinton, G., 2018. Deep learning—a technology with the potential to transform health care. *JAMA* 320 (11), 1101–1102.
- Hoo-Chang, S., Roth, H.R., Gao, M., Lu, L., Xu, Z., Nogues, I., Yao, J., Mollura, D., Summers, R.M., 2016. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Trans. Med. Imag.* 35 (5), 1285.
- Hoover, A., 1975. Stare Database. Available. Available: <http://www.ces.clemson.edu/ahoo/stare>
- Howard, A.G., 2013. Some improvements on deep convolutional neural network based image classification. arXiv preprint arXiv:1312.5402.
- Iandola, F., Moskewicz, M., Karayev, S., Girshick, R., Darrell, T., Keutzer, K., 2014. Densenet: implementing efficient convnet descriptor pyramids. arXiv preprint arXiv:1404.1869.
- ICO, 2017. Guidelines for diabetic eye care, 2nd edn. Int. Council Ophthalmol. (ICO). Jaccard, P., 1908. Nouvelles recherches sur la distribution florale. *Bull. Soc. Vaud. Sci.* Nat. 44, 223–270.
- Javid, M., Pourreza, H.-R., Harati, A., 2017. Vessel segmentation and microaneurysm detection using discriminative dictionary learning and sparse representation. *Comput. Method. Progr. Biomed.* 139, 93–108.
- Jelinek, H., Cree, M.J., 2009. Automated Image Detection of Retinal Pathology. Crc Press.
- Jonas, R.A., Wang, Y.X., Yang, H., Li, J.J., Xu, L., Panda-Jonas, S., Jonas, J.B., 2015. Optic disc-fovea distance, axial length and parapapillary zones. the beijing eye study 2011. *PLoS One* 10 (9), e0138701.
- Jones, S., Edwards, R.T., 2010. Diabetic retinopathy screening: a systematic review of the economic evidence. *Diabet. Med.* 27 (3), 249–256.
- Jordan, K.C., Menolotto, M., Bolster, N.M., Livingstone, I.A.T., Giardini, M.E., 2017. A review of feature-based retinal image analysis. *Expert Rev. Ophthalmol.* 12 (3), 207–220.
- Joshi, S., Karule, P.T., 2019. Mathematical morphology for microaneurysm detection in fundus images. *Eur. J. Ophthalmol.* 1120672119843021
- Kamble, R., Kokare, M., Deshmukh, G., Hussin, F.A., Mériauadeau, F., 2017. Localization of optic disc and fovea in retinal images using intensity based line scanning analysis. *Comput. Biol. Med.* 87, 382–396.
- Kao, E.-F., Lin, P.-C., Chou, M.-C., Jaw, T.-S., Liu, G.-C., 2014. Automated detection of fovea in fundus images based on vessel-free zone and adaptive gaussian template. *Comput. Method. Progr. Biomed.* 117 (2), 92–103.
- Kauppi, T., Kamarainen, J.-K., Lensu, L., Kalesnykiene, V., Sorri, I., Uusitalo, H., Kälviäinen, H., 2012. A framework for constructing benchmark databases and protocols for retinopathy in medical image analysis. In: International Conference on Intelligent Science and Intelligent Data Engineering. Springer, pp. 832–843.
- Ker, J., Wang, L., Rao, J., Lim, T., 2018. Deep learning applications in medical image analysis. *IEEE Access* 6, 9375–9389.
- Khojasteh, P., Júnior, L.A.P., Carvalho, T., Rezende, E., Aliahmad, B., Papa, J.a.P., Kumar, D.K., 2018. Exudate detection in fundus images using deeply-learnable features. *Comput. Biol. Med.*
- Kim, J., Hong, J., Park, H., Kim, J., Hong, J., Park, H., 2018. Prospects of deep learning for medical imaging. *Precis. Future Med.* 2 (2), 37–52.
- Kingma, D.P., Ba, J., 2014. Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980.
- Koliias, A.N., Ulbig, M.W., 2010. Diabetic retinopathy: early diagnosis and effective treatment. *Deutsch. Arzteblatt Int.* 107 (5), 75.
- Krause, J., Gulshan, V., Rahimy, E., Karth, P., Widner, K., Corrado, G.S., Peng, L., Webster, D.R., 2018. Grader variability and the importance of reference standards for evaluating machine learning models for diabetic retinopathy. *Ophthalmology* 125 (8), 1264–1272.
- Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems, pp. 1097–1105.
- Lam, C., Yu, C., Huang, L., Rubin, D., 2018. Retinal lesion detection with deep learning using image patches. *Invest. Ophthalmol. Vis. Sci.* 59 (1), 590–596.
- Li, H., Chutatape, O., 2004. Automated feature extraction in color retinal images by a model based approach. *IEEE Trans. Biomed. Eng.* 51 (2), 246–254. Lin, S., Ramulu, P., Lamoureux, E.L., Sabanayagam, C., 2016. Addressing risk factors, screening, and preventative treatment for diabetic retinopathy in developing countries: a review. *Clin. Exper. Ophthalmol.* 44 (4), 300–320.
- Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciampi, F., Ghafoorian, M., Van Der Laak, J.A., Van Ginneken, B., Sánchez, C.I., 2017. A survey on deep learning in medical image analysis. *Med. Image Anal.* 42, 60–88. Long, J., Shelhamer, E., Darrell, T., 2015. Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 3431–3440.
- Lynch, S.K., Shah, A., Folk, J.C., Wu, X., Abramoff, M.D., 2017. Catastrophic failure in image-based convolutional neural network algorithms for detecting diabetic retinopathy. *Investig. Ophthalmol. Vis. Sci.* 58 (8), 3776–3776. Marin, D., Gegundez-Arias, M.E., Ponte, B., Alvarez, F., Garrido, J., Ortega, C., Vasallo, M.J., Bravo, J.M., 2018. An exudate detection method for diagnosis risk of diabetic macular edema in retinal images using feature-based and supervised classification. *Med. Biol. Eng. Comput.* 56 (8), 1379–1390.
- Marin, D., Gegundez-Arias, M.E., Suero, A., Bravo, J.M., 2015. Obtaining optic disc center and pixel region by automatic thresholding methods on morphologically processed fundus images. *Comput. Method. Progr. Biomed.* 118 (2), 173–185. Mary, M.C.V.S., Rajasingh, E.B., Jacob, J.K.K., Anandhi, D., Amato, U., Selvan, S.E., 2015. An empirical study on optic disc segmentation using an active contour model. *Biomed. Signal Process. Control* 18, 19–29. Medhi, J.P., Dandapat, S., 2014. Analysis of maculopathy in color fundus images. In: 2014 Annual IEEE India Conference (INDICON). IEEE, pp. 1–4. Mendonça, A.M., Sousa, A., Mendonça, L., Campilho, A., 2013. Automatic localization of the optic disc by combining vascular and intensity information. *Comput. Med. Imag. Graph.* 37 (5–6), 409–417.
- Mookiah, M.R.K., Acharya, U.R., Chandran, V., Martis, R.J., Tan, J.H., Koh, J.E.W., Chua, C.K., Tong, L., Laude, A., 2015. Application of higher-order spectra for automated grading of diabetic maculopathy. *Med. Biol. Eng. Comput.* 53 (12), 1319–1331.
- Mookiah, M.R.K., Acharya, U.R., Chua, C.K., Lim, C.M., Ng, E., Laude, A., 2013. Computer-aided diagnosis of diabetic retinopathy: a review. *Comput. Biol. Med.* 43 (12), 2136–2155.
- Mookiah, M.R.K., Acharya, U.R., Martis, R.J., Chua, C.K., Lim, C.M., Ng, E., Laude, A., 2013. Evolutionary algorithm based classifier parameter tuning for automatic diabetic retinopathy grading: a hybrid feature extraction approach. *Knowl.-Based Syst.* 39, 9–22.
- Morales, S., Engan, K., Naranjo, V., Colomer, A., 2017. Retinal disease screening through local binary patterns. *IEEE J. Biomed. Health Informat.* 21 (1), 184–192. Morales, S., Naranjo, V., Angulo, J., Alcaráz, M., 2013. Automatic detection of optic disc based on PCA and mathematical morphology. *IEEE Trans. Med. Imag.* 32 (4), 786–796.
- Murphy, K.P., 2012. Machine learning: a probabilistic perspective. MIT press. Nagy, B., Harangi, B., Antal, B., Hajdu, A., 2011. Ensemble-based exudate detection in color fundus images. In: Image and Signal Processing and Analysis (ISPA), 2011 7th International Symposium on. IEEE, pp. 700–703. Naqvi, S.A.G., Zafar, H.M.F., et al., 2018. Automated system for referral of cotton-wool spots. *Current Diabetes Rev.* 14 (2), 168–174. Niemeijer, M., Abramoff, M.D., Van Ginneken, B., 2009. Information fusion for diabetic retinopathy CAD in digital color fundus photographs. *IEEE Trans. Med. Imag.* 28 (5), 775–785.
- Niemeijer, M., Van Ginneken, B., Cree, M.J., Mizutani, A., Quellec, G., Sánchez, C.I., Zhang, B., Hornero, R., Lamard, M., Muramatsu, C., et al., 2010. Retinopathy online challenge: automatic detection of microaneurysms in digital color fundus photographs. *IEEE Trans. Med. Imag.* 29 (1), 185–195.
- Niemeijer, M., Van Ginneken, B., Staal, J., Suttorp-Schulten, M.S.A., Abramoff, M.D., 2005. Automatic detection of red lesions in digital color fundus photographs. *IEEE Trans. Med. Imag.* 24 (5), 584–592.
- Nørgaard, M.F., Grauslund, J., 2018. Automated screening for diabetic retinopathy—a systematic review. *Ophthalmic Res.*
- Orlando, J.I., Prokofyeva, E., del Fresno, M., Blaschko, M.B., 2018. An ensemble deep learning based approach for red lesion detection in fundus images. *Comput. Method. Progr. Biomed.* 153, 115–127.
- Osareh, A., Shadgar, B., Markham, R., 2009. A computational-intelligence-based approach for detection of exudates in diabetic retinopathy images. *IEEE Trans. Inf. Technol. Biomed.* 13 (4), 535–545.
- Patton, N., Aslam, T.M., MacGillivray, T., Deary, I.J., Dhillon, B., Eikelboom, R.H., Yeganeh, K., Constable, I.J., 2006. Retinal image analysis: concepts, applications and potential. *Progress Retinal Eye Res.* 25 (1), 99–127.
- Perdomo, O., Otalora, S., Rodríguez, F., Arevalo, J., González, F. A., 2016. A novel machine learning model based on exudate localization to detect diabetic macular edema.
- Pereira, C., Gonçalves, L., Ferreira, M., 2015. Exudate segmentation in fundus images using an ant colony optimization approach. *Inf. Sci.* 296, 14–24. Pereira, C.S., Mendonça, A.M., Campilho, A., 2007. Evaluation of contrast enhancement filters for lung nodule detection. In: International Conference Image Analysis and Recognition. Springer, pp. 878–888. Pires, R., Avila, S., Jelinek, H.F., Wainer, J., Valle, E., Rocha, A., 2017. Beyond lesion-based diabetic retinopathy: a direct approach for referral. *IEEE J. Biomed. Health Informat.* 21 (1), 193–200.
- Porwal, P., Pachade, S., Kamble, R., Kokare, M., Deshmukh, G., Sahasrabuddhe, V., Mériauadeau, F., 2018. Indian diabetic retinopathy image dataset (IDRid). *IEEE Dataport doi: 10.21227/H25W98*.
- Porwal, P., Pachade, S., Kamble, R., Kokare, M., Deshmukh, G., Sahasrabuddhe, V., Mériauadeau, F., 2018. Indian diabetic retinopathy image dataset (IDRid): a database for diabetic retinopathy screening research. *Data* 3 (3/25). doi: [10.3390/data3030025](https://doi.org/10.3390/data3030025).

- Ponwal, P., Pachade, S., Kokare, M., Giancardo, L., Mériadeau, F., 2018. Retinal image analysis for disease screening through local tetra patterns. *Comput. Biol. Med.* 102, 200–210.
- Quellec, G., Charière, K., Boudi, Y., Cochener, B., Lamard, M., 2017. Deep image mining for diabetic retinopathy screening. *Med. Image Anal.* 39, 178–193.
- Quellec, G., Lamard, M., Erginay, A., Chabouis, A., Massin, P., Cochener, B., Cazuguel, G., 2016. Automatic detection of referral patients due to retinal pathologies through data mining. *Med. Image Anal.* 29, 47–64.
- Optimal wavelet transform for the detection of microaneurysms in retina photographs. *IEEE Trans. Med. Imag.* 27 (9), 1230–1241.
- Raman, R., Gella, L., Srinivasan, S., Sharma, T., 2016. Diabetic retinopathy: an epidemic at home and around the world. *Indian J. Ophthalmol.* 64 (1), 69.
- Rangree, S.B., Sivaswamy, J., 2017. Automatic detection of fundus image readers. *J. Med. Imag.* 4 (2), 024503.
- Ravi, D., Wong, C., Deligianni, F., Berthelot, M., Andreu-Perez, J., Lo, B., Yang, G.-Z., 2017. Deep learning for health informatics. *IEEE J. Biomed. Health Informat.* 21 (1), 4–21.
- Reichel, E., Salz, D., 2015. Diabetic retinopathy screening. In: *Managing Diabetic Eye Disease in Clinical Practice*. Springer, pp. 25–38.
- Rocha, A., Carvalho, T., Jelinek, H.F., Goldenstein, S., Wainer, J., 2012. Points of interest and visual dictionaries for automatic retinal lesion detection. *IEEE Transactions on biomedical engineering* 59 (8), 2244–2253.
- Romero-Oraá, R., Jiménez-García, J., García, M., López-Gálvez, M.I., Oraá-Pérez, J., Hornero, R., 2019. Entropy rate superpixel classification for automatic red lesion detection in fundus images. *Entropy* 21 (4), 417.
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. In: *International Conference on Medical image computing and computer-assisted intervention*. Springer, pp. 234–241.
- Roychowdhury, S., Koozekanani, D.D., Parhi, K.K., 2014. Dream: diabetic retinopathy analysis using machine learning. *IEEE J. Biomed. Health Informat.* 18 (5), 1717–1728.
- Saito, T., Rehmsmeier, M., 2015. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS One* 10 (3), e0118432.
- Sánchez, C.I., García, M., Mayo, A., López, M.I., Hornero, R., 2009. Retinal image analysis based on mixture models to detect hard exudates. *Med. Image Anal.* 13 (4), 650–658.
- Sánchez, C.I., Niemeijer, M., Işgum, I., Dumitrescu, A., Suttorp-Schulten, M.S.A., Abramoff, M.D., van Ginneken, B., 2012. Contextual computer-aided detection: improving bright lesion detection in retinal images and coronary calcification identification in CT scans. *Med. Image Anal.* 16 (1), 50–62.
- Seoud, L., Hurtut, T., Chelbi, J., Cheriet, F., Langlois, J.M.P., 2016. Red lesion detection using dynamic shape features for diabetic retinopathy screening. *IEEE Trans. Med. Imag.* 35 (4), 1116–1126.
- Shah, M.P., Merchant, S.N., Awate, S.P., 2018. Abnormality detection using deep neural networks with robust quasi-norm autoencoding and semi-supervised learning. In: *Biomedical Imaging (ISBI) 2018*, 2018 IEEE 15th International Symposium on. IEEE, pp. 568–572.
- Shen, D., Wu, G., Suk, H.-I., 2017. Deep learning in medical image analysis. *Annu. Rev. Eng.* 19, 221–248.
- Shi, W., Caballero, J., Huszár, F., Totz, J., Aitken, A.P., Bishop, R., Rueckert, D., Wang, Z., 2016. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1874–1883.
- Shortiffe, E.H., Blois, M.S., 2006. The computer meets medicine and biology: emergence of a discipline. In: *Biomedical Informatics*. Springer, pp. 3–45.
- Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Sivaswamy, J., Krishnadas, S.R., Joshi, G.D., Jain, M., Tabish, A.U.S., 2014. Drishti-gs: Retinal image dataset for optic nerve head (ohn) segmentation. In: *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*. IEEE, pp. 53–56.
- Son, J., Bae, W., Kim, S., Park, S.J., Jung, K.-H., 2018. Classification of findings with localized lesions in fundoscopic images using a regionally guided CNN. In: *Computational Pathology and Ophthalmic Medical Image Analysis*. Springer, pp. 176–184.
- Son, J., Park, S.J., Jung, K.-H., 2017. Retinal vessel segmentation in fundoscopic images with generative adversarial networks. *arXiv preprint arXiv:1706.09318*.
- Son, J., Shin, J.Y., Kim, H.D., Jung, K.-H., Park, K.H., Park, S.J., 2019. Development and validation of deep learning models for screening multiple abnormal findings in retinal fundus images. *Ophthalmology*.
- Sopharak, A., Uyyanonvara, B., Barman, S., Williamson, T.H., 2008. Automatic detection of diabetic retinopathy exudates from non-dilated retinal images using mathematical morphology methods. *Comput. Med. Imag. Graph.* 32 (8), 720–727.
- Sreng, S., Maneerat, N., Hamamoto, K., Panjaphongse, R., 2019. Cotton wool spots detection in diabetic retinopathy based on adaptive thresholding and ant colony optimization coupling support vector machine. *IEEE Trans. Electr. Electron. Eng. Srivastava, R., Duan, L., Wong, D.W.K., Liu, J., Wong, T.Y., 2017. Detecting retinal microaneurysms and hemorrhages with robustness to the presence of blood vessels. Comput. Method. Progr. Biomed.* 138, 83–91.
- Staal, J., Abramoff, M.D., Niemeijer, M., Viergever, M.A., van Ginneken, B., 2004. Ridge based vessel segmentation in color images of the retina. *IEEE Trans. Med. Imag.* 23 (4), 501–509.
- Sudre, C.H., Li, W., Vercauteren, T., Ourselin, S., Cardoso, M.J., 2017. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Springer, pp. 240–248.
- Suzuki, K., 2017. Overview of deep learning in medical imaging. *Radiol. Phys. Technol.* 10 (3), 257–273.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., 2015. Going deeper with convolutions. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–9.
- Tajbakhsh, N., Shin, J.Y., Gurudu, S.R., Hurst, R.T., Kendall, C.B., Gotway, M.B., Liang, J., 2016. Convolutional neural networks for medical image analysis: full training or fine tuning? *IEEE Trans. Med. Imag.* 35 (5), 1299–1312.
- Tan, J.H., Acharya, U.R., Bhandary, S.V., Chua, K.C., Sivaprasad, S., 2017. Segmentation of optic disc, fovea and retinal vasculature using a single convolutional neural network. *J. Comput. Sci.* 20, 70–79.
- Tan, J.H., Fujita, H., Sivaprasad, S., Bhandary, S.V., Rao, A.K., Chua, K.C., Acharya, U.R., 2017. Automated segmentation of exudates, haemorrhages, microaneurysms using single convolutional neural network. *Inf. Sci.* 420, 66–76.
- Tang, L., Niemeijer, M., Reinhardt, J.M., Garvin, M.K., Abramoff, M.D., 2013. Splat feature classification with application to retinal hemorrhage detection in fundus images. *IEEE Trans. Med. Imag.* 32 (2), 364–375.
- Thakur, N., Juneja, M., 2017. Clustering based approach for segmentation of optic cup and optic disc for detection of glaucoma. *Current Med. Imag. Rev.* 13 (1), 99–105.
- Ting, D.S.W., Cheung, G.C.M., Wong, T.Y., 2016. Diabetic retinopathy: global prevalence, major risk factors, screening practices and public health challenges: a review. *Clin. Exper. Ophthalmol.* 44 (4), 260–277.
- Trucco, E., Ruggeri, A., Karnowski, T., Giancardo, L., Chaum, E., Hubschman, J.P., Al-Diri, B., Cheung, C.Y., Wong, D., Abramoff, M., et al., 2013. Validating retinal fundus image analysis algorithms: issues and a proposal. *Invest. Ophthalmol. Visual Sci.* 54 (5), 3546–3559.
- Uribe-Valencia, L.J., Martínez-Carballido, J.F., 2019. Automated optic disc region location from fundus images: using local multi-level thresholding, best channel selection, and an intensity profile model. *Biomed. Signal Processing and Control* 51, 148–161.
- Voulodimos, A., Doulaimes, N., Bebis, G., Stathaki, T., 2018. Recent Developments in deep learning for engineering applications. *Comput. Intell. Neurosci.* 2018. Walter, T., Klein, J.-C., Massin, P., Erginay, A., 2002. A contribution of image processing to the diagnosis of diabetic retinopathy-detection of exudates in color fundus images of the human retina. *IEEE Trans. Med. Imag.* 21 (10), 1236–1243.
- Welfer, D., Scharcanski, J., Marinho, D.R., 2011. Fovea center detection based on the retina anatomy and mathematical morphology. *Comput. Method. Progr. Biomed.* 104 (3), 397–409.
- Winder, R.J., Morrow, P.J., McRitchie, I.N., Bailie, J.R., Hart, P.M., 2009. Algorithms for digital image processing in diabetic retinopathy. *Comput. Med. Imag. Graph.* 33 (8), 608–622.
- Wong, T.Y., Cheung, C.M.G., Larsen, M., Sharma, S., Simo, R., 2016. Diabetic retinopathy. *Nature Rev. Disease Primers*.
- Wu, B., Zhu, W., Shi, F., Zhu, S., Chen, X., 2017. Automatic detection of microaneurysms in retinal fundus images. *Comput. Med. Imag. Graph.* 55, 106–112.
- Wu, L., Fernandez-Loaiza, P., Sauma, J., Hernandez-Bogantes, E., Masis, M., 2013. Classification of diabetic retinopathy and diabetic macular edema. *World J. Diabetes* 4 (6), 290.
- Wu, X., Dai, B., Bu, W., 2016. Optic disc localization using directional models. *IEEE Trans. Image Process.* 25 (9), 4433–4442.
- Yu, F., Wang, D., Shelhamer, E., Darrell, T., 2017. Deep layer aggregation. *arXiv preprint arXiv:1707.06484*.
- Yu, H., Barriga, E.S., Agurto, C., Echegaray, S., Paticchis, M.S., Bauman, W., Soliz, P., 2012. Fast localization and segmentation of optic disk in retinal images using directional matched filtering and level sets. *IEEE Trans. Inf. Technol. Biomed.* 16 (4), 644–657.
- Yun, W.L., Acharya, U.R., Venkatesh, Y.V., Chee, C., Min, L.C., Ng, E.Y.K., 2008. Identification of different stages of diabetic retinopathy using retinal optical images. *Inf. Sci.* 178 (1), 106–121.
- Zhang, B., Karray, F., Li, Q., Zhang, L., 2012. Sparse representation classifier for microaneurysm detection and retinal blood vessel extraction. *Inf. Sci.* 200, 78–90.
- Zhang, X., Thibault, G., Decencière, E., Marcotegui, B., Lay, B., Danno, R., Cazuguel, G., Quellec, G., Lamard, M., Massin, P., et al., 2014. Exudate detection in color retinal images for mass screening of diabetic retinopathy. *Med. Image Anal.* 18 (7), 1026–1043.
- Zhou, W., Wu, C., Yi, Y., Du, W., 2017. Automatic detection of exudates in digital color fundus images using superpixel multi-feature classification. *IEEE Access* 5, 17077–17088.
- Zilly, J., Buhmann, J.M., Mahapatra, D., 2017. Glaucoma detection using entropy sampling and ensemble learning for automatic optic cup and disc segmentation. *Comput. Med. Imag. Graph.* 55, 28–41.