**电子科技大学**

**大学生创新实验计划项目**

**结题报告书**

项目名称： 医疗行业机器读片平台

申 请 人： 黄璞

申请金额： 5000

指导老师： 王一峰

联系导师： 庞晋雁

联系电话： 13281213115

电子信箱： [awsomekde@gmail.com](mailto:awsomekde@gmail.com)

结题日期： 2016-11-20

**一、基本情况**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **1．项目情况** | | | | | | | | | | | | | | | |
| 项目名称 | | | | | | | 医疗行业机器读片平台 | | | | | | | | |
| 项目性质 | | | | | | | □应用性项目 | | | | | | | | |
| 项目来源 | | | | | | | □自主立题 | | | | | | | | |
| 项目起止时间 | | | | | | | 2016年 5 月 至 2016 年 11 月 | | | | | | | | |
| 资助金额（元） | | | | | | | 5000 | | | | | | | | |
| **2．申请者情况** | | | | | | | | | | | | | | | |
| 姓名 | 黄璞 | | | | 性别 | | | 男 | | | 所在学院 | | 机械电子工程学院 | | |
| 学号 | 201621080212 | | | | 专业 | | | 机械工程 | | | 联系方式 | | 13281213115 | | |
| **3．项目组成员情况(包括组长，共5人)** | | | | | | | | | | | | | | | |
| 姓名 | | | 性别 | 学号 | | | | | | 主要负责工作 | | | | | 签字 |
| 黄璞 | | | 男 | 201621080212 | | | | | | 主要代码编写，项目进度统筹 | | | | |  |
| 刘旭 | | | 男 | 201621080131 | | | | | | 算法讨论，测试代码编写，文档撰写 | | | | |  |
|  | | |  |  | | | | | |  | | | | |  |
|  | | |  |  | | | | | |  | | | | |  |
|  | | |  |  | | | | | |  | | | | |  |
| **4．指导老师情况** | | | | | | | | | | | | | | | |
| 姓名 | | | 王一峰 | | | 职称 | | | |  | | 年龄 | |  | |
| 所在学院 | | |  | | | | | | | 联系方式 | |  | | | |
| **5．联系导师情况** | | | | | | | | | | | | | | | |
| 姓名 | | 庞晋雁 | | | | | | | 联系方式 | | |  | | | |
| 1. **项目完成情况及取得的成果情况： 1.项目分析**   本次“银杏黄”创新创业基金项目添加了极具特色的“黄氏高阶命题”--“认知计算”领域，出于理论实践意义与项目创新性的要求，我们选取了其中的“医疗行业机器读片平台”作为我们的立项项目。此项目要求针对某特定病种设计医疗图像辅助诊断程序。具体地，对于此次项目，基金会提供了眼科视网膜彩色图像，即要求开发可以针对眼底病变图像进行分类以辅助医疗诊断的程序。目前，针对眼底病变（Retinopathy）的机器分类方法的研究较少，在SCI网站以主题关键字TS=(retinopathy and machine learning)进行索引，仅有数十篇相关研究。例如Bolon-Canedo, V.[1]等人的研究是基于提取眼底血管网络特征，并且只针对糖尿病性视网膜病（DR）和青光眼进行病变程度的分类；Jiayi Wu[2]等人的研究则是基于DR的微动脉瘤特征，Akyol, Kemal[3]等人的研究则是基于血管渗出物特征结合机器学习进行分类。可以看出，已有的研究大多是基于病理机理本身，即需要用到相关医学专业知识。虽然通过这种途径可以获得不错的分类准确度，但是显然这种方法需要相关疾病的专业从业人员。同时，通过此种途径开发的分类算法不具有通用性，即每种不同的疾病识别均需要经过“人工分析-提取特征-建模”的步骤，加大了推广机器分类方法的难度。  随着近几年深度学习（Deep Learning）的逐渐流行与成熟，为解决此类图像分类问题提供了一种全新的思路，即通过深度卷积神经网络对训练图像进行自动特征提取以达到“通用建模”的目的。关于这方面的研究，van Grinsven, Mark J. J. P.[4]等人利用改进的卷积神经网络（CNN）在他们给出的两个测试集上达到了89%与97%的准确率，这几乎与专家具有相同的分类正确率。但是van等人的研究建立在他们独自开发出的CNN模型与训练方式之上，结构上与方法上依旧不具备通用性。并且深度卷积神经网络的训练是极其耗时并且其参数的调整也具有挑战性，因此如何开发结构与方法上均具有通用性的病理图像分类方法是一个具有挑战性与创新新的课题。  我们结合现有的深度学习成果，决定基于Google开发的开源机器学习框架“TensorFlowTM”[5]搭建深度卷及神经网络模型，模块化开发用于眼底病变图像识别的分类程序。TensorFlow是一个高伸缩性、高可移植性与高性能的跨平台机器学习通用建模框架。它简化了开发机器学习模型的复杂度，并且可以运行在基于CPU、GPU的桌面计算机、服务器甚至移动设备之上。TensorFlow已经用于“Google 大脑”项目并获得了巨大的成功。  以上是我们课题背景及特色，接下来将详细阐述整个分类机器学习算法的开发过程与细节。   1. **数据准备**   基金会提供了7种共372张高分辨的眼底图像，这其中包括：   |  |  | | --- | --- | | BRVO\_CRVO | 视网膜分支静脉阻塞&视网膜中央静脉阻塞 | | CSC | 中心性浆液性视网膜脉络膜病变 | | DR | 糖尿病视网膜病变 | | PM | 高度近视 | | AMD | 湿性老年黄斑变性 | | Others | 其他非正常眼底图片 | | Normal | 正常眼底图片 |   表1. 眼底图像分类对应表  这些原始图像的分辨率都较高，其中分辨率最大的达到了2592x1728，大小也接近13M/张，这样的尺寸与大小是不适合直接用来作为深度卷积神经网络的输入的，因为单张的特征维度就可以高达2592x1728=4478976！并且基金会给出的数据还具有命名不统一（以Normal分类为例如图1所示）、图像存在标注（如图2、3所示）、存在黑边、格式不统一（tif与jpg）、大小不统一、明亮度对比度有差异等等诸多问题，这是实际生产情况中不可避免存在的问题，如果直接使用这些数据进行特征提取训练会引入诸多噪声并且会对结果的收敛产生不利影响。因此在使用深度学习技术之前，数据需要进行预处理。  Screenshot from 2016-11-19 23-32-53  图1. 原始图像数据（Normal）  image_out  图2. 图像标注1  image_out2  图3. 图像标注2  首先需要对数据进行整理的步骤是规则化重命名。虽然全部数据不算很多，仅有近400张，但是进行手工重命名也是一件效率极其低下并且不利于以后数据集拓展的方法，因此需要编写程序自动化重命名。我们采用“[类别]\_[编号]”的方式编写了Linux Shell脚本进行统一命名处理。  随后就是对图像内容的预处理，这里我们使用了ImageMagic工具，这是一个在Linux平台非常流行的命令行图像批处理程序，可以方便地进行丰富的图像处理操作。首先是需要对带有标签的图像进行切割处理。我们使用了如下方法：通过观察，图像标注文字为白色，图像背景为黑色，因此可以通过对白色部分的起始部分进行垂直与水平切割达到去除的目的。其次还需要对图像进行去黑边处理，因为图像的每个输入像素均为需要学习的特征，输入过多重复的背景噪声会导致神经网络对这些背景的“过度学习”而忽略了真正有意义的特征。最后就是对图像格式与大小的归一化处理。由于TIFF格式是一种灵活的位图格式，它的可扩展性带来了很多兼容问题，例如，可以随意定义新的标签和选项，因此并不是所有的实现程序都能支持这些所有这些创造出的标签[6]。所以我们统一采用PNG格式，这是一种被标准化的无损压缩的位图图形格式，支持索引、灰度、RGB三种颜色方案以及Alpha通道等特性[7]。同时，将图片重新调整大小为96x96像素（原因见第二部分经验教训及自我评价）。在图像调整的过程中，我们还发现了几张畸形图像，虽然数目不多，并且对于深度学习来说，尽可能地人工逐一剔除数据是非常不鼓励的（会带来通用性以及拓展性问题），但对于我们这个仅仅有不到400张的训练集来说，小小的噪声对于结果的影响太大，不得不人工剔除这几张畸形图像数据（现代的典型深度学习神经网络训练集至少在10^3数量级！）。经过上述数据预处理技术步骤后，得到的图像如图4所示（以Normal分类为例）：  Screenshot from 2016-11-19 23-59-02  图4. 经过预处理的图像  得到预处理的图像后，还应对已有的数据进行分类数值化（打标签），分布统计等操作，以方便后面的训练过程。图像一共有7个类别，因此标签范围为[0..6]，总结如下：   |  |  |  |  | | --- | --- | --- | --- | | 标签 | 类别 | 数量 | 占比 | | 0 | AMD | 33 | 9.43% | | 1 | BRVO\_CRVO | 66 | 18.86% | | 2 | CSC | 33 | 9.43% | | 3 | DR | 33 | 9.43% | | 4 | Normal | 112 | 32% | | 5 | Others | 40 | 11.42% | | 6 | PM | 33 | 9.43% |   表2. 标签与对应类别数量分布  通过分析图像的数量统计分布可以发现，其分布存在相当大的不平衡现象，即Normal部分远多于其他部分。这也与实际相符，即病理的比例低于正常的比例。这一点在训练过程是需要考虑的，否则很容易导致“过采样”，即神经网络总是学习某一个类别，导致对这个类别的特征学习率很高而忽略了别的类别特征进而导致分类准确度的下降。  关于图像处理这部分还有两点值得说明：  1.没有对图像的亮度、对比度进行处理。因为这部分处理放在了训练阶段的数据“实时扩增”（Data augment）操作。详见第3、4“方案设计”与“模型结构”部分。  2.没有对图像进行灰度或者二值化处理。因为病理图像的分类不同于普通的图像识别或分类，其中的颜色信息十分关键，甚至可能是决定因素，因此我们决定直接输入原始图像色彩，即三通道图像（RGB通道）进行训练。  至此，图像的预处理基本完成，具体处理程序参见附件。   1. **方案设计**   对于典型的图像识别分类，已经有许多优秀的卷积神经网络结构，甚至是带有已经训练好的参数的模型。例如Yann LeCun提出的用于邮政编码以及手写数字识别的LeNet[8]，Alex Krizhevsky提出的用于ImageNet-2012比赛的AlexNet[9]，Matthew Zeiler提出的用于ImageNet-2013比赛的ZF Net[10]，Google提出的用于ImageNet-2014比赛的GoogLeNet[11]，Karen Simonyan提出的用于ImageNet-2014比赛的VGGNet[12]，Kaiming He提出的用于ImageNet-2015比赛的ResNet[13]，以及Google的Inception[14]。  然而，这些已经被证明的高准确率深度卷积神经网络结构是建立在海量数据（10^5数量级）以及强大的计算资源支持之上的，并不完全适合此次小样本数据集和低计算力支持的实例。因此我们在借鉴CIFAR-10Net两层卷积神经网络的基础上设计了11层卷积神经网络用于此次项目的图像识别分类。  同时，我们经过深入论证还发现，这些网络结构倾向于优化图像在复杂背景环境下的物体识别，因此这些网络结构倾向于抛弃背景噪声，然而对于病理图像来说，不存在背景噪声的问题，图像中的几乎各个部分均可能存在疾病特征。这是我们没有采用主流的网络结构的另一个原因。  另外，通过这次项目实践，关于卷及神经网络的输入维度我们有了一些心得体验。一开始我们将图像的原始输入大小调整为512x512，训练输入大小为256x256，但是发现这样的大小设置导致训练极其缓慢，在我们的笔记本计算机上的训练速度约为10min/batch，完成10000个steps可能需要三个月之久！但是值得说明的一点是这里也可以理解为计算机性能不足，因为在支持多GPU并行加速的计算机设备上，对比顶级CPU加速比可以达到70倍，也就是说相对于我们使用的计算机加速比可能达到500倍之多！因此，在有海量数据以及大量计算资源支持的情况下，输入维数越大理论上获得的信息越多，训练结果越好。这里着重在于说明基于已有的计算资源以及数据量下如何权衡获得最优效果。  综上，基于权衡已有计算资源、已有数据量以及数据的特征特点，我们采用了自己设计11层卷积神经网络，具体的，整体以及各个部分的TensorFlow结构图参见第4部分“模型结构”。   1. **模型结构**   首先，给出我们设计的网络在TensorBoard中的全貌图，如图5所示，由于整体图像过大，接下来将会对各个部分进行详细阐述。  graph  图5. 网络全貌图  图中，圆角方框为主要操作节点，椭圆框图为基本单元操作节点，细线为原始数据（3维图像数据[height, width, channels]）流向连接线，粗线为4维Tensor数据（[batch, height, width, channels]）流向连接线。  可以看出，整个结构图人为地分为左右两个部分，左边为主体部分，即整个网络最核心的数据流图；右边为辅助部分，即方便训练过程的操作节点，这些节点有些是为了记录过程数据和最终模型，有些则是为了方便可视化训练过程。  整个网络分为4大部分，分别为：数据输入及预处理层，卷积层，全连接层，softmax分类层。  数据输入及预处理层，如图6所示：  graph_input  图6. 数据输入及预处理层  可以看出这部分可以分为四个部分。第一个部分为随机队列（random\_queue），在第2部分“数据准备”阶段曾提到数据分布不均匀这个问题，因此需要设计随机队列每次从已有数据中随机抽取若干数据进行训练，保证每次训练的数据已经充分混合，防止“过采样”。随机队列的实现原理[15]如图7所示：  graph_queue  图7. 随机队列原理图  从图中可以看出，随机队列的存在保证了每次提供给网络参与训练的数据的随机均匀性，同同时也为下一步“数据扩增（Data augment）”提供了随机数据来源。第二部分就是“数据扩增（Data augment）”，从图中可以看出这个部分包括了   1. **训练调参** 2. **结果评估** 3. **成果情况** 4. **其他说明** 5. 由于在项目期间有许多值得记录与分享的心得经验，因此撰写了若干相关博客，记录了项目期间的点点滴滴，也算是一些额外成果收获。地址：<http://hp.stuhome.net/> 6. 出于方面代码管理与分享的目的，项目全部代码及文档均托管在GitHub，地址：<https://github.com/callofdutyops/YXH2016724098982>   参考文献  [1] Bolon-Canedo V. Dealing with inter-expert variability in retinopathy of prematurity: A machine learning approach.  [2] Jiayi Wu. New hierarchical approach for microaneurysms detection with matched filter and machine learning.  [3] Akyol Kemal. Automatic Detection of Optic Disc in Retinal Image by Using Keypoint Detection, Texture Analysis, and Visual Dictionary Techniques.  [4] van Grinsven MJJP. Fast Convolutional Neural Network Training Using Selective Data Sampling: Application to Hemorrhage Detection in Color Fundus Images.  [5] <https://www.tensorflow.org/>  [6] <https://zh.wikipedia.org/wiki/TIFF>  [7] <https://zh.wikipedia.org/wiki/PNG>  [8] Yann LeCun. Gradient-Based Learning Applied to Document Recognition.  [9] Alex Krizhevsky. ImageNet Classification with Deep Convolutional Neural Networks.  [10] Matthew D Zeiler. Visualizing and Understanding Convolutional Networks.  [11] Christian Szegedy. Going Deeper with Convolutions.  [12] Karen Simonyan. Very Deep Convolutional Networks for Large-Scale Visual Recognition.  [13] Kaiming He. Deep Residual Learning for Image Recognition.  [14] Christian Szegedy. Rethinking the Inception Architecture for Computer Vision.  [15] <https://www.tensorflow.org/versions/r0.11/how_tos/reading_data/index.html#creating-threads-to-prefetch-using-queuerunner-objects> | | | | | | | | | | | | | | | |
| 1. **经验教训及自我评价：** 虽然这次“银杏黄”项目只有短短的六个月时间，但是我们还是从中学到了很多平时接触不到的知识，尤其是在机器学习深度学习这个领域。但这其中也包括许多经验教训。例如：对深度学习计算量的乐观估计以及对数据量的把握不周，导致一开始盲目设计了过深的（多达21层）神经网络。   但是不管怎样，我们还是对这次项目还是比较满意的，因为基本上完成了预期的目标，并且也做了许多自己的尝试，无愧“银杏黄”项目基金组以创新创业为目的的初衷。  一些无关项目但是值得注意的其他说明：  1.关于调整图像的大小：   1. 关于卷积网络的大小与深度： 2. 关于计算机性能的说明： | | | | | | | | | | | | | | | |
| 1. **经费使用情况：** 1.资料文献费：700元 2. 文印费：100元 3. 交通费：150元 4. 杂务费：150元 | | | | | | | | | | | | | | | |
| **四、指导老师意见：**  签字：  　　　　　 年　　月　　日 | | | | | | | | | | | | | | | |
| **五、联系导师意见**  签字：  　　　　　 年　　月　　日 | | | | | | | | | | | | | | | |
| **六、专家组评价意见：**    签字：  年 月 日 | | | | | | | | | | | | | | | |
| **七、学院审核意见：**  　  　　　　　　　　　　　 签章：  年 月 日 | | | | | | | | | | | | | | | |

注：附有关论文、成果实物及其证明材料等。