

# 基于深度学习的帕金森病症早期诊断<sup>①</sup>

张巧丽<sup>1,2</sup>, 迟学斌<sup>1,2</sup>, 赵 地<sup>1</sup>

<sup>1</sup>(中国科学院 计算机网络信息中心, 北京 100190)

<sup>2</sup>(中国科学院大学, 北京 100049)

通讯作者: 迟学斌, E-mail: [chi@sccas.cn](mailto:chi@sccas.cn)

**摘 要:** 在世界上, 约有七百万到一千万老年人正在承受着帕金森 (Parkinson's Disease, PD) 疾病带来的困扰. 帕金森疾病是一种常见的神经系统变性疾病, 它的临床特征为震颤、强直、运动迟缓以及自主能力的下降. 其临床表现和多系统萎缩 (Multiple System Atrophy, MSA) 病症极为相似. 研究表明, 帕金森病患者在确诊时, 往往已经到了无法挽回的境地, 所以对于帕金森病症能够区别于 MSA 病症并且得到早期诊断, 人们在不断探索新的方法. 随着大数据时代的到来, 深度学习在图像识别和分类方面取得了重大性突破. 所以, 本研究提出使用深度学习方法实现对帕金森疾病、多系统萎缩症和健康人群的诊断. 本数据来源北京 301 医院. 原始核磁共振图像 (Magnetic Resonance Image, MRI) 的处理得到北京 301 医院医生的指导. 本实验重点在于改进现有神经网络, 使其在医学图像识别和诊断中获得良好的效果. 本实验依据帕金森病症的病理特点提出了改进算法, 通过对比模型损失、准确率等指标获得了较好的实验结果.

**关键词:** 帕金森病症; 多系统萎缩症; 深度学习; 核磁共振图像; 神经网络

引用格式: 张巧丽, 迟学斌, 赵地. 基于深度学习的帕金森病症早期诊断. 计算机系统应用, 2018, 27(9): 1-9. <http://www.c-s-a.org.cn/1003-3254/6538.html>

## Early Diagnosis of Parkinson's Disease Based on Deep Learning

ZHANG Qiao-Li<sup>1,2</sup>, CHI Xue-Bin<sup>1,2</sup>, ZHAO Di<sup>1</sup>

<sup>1</sup>(Computer Network Information Center, Chinese Academy of Sciences, Beijing 100190, China)

<sup>2</sup>(University of Chinese Academy of Sciences, Beijing 100049, China)

**Abstract:** In the world, about seven to ten million elderly people are suffering from the Parkinson's Disease (PD). PD is a common degenerative nervous system disease. Its clinical characters are tremor, muscle rigidity, bradykinesia, and the depression of independent ability. These characters are similar with the Multiple System Atrophy (MSA). Research shows that patients with PD are often irreparably diagnosed, so people are constantly exploring new ways to differentiate PD with MSA and get early diagnosis. With the advent of the big data era, deep learning has made major breakthroughs in image recognition and classification. Therefore, the study uses the deep learning methods to differentiate PD, MSA, and healthy people. The data is from 301 Hospital of Beijing. The pre-treatment of the original Magnetic Resonance Image (MRI) is directed by the physicians of 301 Hospital of Beijing. The focus of this experiment is to optimize the neural network and make it get good results in medical image recognition and diagnosis. Based on the pathological characteristics of PD, the study proposed an improved algorithm, and it gets the better experimental results in loss, accuracy, and other indicators.

**Key words:** Parkinson's Disease (PD); deep learning; Multiple System Atrophy (MSA); Magnetic Resonance Image (MRI); neural network

① 基金项目: 国家重点研发计划 (2016YFB0201404); 国家自然科学基金 (91530324)

Foundation item: National Key R & D Projects of China (2016YFB0201404); National Natural Science Foundation of China (91530324)

收稿时间: 2018-01-27; 修改时间: 2018-02-27; 采用时间: 2018-03-06; csa 在线出版时间: 2018-07-26

## 1 引言

### 1.1 研究背景

帕金森疾病 (Parkinson's Disease, PD) 又名震颤麻痹, 老年人中枢神经系统第二大常见退行性疾病, 其特点是黑质致密部多巴胺能神经元逐渐丧失, 运动功能受损<sup>[1]</sup>, 并且表现为多巴胺能 (DopAminergic, DA) 神经元进行性丧失的特征<sup>[2]</sup>. 该病症的平均发病年龄大概在 60 岁左右, 40 岁以下发病的青年帕金森病较少见. 据统计, 在全世界大约有七百万到一千万的老年人正在承受着该疾病的困扰. 我国 65 岁以上人群 PD 的患病率大约是 1.7%<sup>[3]</sup>, 大部分帕金森患者为散发病例, 仅有不到 10% 的患者有家族史, 所以到目前为止 PD 的病因和发病机制还未明确.

多系统萎缩症 (Multiple System Atrophy, MSA) 是一种缓慢进行性的神经退行性疾病<sup>[4]</sup>, 其特征以帕金森综合征为主的 MSA-P 亚型运动减退性障碍和以小脑性共济失调 (cerebellar ataxi) 为主的 MSA-C 亚型肢体共济失调.

核磁共振成像 (Magnetic Resonance Imaging, MRI) 是断层成像的一种, 它利用磁共振现象从人体中获得电磁信号, 并重建出人体信息. 目前该方法已经被广泛的应用于医学成像. 并且在核磁共振图像中我们可以获取物质的多种物理特性参数, 如质子密度, 自旋-晶格弛豫时间 T1, 自旋-自旋弛豫时间 T2, 扩散系数, 磁化系数, 化学位移等等. 在本实验中根据医生建议采用磁共振扩散加权成像 (Diffusion Weighted Imaging, DWI)、T2 和冠状面 T2 水抑制序列 (CorT2) 三种 MRI 图像.

### 1.2 研究现状

目前医学科技发展迅速, 但是对于帕金森疾病的诊断多数通过病人的临床症状才能确诊; 然而, 国内外专家通过大量临床经验及实验证明, 帕金森的主要病理改变是黑质纹状细胞的进行性损失和细胞内路易小体的聚集, 从黑质、纹状体 DA 能神经元变性-丢失-出现临床症状要经历一个漫长的临床前过程, 潜伏期约 5 年, 黑质 DA 能神经元丢失 < 50% 的患者临床症状不明显. 当患者出现 PD 临床症状时, 脑内 DA 能神经元丢失 70%~80%<sup>[5]</sup>.

在 MR 高分辨率 T2WI 加权像/磁敏感加权成像 (SWI) 上正常黑质核团-1 轴位形似燕尾, 称为燕尾征. PD 病理特点是存在于黑质中的多巴胺能神经元进行

性缺失. 先前研究发现黑质种存在 5 个黑质小体, 且最大的黑质小体-1 是主要影响 PD 病理改变的结构. 黑质小体-1 位于黑质后 1/3, 轴位 SWI 上表现为条状或者逗号形的高信号, 形似燕尾. 黑质小体所表现的“燕尾”前方、侧面和内侧被 SWI 低信号围绕, 可见分叉. “燕尾征”消失用于诊断帕金森病准确率大约为 90%, 对于帕金森患者而言, 黑质小体-1 信号较低, 表现为燕尾征消失<sup>[6]</sup>, 如图 1 所示.

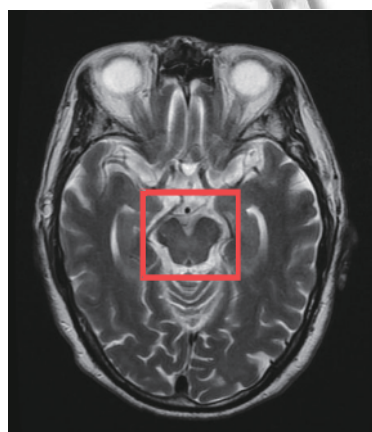


图 1 PD 患者脑图 (中心标注“燕尾征”区)

常规 MRI 检查可见: 1) 脑萎缩: 主要是锥体外系萎缩引起第三脑室增宽, 弥漫性脑皮层萎缩所致的脑沟增宽. 2) 黑质致密带萎缩: 在 T2 加权像/质子密度加权像上, 正常脑组织黑质网状带和红核中存在高浓度铁, 呈低信号; 致密带铁浓度较低音器局部呈等信号. 此外, 还可见 PD 患者因黑质细胞变性坏死和铁代谢已成引起的致密带变窄、边缘模糊等表现. 通过观察黑质致密带形态、信号变化、测量黑质致密带宽度以及黑质致密带宽度与中脑的比值, 为诊断 PD 和鉴别诊断 PD 与血管性帕金森综合征提供客观依据. 3) 由壳核后外侧部铁沉积引起 T2 加权像上纹状体区呈低信号<sup>[7]</sup>.

到目前为止, 将深度学习方法和帕金森病症诊断相结合的研究主要在以下方面. Al-Fatlawi 和 Jabardi 等提出使用深度信念网络 (Deep Belief Network, DBN)<sup>[8]</sup> 进行帕金森病症诊断, 其中分析的信息为病人的语音信号. 该深度信念网络由两个受限玻尔兹曼机<sup>[9]</sup>和一个输出层构成, 第一个进行无监督学习, 第二个是进行反向传播微调的监督学习. 在该研究中测试的准确度达到 94%. Shamir 和 Dolber<sup>[10]</sup>等提出使用深度学习方

检测病人的肢体迟缓程度, 然后进行分类诊断. 在比较传统机器学习方法和基于卷积神经网络的深度学习方法中, 深度学习在准确率方面优于其他机器学习方法 4.6 个百分点. 以上研究均是基于将深度学习方法和帕金森诊断相结合的. 然而本实验采用帕金森患者的脑图作为病症诊断依据在以上帕金森病症病理诊断中已经做了充分调查研究, 实际效果也体现了实验可行性. 这也是本实验得创新之处.

本实验采用对图像识别有良好效果的深度学习 (Deep Learning, DL) 方法, 通过深度神经网络模型训练大批量的 MRI 图像, 并学习图像的特征, 然后进行病症的预测和诊断. 实验中采用的网络是基于 AlexNet 网络的优化网络. AlexNet 是 2012 年在 ImageNet 比赛中脱颖而出的网络, 其良好的分类效果斩获了当年的比赛冠军. GoogleNet 是 2014 年 ImageNet 的比赛冠军. 本实验采用的基础对照模型为 AlexNet 和 GoogleNet, 然后基于 AlexNet 模型进行优化, 优化后的模型在实验中获得优于原始模型的效果, 并且也优于经典网络 GoogleNet.

## 2 深度学习

### 2.1 深度学习概况

2006 年, 深度学习以机器学习领域的一个分支呈现给人们, 它采用多层复杂结构或者采用多重非线性变换构成的多个层进行数据处理<sup>[11]</sup>. 到目前为止, 深度学习在自然语言处理、语音识别特别是计算机视觉方面取得了突破性进展<sup>[12]</sup>. 深度学习的优点在于应用分层的高效特征提取方法来代替手工获取特征, 这有效的解决了大批量的人工标注工作. 深度神经网络层次结构如图 2.

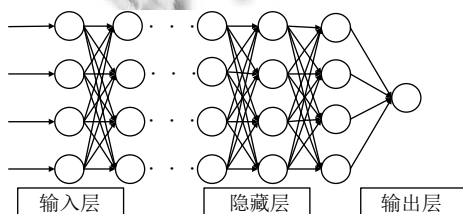


图 2 深度神经网络多层结构

深度学习的基本思想是通过构建如上图所示多层次神经网络, 从底层向高层逐步学习提取特征, 最终通过大批量数据的训练学习, 构建对应的网络模型, 学习

训练对象的相关特征.

### 2.2 深度神经网络结构

目前常见深度神经网络主要包括: 卷积层、池化层、激活层等.

1962 年, Hubel 和 Wiesel 通过研究猫眼的瞳孔区域和大脑皮层神经元, 提出了感受野 (receptive field) 的概念<sup>[13]</sup>. 后来学者 Fukushima 又基于此概念提出神经认知机 (neocognitron), 这是感受野概念在人工神经网络领域的第一次应用.

含有卷积层的神经网络为多层神经网络, 该卷积层有多个二维矩阵构成, 每个矩阵有多个独立神经元构成. 卷积层的核心在于通过感受野和权值共享的应用减少了深度神经网络要训练的参数个数. 如图 3 所示.

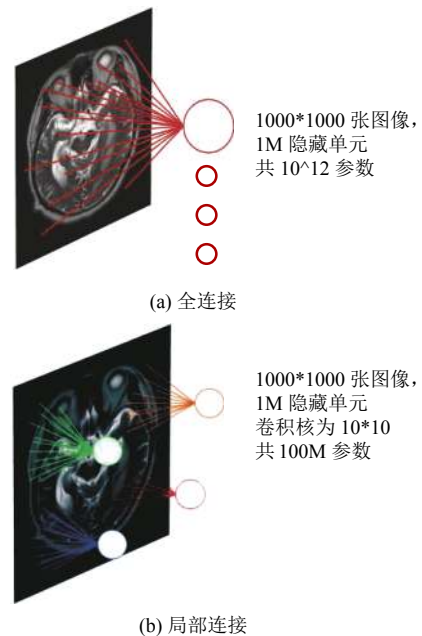


图 3 全连接和局部连接

权值共享为设每个神经元的参数相同, 感受野即对应卷积核局部学习的概念, 通过上图参数对比可知该计算减少了 4 个数量级<sup>[14]</sup>.

池化层包含两种: 一种是平均值池化, 另一种是最大值池化.

池化操作是特征图缩小操作, 会在原特征图中提取主要特征.

因为线性模型的表达能力不够, 所以引入了非线性模型. 激活层实现的是对输入数据的激活即非线性函数变换. 常用的激活函数有 Sigmoid, tanh, ReLU<sup>[15]</sup>



等, 可以根据模型效果选择不同的激活, 通过激活层可以实现数据的更好分类.

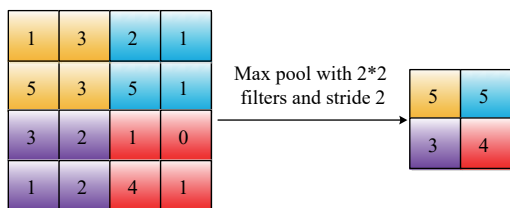


图4 最大值池化

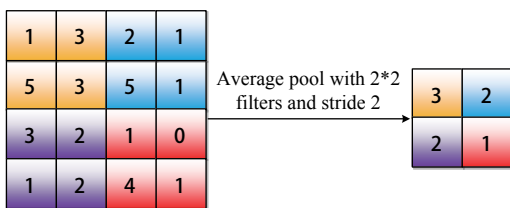


图5 平均值池化

### 3 数据集和预处理

#### 3.1 数据集

本实验数据来自北京 301 医院, 其原始数据为 DICOM (Digital Imaging and Communications in Medicine) 图像, 通过 RadiAntDICOMViewer 软件将病人信息删除并导出 img 格式. 其中详细为: 训练集: 13 571; 验证集: 2396(占训练集的 15%); 测试集: 2237(占总数据的 10%).

图 6 分别为 PD、MSA 和 Normal(正常人) 脑图.

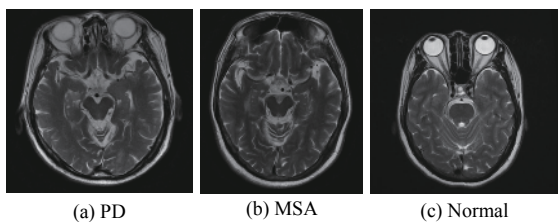


图6 PD, MSA 和 Normal 的脑图

#### 3.2 数据预处理

由于神经网络训练需要批量数据学习特征, 所以实验时需要将原始数据进行扩充, 本实验采用以下两种方式进行数据扩充.

##### 3.2.1 图像的旋转

图像旋转是指图像以某一点为中心旋转一定的角度形成一幅新的图像的过程, 并且旋转前后的点离中

心的位置不变. 假设点  $(x_0, y_0)$  距离原点的距离为  $r$ , 点与原点之间的连线与  $x$  轴的夹角为  $b$ , 旋转的角度为  $a$  旋转后的点为  $(x_1, y_1)$ , 那么:

原始点的位置:

$$\begin{cases} x_0 = r \cos b \\ y_0 = r \sin b \end{cases}$$

旋转后点的位置:

$$\begin{cases} x_1 = r \cos(b - a) = r \cos b \cos a + r \sin b \sin a \\ \quad = x_0 \cos a + y_0 \sin a \\ y_1 = r \sin(b - a) = r \sin b \cos a - r \cos b \sin a \\ \quad = -x_0 \sin a + y_0 \cos a \end{cases}$$

得到旋转后的坐标, 旋转后的图像的长和宽会发生变化, 要重新计算新图像的长和宽, 计算方法如下: 设原始图像长为  $srcH$ , 宽为  $srcW$ , 以图像中心为原点左上角, 右上角, 左下角和右下角的坐标用于计算旋转后的图像的高和宽, 它们大小分别为:

$$\begin{aligned} pLT.x &= -srcW/2; pLT.y = srcH/2 \\ pRT.x &= -srcW/2; pRT.y = srcH/2 \\ pLB.x &= -srcW/2; pLB.y = -srcH/2 \\ pRB.x &= srcW/2; pRB.y = -srcH/2 \end{aligned}$$

旋转之后的坐标分别设为  $pLTN, pRTN, pLBN, pRBN$ , 大小分别为:

$$\begin{aligned} pLTN.x &= pLT.x * \cos a + pLT.y * \sin a \\ pLTN.y &= -pLT.x * \sin a + pLT.y * \cos a \\ pRTN.x &= pRT.x * \cos a + pRT.y * \sin a \\ pRTN.y &= -pRT.x * \sin a + pRT.y * \cos a \\ pLBN.x &= pLB.x * \cos a + pLB.y * \sin a \\ pLBN.y &= -pLB.x * \sin a + pLB.y * \cos a \\ pRBN.x &= pRB.x * \cos a + pRB.y * \sin a \\ pRBN.y &= -pRB.x * \sin a + pRB.y * \cos a \end{aligned}$$

旋转后的长和宽分别设为  $desHeight, desWidth$ , 大小为:

$$\begin{aligned} desWidth &= \max(\text{abs}(pRBN.x - pLTN.x), \text{abs}(pRTN.x - pLBN.x)) \\ desHeight &= \max(\text{abs}(pRBN.y - pLTN.y), \text{abs}(pRTN.y - pLBN.y)) \end{aligned}$$

原始图和旋转  $90^\circ$  后的图像如图 7 所示.

##### 3.2.2 图像的镜像处理

图像的镜像分为水平镜像和垂直镜像. 设图像的宽为  $width$ , 长度为  $height$ ,  $(x, y)$  为变换后的坐标,  $(x_0, y_0)$  为原图像的坐标,

垂直镜像变换:

$$\begin{cases} x = x_0 \\ y = height - y_0 - 1 \end{cases}$$

其逆变换:

$$\begin{cases} x_0 = x \\ y_0 = height - y - 1 \end{cases}$$

水平镜像:

$$\begin{cases} x = width - x_0 \\ y = y_0 \end{cases}$$

其逆变换:

$$\begin{cases} x_0 = width - x \\ y_0 = y \end{cases}$$

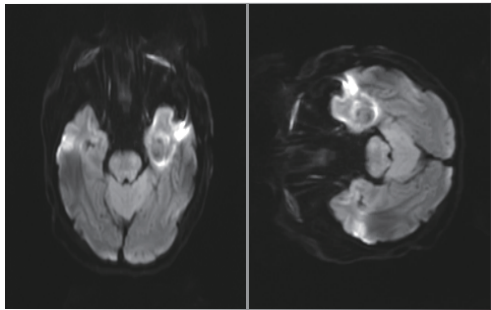


图7 原始图(左)旋转 90°后(右)

在本实验中采用的是垂直镜像,实现左右脑图像的对称交换.垂直镜像图像对比如图8.

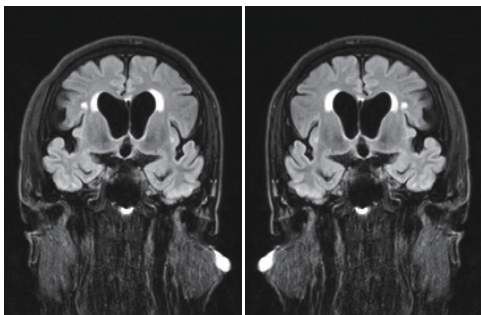


图8 原始图(左)镜像处理后(右)

## 4 实验及结果

### 4.1 优化的 AlexNet 神经网络

AlexNet 神经网络主要包括八个网络层,5个卷积层,3个全连接,在每个卷积层内部有更加细致的层次划分,如图9所示.优化后网络如图10.

在第五层的池化层后,添加归一层:

```
layer{
```

```
name: "norm5"
type: "LRN"
bottom: "pool5"
top: "fc6"
lrn_param{
  local_size: 5
  alpha: 0.0001
  beta: 0.75
}
```

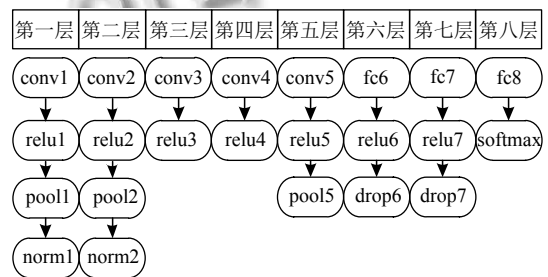


图9 AlexNet 神经网络结构

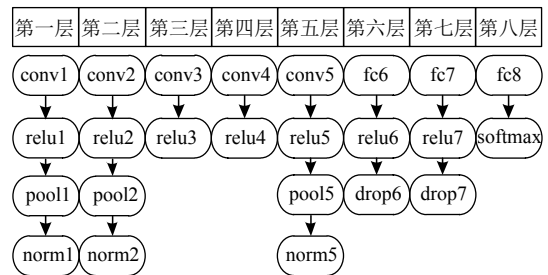


图10 优化的神经网络结构

对比如上网络结构可知在原有 AlexNet 神经网络的第五层,本实验添加了 norm5 网络层结构.对于加 norm5 实现的是 Batch Normalization 操作,在训练深度神经网络的过程中,其复杂性在于随着前几层参数的变化,每层输入的分布在训练过程中也发生改变,所以导致了在训练的过程中学习率 (learning rate) 要设置的非常小,减慢了训练速度,该现象被称为 internal covariate shift<sup>[16]</sup>.由于本实验图像的多样性(脑部的不同位置截图),所以通过在第五层连接全连接层的位置添加归一化层,通过归一化输入全连接层的数据,将其规范化为模型架构的一部分,使得模型可以使用更高的学习率,加速收敛,并且提升了模型效果.算法实现如下.

Input: Values of  $x$  over a mini-batch:  $B=\{x_{1..m}\}$ ; Parameters to be learned:  $\gamma, \beta$

Output:  $\{y_i=BN_{\gamma,\beta}(x_i)\}$

$$\mu_\beta \leftarrow \frac{1}{m} \sum_{i=1}^m x_i \quad // \text{ mini-batch mean}$$

$$\sigma_\beta^2 \leftarrow \frac{1}{m} \sum_{i=1}^m (x_i - \mu_\beta)^2 \quad // \text{ mini-batch variance}$$

$$\hat{x}_i \leftarrow \frac{x_i - \mu_\beta}{\sqrt{\sigma_\beta^2 + \epsilon}} \quad // \text{ normalize}$$

$$y_i \leftarrow \gamma \hat{x}_i + \beta \equiv BN_{\gamma,\beta}(x_i) \quad // \text{ scale and shift}$$

以上算法实现了对  $x$  的小批量激活转换。

2015年 Ioffe 和 Szegedy, 将该方法应用在 Inception network<sup>[17]</sup>用来进行 ImageNet classification 分类, 取得了第五名 4.82% 的错误率, 超出了人类的准确率。

GoogleNet 以其较深且复杂的网络结构在 2014 年脱颖而出, 其主要创新在于根据深度和宽度受限设计的, 并且设计了两个辅助 loss, 是当下比较成熟的深度神经网络, 所以以其作为对照实验。

### 4.2 实验结果

针对以上两种模型进行了四组实验, 分别是 PD vs Normal, PD vs MSA, MSA vs Normal 和 PD vs MSA vs Normal. 以上模型训练的 GPU 设备配置了 13 块 NVIDIA Tesla K80, Intel(R) Xeon(R) CPU E5-2640 v4 6 核处理器 (2.40 GHz). 四组实验 12 个模型可以同时训练, 所有模型训练完成不超过 20 min. 实验结果如图 11 至图 22.

#### 4.2.1 PD vs Normal (PN)

对于 PD 和 Normal 即帕金森病症和正常人进行分类实验中 (如图 11 至图 13), 在原有 AlexNet 实验的基础上准确率获得了 0.2% 的提升, 验证集损失获得了 0.01 的降低, 训练集损失获得了 0.04 的降低。

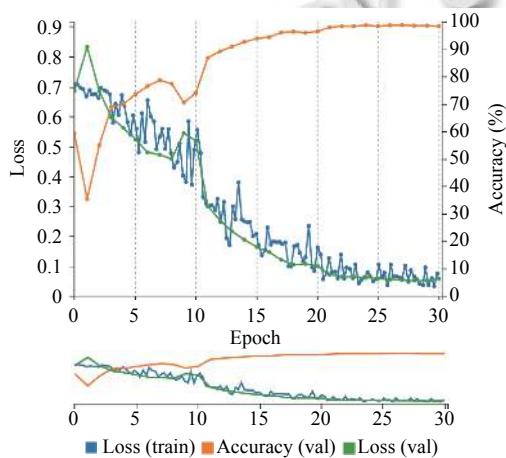


图 11 优化网络结果

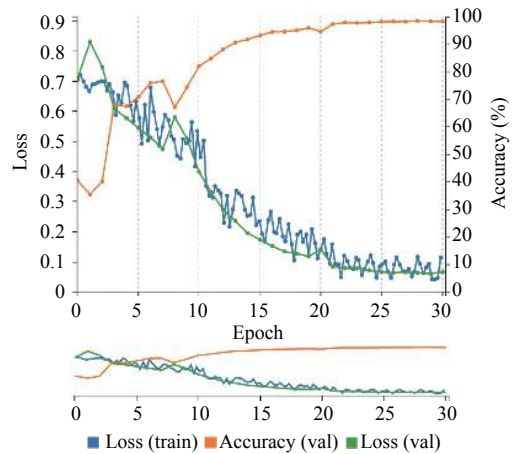


图 12 原 AlexNet 网络结果

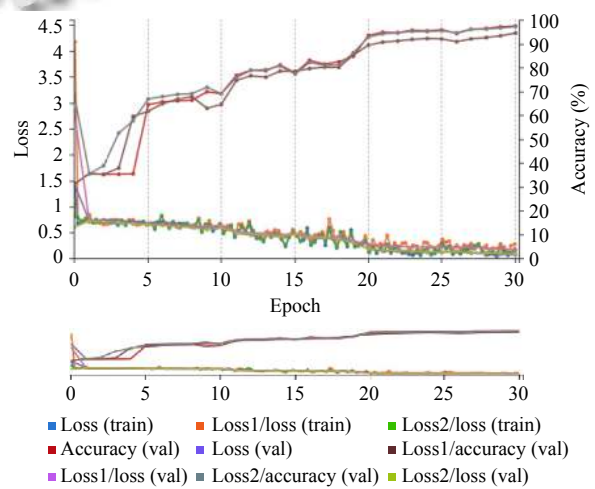


图 13 GoogleNet 网络结果

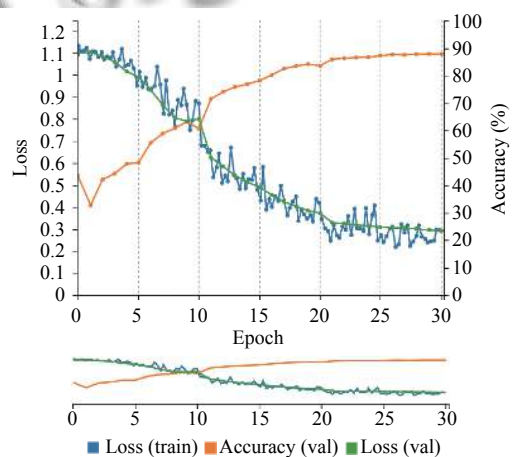


图 14 优化网络结果

#### 4.2.2 PD vs MSA (PM)

对于 PD 和 MSA 即帕金森病症和多系统萎缩进

行分类实验中(如图 14 至图 16), 在原有 AlexNet 实验得基础上准确率获得了 1% 的提升, 验证集损失保持持平, 训练集损失获得了 0.01 的降低.

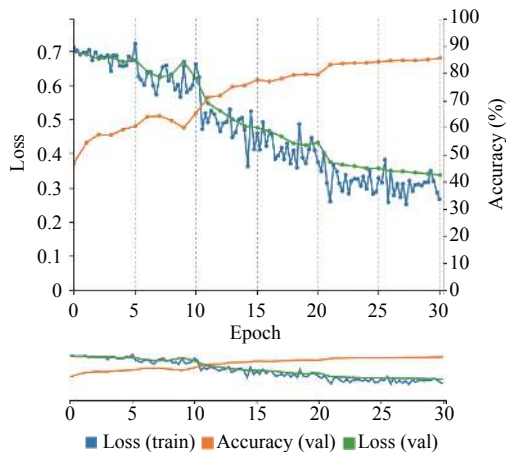


图 15 原 AlexNet 网络结果

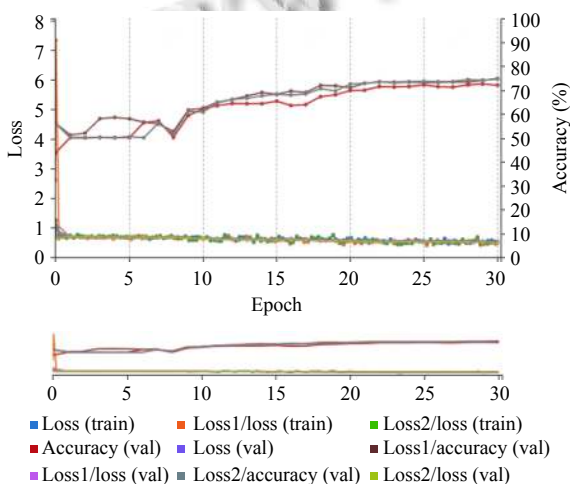


图 16 GoogleNet 网络结果

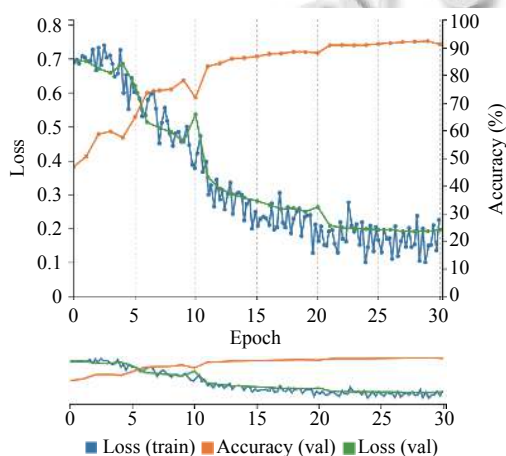


图 17 原 AlexNet 网络结果

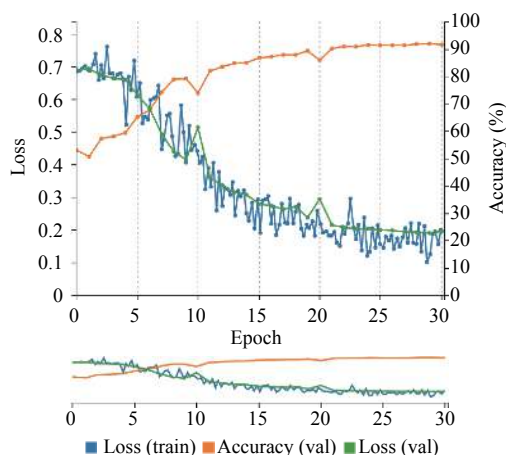


图 18 优化网络结果

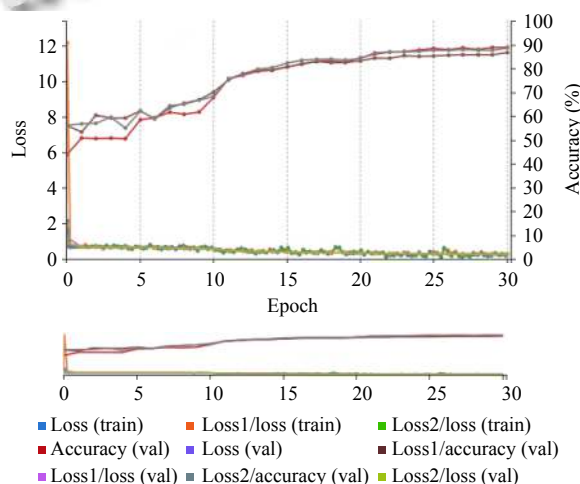


图 19 GoogleNet 网络结果

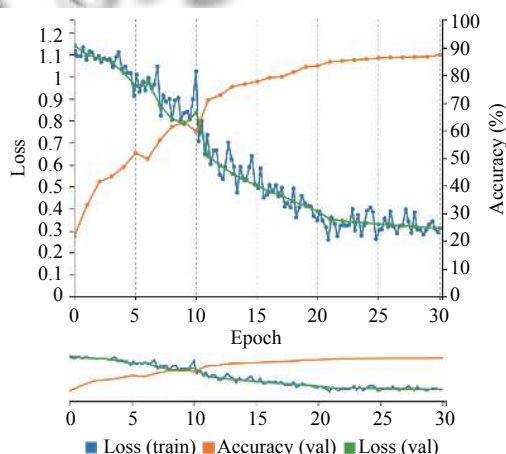


图 20 优化网络结果

### 4.2.3 MSA vs Normal (MN)

对于 MSA 和 Normal 即多系统萎缩病症和正常人



进行分类验证的实验中(如图 17 至图 19), 在原有 AlexNet 的实验基础上准确率获得了 0.3% 的提升, 验证集损失获得了 0.01 的降低, 训练集损失获得了 0.03 的降低。

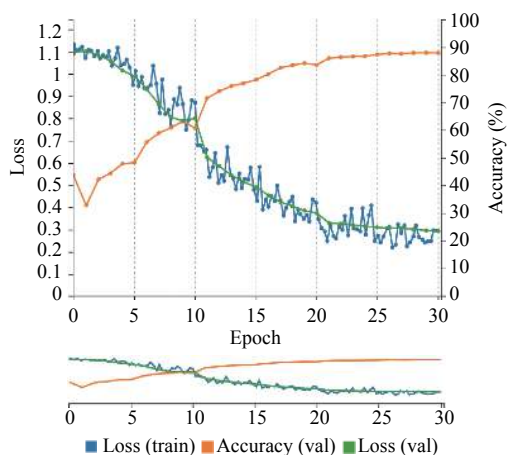


图 21 原 AlexNet 网络结果

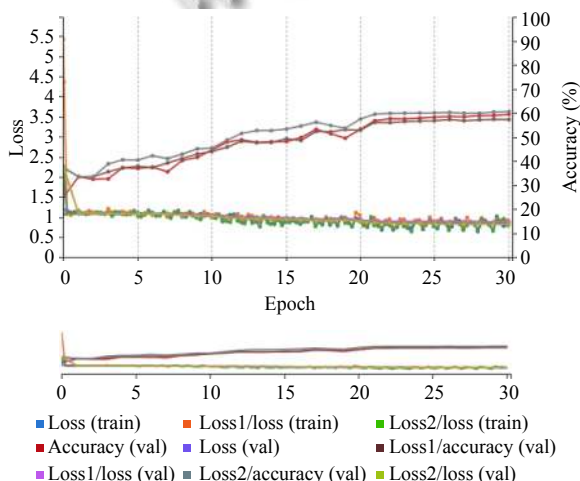


图 22 GoogleNet 网络结果

#### 4.2.4 PD vs MSA vs Normal (PMN)

对于 PD、MAS 和 Normal 即帕金森病症、多系统萎缩症和正常病人三种图像分类的实验中(如图 20 至图 22), 在原有 AlexNet 的实验基础上准确率获得了 0.6% 的提升, 验证集损失获得了 0.12 的降低, 训练集损失与原实验保持持平。

对照原 AlexNet 实验结果和优化网络的实验结果以及 GoogleNet 实验结果汇总如表 1 至表 3。

由上表数据分析, 优化的网络结构除了比原 AlexNet 网络结构较好之外, 通过和 GoogleNet 的 Acc 和 Loss 的相关数据对比可知, 优化的网络依然优

于 GoogleNet。

表 1 原 AlexNet 实验指标汇总

	PN	PM	MN	PMN
Acc (val)	98.2	84.7	91.3	87.4
Loss (val)	0.06	0.33	0.20	0.31
Loss (train)	0.11	0.27	0.23	0.29

表 2 优化网络实验指标汇总

	PN	PM	MN	PMN
Acc (val)	98.43	85.66	91.6	88
Loss (val)	0.05	0.33	0.19	0.29
Loss (train)	0.07	0.26	0.20	0.29

表 3 GoogleNet 实验指标汇总

	PN	PM	MN	PMN
Acc (val)	97.6	72.2	89	59.6
Loss (val)	0.07	0.54	0.26	0.84
Loss1/Acc (val)	94.6	75	86.8	57.5
Loss1/loss (val)	0.17	0.51	0.30	0.88
Loss2/Acc (val)	97.3	74.7	88.7	60.7
Loss2/loss (val)	0.09	0.51	0.27	0.82

## 5 结束语

本实验基于 AlexNet 神经网络结构设计了优化版的神经网络. 通过对改进的 AlexNet 网络和原始网络分别在四组数据中的实验结果进行分析, 改进版的 AlexNet 展示出了较好的分类效果, 并且优化版的神经网络和 GoogleNet 实验结果指标对比中依然占优势. 由于本实验数据量有限, 可能存在些许误差, 但本实验为今后的网络优化提供了参考因素, 并且对于医学图像的分类实现了全自动化, 避免了人工筛选造成的误差. 并且对 PD 的早期诊断和区别 PD 和 MSA 病症提供了研究意义。

## 致谢

特别感谢北京 301 医院提供的实验数据以及相关指导。

## 参考文献

- 1 Tsai CW, Tsai RT, Liu SP, *et al.* Neuroprotective effects of betulin in pharmacological and transgenic caenorhabditis elegans models of Parkinson's disease. *Cell Transplantation*, 2017, 26(12): 1903–1918. [doi: 10.1177/0963689717738785]
- 2 Scott L, Dawson VL, Dawson TM. Trumping neurodegeneration: Targeting common pathways regulated by autosomal recessive Parkinson's disease genes. *Experimental Neurology*, 2017, (298): 191–201. [doi: 10.1016/



- [j.expneurol.2017.04.008](#)]
- 3 朱亨焯. 大定风珠加味治疗帕金森病 48 例. 中国医药学报, 2001, 16(6): 75. [doi: [10.3321/j.issn:1673-1727.2001.06.028](#)]
  - 4 Levin J, Maaß S, Schuberth M, *et al.* Multiple system atrophy. In: Falup-Pecurariu C, Ferreira J, Martinez-Martin P, *et al.*, eds. Movement Disorders Curricula. Springer, Vienna. 2017. 183–192. [doi: [10.1007/978-3-7091-1628-9\\_17](#)]
  - 5 Fearnley JM, Lees AJ. Ageing and Parkinson's disease: Substantia nigra regional selectivity. *Brain*, 1991, 114(5): 2283–2301. [doi: [10.1093/brain/114.5.2283](#)]
  - 6 Gao P, Zhou PY, Wang PQ, *et al.* Universality analysis of the existence of substantia nigra “swallow tail” appearance of non-Parkinson patients in 3T SWI. *European Review for Medical and Pharmacological Sciences*, 2016, 20(7): 1307–1314.
  - 7 李坤成, 杨小平. 帕金森病的影像学诊断. *诊断学理论与实践*, 2005, 4(4): 273–274. [doi: [10.3969/j.issn.1671-2870.2005.04.005](#)]
  - 8 Al-Fatlawi AH, Jabardi MH, Ling SH. Efficient diagnosis system for Parkinson's disease using deep belief network. *Proceedings of 2016 IEEE Congress on Evolutionary Computation*. Vancouver, BC, Canada. 2016. 1324–1330.
  - 9 Hinton GE, Sejnowski TJ. Learning and relearning in Boltzmann machines. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge, MA, USA. 1986. 282–317.
  - 10 Shamir RR, Dolber T, Noecker AM, *et al.* Machine learning approach to optimizing combined stimulation and medication therapies for Parkinson's disease. *Brain Stimulation*, 2015, 8(6): 1025–1032. [doi: [10.1016/j.brs.2015.06.003](#)]
  - 11 LeCun Y, Bengio YA, Hinton G. Deep learning. *Nature*, 2015, 521(7553): 436–444. [doi: [10.1038/nature14539](#)]
  - 12 Deng L, Yu D. Deep learning: Methods and applications. *Foundations & Trends in Signal Processing*, 2014, 7(3-4): 197–387.
  - 13 Hubel DH, Wiesel TN. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, 1962, 160(1): 106–154. [doi: [10.1113/jphysiol.1962.sp006837](#)]
  - 14 张巧丽, 赵地, 迟学斌. 基于深度学习的医学影像诊断综述. *计算机科学*, 2017, 44(S2): 1–7.
  - 15 Zhang C, Woodland PC. Parameterised sigmoid and ReLU hidden activation functions for DNN acoustic modelling. *Sixteenth Annual Conference of the International Speech Communication Association*. Dresden, Germany. 2015. 3224–3228.
  - 16 Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *Proceedings of the 32nd International Conference on Machine Learning*. Lille, France. 2015. 448–456.
  - 17 Szegedy C, Liu W, Jia YQ, *et al.* Going deeper with convolutions. *Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition*. Boston, MA, USA. 2015.