

# 基于深度学习的糖尿病视网膜病变诊断方法研究

黄小如, 尚 勋, 何 飞, 谢毅飞

贵州大学, 大数据与信息工程学院, 贵州 贵阳

收稿日期: 2022年3月20日; 录用日期: 2022年4月21日; 发布日期: 2022年4月28日

## 摘 要

由于计算机技术以及人工智能的崛起和高速发展完善, 图像处理和计算机视觉一类的高新技术开始被应用到医疗诊断的领域。传统对糖尿病视网膜病变(糖网病)的诊断受到医疗资源分布不均、医生的个人主观因素等影响较大, 并且诊断过程耗时较长, 本项目是以研究基于深度学习神经网络的方法来预测视网膜是否有病变, 或者有轻度、高度、重度、增殖型病。预期通过研究经典的神经网络模型, 寻找一种基于深度学习智能算法的糖网病病变程度诊断方法, 并搭建本地PC端可视化界面, 使其具有易操作、普及度高的优势, 实现对糖尿病视网膜病变的初步预测, 辅助医生进行诊断治疗。

## 关键词

糖尿病视网膜病变(DR), 计算机视觉, 深度学习, 图像处理

## Research on the Detection Method of Diabetic Retinopathy Based on Deep Learning

Xiaoru Huang, Xun Shang, Fei He, Yifei Xie

College of Big Data and Information Engineering, Guizhou University, Guiyang Guizhou

Received: Mar. 20<sup>th</sup>, 2022; accepted: Apr. 21<sup>st</sup>, 2022; published: Apr. 28<sup>th</sup>, 2022

## Abstract

Due to the rise and rapid development of computer technology and artificial intelligence, new technologies such as image processing and computer vision are beginning to be applied to the

field of medical diagnosis. Traditional diagnostic methods of diabetic retinopathy (DR) are greatly affected by the uneven distribution of medical resources and doctors' personal subjective factors. In addition, the diagnostic process takes a long time. Based on the deep learning and neural network, this paper searches for the best way to predict the presence of lesions in the retina or mild, high-grade, severe, or value-added disease. By studying the classic neural network model, we look for a diagnostic method for the degree of diabetic retinopathy based on deep learning intelligent algorithm, and build a local PC-side visualization interface and physical display module, so that it has the advantages of easy operation and high popularity, realizes the preliminary prediction of diabetic retinopathy, and assists doctors in diagnosis and treatment.

## Keywords

Diabetic Retinopathy (DR), Computer Vision, Deeping Learning, Image Processing

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 课题研究背景及意义

糖尿病视网膜病变(diabetic retinopathy, DR)也称“糖网病”，是糖尿病最常见的微血管并发症之一，是慢性进行性糖尿病导致的视网膜微血管渗漏和阻塞而引起一系列的眼底病变，也是当今世界上造成视力损害和失明的主要原因。糖网病如果没有及时发现并治疗，任其发展到后期会导致患者的视力急剧下降，最终致患者失明。在临床上诊断糖尿病视网膜病变的传统方法是进行全面的眼科检测，检测流程复杂，需要耗费一定的人力物力。在早期筛查过程中，如果能对眼底图像进行自动分类，达到判断视网膜是否有病变以及针对轻度、高度、重度、增殖性病变进行筛选，就可以解决现存检测办法既费时又耗力，可能延误患者病情，检测效果不好等问题。目前主要依靠人工依次评估每幅视网膜图像来检测糖尿病视网膜病变，对医生的要求高并且所费时间较长，显然不利于患者的尽快就诊和治疗。随着科技化时代的到来，深度学习的重要性和优势不言而喻，而目前关于深度学习在众多其他行业的应用尚且不饱和，具有较高的研究价值[1]。

本项目采用深度学习卷积神经网络(CNN)技术研究筛选糖网病病变类型的有效分类方法，解决人工诊断的高耗时和检测效果不够好的问题，从而提高糖网病的诊断效率，提高糖尿病视网膜病变的筛选正确率，让尽可能多的糖尿病患者远离失明危险。

## 2. 卷积神经网络

### 2.1. 神经网络

在传统上，人类的认知是基于信息中的模式；这些模式可以被表示成符号，并且通过这些符号，使用逻辑规则进行计算与推理。基于统计的连接实现的模型是从脑神经科学中获得启发，实图将认知所需要的功能属性结合到模型中来，通过模拟生物神经网络的信息处理方式来构建具有认知功能的模型。主要特点有：

- 1) 拥有处理信号的基本单元。
- 2) 处理单元之间连接的方式为并行连接。

3) 每一个处理单元之间的连接都伴随一个连接权重值。

这就是人工神经网络的基础模型。

神经元：神经元是基本的信息炒作和处理单元。它负责接受一组输入，并将这组输入与权重相乘后再求和，最后由激活函数来计算这个神经元的输出。基本神经元结构如图 1。

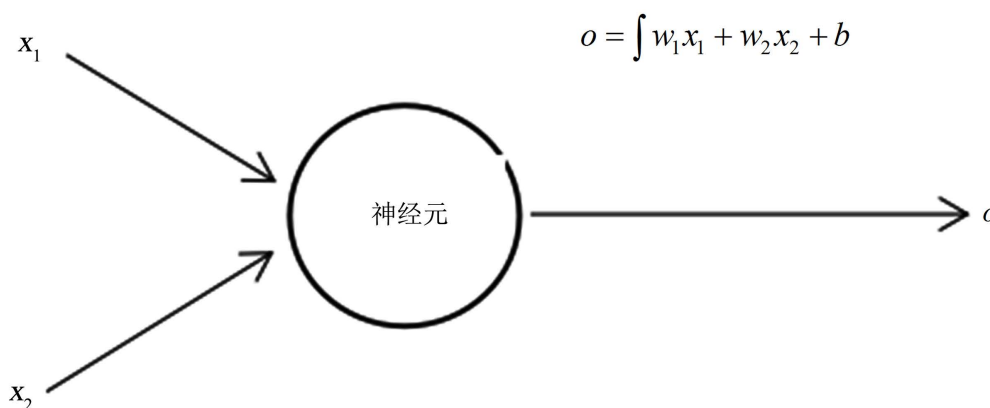
输入：一组张量可以作为一个神经元的输入： $x = \{x_1, x_2, \dots, x_n\}$ 。

连接权值：连接权值向量为一组张量  $W = \{w_1, w_2, \dots, w_n\}$ ，其中对应输入  $x_i$  的连接权值；神经元将输入进行加权求和：

偏置：在有的时候甲醛求和时会加上一项常数项  $b$  作为偏执：其中张量  $b$  的形状要与  $w_x$  的形状保持一致。

$$\text{sum} = W_x + b \quad (1)$$

激活函数：激活函数  $f(\cdot)$  被施加到输入加权和  $\text{sum}$  上，产生神经元的输出；如  $\text{sum}$  是大于二阶的张量，则  $f(\cdot)$  被施加到  $\text{sum}$  的每一个元素上。



**Figure 1.** Basic neurons  
**图 1.** 基本神经元

$$o = f(\text{sum}) \quad (2)$$

## 2.2. 卷积神经网络

一般来说，卷积神经网络(Convolutional Neural Network, CNN)是由卷积层、采样层(池化层)、全连接层(非线性激活函数层)组成[2]。如图 2。多层神经层组成神经网络，每层神经层又包含众多神经元，而神经元就是神经网络识别事物的关键。由于全连接神经网络在处理大量数据时，参数量巨大，这导致训练十分缓慢，泛化性差。为了实现利用神经网络识别大量数据的目的，卷积神经网络应运而生。异于常规神经网络的是，卷积神经网络的各层中的神经元是 3 维排列的：宽度、高度和深度。就图像识别来说，神经网络对图片上每一小块像素区域进行加权，从而取代对每个像素的输入信息做处理，其中的小块区域叫做局部感受野，区域的权值叫做卷积核[3]。过滤器通过持续不断的在图片上滚动来收集一块块像素区域内图片里的信息，重复上述操作，神经网络就能利用边缘信息总结出更高层的信息结构，最后再把信息套入普通的全连接神经网络进行分类。例如图 3 的例子，卷积神经网络收集到的边缘信息使神经网络从线条看到部分五官再到全部面部。

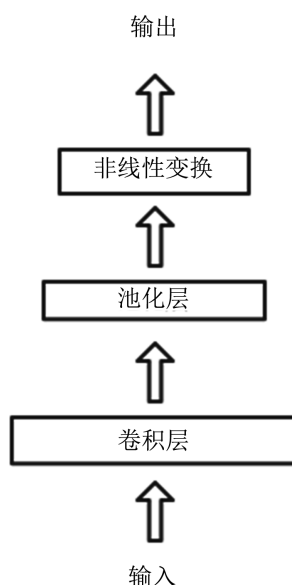


Figure 2. Structure of convolutional neural networks  
图 2. 卷积神经网络结构

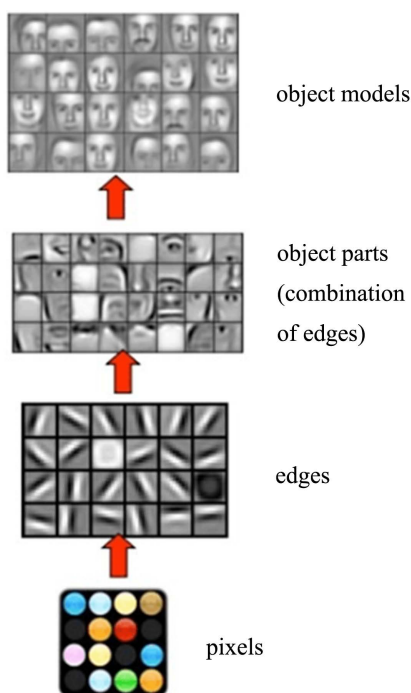


Figure 3. An example of a convolutional neural network  
图 3. 卷积神经网络实例

图片有长、宽、深度(计算机产生颜色所使用的信息, 即 RGB 通道: 红、绿、蓝三种颜色通道)三个参数。如若是黑白照片, 则深度为 1。研究表明, 在每次卷积的时候, 神经层存在丢失信息的问题, 为了解决该问题提出了池化层(pooling), 即卷积的时候不压缩长、宽, 从而尽量地保留更多信息, 交由池化层来负责压缩, 从而提高准确性, 提高计算速度, 同时提高所提取特征的鲁棒性。最后由全连接层进行分类和输出。

### 2.2.1. 输入层

卷积神经网络的输入层可以处理多维数据。输入层的主要作用是输入图像和进行卷积操作。为保障特征提取的有效性，需要对图像进行预处理，包括通道分离，亮度归一化、对比度增强等操作。原始图像像素大，训练时对内存要求高，所以需要设置图像的大小，比如，常用的就是将图片大小设置为  $224 \times 224$  的像素大小。也可以对图像进行平移、旋转、裁切等达到数据增强的目的。

### 2.2.2. 卷积层

先了解卷积操作的运算过程。用二维数组来解释一次卷积操作的运算过程，如图4。将  $3 \times 3$  矩阵按照一定的步骤与左边矩阵相对应位置进行卷积操作，得出最后的结果。在此次操作过程中，等号左边  $3 \times 3$  矩阵叫做卷积核(Convolutional kernel)或者滤波器(Filter)，卷积核的大小不确定。卷积层的作用是提取输入数据的特征[4]。图4中左边的矩阵为输入的矩阵，而等号之后的一则为输出的矩阵，输出的矩阵作为下一层卷积的输入。卷积步长(Strided convolutions)指的是卷积核扫描两次相邻特征图时位置之间的距离。卷积核上的参数完成对卷积层的特征提取，因此，卷积核的参数决定该卷积层的作用[5]。在卷积层中，输出特征图的大小由卷积核的大小、步长和填充三者共同决定。卷积核尺寸越大，可提取的输入特征越复杂[6]由上述例子可以看出，输入的图像为  $5 \times 5$  大小的矩阵，卷积核是  $3 \times 3$  的矩阵，将过卷积操作之后的输出矩阵的大小为  $3 \times 3$ ，由此可见，随着卷积核的交叉计算和网络层数的增加，特征图的尺寸会随之减小。为了避免图像尺寸减小，对图像进行填充。常见的矩阵填充方法有0填充(Zero Padding)和重复边界值填充(Replication Padding) [7]。输出图像大小的公式为：

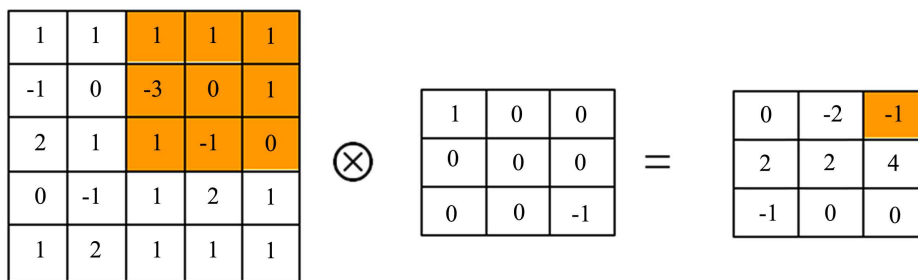


Figure 4. The process of two-dimensional convolution

图4. 二维卷积的过程

$$o = \lceil (1 - k + 2p) / s \rceil + 1 \quad (3)$$

其中，图像输出为  $O$ ，输入为  $i$ ，代表卷积核大小， $s$  表示步长。执行卷积操作的位置叫做卷积层。卷积神经网络的计算可涉及多维。其中，二维图像的卷积计算公式如下所示：

$$O^*(i, j) = \sum_x \sum_y O(i+x, j+y) \times w(x, y) \quad (4)$$

其中， $o^*(i, j)$  是特征图中位于  $(i, j)$  处的像素值， $w(x, y)$  表示卷积核位于  $(x, y)$  的权重值。

### 2.2.3. 激活函数

在卷积层处理图像的过程中，对每个像素点赋予了一个权值，这个操作是线性的。但是样本不一定是线性可分的，为此引入了非线性因素，解决线性模型所不能解决的问题。即引入激活函数来增加非线性从而赋予神经网络表达更复杂特征的能力。激活函数具有单调性和可微性。单调性保证了单层网络模型具有凸函数的性能，可微性是指可以使用误差梯度对模型权重进行微调[8]。常见的激活函数有 Sigmoid、Tanh 和 ReLu 三种，图像分别为图5~7。其中 Sigmoid 函数的表达式为：

$$f(x) = \frac{1}{1+e^{-x}} \quad (5)$$

Sigmoid 函数和 Tanh 函数图像的对称点分别为(0, 0.5)和(0, 0)。目前最常用的激活函数是 ReLu 函数，其计算公式如下式：

$$f(x) = \max(0, x) \quad (6)$$

ReLu 激活函数在大于零的部分，保持原始值不变，在小于零时值为零，解决了正区间梯度消失问题，也减少了运算量。

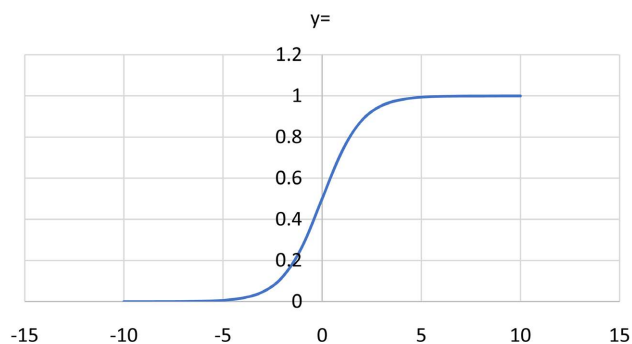


Figure 5. Function image of Sigmoid

图 5. Sigmoid 函数图像

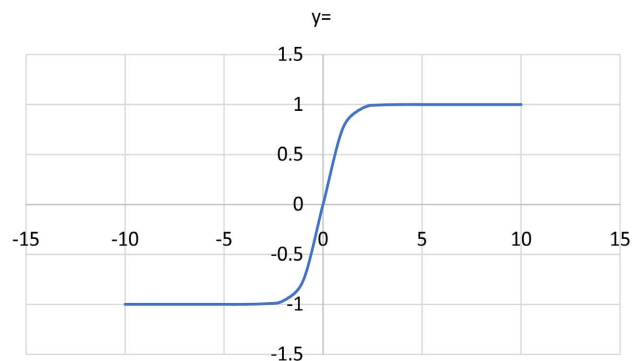


Figure 6. Function image of Tanh

图 6. Tanh 函数图像

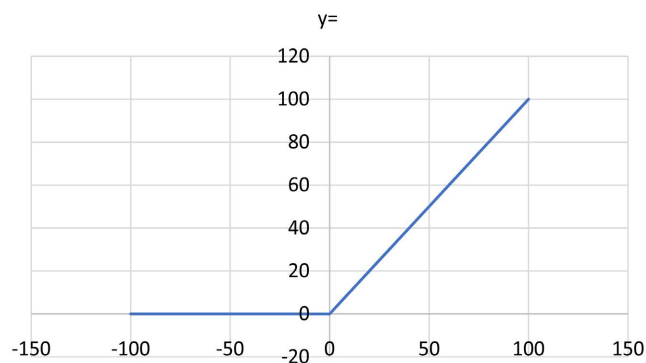
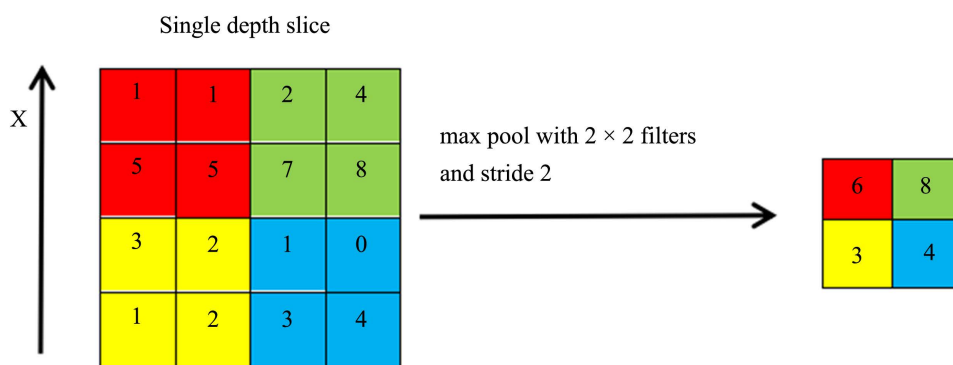


Figure 7. Function image of ReLu

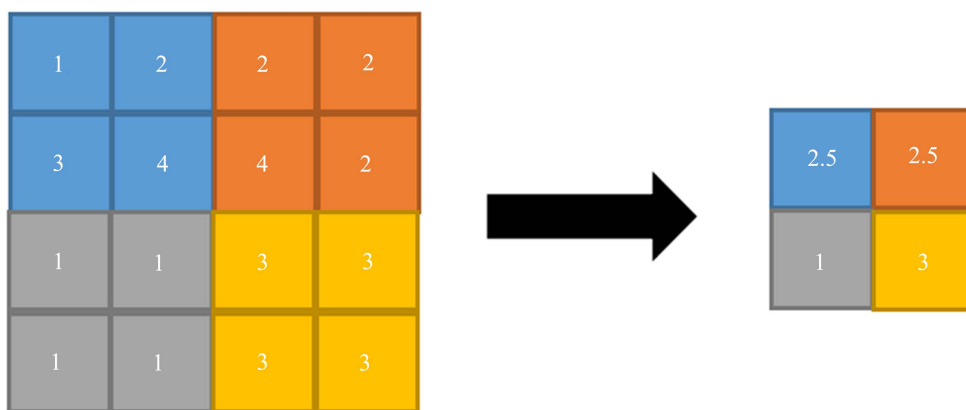
图 7. ReLu 函数图像

### 2.2.4. 池化层

经过卷积层的特征提取之后, 需要进一步通过池化来进行特征的选择与信息的过滤, 池化主要负责进行下采样降维操作。下采样层一般没有权重更新, 常有的下采样有提取局部平均值或者最大值作为提取的特征, 这样做的目的是降低特征空间的维度, 之抽取局部最显著的特征, 同时这些特征出现的具体位置也被忽略。常见的池分最大池化(max-pooling)、平均池化 mean-pooling 两种, 其中, 最大池化指的是取局部接受域中值最大的点。如图 8, 对一个  $4 \times 4$  的特征图(feature map)领域内的值, 用一个  $2 \times 2$  的过滤器(filter), 以 2 为步长进行图片特征的扫描和提取, 每一块  $2 \times 2$  的区域都取区域内的最大值, 整理后将提取出的最大值输出到下一层, 这种做法能有效缩小特征图的尺寸, 能很好的保留纹理特征。平均池化指的是将图片按照固定大小网格分割, 网格内的像素值取网格内所有像素的平均值作为特征值。例如图 9, 对于一个  $4 \times 4$  的特征图(feature map), 同样用一个  $2 \times 2$  的过滤器(filter), 以 2 为步长进行图片特征的扫描和提取, 将每一块  $2 \times 2$  区域中四个值计算平均值并整理, 输出到下一层。相比于最大池化, 平均池化能更好的保留背景。



**Figure 8.** The process of maximizing pooling  
**图 8.** 最大池化过程



**Figure 9.** The process of averaging pooling  
**图 9.** 平均池化过程

### 2.2.5. 全连接层

全连接层在卷积网络的最后一层, 经过上述的一系列操作之后, 通过卷积神经网络对结果进行分类, 起到分类器的作用。全连接层的实质就是把卷积操作后的多通道图像转换成一维图像的过程。由于全连接层存在参数过多, 易导致过拟合的问题, 所以应该尽量避免使用全连接层而改用全局平均值进行替换。



### 3. 基于 ResNet 算法的糖尿病性视网膜病变区域检测

#### 3.1. ResNet 网络算法思想

ResNet 在 2015 年被微软研究院的提出。提出，在检测、分割、识别等领域被广泛使用。随着计算机行业的不断发展与变革，网络的深度也随之在不断地增加。但相关研究调查结果显示，随着网络的加深，反而出现了训练集准确率下降的现象，如图 1 所示。出现这种情况的原因是卷积网络是由输入、隐藏层、输出三部分组成的，而要确定某一训练模型的效果，就是通过训练后数据集的准确度或通过求取偏差值来判断的。然而，网络的不断迭代以及参数的不断调整改变，使得输出结果的偏差值越来越小。在反向传播的过程中，神经网络需要不断传播梯度，而网络层数的增加使得梯度在传播过程中出现梯度消失、爆炸和深层网络的退化问题的概率增加，导致深度网络模型无法达到预期效果。在此背景下，微软研究院的何恺明、张祥雨、任少卿、孙剑等人在 2015 年提出了 ResNet 网络，即深度残差网络[9]。它能够使网络尽可能加深的同时保证模型的训练不会出现偏差值上升的现象。利用同余映射的思想，将上层的输出直接作为下层的输入。其中引入了全新的结构如图 1：将深层网络的后面若干层学习成恒等映射  $h(x)=x$ ，那么模型就退化为浅层网络，也就避免了深层网络的退化问题。把网络设计成：

$$H(x) = F(x) + x \Rightarrow H(x) - x \tag{7}$$

其中  $F(x)$ 为残差，那么当残差  $F(x)$ 为 0 时， $H(x) = x$ ，也就构成了上面所说的恒等映射，实际上残差不会为 0，这保证了堆积层在输入特征的基础上学习到新的特征，从而拥有更好的性能。残差学习的结构如图 10 所示， $F(x) + x$  构成的 block 称作残差块(Residual Block)，多个残差块串联形成 ResNet， $x$  路径称为 shortcut， $F(x)$ 称为残差路径。短路连接(shortcut connection)与前向神经网络相链接。

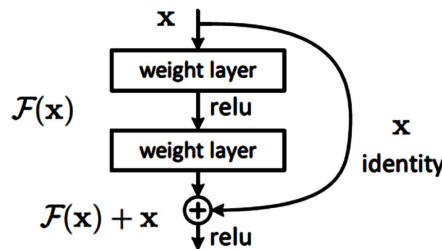


Figure 10. Residual learning units  
图 10. 残差学习单元

残差路径可分为如图 11 所示的两种情况，即 bottleneck 和 basic block。区别在于 bottleneck 为了降低计算的复杂度增加了一层  $1 \times 1$  的卷积层。Shortcut 路径也大致可分为两种，如图 12。左图的路径将输入  $x$  原样输出，右图的路径通过最右边的  $1 \times 1$  卷积来升降维。

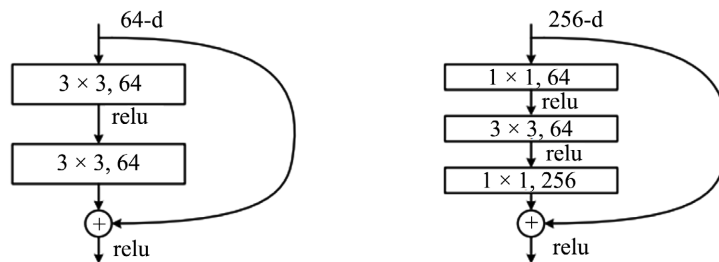


Figure 11. Residual path  
图 11. 残差路径



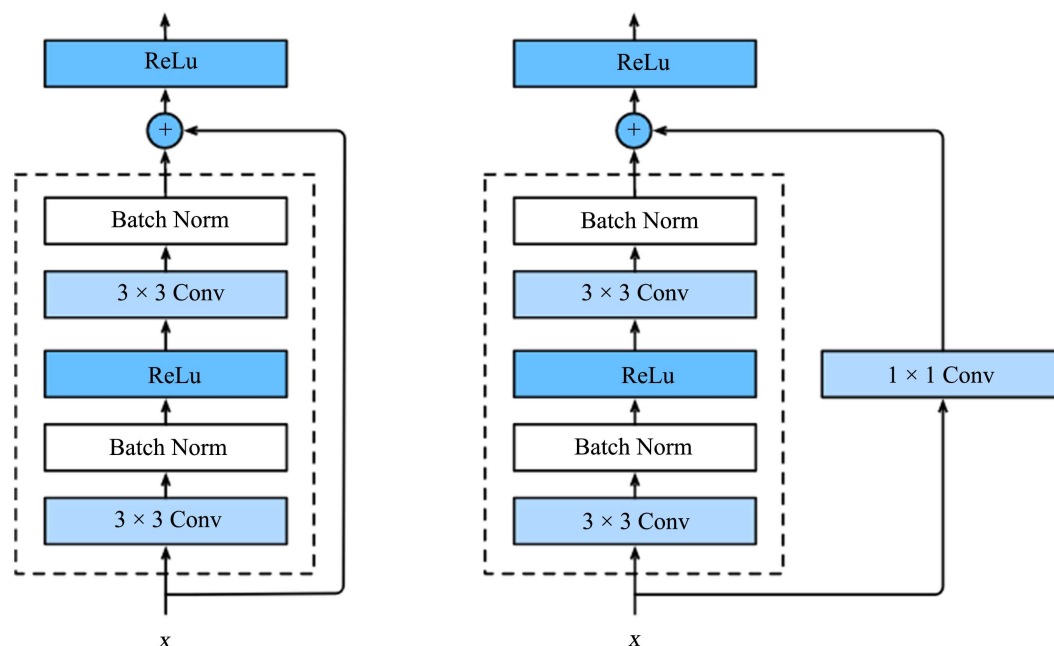


Figure 12. Shortcut path

图 12. 短路路径

ResNet 网络探求更易于优化的模型结构，通过短路机制加入了残差单元。ResNet 直接使用  $\text{stride} = 2$  的卷积做下采样，并且用 global average pool 层(全局池化层)替换了全连接层。ResNet 的一个重要设计原则是：当 feature map 大小降低一半时，feature map 的数量增加一倍，这保持了网络层的复杂度[10]。图 13 展示了 ResNet-34 与 34-layer plain net 和 VGG 三种网络结构的对比。ResNet 通过残差学习解决了深度网络的退化问题，让训练更深的网络成为可能[11]。通过堆叠不同数量的残差块可以得到多种不同的 ResNet。

## 3.2. 基于 ResNet 算法的检测实验结果与分析

### 3.2.1. 实验数据说明

本实验将在检测糖尿病视网膜病变的数据库进行训练和测试。数据库有彩色眼底图像 44,349 张，正常图像 29,804 张，含有不同病变程度特征的图像 14,545 张。数据集根据视网膜病变的程度划分为正常、轻度、中度、重度和增殖型糖尿病视网膜病变五个等级，其病变特征主要表现为微动脉血管瘤、硬性渗出、棉絮斑、出血点等[12]。0 对应于正常健康状态，4 是最严重的状态。经过数据增强之后，该数据集一共有 35,108 张图片，正常的有 25,802 张，轻度病变 2438 张，中度病变 5288 张，重度病变 872 张，增殖型病变 708 张。

### 3.2.2. 数据预处理

通过对图像的翻转、旋转、裁剪、改变光照强度等方式，每一种的处理方式都使得数据集增加一倍[13]，对图像进行数据增强后的效果图如图 14 所示。

### 3.2.3. 实验环境及参数设置

对实验数据集的参数设置：输入图像的大小为  $224 \times 224$ 。数据批处理的大小设置为 15。开始时，将学习率的大小设置为 0.0001。实验环境及配置如表 1：

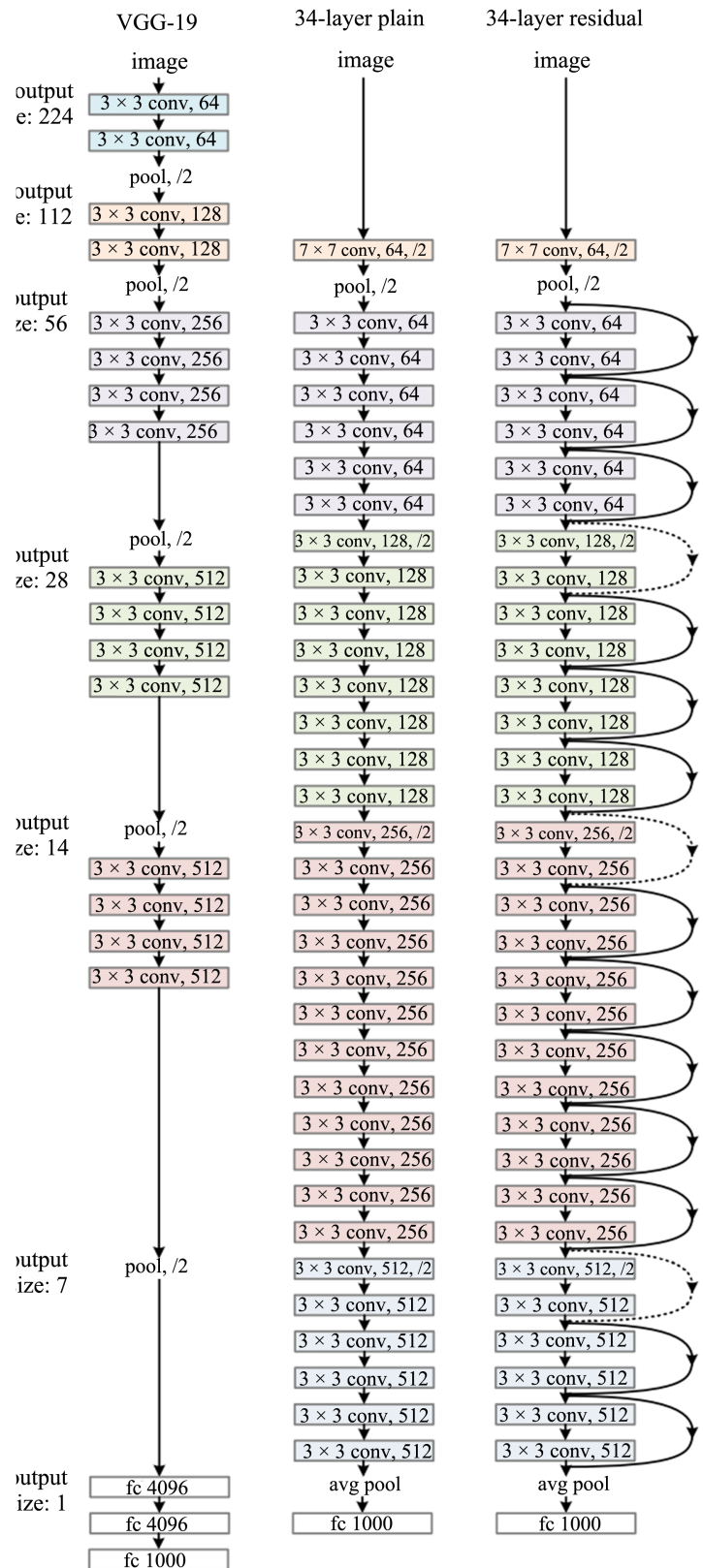
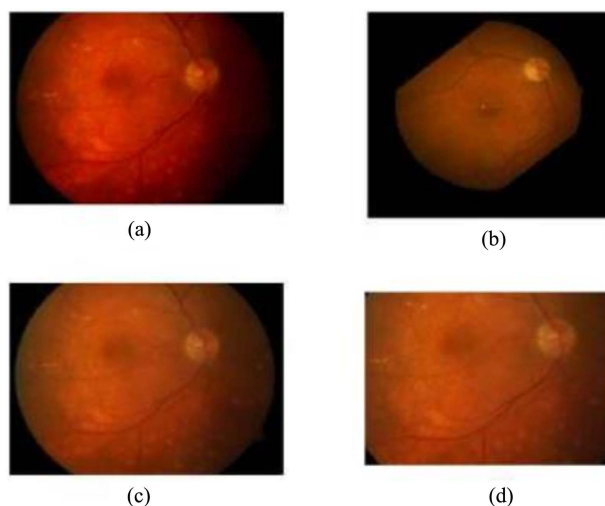


Figure 13. Comparison of three network structures  
图 13. 三种网络结构对比



**Figure 14.** (a) Original image, (b) Image rotation, (c) Increased image contrast, (d) Image random clipping  
**图 14.** (a) 原始图像, (b) 图像旋转, (c) 图像对比度增加, (d) 图像随机裁剪

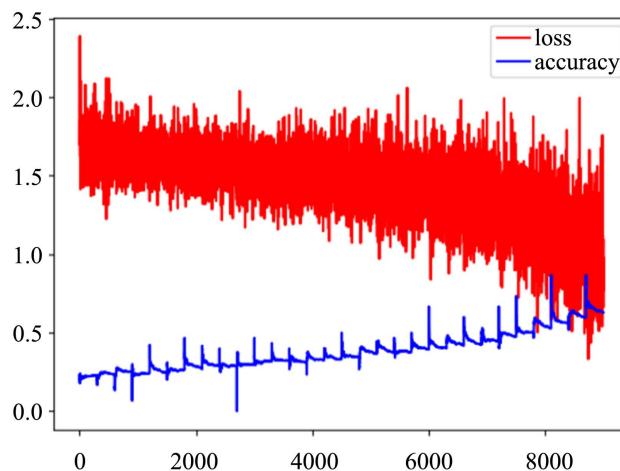
**Table 1.** Experimental environment and configuration

**表 1.** 实验环境及配置

实验环境	配置
操纵系统	Windows11
CPU	Intel CORE i5 9 <sup>th</sup> Gen
GPU	NVIDIA GEFORCE GTX 4G
内存	8G
深度学习框架	Pytorh
编程语言	Python3.7

### 3.2.4. 实验结果

如图 15。



**Figure 15.** Loss rate and accuracy curve of ResNet50

**图 15.** ResNet50 损失率与准确率变化曲线图

## 4. 基于 MobileNetV3 算法的糖尿病性视网膜病变区域检测方法研究

### 4.1. MobileNetV3 (轻量化卷积神经网络)算法思想

MobileNet 是 Google 公司推出的轻量化系列网络，2019 年，Google 发布了第三代 MobileNet，即 MobileNetV3 [14]。MobileNetV3 主要的贡献就是使用科学的理论重新设计了网络结构，并且引入了 H-Swish 激活函数与 Relu 搭配使用。另外在网络中还引入了 Squeeze-And-Excite 模块。MobileNetV3 中加入了 SE 模块，即注意力机制，其次是更新了激活函数，如图 16。注意力机制就是针对每一个 channel 进行池化处理，得到 channel 个数个元素，通过两个全连接层，得到输出的这个向量。第一个全连接层的节点个数等于 channel 个数的 1/4，第二个全连接层的节点和 channel 保持一致，得到的输出就相当于对原始的特征矩阵的每个 channel 分析出其重要程度，越重要的赋予越大的权重，反之越小。如下图 17 所示，先采用平均池化将每一个 channel 变为一个值，然后经过两个全连接层之后得到通道权重的输出，第二个全连接层使用 Hard-Sigmoid 激活函数，然后将通道的权重乘回原来的特征矩阵就得到了新的特征矩阵。

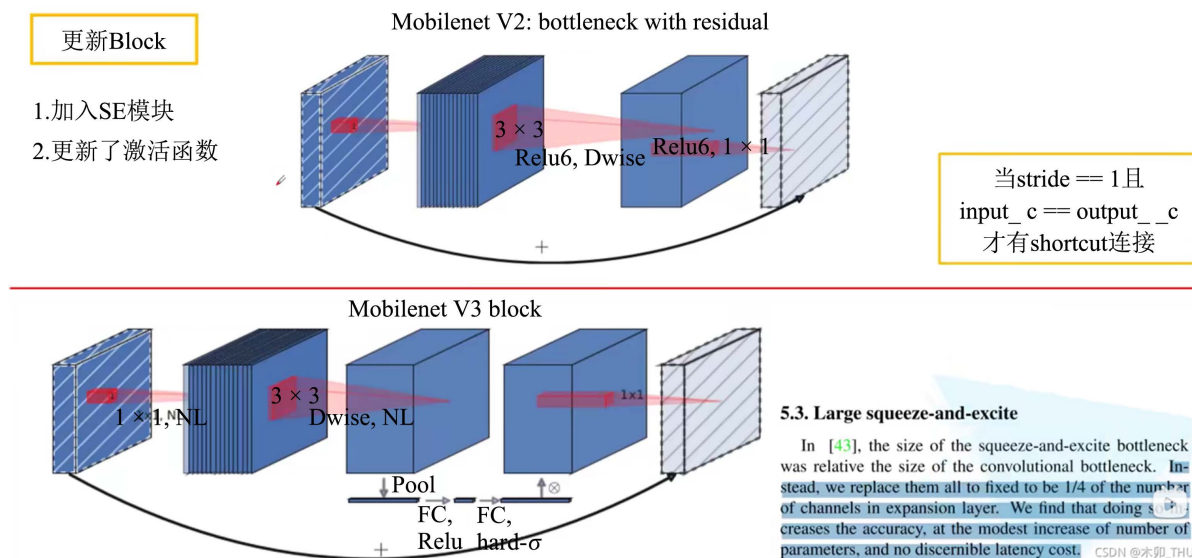


Figure 16. The renewal of the MobileNetV3

图 16. MobileNetV3 的更新

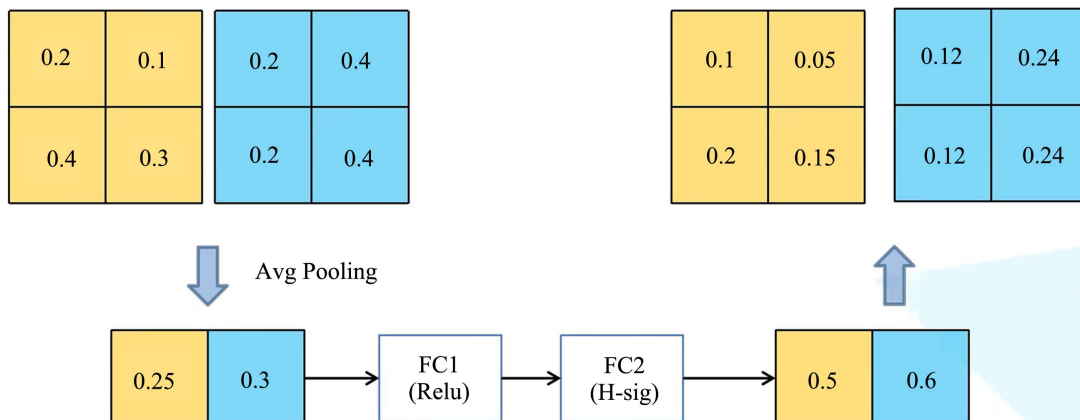


Figure 17. Squeeze-and-excitation network

图 17. 注意力机制

激活函数部分, 在 MobileNetV3 出现之前, 比较常用的是 swish 激活函数, 即  $x$  乘上 sigmoid 函数, 使用此方法可以提高网络的准确率, 但其计算和求导时间复杂, 且对量化过程不友好, 特别是对于移动设备, 一般情况下为了加速都会进行量化操作。为了解决计算复杂量化难度高的问题, 轻量级网络的作者提出了一个叫做 h-swish 的激活函数。由于 swish 是  $x$  乘上 sigmoid, 故 h-swish 是  $x$  乘上 h-sigmoid。h-sigmoid 函数计算公式如下式:

$$\text{H-sigmoid} = \text{ReLU6}(x+3)/6 \quad (8)$$

其中  $\text{ReLU6}(x) = \min(\max(x, 0), 6)$ 。

## 4.2. 基于 MobileNetV3 算法的检测实验结果与分析

### 实验结果

实验环境及参数设置与表 1 相同。实验结果如下图 18:

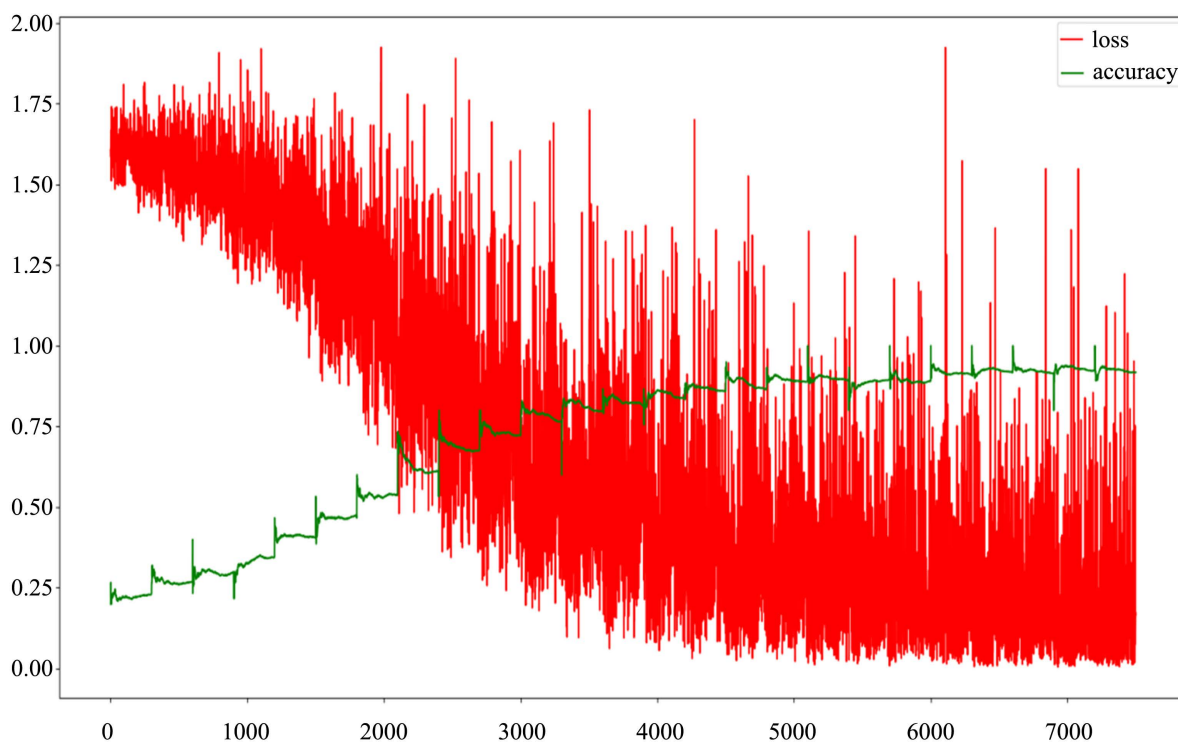


Figure 18. Loss rate and accuracy curve of MobileNetV3

图 18. MobileNetV3 损失率与准确率变化曲线

分析: 使用 MobileNetV3 模型训练所得的结果明显比用 ResNet 模型训练效果好得多, 在这个模型迭代 25 次时损失率已经下降到了 0.09 附近, 准确率也达到了 0.91 附近, 通过损失率变化曲线可以看出在从刚开始的损失率为 1.6 左右下降到 0.75 值的附近时损失率的下降开始变缓, 当继续训练下降到 0.25 附近时开始趋于平稳, 模型训练的结果较为可观。

鉴于 mobilenetv3 的模型收敛比较快, 以下数据全部采用 mobilenetv3 进行研究和测试。

测试集测试效果: 第一次测试结果如图 19, 第二次测试结果如图 20, 病变类型及对应图片数量如表 2。

分析: 通过两次对训练好的模型进行测试测试, 此模型在测试集上的预测准确率几乎保持在 0.73 左右, 如果增加迭代次数, 模型的预测效果会更加准确。

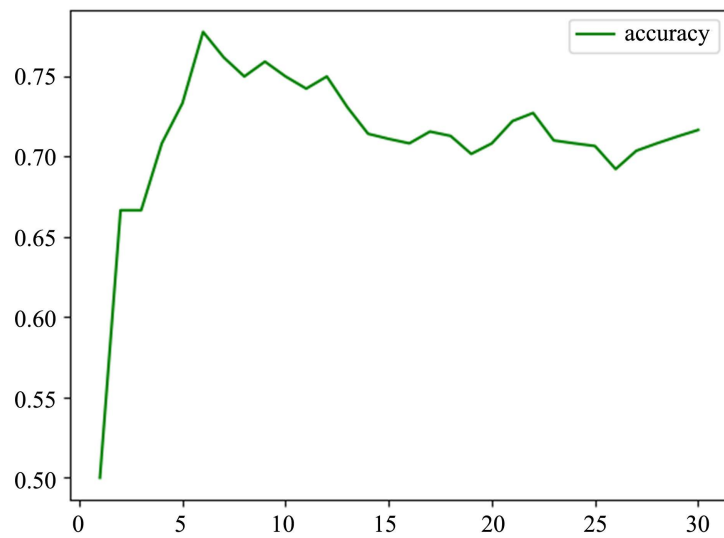


Figure 19. The results of the first test

图 19. 第一次测试结果

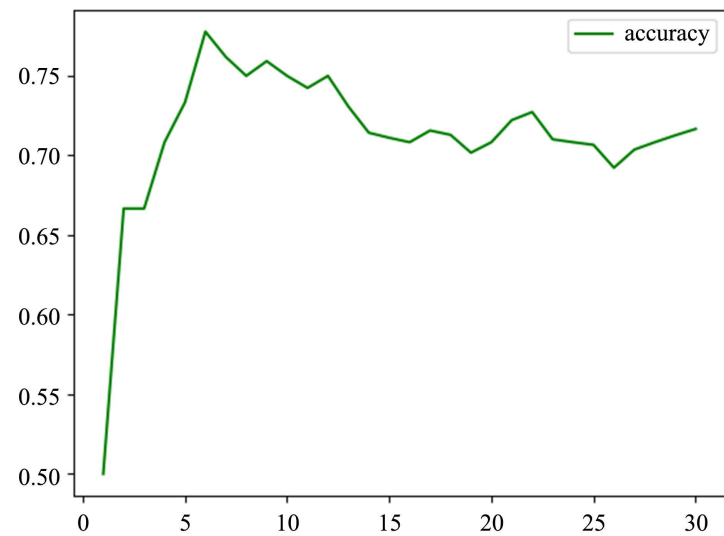


Figure 20. The results of the second test

图 20. 第二次测试结果

Table 2. Lesion type and number of images

表 2. 病变类型及图片数量

病变类型	数量(张)
无病(0)	50
初级(1)	40
中级(2)	40
严重(3)	40
增值型(4)	40
总计	210

## 5. 可视化平台搭建

可视化界面的作用主要是加强计算机和人之间的交互，一个操作便捷的可视化界面有益于用户更加轻松的使用。如下图是基于 HTML5 与 CSS 搭建的人机交互网页，通过上传待预测的眼底图像，后端获取眼底图像后使用模型进行预测，然后返回预测结果。给用户提供一个提交眼底图片的平台，并将检验结果快速准确的展示在电脑终端。网页提供历史查询、专家咨询和快速就诊功能，能实现用户自行上传眼底图片之后返回预测结果，查询账户历史使用记录以及远程与专家交流。网页如下图 21：

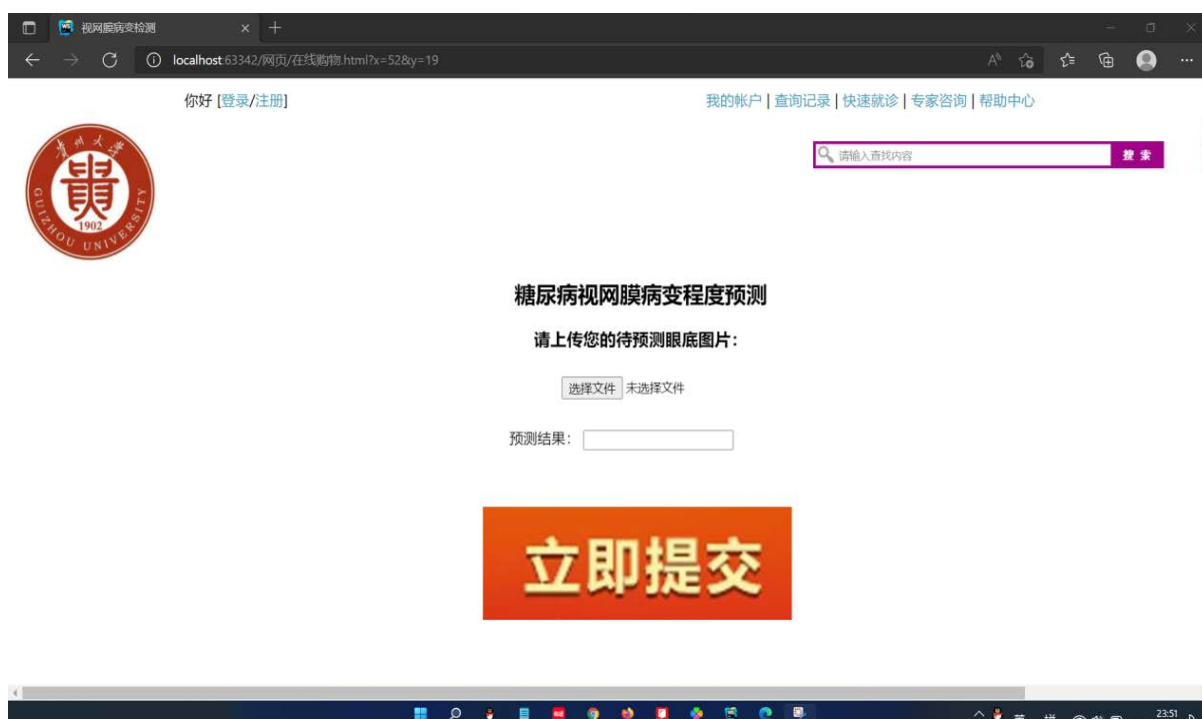


Figure 21. Visual platform

图 21. 可视化平台

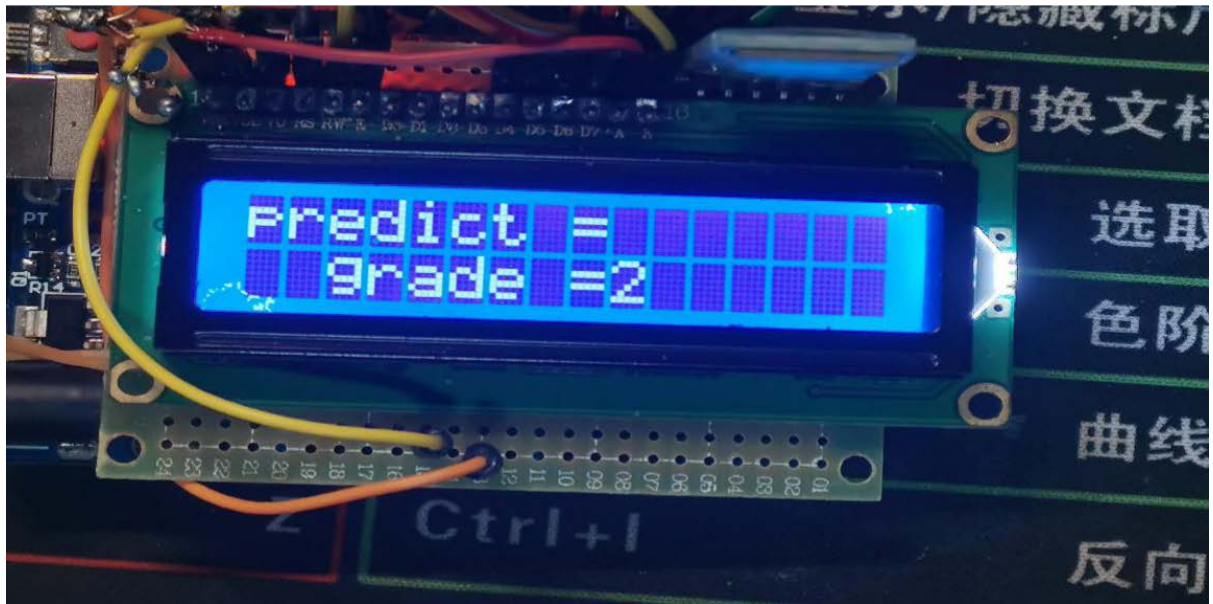
## 6. 实物设计与实现

在本次研究中，为了能够更加直观地观测到模型预测的结果，设计了一款基于嵌入技术的辅助显示硬件单元，该硬件主要包含了蓝牙接受模块、液晶显示模块、和主控模块，蓝牙模块通过蓝牙与 PC 端进行连接，通过串口数据透传实现数据在单片机和电脑之间的传递。蓝牙模块采用的是 HC05 蓝牙模块、显示模块使用的是 1602 液晶显示器、主控模块使用的是 Arduino 控制板。这款辅助显示单元会在 PC 端的代码运行完成以后将预测的分类结果显示在液晶显示屏上面、分别是数字 0 代表无病、数字 1 代表轻度患病、数字 2 代表高度患病、数字 3 代表重度患病、数字 4 代表增值型。其实物如下图 22、图 23 所示。

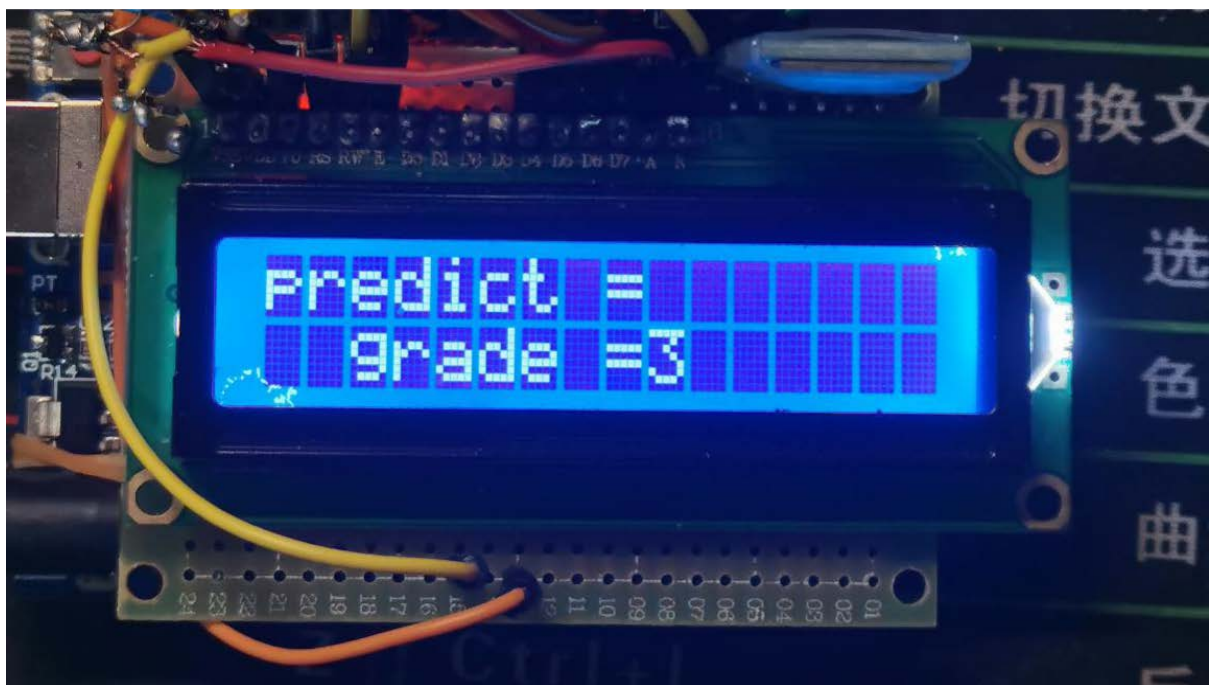
## 7. 总结

本文主要通过深度卷积神经网络算法的学习，对糖尿病视网膜病变的检测与分类进行了研究。首先，就糖尿病视网膜病变的检测与分类的现状背景及研究意义进行阐述。其次，使用并比较了 ResNet 网络算法和 MobileNetV3 两种算法的实验结果和训练效果。最后，搭建了电脑终端的可视化页面以及完成了





**Figure 22.** The classification grade was 2; corresponding to the predicted severity of the disease is moderate  
**图 22.** 显示分类等级为 2，对应预测患病程度为中级



**Figure 23.** The classification grade was 3; corresponding to the predicted severity of the disease is severe  
**图 23.** 显示分类等级为 3，对应预测患病程度为严重

实物的设计和实现。本文研究的主要工作如下：建立一套端对端、高效、自主学习的视网膜图像特征分析、提取、分类识别深度卷积神经网络算法。利用 DCNN 实现对视网膜图像特征的自主学习提取，对提取的特征进行转换融合，最后实现对病变细胞的五分类诊断，并通过 Caffe 这种基于卷积神经网络框架的开源深度学习平台验证算法的有效性，来确定适合糖尿病视网膜病变诊断的 DCNN 网络模型最优结构参数。

## 参考文献

- [1] 庞浩, 王枏. 用于糖尿病视网膜病变检测的深度学习模型[J]. 软件学报, 2017, 28(11): 3018-3029.
- [2] 李策, 陈海霞, 汉语, 左胜甲, 赵立刚. 深度学习算法中卷积神经网络的概念综述[J]. 电子测试, 2018(23): 61-62.
- [3] 张慧. 基于深度网络架构的椭圆检测[D]: [硕士学位论文]. 大连: 大连理工大学, 2019.  
<https://doi.org/10.26991/d.cnki.gdllu.2019.001980>
- [4] 邓瑞. 基于深度学习的建筑物识别[D]: [硕士学位论文]. 重庆: 重庆邮电大学, 2019.  
<https://doi.org/10.27675/d.cnki.gcydx.2019.000210>
- [5] Howard, A., Sandler, M., Chu, G., Chen, L.-C., Chen, B., Tan, M.X., Wang, W.J., Zhu, Y.K., Pang, R.M., Vasudevan, V., Le, Q.V. and Adam, H. (2019) Searching for MobileNetV3. CoRR.
- [6] 杨子贤. 基于深度学习卷积神经网络图像识别技术的研究与应用[J]. 信息与电脑(理论版), 2020, 32(8): 41-43.
- [7] 彭博. 基于深度学习的遥感图像道路信息提取算法研究[D]: [硕士学位论文]. 成都: 电子科技大学, 2019.
- [8] Albawi, S., Mohammed, T.A. and Al-Zawi, S. (2017) Understanding of a Convolutional Neural Network. 2017 *International Conference on Engineering and Technology (ICET)*, Antalya, 21-23 August 2017, 1-6.  
<https://doi.org/10.1109/ICEngTechnol.2017.8308186>
- [9] 张猛, 钱育蓉, 杜娇, 范迎迎. 基于改进的 LeNet-5 卷积神经网络交通标志的识别[J]. 东北师大学报(自然科学版), 2020, 52(1): 92-97. <https://doi.org/10.16163/j.cnki.22-1123/n.2020.01.018>
- [10] 唐旭, 谭继文, 徐卫晓, 井陆阳. 基于卷积神经网络的数控机床滚珠丝杠副故障诊断研究[J]. 煤矿机械, 2019, 40(1): 141-143. <https://doi.org/10.13436/j.mkjx.201901050>
- [11] 林宇, 陈晓勇. 基于 inceptionv3 模型的道路交通标志识别研究[J]. 江西科学, 2018, 36(5): 849-852.  
<https://doi.org/10.13990/j.issn1001-3679.2018.05.026>
- [12] He, K., Zhang, X., Ren, S., et al. (2016) Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, 27-30 June 2016, 770-778.  
<https://doi.org/10.1109/CVPR.2016.90>
- [13] 范家伟, 张如如, 陆萌, 何佳雯, 康霄阳, 柴文俊, 石坤达, 宋美娜, 邬海红, 欧中洪. 深度学习方法在糖尿病视网膜病变诊断中的应用[J/OL]. 自动化学报, 2021: 1-21.
- [14] 高瑶瑶. 基于深度学习的糖尿病性视网膜病变检测方法研究[D]: [硕士学位论文]. 太原: 中北大学, 2021.