

# 基于自动驾驶系统的轻量型卷积神经网络优化<sup>①</sup>



高秀龙, 葛动元

(广西科技大学 机械与交通工程学院, 柳州 545000)

通讯作者: 葛动元, E-mail: [616533142@qq.com](mailto:616533142@qq.com)

**摘要:** 计算机视觉技术大量应用于自动驾驶系统, 主要解决物体识别与物体分类问题, 本文根据任务提出了一种轻量化的神经网络结构. 为解决训练数据规模不足的问题, 采用了改进型数据增强算法, 使训练数据成倍增加. 同时为解决使用数据生成器作为验证集, 无法使用 tensorboard 的问题, 提出了解决方案, 通过卷积网络可视化方法详细研究了神经网络处理图像信息的原理并提出了优化方法. 训练后的模型在验证集上准确率达到了 97.5%, 满足了自动驾驶系统对分类任务准确率的要求.

**关键词:** 计算机视觉; 分类网络; 数据增强; 卷积网络可视化

引用格式: 高秀龙, 葛动元. 基于自动驾驶系统的轻量型卷积神经网络优化. 计算机系统应用, 2020, 29(3): 93-99. <http://www.c-s-a.org.cn/1003-3254/7320.html>

## Optimization of Lightweight Convolution Neural Network Based on Automatic Driving System

GAO Xiu-Long, GE Dong-Yuan

(School of Mechanical and Transportation Engineering, Guangxi University of Science and Technology, Liuzhou 545000 china)

**Abstract:** Computer vision technology is widely used in autopilot system, which mainly solves the problem of object recognition and object classification. In this study, a lightweight neural network structure is proposed according to the task. In order to solve the problem of insufficient training data, an improved data enhancement algorithm is adopted to double the training data. At the same time, in order to solve the problem of using data generator as verification set and unable to use tensorboard, a solution is proposed. The principle of neural network processing image information is studied in detail by convolution network visualization method, and the optimization method is put forward. The accuracy of the trained model is 97.5% on the verification set, which meets the accuracy needs of autopilot system for classification task.

**Key words:** computer vision; classification network; data enhancement; convolution network visualization

自动驾驶系统主要应用车载智能设备通过互联网与其它交通工具进行实时交互. 所以采集道路上的物体来进行分类识别是自动驾驶系统最主要的任务之一.

近年来百度, 特斯拉等公司对自动驾驶汽车进行了许多的研究, 自动驾驶技术需要车载智能系统具有高稳定性, 处理信息准确性高, 和能够实时进行信息处理和网络交互的能力. 当前计算机视觉技术是该领域的主要方法, 它已经广泛的应用于道路分割, 物体分

类<sup>[1,2]</sup>, 物体检测与追踪的任务中, 因此研究计算机视觉技术是当前主要任务.

通过大量的研究可知, 在卷积神经网络<sup>[3]</sup>结构中, 网络的层数与卷积核的规模对网络分类的能力有决定性的作用, 网络结构越复杂, 权值参数越多分类就更加准确. 但随着网络的层数不断增加, 网络中需要保存的权值参数变得巨大, 而随之内存容量要求变高, 进而会导致训练成本增大, 网络处理信息的实时性也随网络

① 基金项目: 国家自然科学基金 (51765007); 广西自然科学基金 (2016GXNSFAA380111)

Foundation item: National Natural Science Foundation of China (51765007); Natural Science Foundation of Guangxi (2016GXNSFAA380111)

收稿时间: 2019-08-01; 修改时间: 2019-09-02; 采用时间: 2019-09-18; csa 在线出版时间: 2020-02-28

规模的增大而下降. 在网络的训练阶段, 由于训练数据的不足, 常常会导致网络训练出的模型泛化能力弱, 使网络的分类识别准确度下降. 为了解决以上问题, 本文提出了一种轻量化的网络结构, 在保证准确度的同时, 也能使网络处理信息的实时性得到增强.

## 1 轻量级网络架构

结合 AlexNet、VGG16、Inception、ResNet 等分类网络的优点, 开发出一个适合于自动驾驶领域的六分类网络<sup>[4]</sup>, 包括汽车、自行车、人、路灯、交通灯、树木. 采用 13 个卷积层来对训练图像做信息提取, 3 个全连接层对卷积层提取的信息进行重新整合.

### 1.1 卷积层设计

卷积层在网络中起到提取图像局部信息的作用, 其结构复杂, 参数规模较大, 直接影响分类精度的高低, 网络结构如表 1.

表 1 卷积层结构

Layer	Output_shape	Kernel_size	Filter
Conv_1	224×224×64	3×3	64
Conv_2	224×224×64	3×3	64
Pool_1	112×112×64	2×2	64
Conv_3	112×112×128	3×3	128
Conv_4	112×112×128	3×3	128
Pool_2	56×56×128	2×2	128
Conv_5	56×56×128	3×3	128
Conv_6	56×56×128	3×3	128
Conv_7	56×56×128	3×3	128
Pool_3	28×28×128	2×2	128
Conv_8	28×28×256	3×3	256
Conv_9	28×28×256	3×3	256
Conv_10	28×28×256	3×3	256
Pool_4	14×14×256	2×2	256
Conv_11	14×14×256	3×3	256
Conv_12	14×14×256	3×3	256
Conv_13	14×14×256	3×3	256
Pool_5	7×7×256	2×2	256

在卷积神经网络中, 卷积层通过卷积核的权值矩阵<sup>[5]</sup>来对输入的图像进行局部信息提取, 所以在网络中卷积核的数量与卷积核中权值是卷积层主要的参数. 鉴于原始图像像素矩阵维度较低, 少量的卷积核便可以有效的对图像信息进行特征提取, 所以对网络的前两层卷积采用 32 个卷积核. 根据前两层网络处理得出的高维数据, 第三, 四层网络采用更多的 64 个卷积核, 对高维数据进行特征提取, 不断的分解细化图像信息. 由于浅层网络不断向深层网络输入高维数据, 所以网络中需要的卷积核数逐层递增. 卷积核的尺寸不易过

大, 经实验与理论分析证明, 当尺寸过大时, 提取到的特征信息不够精确, 进而使网络的性能下降.

### 1.2 全连接层设计

全连接层将卷积层提取的局部特征重新通过权值矩阵整合成完整的特征图, 在全连接层的底部通过损失函数与训练数据标签比较. 根据工程任务设定 3 层全连接网络, 前两层选用 120 个神经元负责整合卷积网络输入的局部特征信息, 最后一层全连接层 6 个神经元, 负责分类任务.

## 2 神经网络优化

### 2.1 激活函数的选择

Tanh 激活函数<sup>[6]</sup>: 当训练数据分布较大或较小时, 函数梯度接近于 0, 易造成梯度消失, 故应用 Tanh 函数时, 要先将训练数据进行归一化. Tanh 函数如:

$$f(x) = \frac{1 - e^{-2x}}{1 + e^{2x}} \quad (1)$$

Relu 激活函数: 当训练数据分布大于 0 时, 其梯度不会消失且易于计算, 当训练数据分布小于 0 时, 神经元不被激活, 这种特性有助于网络稀疏性的建立, Relu 函数如:

$$f(x) = \max(x, 0) \quad (2)$$

通过在 playground-tensorflow 系统中的实验对比发现, 当选用 Relu 激活函数时, 如图 1(a) 网络进行了 60 次迭代后, 模型拟合完毕且损失函数值比较小. 选用 Tanh 激活函数时, 网络进行 110 次的迭代后, 模型拟合完毕, 但是损失值较大. 本文选用 Relu 函数作为网络的激活函数.

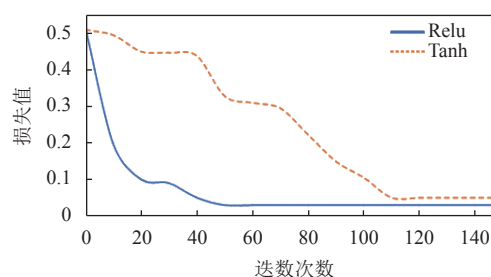


图 1 激活函数对比

### 2.2 梯度下降算法的选择与优化

神经网络的模型拟合过程可以分为两个阶段, 第一阶段先通过正向传播算法得到预测值, 并将预测值和训练数据标签作对比得出两者之间的差距. 第二阶段通过反向传播算法计算损失函数对每一个参数的梯

度,再使用梯度下降算法更新所有权值。

SGD<sup>[7,8]</sup>算法运算速度快并且占用硬件资源少,是当前深度学习领域主要的优化算法,公式为:

$$g(\varphi) = \sum_{j=0}^n \varphi_j x_j \quad (3)$$

$$h(\varphi) = \frac{1}{2m} \sum_{i=1}^m (y^i - g_{\varphi}(x^i))^2 \quad (4)$$

$$\varphi = \varphi - \eta \nabla_{\varphi} h(\varphi) \quad (5)$$

式中,  $\varphi$  代表网络中权值,  $\nabla \varphi$  表示梯度值,  $h(\varphi)$  代表 loss 函数,  $g(\varphi)$  代表目标函数,  $y_i$  代表第  $i$  个样本的样本值,  $m$  代表迭代次数,  $\eta$  表示学习率,  $j$  表示网络中参数的总数目。

但是 SGD 执行更新时,所伴随的高方差导致损失值的剧烈波动,如果把学习率的值设置的过大会导致始终无法达到最优解,也会可能会直接造成进入局部最优解,如图 2 中马鞍面 A 点,如果全局使用过小的学习率会导致模型拟合速度过慢,迭代次数增多.为此本文提出了一种基于 SGD 的自适应学习率优化算法<sup>[9]</sup>,见算法 1,随着迭代进行学习率有规律的由大变小。

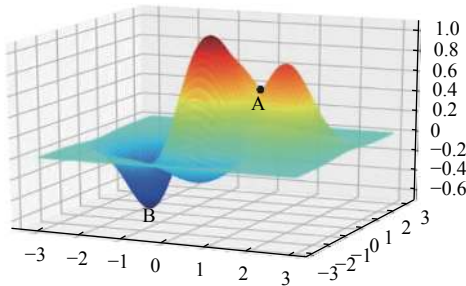


图 2 学习率测试

### 2.3 过拟合问题处理

权值正则化<sup>[10]</sup>:当模型非常复杂之后,它可以准确地记忆每一个训练数据,包括随机噪音的部分,但模型学习的不是训练数据中通用的趋势.为了防止模型从训练数据中学到错误和无关的特征信息,最优的解决方式是获得更多的训练数据,来训练模型.如果无法获得更多的训练数据数据,次优解决方法是增大模型允许存储的信息容量,或对模型允许存储的信息大小加以约束.如果一个网络只能记住几个模式,那么优化过程会迫使模型集中学习最重要的模式,这样就能得到良好的泛化.根据以上解决过拟合思想,结合本文任务,

引入权值正则化算法,根据给定的数据和网络架构强制让模型的权值只能取较小的值,限制模型的复杂度。

算法 1. SGD 自适应学习率优化算法

输入: 当前的迭代次数  $epochs$

输出: 学习率  $\eta$

**Begin**

**While**  $epochs < all\_epochs$  // 判断当前迭代次数是否小于总的迭代次数

**If**  $epochs \geq begin\_epoch$  **then** // 当迭代次数大于设置的开始更新数值时

$epochs = epochs \% begin\_epoch$  // 将迭代次数与初始迭代次数取模运算

$\eta = drop * (in\_rate / \gamma) + (\sigma / \gamma)$  //  $\eta$  更新方法 1

**else**

$\eta = drop * in\_rate + \sigma$  //  $\eta$  更新方法 2

**end if**

**return**  $\eta$  // 输出学习率

**END**

参数介绍:

$in\_rate$ : 初始学习率;

$drop$ : 更新幅度;

$all\_epochs$ : 总迭代次数;

$begin\_epoch$ : 开始更新  $\eta$  时的迭代次数;

$\sigma$ : 保护参数,防止学习率小于 0;

$\gamma$ : 学习率下降幅度

Dropout 正则化: 本文在全连接层中使用 Dropout<sup>[11,12]</sup> 算法,该算法能够使全连接层中的一部分神经元以 0.5 的概率失活,失活的神经元激活函数输出为 0.此时神经元在网络中没有任何贡献.在测试阶段在乘以一定的系数并重新激活,这种正则化方法由于其随机性,可以为网络引入一些随机噪声,这样可以使模型的泛化能力变强,因为它不会太依赖某些局部的特征,从而减少模型的过拟合,如图 3(a) 是没有引入正则化与 Dropout 处理方法而进行的分类网络实验,可以看到模型过度拟合了数据,包括噪音,而图 3(b) 是引入权值正则化和 Dropout 的网络,可看到网络自动将两个红色的噪音屏蔽掉,从而使训练出的模型分类准确率得到提高。

## 3 实验结果与分析

### 3.1 实验平台的搭建

本文网络使用 Keras 作为前端, tensorflow 为后端进行构建,计算机 CPU 为 6 核 12 线程,主频为 2.2 GHz, GPU: NVIDIA Geforce GTX1080TI 显存 11 GB,内存 32 GB,操作系统 Ubuntu14.02.



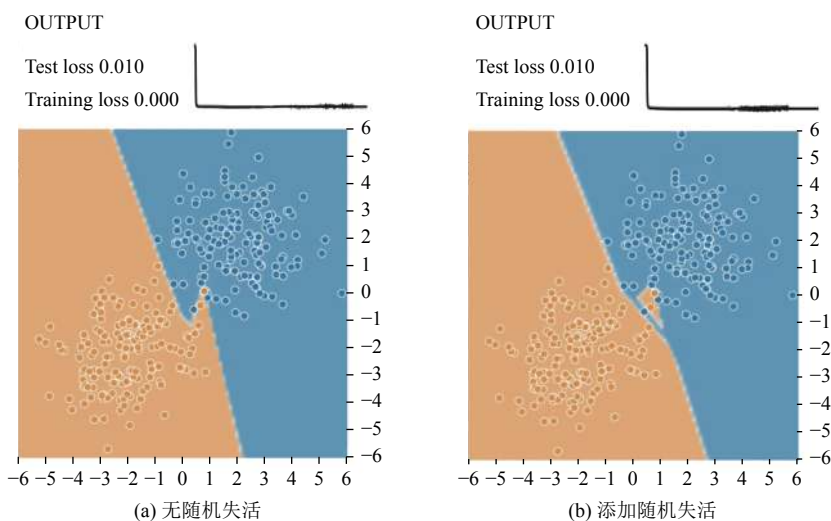


图3 随机失活测试

### 3.2 数据收集与数据增强

通过网络爬虫获取实验数据,共收集6种类别,共12 300张各种尺寸的图像.经大量理论分析与实验证明,模型过拟合的原因之一是训练数据规模小,导致无法训练出能够泛化到新数据的模型,鉴于此问题本文采用 tensorflow 的数据生成器来进行数据增强<sup>[13]</sup>,从现

有的训练样本中生成更多的训练数据,利用多种能够生成可信图像的随机变换来增加样本,具体方式为图像随机旋转、水平和垂直方向上平移、随机错切变换、随机缩放、随机翻转图像一半、随机像素填.如图4所示代表原图与生成的新图像.将增强后得到的数据,分成训练集与测试集,比例分别占80%和20%.



图4 数据增强

### 3.3 实验验证

由于 Keras 框架在网络训练过程中,使用训练生成器对数据进行增强便无法得到网络的训练日志,进而无法调用 tensorboard 模块来观察网络中卷积层与全连接层中的权值和偏置值的分布变化,所以针对此问题本文提出了一种解决方法,代码如下:

```
if python_version_tuple()[0] == '3':
    xrange = range
    imap = map
else:
```

```
x, y = izip (*(validation_generator[i] for i in xrange
              (len(validation_generator))))
```

```
x_val, y_val = np.vstack(x), np.vstack(y)
```

网络经过200次迭代后,模型准确率变化情况如图5(a),损失值变化情况如图5(b),模型在验证集上的准确率达到97.5%.分析结果可知,随着网络不断迭代,损失值变化平稳,学习率与权值初始化参数设置合理,高精度的准确率证明网络深度与卷积核数量满足本文要求.从准确率曲线变化可知损失值还有下降空间,后期可增加迭代次数再次提高模型准确率.卷积层及全

连接层的权值分布变化分别如图6和图7,可知权值参数变化有序平稳,没有出现剧烈波动.偏置值分布变化不明显,由此可知偏置值对模型分类准确度的贡献度

远小于权值,通过研究最终参数的分布规律,可以在升级改造的过程中,有针对性的对权值参数进行初始化操作,进而降低训练周期.

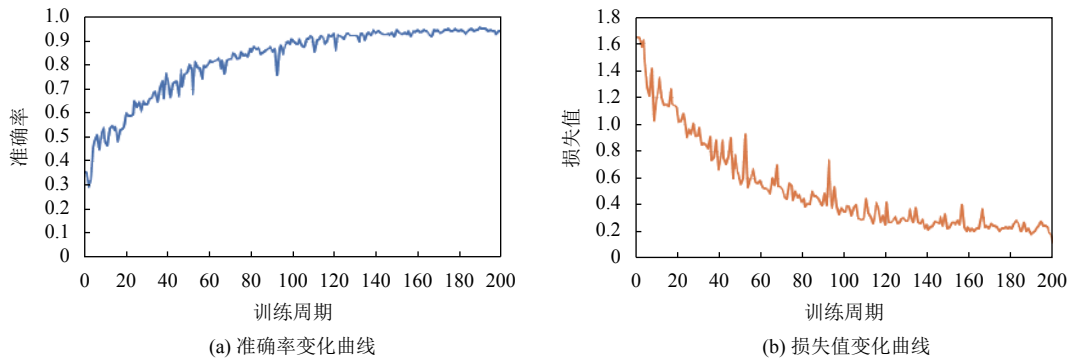


图5 准确率与损失值变化曲线

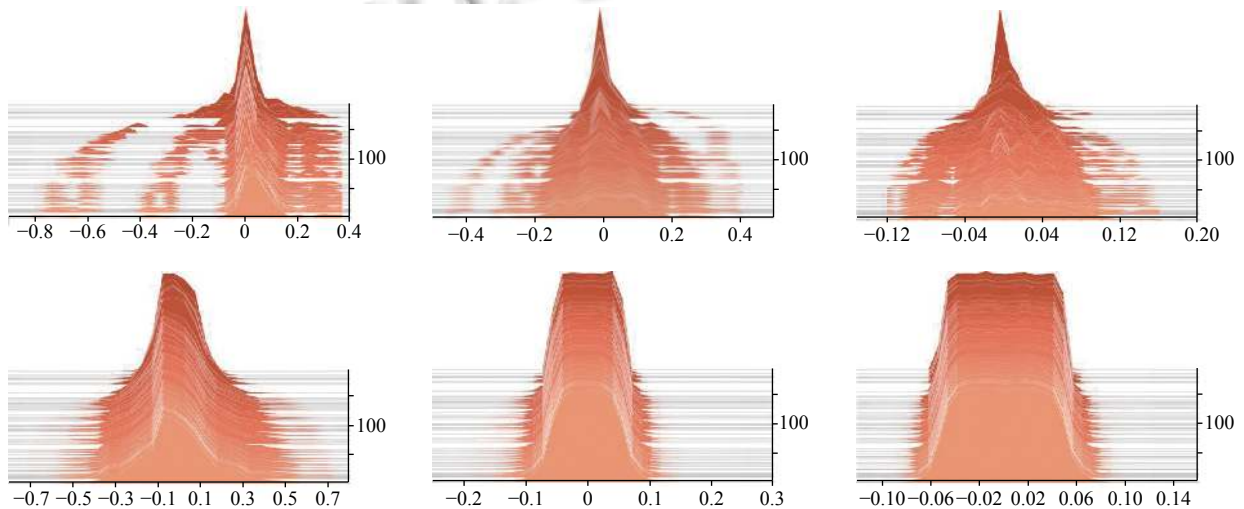


图6 卷积层权值与偏置值分布变化

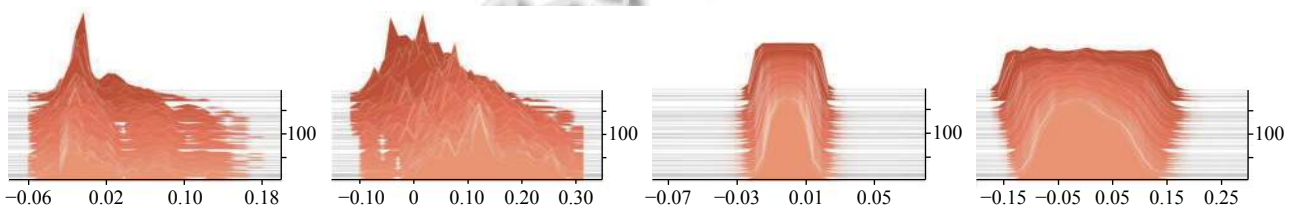


图7 全连接层权值与偏置值分布变化

## 4 卷积神经网络的可视化

### 4.1 可视化卷积神经网络的特征输出

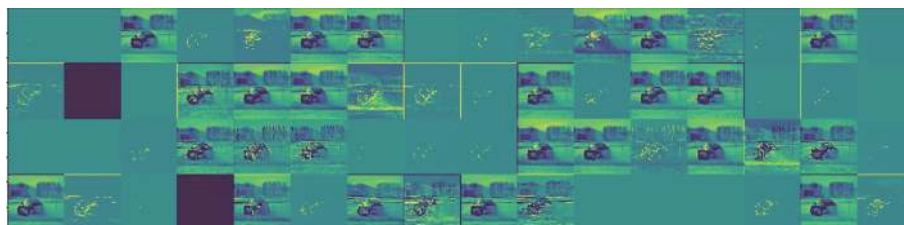
对卷积神经网络输出的特征图进行可视化<sup>[14]</sup>,来研究网络中卷积核如何对输入进行处理与变换.如图8所示,通过检查卷积核的输出,发现浅层网络中卷积核的特征输出保留了原始图像中的所有信息,而深层网

络中卷积核的输出是抽象特征,如车窗与轮胎等,这些特征信息更加有利于对识别,分类任务等决策任务形成贡献.

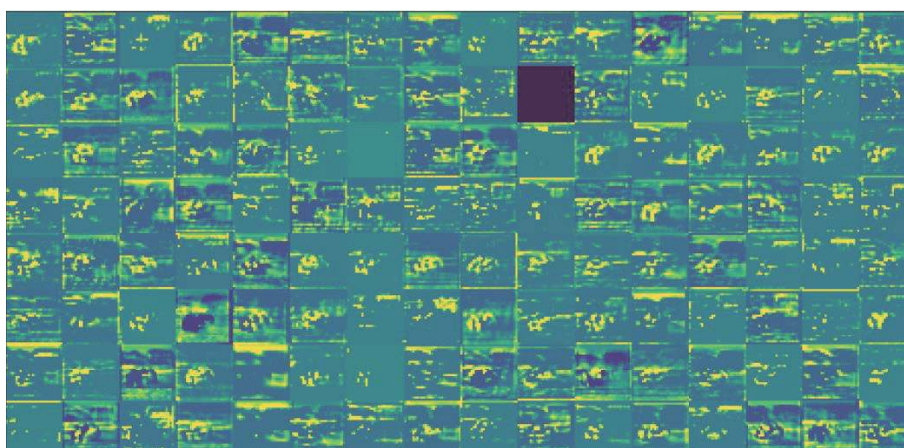
通过实验中卷积层1与卷积层7对比发现,随着网络层数加深,每层网络卷积核的特征输出不断增加,但是无意义的特征输出也在不断增加<sup>[15-17]</sup>,如图8中

无图像的灰绿色部分. 根据本文的网络特点, 经过理论分析可知, 引起这种现象的原因是由于卷积核权值处理过的训练数据分布小于 0, 以这种分布作为 Relu 激

活函数的输入, 引起了神经元的假性失活. 针对此问题带来的神经网络参数冗余, 引入剪枝算法来屏蔽假性失活的卷积核, 使模型既能满足准确度不变, 又响应速度变快.



(a) Conv\_1 层可视化



(b) Conv\_7 层可视化

图 8 卷积层可视化测试

### 4.2 类激活热力图

通过 Grad-CAM 可视化算法<sup>[18]</sup>, 本文方法研究出图像中哪些部位特征, 使神经网络做出最终的分类决策. 如图 9 所示的神经网络, 在识别汽车过程中是以车轮和车窗为主要特征进行判断的. 由理论分析可知, 这

些图像特征经过神经网络的数字化处理后形成特征矩阵, 训练后的神经网络便以这些特征矩阵作为分类决策. 由此我们在训练网络的过程中, 可以为某一物体准备大量具有不同特征的图像数据, 来增加网络在测试集上的准确率.



图 9 热力激活测试



## 5 结束语

本文对自动驾驶领域物体分类与识别问题,提出了基于深度学习的一种轻量化改进型网络结构,在保证分类准确度和网络实时性的情况下,使网络的参数规模显著减少.通过大量实验和理论分析选择出了适合此网络的激活函数,并证明了权值正则化与Dropout在网络中至关重要的作用,选择了适合此网络的过拟合处理方法.通过tensorboard对卷积核中的权值分布变化进行了详细直观的描述.利用Grad-CAM等卷积可视化算法,直观的阐述了神经网络在分类任务中的工作原理和特点,为网络的结构的搭建和后期改进提出了理论依据和方向.

### 参考文献

- 1 林景栋,吴欣怡,柴毅,等.卷积神经网络结构优化综述.自动化学报,2019: 1-14.
- 2 高磊,范冰冰,黄穗.基于残差的改进卷积神经网络图像分类算法.计算机系统应用,2019,28(7): 139-144. [doi: 10.15888/j.cnki.csa.006997]
- 3 Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Proceedings of the 25th International Conference on Neural Information Processing Systems. Lake Tahoe, NV, USA. 2012. 1097-1105.
- 4 Wang ZQ, Liu J. A review of object detection based on convolutional neural network. Proceedings of the 36th Chinese Control Conference. Dalian, China. 2017. 11104-111092.
- 5 LeCun Y, Bengio Y, Hinton G. Deep learning. Nature, 2015, 521(7553): 436-444. [doi: 10.1038/nature14539]
- 6 田娟,李英祥,李彤岩.激活函数在卷积神经网络中的对比研究.计算机系统应用,2018,27(7): 43-49. [doi: 10.15888/j.cnki.csa.006463]
- 7 王功鹏,段萌,牛常勇.基于卷积神经网络的随机梯度下降算法.计算机工程与设计,2018,39(2): 441-445, 462.
- 8 史加荣,王丹,尚凡华,等.随机梯度下降算法研究进展.自动化学报,2019. [doi: 10.16383/j.aas.c190260]
- 9 蒋文斌,彭晶,叶阁焰.深度学习自适应学习率算法研究.华中科技大学学报(自然科学版),2019,47(5): 79-83.
- 10 Wan L, Zeiler M, Zhang SX, et al. Regularization of neural networks using dropout. Proceedings of the 30th International Conference on Machine Learning. Atlanta, GA, USA. 2013. 1058-1066.
- 11 Srivastava N, Hinton G, Krizhevsky A, et al. Dropout: A simple way to prevent neural networks from overfitting. Journal of Machine Learning Research, 2014, (15): 1929-1958.
- 12 Wager S, Wang SD, Liang P. Dropout training as adaptive regularization. Proceedings of the 26th International Conference on Neural Information Processing Systems. Lake Tahoe, NV, USA. 2013. 351-359.
- 13 吴睿曦,肖秦琨.基于深度网络和数据增强的多物体图像识别.国外电子测量技术,2019,38(5): 86-90.
- 14 俞海宝,沈琦,冯国灿.在反卷积网络中引入数值解可视化卷积神经网络.计算机科学,2017,44(S1): 146-150.
- 15 Zeiler MD, Fergus R. Visualizing and understanding convolutional networks. Proceedings of the 13th European Conference on Computer Vision. Zurich, Switzerland. 2014. 818-833.
- 16 Yosinski J, Clune J, Nguyen A, et al. Understanding neural networks through deep visualization. arXiv: 1506.06579, 2015.
- 17 Erhan D, Bengio Y, Courville A, et al. Visualizing higher-layer features of a deep network. Technical Report 1341. Montreal: University of Montreal, 2009.
- 18 Selvaraju RR, Cogswell M, Das A, et al. Grad-CAM: Visual explanations from deep networks via gradient-based localization. Proceedings of 2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy. 2017. 618-626.