



南京大學

研究生畢業論文 (申請碩士學位)

論文題目 卷积神经网络的模型压缩算法

作者姓名 李悦

学科、专业方向 计算机技术

指导教师 商琳 教授

研究方向 深度学习

2019 年 5 月

学 号 : MF1633021

论文答辩日期 : 2019 年 6 月 1 日

指 导 教 师 : (签字)

L^AT_EX Thesis Template of Nanjing University

by

Wenbo Yang

Directed by

Professor XX XXX

Mathematics Department

Nanjing University

May 2010

*Submitted in partial fulfilment of the requirements
for the degree of PhD in Probability Theory and Mathematical Statistics*

南京大学研究生毕业论文中文摘要首页用纸

毕业论文题目： 卷积神经网络的模型压缩算法
计算机技术 专业 2016 级硕士生姓名： 李悦
指导教师（姓名、职称）： 商琳 教授

摘 要

深度学习模型近年来在图像识别、自然语言处理等方向上取得了极好的应用成绩，如图像分类领域的[33]，图像描述生成领域的[50]，[31]，领域自适应领域的[10]等。其中，在图像分类领域，各种优秀的模型如AlexNet, VGGNet, ResNet等，层出不穷。伴随着这些模型在imagenet等数据集上精度的不断提升，模型自身的结构日趋复杂，模型层数也越来越深。更大、更深的网络结构导致深度学习模型在实际应用时需要进行大量的浮点数运算。在一些实时性要求比较高的场景下，深度学习模型的复杂度对设备的计算能力提出了较高的要求。为了降低深度学习模型的复杂度，研究者们主要进行了三个方向的研究：模型参数的量化（Quantization），模型参数、结构的剪枝（Pruning）和轻量化模型（如SqueezeNet等）。本文的工作主要集中于量化算法和模型结构剪枝上。

现有的深度学习量化模型分为训练中量化和预训练后量化。目前能做到无损模型精度的，主要是预训练模型之后，再量化的算法。这类算法的目的是在保证模型精度的同时，将CNN卷积核中参数替换为量化目标值（2的幂的形式），这样在特定硬件上即可将卷积运算中的浮点数乘法变为bit位移运算。同时，量化目标中的2的幂值种类越少，表示一个参数所需要的bit位就越少，模型压缩程度越大。目前这类算法在选择量化时，往往是凭借先验知识（超参数）来确定的，但这样找出的量化目标值会有冗余，有进一步压缩的空间。我们使用聚类算法自适应地去确定量化目标值，实验结果表明，在保证模型精度的同时，能以更少的量化目标值表示模型参数。

剪枝算法近两年从对模型参数剪枝，转到了对模型结构的剪枝上。通过对模型各层通道数（channels/filters）的修剪，使得模型每一层的卷积核数量更少。剪枝后的模型参数量减少，因此模型的存储空间以及应用时需要的计算资源都大大减少。而且，这些剪枝后模型可以在现有的深度学习框架上直接运行。目前几乎所有的剪枝算法，在进行剪枝决策时，都依赖大量的预训练过程。而我们认为，即使基于少量的预训练，也可以做出有效的剪枝决策。我们提出了基于少量预训练的剪枝思想，并通过实验证明了该想法下的剪枝策略，对模型的压缩效果不输于基于大量预训练的剪枝策略。

现有的剪枝策略总是根据

这样的转化，在理论上有，卷积核中出现的2的幂值种类越少，那么模型表示每个参数所需要的比特位越少；

关键词： 南京大学; 学位论文; L^AT_EX 模板

南京大学研究生毕业论文英文摘要首页用纸

THESIS: L^AT_EX Thesis Template of Nanjing University

SPECIALIZATION: Probability Theory and Mathematical Statistics

POSTGRADUATE: Wenbo Yang

MENTOR: Professor XX XXX

Abstract

This paper is a thesis template of Nanjing University. Besides that the usage of the L^AT_EX document class NJUthesis, a brief guideline for writing the thesis is also included.

Keywords: Nanjing University (NJU), Thesis, L^AT_EX Template

目 录

目录	iv
第一章 绪论	1
1.1 研究背景	1
1.2 研究现状	2
1.2.1 CNN量化算法	2
1.2.2 剪枝算法	3
1.2.3 待研究问题	5
第二章 相关技术背景	6
2.1 深度CNN网络常见技术	6
2.1.1 卷积操作	6
2.1.2 其他常见CNN网络技术	6
2.2 聚类（Clustering）技术	8
2.3 算法验证数据集、模型的选择	8
2.3.1 数据集介绍	8
2.3.2 算法模型介绍	9
2.3.3 本文验证数据集和验证模型的选择	9
2.4 注意力机制（Attention）	9
第三章 量化算法目标值的优化	10
3.1 引言	10
3.2 用聚类算法优化量化算法目标值	11
3.2.1 优化量化算法的目标值的意义	11
3.2.2 使用聚类优化目标值的具体流程	12
第四章 基本使用帮助	13
4.1 系统要求	13

4.2	下载与安装	13
4.3	本模板打包的文件	14
4.4	编译方法	14
4.4.1	Windows 平台下	14
4.4.2	Linux 平台下	17
4.5	问题反馈	17
第五章	论文撰写	18
5.1	文档类选项	18
5.2	封面和摘要页的字符串	18
5.3	内容撰写	18
5.4	插图	18
5.5	数学公式	19
5.6	参考文献	19
5.7	其它问题	20
第六章	常见问题	21
6.1	表格	21
6.2	脚注	21
第七章	总结与展望	22
附录 A	博士(硕士)学位论文编写格式规定(试行)	23
A.1	适用范围	23
A.2	引用标准	23
A.3	印制要求	23
A.4	编写格式	23
A.5	前置部分	23
A.5.1	封面（博士论文国图版用）	23
A.5.2	题名	24
A.5.3	前言	24
A.5.4	摘要	25

A.5.5	关键词	25
A.5.6	目次页	25
A.5.7	插图和附表清单	25
A.5.8	符号、标志、缩略词、首字母缩写、计量单位、名词、术语等的注释表符号、标志、缩略词、首字母缩写、计量单位、名词、术语等的注释说明汇集表，应置于图表清单之后。	25
A.6	主体部分	25
A.6.1	格式	25
A.6.2	序号	26
A.6.3	绪论	26
A.6.4	正文	26
A.6.5	结论	28
A.6.6	致谢	29
A.6.7	参考文献表	29
A.7	附录	29
A.8	结尾部分(必要时)	30
参考文献		31
简历与科研成果		37
致谢		38

表 格

插图

1.1	量化算法的简单原理示意图：将浮点数参数替换为2的幂值。	3
1.2	两类剪枝算法的对比：参数剪枝和filters（channels通道）剪枝.图中虚线部分代表被剪去的参数；圆代表被剪去的filters（channels）。	4
2.1	卷积操作示意图.	7
2.2	四种常见的激活函数.	8
2.3	两种池化技术：max pooling和average pooling.	9
3.1	图中我们对比了传统量化算法目标值的确定方法，以及通过分布计算目标值的方法。我们能很显然地发现，利用参数分布得到的量化目标值更有代表性，同时数值种类更少，可以用更少的bit位表示。图中下半部分通过流程图展示了大致的聚类、计算量化中心的过程。	12

第一章 绪论

1.1 研究背景

人工神经网络ANN(Artificial Neural Network)作为人工智能领域的一类算法，有着悠久的发展历史。

1958年，认知心理学家Frank提出了感知机模型。这在当时引起一股研究热潮，但是随后Marvin Minsky和Seymour Papert发现感知机不能处理异或问题，研究陷入低潮。1986年，Hinton提出反向训练算法来训练MLP（Multi-Layer Perception）。1998年，Yann LeCun等研究员提出了七层的CNN模型LeNet [35]，以识别手写数字，但是因为SVM方法的兴起，当时并没有引起重视。伴随着计算机算力的进步，神经网络的模型结构可以设计得更大，这使得神经网络得以在图像识别领域发挥更大的作用。2012年，Hinton等研究者提出的AlexNet [34]模型在ImageNet数据集上，以巨大的优势获得了冠军，这掀起了对深度ANN，即深度学习（Deep Learning）的研究热潮。在12年之后，历年ImageNet比赛的冠军基本都由深度学习模型包揽，从14年的GoogleNet[48]，到15年的ResNet[19]。再到17年的Se-Net[23]。伴随着模型精度的提升，深度学习已经不仅仅停留在理论研究上，而是越来越多的进入到我们的生活中。从抖音上人脸测试；各类app的颜值、微笑打分系统；到乘坐高铁、飞机，我们需要刷身份证进行的人脸认证；再到自动驾驶系统中的路况识别。

但是这些知名CNN模型的高精度并不是没有代价的，深度学习模型往往内存空间占用大，处理图片的计算量也很庞大。如AlexNet模型内存约200MB，计算量约720Mflops；VGG16模型存储约500MB，计算量约15300Mflops。CNN模型巨大的存储和计算量消耗对硬件提出了更高的要求。而实际应用场景中，CNN模型往往面对一些限制，如：1.硬件的存储空间、计算能力往往受限的，如移动端的app，手机的存储和计算性能显然远远不如PC端；2.部分应用场景实时性要求很高，如果自动驾驶中的场景识别。早在2013年，研究者[7]发现，深度学习模型中的参数往往有巨大的冗余，所以很多研究者们将目光转向对深度学习模型的压缩、加速的研究。深度学习模型目前大致分为两类：CNN和RNN，目前学术界的研究主要集中在对CNN的压缩、加速上，我们的工作也集中在对CNN的研究上。

1.2 研究现状

目前对CNN模型的压缩、加速方法主要分为三大类：量化、剪枝和设计轻量化模型。CNN的量化算法的特点是对CNN卷积核的参数数值添加了一定的限制条件。早年的量化算法更着重于模型的压缩，[3]，[11]通过共享权重、压缩参数取值范围等方法让每个参数所占有的bit位数尽可能少；近年的工作不仅在考虑压缩模型，更考虑优化CNN模型的计算，如[28]，[56]等，通常的做法是将参数限制为定点数甚至 2^n 的形式，这样浮点数乘法运算可以转化为定点数运算，甚至在特定硬件上能转化为位移运算，不仅优化了存储同时能大大降低计算代价。剪枝算法历史悠久，主要有两个发展阶段：16年之前的剪枝算法主要是对CNN卷积核的参数进行修剪，构建稀疏的卷积核[16]，[30]；16年之后，研究者开始对卷积层的[1]，filters[49]，[8]，[41]（也有叫，通道channels）进行修剪。轻量化模型也是近年的热点，经典的轻量化模型有MobileNet[22]，SqueezeNet[26]，ShuffleNet[54]，DenseNet[24]。本文的工作主要集中在量化算法和剪枝算法领域，因此我们重点介绍这两个领域的研究现状。

1.2.1 CNN量化算法

计算机中表示数字的方式有两种：定点数和浮点数。然而即使是浮点数表示出来的数字，相对于整个数域而言也是离散的。量化（Quantization）算法将这种离散化的思想发挥到了极致，试图用特定的、有限的、更离散的数值来替代现有CNN网络中的浮点数参数。这样的做法有两个好处：1.参数的变化范围更有限，意味着我们在表示每个参数时，可以用更少的bit位，这样即使在模型参数不变的情况下也能起到对模型存储压缩的效果；2.如图1.1，参数量化目标数值一般选为2的幂值形式（ 2^n ），这样我们可以将CNN网络卷积操作中的浮点数乘法运算简化为位移运算——特征图（feature maps）中的参数是浮点数形式，而卷积网络经过量化操作，参数都变成了2的幂值形式。CNN卷积运算被优化，起到了对模型运算加速的作用。

早期的量化算法将经理集中在模型压缩上，[3]通过hash函数，使得CNN模型得以共享参数，实现了存储的优化。[11]这篇文章则将参数限定在特定取值上，因为整个网络的参数数值种类有限，所以只需要有限的bit位作为编码就能表示出所有的参数。但是[3]，[11]这两种方法在实际应用时，还需要将编码转化为实际数值，因此并不能起到对模型加速的效果。[13]这篇文章也是较早提到量化操作的文章。该文量化操作之后，模型中的参数只是更稀疏，数值却是任意的浮点数，并不能对模型计算进行优化。[5]这篇文章可以算是开了2的幂值作为模型参数量化目标的先河；这篇文章尝试将模型的参数以及各卷积层的

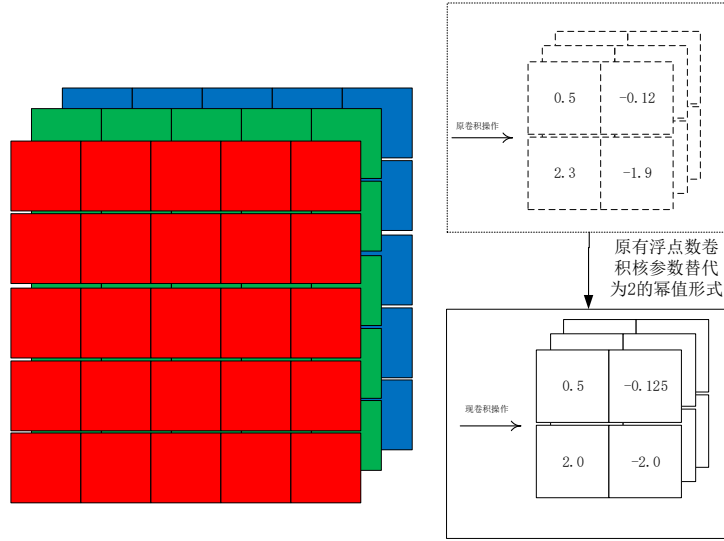


图 1.1: 量化算法的简单原理示意图：将浮点数参数替换为2的幂值。

输出都二值化为 ± 1 ，但是最终在测试集上，二值化的模型的精度损失较大。为了提高网络的信息提取能力，[38]中作者将网络参数范围放宽到了 $-1, 0, 1$ ，显然模型的精度比二值网络有所上升，但误差还是不可避免。

在此之后，研究者们开始分两个方向进行量化的研究。一派如[6]、[44]、[25]、[2]、[57]尝试对网络计算量进行进一步压缩，这派研究者不仅努力将卷积的参数量化为 ± 1 或者 $\pm 2^n$ ，同时试图将卷积之后输出的特征图也进行量化为 ± 1 或者 $\pm 2^n$ 。如[44]，作者尝试将参数，激活函数的输出值都量化为 $0, \pm 1$ ，这种量化的好处是可以将卷积运算从浮点数乘法、加法变成xnor和bitcount计算。毫无疑问，[44]在计算量上的优化的十分有优势的，但是实验结果看来，这样的量化操作会造成模型分类精度的下降。部分研究者，如[57]、[25]尝试放宽卷积输出特征图激活值的量化范围，研究者尝试将卷积中参数量化为 ± 1 ，而激活值量化为 $\pm 2^n$ 。另一派仅仅尝试将参数量化，而不考虑对激活值进行量化，如[58]、[56]，其中[56]是截止到目前为数不多能够保证模型分类精度不下降的量化算法。这篇文章的出发点和之前的量化工作都不太一样，作者并没有尝试训练出一个量化网络，而是训练了一个全精度的网络，然后将这个网络中的参数转化成了目标量化值。

1.2.2 剪枝算法

剪枝算法有着很悠久的历史，早在1990年前后，已经有研究者开始尝试将剪枝算法应用于神经网络[36]、[17]。早期的剪枝算法主要集中于对模型参数的

剪枝，这样的剪枝不仅可以优化模型的存储大小，同时还可以起到防止模型过拟合的效果。总得来说，剪枝算法就是利用某种标准将模型中不重要的参数或者filters选出，然后剪掉，从而在保证模型精度的同时，尽可能压缩模型的存储空间、计算消耗。

截止到今天，CNN模型的剪枝算法大致可以分为两类：第一类以模型的参数修剪作为目标；第二类则以模型的filters（或者说通道数channels）作为剪枝目标。我们在图1.2中简略地展示了两类剪枝的大致原理。参数剪枝算法，着重修剪参数，让卷积核变得稀疏，这样模型需要存储的参数数量会变少，从而能起到对模型存储进行压缩的效果。但是在参数剪枝算法中，卷积核变得稀疏，但是卷积核本身的数量并没有变化。这类剪枝算法在2016年之前比较流行，如[15]中作者根据参数数值判断参数的重要性，该方法被结合到经典文章[14]。在[51]，[46]中，作者结合Lasso思想，借用惩罚函数构造稀疏卷积核。但是不管是怎样的标准，总不可避免会过度剪枝，因此在[12]中作者提出了一种动态剪枝技术，通过标准来恢复部分参数。稀疏卷积核的卷积操作需要依赖特定的库，所以实际上对CNN的卷积操作加速效果并不理想。在16年之后，研究者们改变思路，不再以参数作为剪枝目标，而是开始剪filters。如图1.2下图展示了剪filters的示意图，对filters的剪枝改变了卷积层输出的特征图数。

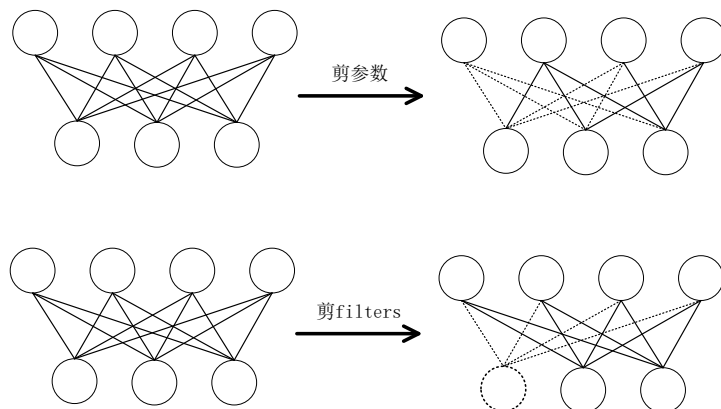


图 1.2: 两类剪枝算法的对比：参数剪枝和filters（channels通道）剪枝.图中虚线部分代表被剪去的参数；圆代表被剪去的filters（channels）。

在剪枝filters的工作中，相当多的作者都在努力探索更优的标准来实现更好的剪枝效果。17年的[39]，作者使用filters的 $l1 - norm$ 来度量filters的重要性，范数值越大的filters越重要，剪枝时优先将不重要的filters剪去。[42]中，作者使用紧跟着卷积层的BatchNorm层的放缩因子 γ 来度量filters或者说通道

(channels) 的重要性, γ 数值越小相应的filters越不重要。[43]提出了利用泰勒展开式来度量filters重要性。也有工作不仅关注模型压缩, 更关心模型的计算量优化: [52]中, 作者根据各个filters的计算量进行了排序, 优先剪去计算代价大的filters。

除了提出各种标准来判断filters的重要性, 还有一些研究者尝试利用现有的模型、算法搜索CNN网络内在的有效结构。[55]中作者利用LSTM提取模型中的有效结构。[53]通过简化CNN网络计算图实现了对CNN网络结构的优化。

1.2.3 待研究问题

在所有量化算法中, 近年的量化算法更着重于对模型计算量的优化。通过将参数优化为量化目标值 $\{0, \pm 2^{n_1}, \pm 2^{n_2}, \pm 2^{n_3}, \dots\}$, 这样不仅能优化存储量, 还想优化计算。但是在量化目标值的选择上, 现有的算法并没有给出很好的解决方案。我们认为量化目标值的选择应当根据不同层做出一定自适应性的优化, 从而在同时现有的量化算法量化目标值的应用范围是整个网络, 如[6]是将整个网络所有参数优化都优化为 ± 1

第二章 相关技术背景

本文完成的工作都是对深度学习模型的压缩、加速工作。现有的压缩、加速算法都集中在对计算量最大的卷积操作的改进，所以在本段中我们首先会重点介绍CNN网络的卷积操作。为了便于读者构建对CNN基本认知，我们会顺带简单介绍CNN网络中的常用技术。本文的工作集中在对CNN网络的量化和剪枝上，因此我们也尝试更细节上说明我们对现有量化算法的第一个改进是利用参数分布来优化量化函数的目标值，这里我们利用了聚类算法，因此会在本段第二部分中介绍一下聚类法。在我们对剪枝预训练量的思考中，我们利用了增量式的思想，所以本段也会简单介绍增量式的思想。我们还在剪枝算法中引入了attention思想来优化现有的剪枝标准，所以我们还会简单介绍attention思想。

2.1 深度CNN网络常见技术

CNN网络主要被应用于图像处理领域，常用的技术有卷积层（Convolution Layer）、激活函数、批标准化Batch Normalization [27]，dropout [21]，[47]，池化层[40]和全连接层。一般卷积层和激活层被我们用来提取图像特征，而全连接层则用来根据提取出来的特征进行分类。而深度CNN网络相较于早期的CNN网络，主要特点就是层数更多，整个网络结构更深。下面我们主要介绍卷积层。

2.1.1 卷积操作

CNN网络的卷积层主要被用于提取图像的特征，因此卷积层的输入是图像，而输出的提取出的特征也是以图像的形式存在。在图2.1中，我们展示了卷积操作的基本流程。图2.1中我们展示了一个输入为RGB图像的CNN网络，第一层的情况。每个卷积层（Convolution Layer）都由多个filters组成，每个filter都会对该层输入的所有特征图进行卷积，一个filter卷积完该层所有输入特征图后会输出一个特征图。通常，每个卷积层都会由多个filters构成。图中的虚线指明了卷积和图像中像素点的对应点乘关系。

2.1.2 其他常见CNN网络技术

CNN网络中常见的技术，还有激活函数、池化（pooling）、批标准化（Batch Normalization）[27]。现在通常在构建深度CNN网络时，我们会将卷积层的输出进行正态标准化处理，同时约束网络、在训练过程中自动调整该

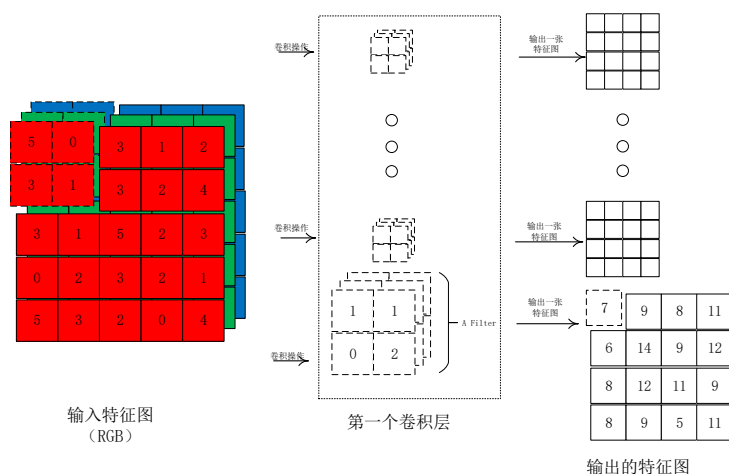


图 2.1: 卷积操作示意图.

标准化的强度，从而加快训练速度并降低权值初始化的成本。激活函数已经成为神经网络的标配。如果没有激活函数，那么神经网络的输入到输出只是一个线性映射，而利用激活函数之后，神经网络可以学到更复杂的映射。在下图2.2中，我们画出了常见的四类激活函数。激活函数一般需要满足四个性质：可微性（便于计算梯度），单调性（保证函数是凸的），非线性（保证数据非线性可分）以及输出值范围有限。2.2中的Sigmoid和Tanh函数在输入值趋向 $\pm\infty$ 时，函数梯度趋向于0，这在利用反向传播算法[45]计算梯度时，造成了梯度消失的问题。因此有研究者提出了激活函数[9]，Relu作为激活函数输入为正时，梯度处处为1，极大缓解了梯度消失的问题。但是Relu函数对负的输入值求梯度时，梯度为0，这很容易造成神经元的“死亡”，梯度不能很好地回传，因此又有研究者提出了[18]。

一般我们在卷积时，会使用padding（在图像周围补0）技术，用来防止卷积造成输出特征图尺寸变小。这样造成了数据维度的提升，为了降低数据维度，研究者提出了池化（pooling）技术。图2.3中我们展示了两种池化技术，池化技术会将图像的临近区域进行合并，合并值会选择区域内最大值（max pooling）或者区域内数值的均值（average pooling）。

一般来说，深度CNN网络中，一般卷积层、批标准化层、激活层是结合在一起使用的。

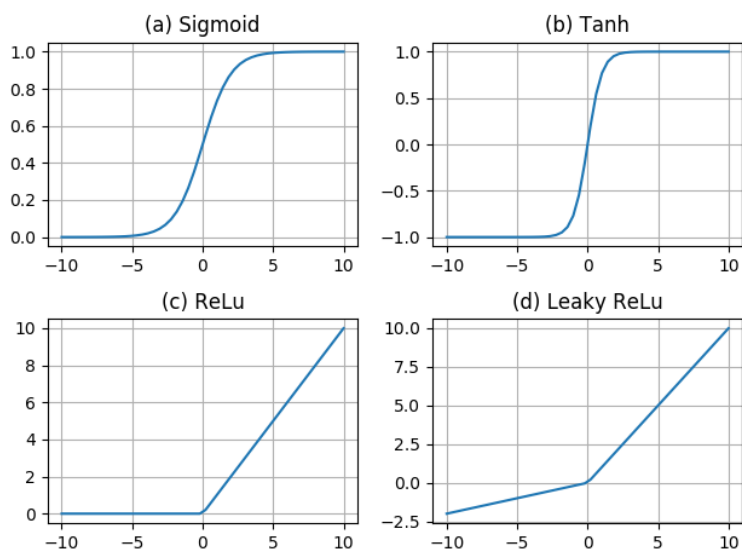


图 2.2: 四种常见的激活函数.

2.2 聚类 (Clustering) 技术

数据聚类技术属于无监督学习的一种。聚类算法会根据一定的标准将数据划分到不同的类别中去。聚类最终期望统一类别内的数据在特性上尽可能的接近；不同类别之间数据特性差异尽可能得大。常见的聚类算法有很多种，有基于划分的，如经典的k-means算法；也有基于层次的，如BIRCH算法、CURE算法等；还有基于密度的算法，如DBSCAN算法。本文应用聚类算法是为了大致探寻一维数据的分布。算法的简便是我们的第一要求，因此我们选择了最经典的k-means算法。

2.3 算法验证数据集、模型的选择

2.3.1 数据集介绍

图像分类领域，比较著名的数据集有MNIST[35]，CIAFR10[32]，CIFAR100和ImageNet[34]。这三个数据集是现行图像分类问题公认的验证数据集。MNSIT数据集共有十个类别，共由50000张训练集和10000张测试集组成，每个类别6000张图片。其中每张图片均为28*28。CIFAR10和MINST数据集的类别、类内图片数量相同，不同的是CIAF10数据集中每张图片大小为32*32。CIFAR100数据集和CIFAR10有相同的图片总数和图片大小，不同的是，CIFAR100中图片共有100个类别，所以每个类别有600张图片。这600张图片中，500张用作训练，100张作为测试。ImageNet数据集共有2万多个类别，一般我们会选择其

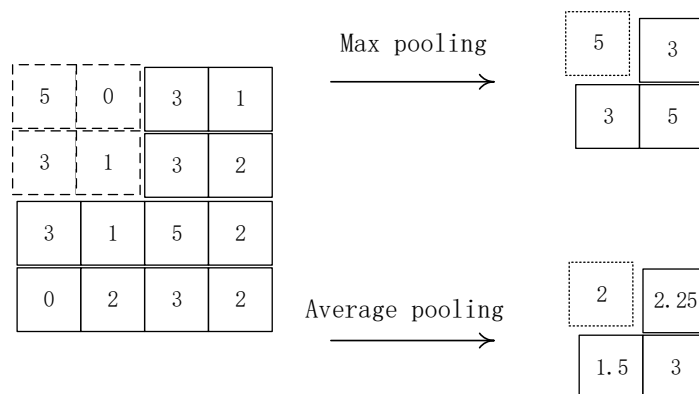


图 2.3: 两种池化技术: max pooling和average pooling.

中的1000个类别来作验证。ImageNet每个类别有数百张图片，每张图片大小大约为 256×256 。一般作者会根据自身的需求选择其中的一个、两个或者全部作为验证数据集。如[39]选择了MNIST、CIFAR10数据集来作为验证集，[55]选择了选择了MNIST、CIFAR10、CIFAR100作为验证数据集。[20]中作者选择了CIFAR10和ImageNet作为验证数据集。

2.3.2 算法模型介绍

现有的比较经典的CNN模型有AlexNet，VGGNet和ResNet。因此现有的量化、剪枝算法一般会选择其中的一到两个模型来作为验证模型。如[39]选择了VGGNet，ResNet验证模型，[55]选择了VGGNet，ResNet作为验证模型；[20]只选择了ResNet作为验证模型。

2.3.3 本文验证数据集和验证模型的选择

现有数据集中，MNSIT，CIFAR10，CIFAR100数据集，数据集的图片数量适中，图片大小较小，所以我们选择了这三个数据集作为验证数据集。而ImageNet数据集规模过大，受限于硬件条件，我们放弃了该数据集。验证模型的选择上，AlexNet，VGGNet和ResNet这三个模型均被我们采用。

2.4 注意力机制（Attention）

第三章 量化算法目标值的优化

3.1 引言

深度学习算法，在诸多领域取得了惊人的应用效果，这大大刺激了人们对深度学习的探索热情。但是深度学习模型高存储消耗、高计算量消耗也束缚着深度学习模型的更广泛应用。因此，对深度学习模型压缩、加速的研究逐渐引起了研究者的兴趣。在所有深度学习模型中，CNN的压缩、加速一直是一大热点。近年来量化算法一直深受研究者们瞩目。早期的量化算法，[4]中，作者尝试利用低bit位数字来替代浮点数的网络参数，将浮点数运算简化为定点数运算。这样的操作可以使用更少的bit位来表示每个参数：之前每个参数都需要32bit来存储一个浮点数，现在是需要8bit甚至更少的bit位数来存一个定点数，整个网络存储消耗降低。同时还能优化计算代价，众所周知，在计算机上定点数计算比浮点数快得多。[11]中作者尝试性地将传统信息理论中的向量量化方法[29]运用到了对CNN参数的优化上。

早期量化算法更多的将注意力集中在对模型的压缩上，近年的量化算法将注意力转向对卷积操作的计算量优化上。最开始研究者们尝试一步到位[6]，将参数变为 ± 1 ；之后的研究不仅想量化卷积核的参数，还想量化卷积核的输出特征图像素值[5]；由于参数量化过度，这类算法往往使模型在分类精度上损失较大。因此[56]中，作者仅尝试将参数量化，取得了很好的量化效果。18年的[37]通过量化算法和scale值的结合，利用ADMM算法，实现了很好的压缩效果。

近年优化计算量的量化算法[5]，[6]，[38]，[56]中，量化的目标值一般为 $\{0, \pm 1\}$ 或者 $\{0 \pm 2^n\}$ 。我们认为这量化算法在存储量上还有进一步优化的空间。在复现[6]算法的时候，我们发现，如果是用该算法处理原始CIFAR10数据集上的AlexNet[34]，会出现网络很难收敛的情况。我们通过添加BatchNorm层解决了这个问题，但是无疑证明如[6]，[38]在量化目标值的选择上，并不具有代表性。现有的工作往往都是根据经验、硬件需要指定量化目标值，我们提出根据参数分布来确定量化目标值的办法。

3.2 用聚类算法优化量化算法目标值

3.2.1 优化量化算法的目标值的意义

量化算法有两个作用：1.压缩模型存储；2.优化计算。现有量化算法的量化目标值集合通常为 $\{0, \pm 2^{n_1}, \pm 2^{n_2}, \pm 2^{n_3}, \dots, \pm 2^{n_k}\}$ ，其中 k 是一个超参，在计算优化上可以将浮点数乘法转化为bit位移运算。但是在压缩模型存储方面，我们认为现有工作尚有改进余地。

CNN网络的参数，往往以浮点数形式存在，我们的目标是使用某些数字取代浮点数参数，并同时想保证网络的精度。在量化目标值集合

$$\{0, \pm 2^{n_1}, \pm 2^{n_2}, \pm 2^{n_3}, \dots, \pm 2^{n_k}\}$$

中，和显然， k 值越小，最终模型中的参数数值种类越少，我们表示一个参数所需要的bit位数也越小。但是 k 值小到极限，那就网络就变成了[38]网络，很显然过度的压缩模型参数数值会造成模型精度的较大损失。所以在模型存储压缩上，我们面临模型精度和模型存储之间的取舍。

我们发现，现有算法在量化目标值的选择上是比较粗糙的。现有的量化算法如[6]、[38]人为规定了量化算法的目标值为 $\{0, \pm 1\}$ ；[56]中作者通过卷积层中参数的最大值和量化参数的bit位数要求确定量化目标值。其中[56]是现有量化网络中为数不多能保证模型量化之后精度的算法。

我们思考了一个问题：如果我们选择的量化目标值集合具有更强的代表性，那么能否较之现有的工作，在保证精度的基础上，实现更高比例的压缩。我们尝试通过根据卷积层参数的分布来确定量化算法的量化目标值集合。我们尝试将这个问题定义为一个优化问题。

首先假设我们的量化目标值集合为 $\{0, \pm 2^{n_1}, \pm 2^{n_2}, \pm 2^{n_3}, \dots, \pm 2^{n_k}\}$ ，其中 k 值已经选定。之后我们定义了量化函数Formula.3.1：

$$\hat{W}_l(i, j) = \begin{cases} 2^{n_1} \text{sgn}(W_l(i, j)) & \text{if } (\alpha + \beta)/2 < \text{abs}(W_l(i, j)) \leq 3\beta/2 \\ 0 & \text{otherwise,} \end{cases} \quad (3.1)$$

上式中， α 和 β 是目标量化值中临近的两个值。在进行量化操作的时候，我们会根据Formula.3.1将浮点数参数量化为目标值。根据Formula.3.1，每个浮点数都会被量化为最接近该数的量化目标值，如 -0.12 会被量化为 -0.125 (-2^{-3})。

我们提出一个度量量化算法误差的度量Formula.3.2：

$$\min \sum_{l,i,j} |W_l(i,j) - \hat{W}_l(i,j)|, \quad (3.2)$$

其中 $W_l(i,j)$ 表示第 l 层的某个参数， $\hat{W}_l(i,j)$ 是该参数量化之后的数值。我们通过累加所有卷积层所有参数的量化误差，即可求出特定量化目标值集合的误差。我们认为， k 值一定时，Formula.3.2值越小的量化目标集合越有效。考虑到该函数是个非凸函数，不好优化。我们考虑从CNN模型的参数分布来入手，选择量化算法的目标值。我们选择了聚类算法来度量CNN网络的参数分布。

3.2.2 使用聚类优化目标值的具体流程

在图.3.1中我们对比了传统的量化目标值确定方法和我们方法的大致流程。

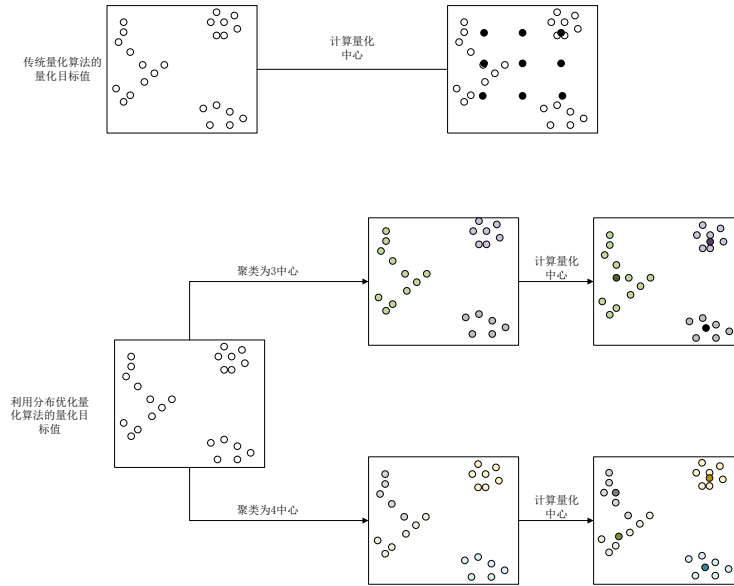


图 3.1: 图中我们对比了传统量化算法目标值的确定方法，以及通过分布计算目标值的方法。我们能很显然地发现，利用参数分布得到的量化目标值更有代表性，同时数值种类更少，可以用更少的bit位表示。图中下半部分通过流程图展示了大致的聚类、计算量化中心的过程。

第四章 基本使用帮助

NJUthesis 宏包针对南京大学对学位论文格式的规定制作的 \LaTeX 宏包，其目的是简化学位论文的撰写，使得论文作者可以将精力集中到论文的内容上而不是浪费在版面设置上。同时宏包在符合学位论文撰写要求的基础上尽可能地进行了美化，其中还参考了出版界的一些排版规范。

下面是 TODO 列表：

- （高优先级）增加本科生论文格式，当 `bachelor` 作为模板参数时生成本科生毕业论文格式。
- （低优先级）提供 UTF-8 格式，在此格式下提供对 Linux 系统的支持。

4.1 系统要求

NJUthesis 宏包可以在目前大多数的 \TeX 系统中使用，例如 \CTeX 、 \MiKTeX 、 \tex 、 \fpTeX 。

NJUthesis 宏包通过 `ctex` 宏包来获得中文支持。`ctex` 宏包提供了一个统一的中文 \LaTeX 文档框架，底层支持 CCT 和 CJK 两种中文 \LaTeX 系统。最新的 `ctex` 宏包可以从 <http://www.ctex.org> 网站下载。

此外，NJUthesis 宏包还使用了宏包 `amsmath`、`amsthm`、`amsfonts`、`amssymb`、`bm`、`geometry`、`enumitem` 和 `hyperref`。目前大多数的 \TeX 系统中都包含有这些宏包。即使您系统中的宏包不全，您也可以到 CTAN 下载到这些宏包。

最新的 \CTeX 套装（2.4.1 以上版本）中包含了以上列出的各种宏包，用户无需额外的设置即可使用。

4.2 下载与安装

NJUthesis 宏包的最新版本可以从 <http://njuthesis.googlecode.com/> 网站或者 <http://share.solrex.org/njuthesis/> 网站下载。

对此宏包最简单的使用方法是直接修改 `template.tex` 和 `template.bib` 文件。如果您仅仅想安装宏包，可以将宏包文件放置到 \TeX 系统的 `localtexmf/tex/latex/njuthesis` 目录下，然后刷新 \TeX 系统的文件名数据库。

关于安装过程的问题可以参考 \CTeX -FAQ 以及其他 \LaTeX 教材。

4.3 本模板打包的文件

关键文件：

NJUthesis.cls 文档类文件，是 NJUthesis 宏包的主要文件，用户在文件头 `documentclass` 声明文档类为 NJUthesis 即可使用此文档类格式化文档。一般不建议用户手工修改该文件。

NJUthesis.cfg 配合文档类的控制文件，里面主要是一些字符串的定义。用户可以编辑修改该文件。

template.tex 随宏包包含的一个模板文件，该模板文件指示了如何使用宏包提供的文档类。用户可以使用该文件作为毕业论文源文件的很好的开端。

template.bib 随宏包包含的模板文件相关的 BibTeX 文献数据库，当用户使用 BibTeX 时可以以其为一个很好的开端；当用户使用自定义的参考文献时，可以在模板文件中注释掉对其的引用并忽略该文件的存在。

辅助文件：

NJUthesis.bst 默认的 BibTeX 样式文件，如果想修改样式，例如使用 IEEEtran.bst 作为 BibTeX 样式，只需要将 IEEEtran.bst 更名为 NJUthesis.bst 即可。

NJUthesis.pdf 宏包帮助文档，其中有对宏包的说明以及对宏包源文件的完全注释，适合高级用户修改文档类文件时参考。

4.4 编译方法

下面以编译模板文件 `template.tex` 生成 `template.pdf` 为例，给出在不同情况下应该使用的编译命令流程。

4.4.1 Windows 平台下

4.4.1.1 标准编译命令

如果您使用 BibTeX 管理参考文献，并且将自己的参考文件条目均添加到 `tempalte.bib` 中，编译命令应该为：

```
set ARTICLE=template
latex %ARTICLE%.tex
bibtex %ARTICLE%
latex %ARTICLE%.tex
gbk2uni %ARTICLE%
latex %ARTICLE%.tex
dvipdfmx %ARTICLE%.dvi
pause
```

其中，`bibtex` 命令完成对参考文献的处理，`gbk2uni` 命令完成对 pdf 书签编码的处理，`latex` 命令生成 dvi 文件，`dvipdfmx` 命令生成 pdf 文件。如果您修改了模板文件名，例如将 `template.tex` 修改成 `thesis-yang.tex`，那么您只需将第一句中的变量定义修改成 `thesis-yang` 即可。

您可以将上面的命令保存成一个文件（并且建议这样做），例如 `make.bat`，这样每次编译时双击该文件，或者在该目录的 Windows 命令行下输入 `make` 即可完成编译。

4.4.1.2 不使用 BibTeX

尽管我们强烈建议您使用 **BibTeX** 管理参考文献，如果您不使用 **BibTeX** 而是自己手写参考文献条目，那么您可以使用以下方式编译模板。

首先在模板文件 `template.tex` 中注释掉引用 **BibTeX** 数据库的语句，即：

```
\bibliography{template}
```

并加入手写的参考文献条目。

其次使用下面命令编译模板文件：

```
set ARTICLE=template
latex %ARTICLE%.tex
latex %ARTICLE%.tex
gbk2uni %ARTICLE%
latex %ARTICLE%.tex
dvipdfmx %ARTICLE%.dvi
pause
```

4.4.1.3 使用 pdflatex 命令

我们强烈不建议这样做，主要原因是本模板中文封面中的南京大学校徽和毛笔校名均使用的是 eps 图片，使用 pdflatex 无法对其进行正常处理。如果您希望使用 pdflatex，那么您至少需要将这两张 eps 图片转成 jpg, png 或者 pdf 图片。

出于对矢量图的偏好，我们目前不愿意提供与 pdflatex 兼容的图片格式，也不提供 pdflatex 的编译流程，因此您需要自己处理与 pdflatex 有关的兼容问题。

4.4.1.4 清理垃圾文件

latex 编译时会生成很多垃圾文件，下面的命令可以清理除 pdf 之外的垃圾文件，您可以将其保存成一个 clean.bat 的批处理文件，即可双击执行。

```
del /Q *.aux *.toc *.lon *.lor *.lof *.ilg *.idx *.ind
del /Q *.out *.log *.exa *.nav *.snm *.bbl *.blg *.spl
del /Q *.lot *.bak *~ *.dvi *.ps
pause
```

如果想同时清除 pdf 文件，您可以在后面加上 *.pdf，并保存成 dist-clean.bat 的批处理文件。

4.4.1.5 使用 cygwin 时的 Makefile

在 Windows 下使用 cygwin 时，用 Makefile 来对项目进行编译是再自然不过的想法。下面给出一个编译本模板可以使用的 Makefile。值得注意的是 Makefile 的缩进需要强制使用制表符。

```
ARTICLE=template
ARTICLE_SRC=$(ARTICLE).tex $(ARTICLE).bib CASThesis.cfg
IMAGES=

all: article

article: $(ARTICLE).pdf
```

```
$(ARTICLE).pdf: $(ARTICLE_SRC) $(IMAGES)
    latex $(ARTICLE).tex
    bibtex $(ARTICLE)
    latex $(ARTICLE).tex
    gbk2uni $(ARTICLE)
    latex $(ARTICLE).tex
    dvipdfmx $(ARTICLE).dvi

clean:
    rm -f *.aux *.toc *.lon *.lor *.lof *.ilg *.idx *.ind
    rm -f *.out *.log *.exa *.nav *.snm *.bbl *.blg *.spl
    rm -f *.lot *.bak *~ *.dvi *.ps

distclean: clean
    rm -f *.pdf
```

4.4.2 Linux 平台下

Void.

4.5 问题反馈

用户在使用中遇到问题、需要增加某种功能或者提供一个 patch，都可以到本项目的 Issue 列表中提出，Issue 列表的地址在：

<http://code.google.com/p/njuthesis/issues/list>

欢迎大家反馈自己的使用情况，使我们可以不断改进宏包，但是在提问前请最好检查自己使用的是否为最新版本，以及列表中类似问题是否已经被提出。总之，感谢您的支持！

第五章 论文撰写

5.1 文档类选项

NJUthesis 文档类提供有以下一些选项：

phd/master 当选择 **phd** 时，生成南大博士论文，包含国家图书馆格式的封面，但不包括书脊，书脊需要单独制作；选择 **master** 时，生成南大硕士论文。

oneside/twoside 当选择 **oneside** 时，生成单面打印格式论文；当选择 **twoside** 时，生成双面打印格式论文。双面打印格式会将封面、摘要起始页、章起始页全部放在奇数页上。

dvipdfm/dvips/pdflatex 兼容的编译命令选项，最好使用默认的 **dvipdfm**。如果您对默认的命令有意见，那么我们认为您充分了解这几个命令的不同，那么您足以自己修改图片格式和决定编译流程，所以使用其它选项时请不要完全使用本文中的帮助。

5.2 封面和摘要页的字符串

NJUthesis 宏包的设置都保存在 `NJUthesis.cfg` 文件中。用户可以在 `.tex` 中通过宏包提供的命令修改设置。对于常用的设置修改，如培养单位名称、专业名称等，可以直接在 `NJUthesis.cfg` 文件中进行。各培养单位可以修改后提供本单位统一的 `NJUthesis.cfg` 文件供本单位用户使用。

5.3 内容撰写

请阅读文献《[The Not So Short Introduction to L^AT_EX 2_ε](#)》[?]，或者其有点儿老的中文翻译版《[一份不太简短的 L^AT_EX 2_ε 介绍](#)》[?]

5.4 插图

请阅读文献《[Using Imported Graphics in L^AT_EX and pdfL^AT_EX](#)》[?]，或者其有点儿老的中文翻译版《[L^AT_EX 2_ε 插图指南](#)》[?]

5.5 数学公式

请阅读文献《The LaTeX Companion》第 8 章“Higher Mathematics”[?]。

5.6 参考文献

本节主要介绍参考文献的格式，在 NJUthesis 中控制参考文献的主要有两个地方：

```
\RequirePackage[sort&compress, numbers]{natbib}
```

和

```
\bibliographystyle{NJUthesis}
```

第一个条目是 natbib 宏包，其规定了参考文献的排序和展示等格式，第二个条目是 NJUthesis.bst 文件，规定了参考文献条目的展示格式。期望使用下面的参考文献格式时，必须使用 BibTeX 来处理参考文献。

在 numeric mode 下（正如本文所用，将 numbers 作为 natbib 参数），参考文献的标准引用方式是使用方括号加文献序号的方式，例如 [?]，这时候使用默认的

```
\cite{key}
```

命令即可。但也可以使用 author year 的方式引用，例如 ?]，这时候使用

```
\citet{key}
```

命令。

但是，如果您希望在 numeric mode 下使用 authoryear mode（将 authoryear 作为 natbib 参数）的标签样式，例如 ? (?)，可以使用我们自定义的

```
\citenju{key}
```

命令，下面是一个例子：

```
\Citenju{Al-Osh1987}~引入了~XXX,
```

```
\citenju{Duffie2001}~介绍了~XXX。
```

```
XXX~可见~\citenjup{Freeland1998},
```

```
XX~可见~(\citenjupr{Xu2007}; \citenjupr{Wang1999})。
```

? (?) 引入了 XXX，? (?) 介绍了 XXX。XXX 可见 (? , ?)，XX 可见 (? , ? ; ? , ?)。

5.7 其它问题

请到南京大学小百合论坛科技文献排版讨论区（网址：<http://bbs.nju.edu.cn/bbsdoc?board=TeX>），或者水木清华论坛 \LaTeX 讨论区（网址：<http://www.newsmth.net/bbsdoc.php?board=TeX>）提问。

第六章 常见问题

6.1 表格

问题 1. 页眉里论文题目和各章标题中的字母均为大写，不能实现大小写的区别，而我写的论文需要在页眉中出现的标题中区分英文字母的大小写比如：YBaCuO而不是YBACUO。

回答 1. 在 NJUthesis.cfg 文件中加上

```
\renewcommand\title[2][\NJUT@value@title]{%
  \def\njut@value@title{#2}
  \def\njut@value@titlemark{#1}}
\def\chaptermark#1{\markboth {{\ifnum \c@secnumdepth>\m@ne
  \if@mainmatter\CTEXthechapter \quad\fi
  \fi #1}}{}}%
\def\sectionmark#1{\markright{{\ifnum \c@secnumdepth >\z@
  \CTEXthesection \quad \fi #1}}}
```

6.2 脚注

问题 2. 如果在章节标题中加入注脚，则不仅会出现在本章首页的页脚，也会出现在目录的页脚，不知是否能够让其不要出现在目录的页脚中。

回答 2. 可以使用如下的命令来定义章节的标题

```
\chapter[出现在目录和页眉的标题]{出现在正文的标题\footnote{这个
不会出现在目录中。}}
```

section、subsection 等命令也有类似的用法。

第七章 总结与展望

附录 A 博士(硕士)学位论文编写格式规定(试行)

A.1 适用范围

本规定适用于博士学位论文编写，硕士学位论文编写应参照执行。

A.2 引用标准

GB7713科学技术报告、学位论文和学术论文的编写格式。

GB7714文后参考文献著录规则。

A.3 印制要求

论文必须用白色纸印刷，并用A4(210mm×297mm)标准大小的白纸。纸的四周应留足空白边缘，上方和左侧应空边25mm以上，下方和右侧应空边20mm以上。除前置部分外，其它部分双面印刷。

论文装订不要用铁钉，以便长期存档和收藏。

论文封面与封底之间的中缝（书脊）必须有论文题目、作者和学校名。

A.4 编写格式

论文由前置部分、主体部分、附录部分(必要时)、结尾部分(必要时)组成。

前置部分包括封面，题名页，声明及说明，前言，摘要(中、英文)，关键词，目次页，插图和附表清单(必要时)，符号、标志、缩略词、首字母缩写、单位、术语、名词解释表(必要时)。

主体部分包括绪论(作为正文第一章)、正文、结论、致谢、参考文献表。

附录部分包括必要的各种附录。

结尾部分包括索引和封底。

A.5 前置部分

A.5.1 封面（博士论文国图版用）

1. 封面是论文的外表面，提供应有的信息，并起保护作用。

2. 封面上应包括下列内容:

- (a) 分类号在左上角注明分类号, 便于信息交换和处理。一般应注明《中国图书资料分类法》的类号, 同时应注明《国际十进分类法UDC》的类号;
- (b) 密级在右上角注明密级;
- (c) “博士学位论文”用大号字标明;
- (d) 题名和副题名用大号字标明;
- (e) 作者姓名;
- (f) 学科专业名称;
- (g) 研究方向;
- (h) 导师姓名, 职称;
- (i) 日期包括论文提交日期和答辩日期;
- (j) 学位授予单位。

A.5.2 题名

1. 题名是以最恰当、最简明的词语反映论文中最重要的特定内容的逻辑组合。

题名所用每一词语必须考虑到有助于选定关键词和编写题录、索引等二次文献可以提供检索的特定实用信息。

题名应避免使用不常见的缩略词、首字母缩写字、字符、代号和公式等。

题名一般不宜超过20字。

论文应有外文题名, 外文题名一般不宜超过10个实词。

2. 可以有副题名。
3. 题名在整本论文中不同地方出现时, 应完全相同。

A.5.3 前言

前言是作者对本论文基本特征的简介, 如论文背景、主旨、目的、意义等并简述本论文的创新性成果。

A.5.4 摘要

1. 摘要是论文内容不加注释和评论的简单陈述。
2. 论文应有中、英文摘要，中、英文摘要内容应相同。
3. 摘要应具有独立性和自含性，即不阅读论文的全文，便能获得必要的信息，摘要中有数据、有结论，是一篇完整的短文，可以独立使用，可以引用，可以用于推广。摘要的内容应包括与论文同等量的主要信息，供读者确定有无必要阅读全文，也供文摘等二次文献引用。摘要的重点是成果和结论。
4. 中文摘要一般在1500字，英文摘要不宜超过1500实词。
5. 摘要中不用图、表、化学结构式、非公知公用的符号和术语。

A.5.5 关键词

关键词是为了文献标引工作从论文中选取出来用于表示全文主题内容信息款目的单词或术语。

每篇论文选取3—8个词作为关键词，以显著的字符另起一行，排在摘要的左下方。在英文摘要的左下方应标注与中文对应的英文关键词。

A.5.6 目次页

目次页由论文的章、节、附录等的序号、名称和页码组成，另页排在摘要的后面。

A.5.7 插图和附表清单

论文中如图表较多，可以分别列出清单并置于目次页之后。

图的清单应有序号、图题和页码。表的清单应有序号、表题和页码。

A.5.8 符号、标志、缩略词、首字母缩写、计量单位、名词、术语等的注释表符号、标志、缩略词、首字母缩写、计量单位、名词、术语等的注释说明汇集表，应置于图表清单之后。

A.6 主体部分

A.6.1 格式

主体部分由绪论开始，以结论结束。主体部分必须由另页右页开始。每一

章必须另页开始。全部论文章、节、目的格式和版面安排要划一，层次清楚。

A.6.2 序号

1. 论文的章可以写成：第一章。节及节以下均用阿拉伯数字编排序号，如1.1，1.1.1等。
2. 论文中的图、表、附注、参考文献、公式、算式等一律用阿拉伯数字分别分章依序连续编排序号。其标注形式应便于互相区别，一般用下例：图1.2；表2.3；附注1)；文献[4]；式(6.3)等。
3. 论文一律用阿拉伯数字连续编页码。页码由首页开始，作为第1页，并为右页另页。封页、封二、封三和封底不编入页码，应为题名页、前言、目次页等前置部分单独编排页码。页码必须标注在每页的相同位置，便于识别。
4. 附录依序用大写正体A、B、C ---编序号，如：附录A。附录中的图、表、式、参考文献等另行编序号，与正文分开，也一律用阿拉伯数字编码，但在数码前题以附条序号，如图A.1；表B.2；式(B.3)；文献[A.5]等。

A.6.3 绪论

绪论（综述）：简要说明研究工作的目的、范围、相关领域的前人工作和知识空白、理论基础和分析，研究设想、研究方法和实验设计、预期结果和意义等。一般在教科书中有的知识，在绪论中不必赘述。

绪论的内容应包括论文研究方向相关领域的最新进展、对有关进展和问题的评价、本论文研究的命题和技术路线等；绪论应表明博士生对研究方向相关的学科领域有系统深入的了解，论文具有先进性和前沿性；

为了反映出作者确已掌握了坚实的基础理论和系统的专门知识，具有开阔的科学视野，对研究方案作了充分论证，绪论应单独成章，列为第一章，绪论的篇幅应达1~2万字，不得少于1万字；绪论引用的文献应在100篇以上，其中外文文献不少于60应按正文中引用的先后排列。

A.6.4 正文

论文的正文是核心部分，占主要篇幅。正文必须实事求是，客观真切，准确完备，合乎逻辑，层次分明，简便可读。

正文的每一章(除绪论外)应有小结, 在小结中应明确阐明作者在本章中所做的工作, 特别是创新性成果。凡本论文要用的基础性内容或他人的成果不应单独成章, 也不应作过多的阐述, 一般只引结论、使用条件等, 不作推导;

A.6.4.1 图

图包括曲线图、构造图、示意图、图解、框图、流程图、记录图、布置图、地图、照片、图版等。

图应具有“自明性”, 即只看图、图题和图例, 不阅读正文, 就可以理解图意。

图应编排序号。每一图应有简短确切的图题, 连同图号置于图下。必要时, 应将图上的符号、标记、代码, 以及实验条件等, 用最简练的文字, 横排于图题下方, 作为图例说明。

曲线图的纵、横坐标必须标注“量、标准规定符号、单位”。此三者只有在不必要标明(如无量纲等)的情况下方可省略。坐标上标注的量的符号和缩略词必须与正文一致。

照片图要求主题和主要显示部分的轮廓鲜明, 便于制版。如用放大缩小的复制品, 必须清晰, 反差适中。照片上应该有表示目的物尺寸的标度。

A.6.4.2 表

表的编排, 一般是内容和测试项目由左至右横读, 数据依序竖排。表应有自明性。

表应编排序号。

每一表应有简短确切的表题, 连同标号置于表上。必要时, 应将表中的符号、标记、代码, 以及需要说明事项, 以最简练的文字, 横排于表题下, 作为表注, 也可以附注于表下。表内附注的序号宜用小号阿拉伯数字并加圆括号置于被标注对象的右上角, 如: $x \times x1$), 不宜用“*”, 以免与数学上共轭和物质转移的符号相混。

表的各栏均应标明“量或测试项目、标准规定符号、单位”。只有在无必要标注的情况下方可省略。表中的缩略词和符号, 必须与正文中一致。

表内同一栏的数字必须上下对齐。表内不宜用“同上”, “同左”和类似词, 一律填入具体数字或文字。表内“空白”代表未测或无此项, “—”或“---”(因“—”可能与代表阴性反应相混)代表未发现, “0”代表实测结果确为零。

如数据已绘成曲线图, 可不再列表。

A.6.4.3 数学、物理和化学式

正文中的公式、算式或方程式等应编排序号，序号标注于该式所在行(当有续行时，应标注于最后一行)的最右边。

较长的式，另行居中横排。如式必须转行时，只能在 $+$ 、 $-$ 、 \times 、 \div 、 $\frac{}{}$ 处转行。上下式尽可能在等号“=”处对齐。

小数点用“.”表示。大于999的整数和多于三位数的小数，一律用半个阿拉伯数字字符的小间隔分开，不用千位撇。对于纯小数应将0列于小数点之前。

示例：应该写成94 652.023 567；0.314 325，不应写成94, 652.023, 567；.314, 325

应注意区别各种字符，如：拉丁文、希腊文、俄文、德文花体、草体；罗马数字和阿拉伯数字；字符的正斜体、黑白体、大小写、上下脚标(特别是多层次，如“三踏步”)、上下偏差等。

A.6.4.4 计量单位

报告、论文必须采用国务院发布的《中华人民共和国法定计量单位》，并遵照《中华人民共和国法定计量单位使用方法》执行。使用各种量、单位和符号，必须遵循附录B所列国家标准的规定执行。单位名称和符号的书写方式一律采用国际通用符号。

A.6.4.5 符号和缩略词

符号和缩略词应遵照国家标准的有关规定执行。如无标准可循，可采纳本学科或本专业的权威性机构或学术团体所公布的规定；也可以采用全国自然科学名词审定委员会编印的各学科词汇的用词。如不得不引用某些不是公知公用的、且又不易为同行读者所理解的、或系作者自定的符号、记号、缩略词、首字母缩写字等时，均应在第一次出现时一一加以说明，给以明确的定义。

A.6.5 结论

报告、论文的结论是最终的、总体的结论，不是正文中各段的小结的简单重复。结论应该准确、完整、明确、精炼。在结论中要清楚地阐明论文中有那些自己完成的成果，特别是创新性成果；

如果不可能导出应有的结论，也可以没有结论而进行必要的讨论。可以在结论或讨论中提出建议、研究设想、仪器设备改进意见、尚待解决的问题等。

A.6.6 致谢

可以在正文后对下列方面致谢：

国家科学基金、资助研究工作的奖学金基金、合作单位、资助或支持的企业、组织或个人；

协助完成研究工作和提供便利条件的组织或个人；

在研究工作中提出建议和提供帮助的人；

给予转载和引用权的资料、图片、文献、研究思想和设想的所有者；

其他应感谢的组织或个人。

A.6.7 参考文献表

A.6.7.1 专著著录格式

主要责任者，其他责任者，书名，版本，出版地：出版者，出版年

例：1. 刘少奇，论共产党员的修养，修订2版，北京：人民出版社，1962

A.6.7.2 连续出版物中析出的文献著录格式

析出文献责任者，析出文献其他责任者，析出题名，原文献题名，版本：文献中的位置。

例：2. 李四光，地壳构造与地壳运动，中国科学，1973 (4)：400—429

A.6.7.3 参考文献采用顺序编码制，按论文正文所引用文献出现的先后顺序连续编码。

A.7 附录

附录是作为报告、论文主体的补充项目，并不是必需的。

1. 下列内容可以作为附录编于报告、论文后，也可以另编成册；
 - (a) 为了整篇论文材料的完整，但编入正文又有损于编排的条理和逻辑性，这一材料包括比正文更为详尽的信息、研究方法和技术更深入的叙述，建议可以阅读的参考文献题录，对了解正文内容有用的补充信息等；
 - (b) 由于篇幅过大或取材于复制品而不便于编入正文的材料；
 - (c) 不便于编入正文的罕见珍贵资料；

- (d) 对一般读者并非必要阅读, 但对本专业同行有参考价值的资料;
- (e) 某些重要的原始数据、数学推导、计算程序、框图、结构图、注释、统计表、计算机打印输出件等。

2. 附录与正文连续编页码。

3. 每一附录均另页起。

A.8 结尾部分(必要时)

为了将论文迅速存储入电子计算机, 可以提供有关的输入数据。可以编排分类索引、著者索引、关键词索引等。

参考文献

- [1] Anwar, S., Sung, W., 2016. Compact deep convolutional neural networks with coarse pruning. arXiv: Learning .
- [2] Cai, Z., He, X., Sun, J., Vasconcelos, N., 2017. Deep learning with low precision by half-wave gaussian quantization. computer vision and pattern recognition , 5406–5414.
- [3] Chen, W., Wilson, J.T., Tyree, S., Weinberger, K.Q., Chen, Y., 2015. Compressing neural networks with the hashing trick. international conference on machine learning , 2285–2294.
- [4] Courbariaux, M., Bengio, Y., David, J.P., 2014. Low precision storage for deep learning .
- [5] Courbariaux, M., Bengio, Y., David, J.P., 2015. Binaryconnect: training deep neural networks with binary weights during propagations. neural information processing systems , 3123–3131.
- [6] Courbariaux, M., Hubara, I., Soudry, D., Elyaniv, R., Bengio, Y., 2016. Binarized neural networks: Training deep neural networks with weights and activations constrained to +1 or -1. arXiv: Learning .
- [7] Denil, M., Shakibi, B., Dinh, L., Ranzato, M., De Freitas, N., 2013. Predicting parameters in deep learning. neural information processing systems , 2148–2156.
- [8] Frankle, J., Carbin, M., 2019. The lottery ticket hypothesis: Finding sparse, trainable neural networks. International Conference on Learning Representations <https://arxiv.org/abs/1904.01303> .
- [9] Glorot, X., Bordes, A., Bengio, Y., 2011a. Deep sparse rectifier neural networks 15, 315–323.
- [10] Glorot, X., Bordes, A., Bengio, Y., 2011b. Domain adaptation for large-scale sentiment classification: A deep learning approach , 513–520.

- [11] Gong, Y., Liu, L., Yang, M., Bourdev, L.D., 2014. Compressing deep convolutional networks using vector quantization. *arXiv: Computer Vision and Pattern Recognition* .
- [12] Guo, Y., Yao, A., Chen, Y., 2016. Dynamic network surgery for efficient dnns. *neural information processing systems* , 1379–1387.
- [13] Han, S., Mao, H., Dally, W.J., 2016a. Deep compression: Compressing deep neural networks with pruning, trained quantization and huffman coding. *international conference on learning representations* .
- [14] Han, S., Mao, H., Dally, W.J., 2016b. Deep compression: Compressing deep neural networks with pruning, trained quantization and huffman coding. *international conference on learning representations* .
- [15] Han, S., Pool, J., Tran, J., Dally, W.J., 2015. Learning both weights and connections for efficient neural networks. *neural information processing systems* , 1135–1143.
- [16] Hanson, S.J., Pratt, L.Y., 1988. Comparing biases for minimal network construction with back-propagation , 177–185.
- [17] Hassibi, B., Stork, D.G., 1992. Second order derivatives for network pruning: Optimal brain surgeon , 164–171.
- [18] He, K., Zhang, X., Ren, S., Sun, J., 2015. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. *international conference on computer vision* , 1026–1034.
- [19] He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. *computer vision and pattern recognition* , 770–778.
- [20] He, Y., Kang, G., Dong, X., Fu, Y., Yang, Y., 2018. Soft filter pruning for accelerating deep convolutional neural networks. *international joint conference on artificial intelligence* , 2234–2240.
- [21] Hinton, G.E., Srivastava, N., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2012. Improving neural networks by preventing co-adaptation of feature detectors. *arXiv: Neural and Evolutionary Computing* .

- [22] Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., Adam, H., 2017. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv: Computer Vision and Pattern Recognition* .
- [23] Hu, J., Shen, L., Sun, G., 2018. Squeeze-and-excitation networks. *computer vision and pattern recognition* , 7132–7141.
- [24] Huang, G., Liu, Z., Der Maaten, L.V., Weinberger, K.Q., 2017. Densely connected convolutional networks. *computer vision and pattern recognition* , 2261–2269.
- [25] Hubara, I., Courbariaux, M., Soudry, D., Elyaniv, R., Bengio, Y., 2017. Quantized neural networks: training neural networks with low precision weights and activations. *Journal of Machine Learning Research* 18, 6869–6898.
- [26] Iandola, F.N., Han, S., Moskewicz, M.W., Ashraf, K., Dally, W.J., Keutzer, K., 2017. Squeezenet: Alexnet-level accuracy with 50x fewer parameters and 0.5mb model size. *arXiv: Computer Vision and Pattern Recognition* .
- [27] Ioffe, S., Szegedy, C., 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *international conference on machine learning* , 448–456.
- [28] Jacob, B., Kligys, S., Chen, B., Zhu, M., Tang, M., Howard, A.G., Adam, H., Kalenichenko, D., 2018. Quantization and training of neural networks for efficient integer-arithmetic-only inference. *computer vision and pattern recognition* , 2704–2713.
- [29] Jegou, H., Douze, M., Schmid, C., 2011. Product quantization for nearest neighbor search. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33, 117–128.
- [30] Jin, X., Yuan, X., Feng, J., Yan, S., 2016. Training skinny deep neural networks with iterative hard thresholding methods. *arXiv: Computer Vision and Pattern Recognition* .
- [31] Karpathy, A., Feifei, L., 2015. Deep visual-semantic alignments for generating image descriptions. *computer vision and pattern recognition* , 3128–3137.

- [32] Krizhevsky, A., Hinton, G., 2009. Learning multiple layers of features from tiny images .
- [33] Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012a. Imagenet classification with deep convolutional neural networks , 1097–1105.
- [34] Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012b. Imagenet classification with deep convolutional neural networks, in: Advances in Neural Information Processing Systems 25: 26th Annual Conference on Neural Information Processing Systems 2012. Proceedings of a meeting held December 3-6, 2012, Lake Tahoe, Nevada, United States., pp. 1106–1114.
- [35] Lecun, Y., Bottou, L., Bengio, Y., Haffner, P., 1998. Gradient-based learning applied to document recognition. Proceedings of the IEEE 86, 2278–2324.
- [36] Lecun, Y., Denker, J.S., Solla, S.A., 1989. Optimal brain damage 2, 598–605.
- [37] Leng, C., Dou, Z., Li, H., Zhu, S., Jin, R., 2018. Extremely low bit neural network: Squeeze the last bit out with admm. national conference on artificial intelligence , 3466–3473.
- [38] Li, F., Liu, B., 2016. Ternary weight networks. arXiv: Computer Vision and Pattern Recognition .
- [39] Li, H., Kadav, A., Durdanovic, I., Samet, H., Graf, H.P., 2017. Pruning filters for efficient convnets. international conference on learning representations .
- [40] Lin, M., Chen, Q., Yan, S., 2014. Network in network. international conference on learning representations .
- [41] Liu, X., Pool, J., 2018. Efficient sparse-winograd convolutional neural networks. international conference on learning representations .
- [42] Liu, Z., Li, J., Shen, Z., Huang, G., Yan, S., Zhang, C., 2017. Learning efficient convolutional networks through network slimming. international conference on computer vision , 2755–2763.
- [43] Molchanov, P., Tyree, S., Karras, T., Aila, T., Kautz, J., 2017. Pruning convolutional neural networks for resource efficient inference. international conference on learning representations .

- [44] Rastegari, M., Ordonez, V., Redmon, J., Farhadi, A., 2016. Xnor-net: Imagenet classification using binary convolutional neural networks. *European conference on computer vision* , 525–542.
- [45] Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1988. Learning representations by back-propagating errors. *Nature* 323, 696–699.
- [46] Scardapane, S., Comminiello, D., Hussain, A., Uncini, A., 2017. Group sparse regularization for deep neural networks. *Neurocomputing* 241, 81–89.
- [47] Srivastava, N., Hinton, G.E., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: a simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research* 15, 1929–1958.
- [48] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S.E., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., 2015. Going deeper with convolutions. *computer vision and pattern recognition* , 1–9.
- [49] Tian, Q., Arbel, T., Clark, J.J., 2017. Efficient gender classification using a deep l1-pruned net. *arXiv: Computer Vision and Pattern Recognition* .
- [50] Vinyals, O., Toshev, A., Bengio, S., Erhan, D., 2015. Show and tell: A neural image caption generator. *computer vision and pattern recognition* , 3156–3164.
- [51] Wen, W., Wu, C., Wang, Y., Chen, Y., Li, H., 2016. Learning structured sparsity in deep neural networks. *neural information processing systems* , 2074–2082.
- [52] Yang, T., Chen, Y., Sze, V., 2017. Designing energy-efficient convolutional neural networks using energy-aware pruning. *computer vision and pattern recognition* , 6071–6079.
- [53] Ye, J., Lu, X., Lin, Z.L., Wang, J.Z., 2018. Rethinking the smaller-norm-less-informative assumption in channel pruning of convolution layers. *international conference on learning representations* .
- [54] Zhang, X., Zhou, X., Lin, M., Sun, J., 2018. Shufflenet: An extremely efficient convolutional neural network for mobile devices. *computer vision and pattern recognition* , 6848–6856.

- [55] Zhong, J., Ding, G., Guo, Y., Han, J., Wang, B., 2018. Where to prune: Using lstm to guide end-to-end pruning , 3205–3211.
- [56] Zhou, A., Yao, A., Guo, Y., Xu, L., Chen, Y., 2017. Incremental network quantization: Towards lossless cnns with low-precision weights. international conference on learning representations .
- [57] Zhou, S., Ni, Z., Zhou, X., Wen, H., Wu, Y., Zou, Y., 2016. Dorefa-net: Training low bitwidth convolutional neural networks with low bitwidth gradients. arXiv: Neural and Evolutionary Computing .
- [58] Zhu, C., Han, S., Mao, H., Dally, W.J., 2017. Trained ternary quantization. international conference on learning representations .

简历与科研成果

基本情况

杨文博，男，汉族，1986 年 11 月出生，河南省民权县人。

教育背景

2007.9~2010.6	中国科学院研究生院信息安全国家重点实验室	硕士
2003.9~2007.6	南京大学数学系	本科

攻读硕士学位期间完成的学术成果

- [1] Wenbo Yang and Wen Tao Zhu, “Voting-on-Grid Clustering for Secure Localization in Wireless Sensor Networks,” in *Proc. IEEE International Conference on Communications (ICC) 2010*, May. 2010.
- [2] Wenbo Yang and Wen Tao Zhu, “Protecting Source Location Privacy in Wireless Sensor Networks with Data Aggregation,” in *Proc. 7th International Conference on Ubiquitous Intelligence and Computing (UIC) 2010*, Oct. 2010.

攻读硕士学位期间参与的科研课题

- 1. 国家自然科学基金面上项目“无线传感器网络在知识获取过程中的若干安全问题研究”（课题年限 2010.1~2012.12），负责位置相关安全问题的研究。
- 2. 中国科学院知识创新工程重要方向项目下属课题“下一代移动通信安全机制研究”（课题年限 2010.1~2010.12），负责 LTE/SAE 认证相关的安全问题研究。

致 谢

首先感谢XXX