

分类号 \_\_\_\_\_

编号 \_\_\_\_\_

U D C \_\_\_\_\_

密级 \_\_\_\_\_



南方科技大学

SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

# 本科生毕业设计（论文）

题目：利用拓扑卷积神经网络实现手写数字识别

姓名：胡子胤

学号：12010623

系别：数学系

专业：数学与应用数学

指导教师：朱一飞 tenure-track Assistant Professor

2024年5月8日

CLC \_\_\_\_\_

Number \_\_\_\_\_

UDC \_\_\_\_\_

Available for reference Yes No



**SUSTech**

Southern University  
of Science and  
Technology

# Undergraduate Thesis

**Thesis Title:** Topological Convolutional Layers  
for Handwritten Digit Recognition

**Student Name:** Ziyin Hu

**Student ID:** 12010623

**Department:** Department of Mathematics

**Program:** Mathematics and Applied mathematics

**Thesis Advisor:** tenure-track Assistant Professor Yifei Zhu

Date: May 8, 2024

# 诚信承诺书

1. 本人郑重承诺所呈交的毕业设计（论文），是在导师的指导下，独立进行研究工作所取得的成果，所有数据、图片资料均真实可靠。

2. 除文中已经注明引用的内容外，本论文不包含任何其他人或集体已经发表或撰写过的作品或成果。对本论文的研究作出重要贡献的个人和集体，均已在文中以明确的方式标明。

3. 本人承诺在毕业论文（设计）选题和研究内容过程中没有抄袭他人研究成果和伪造相关数据等行为。

4. 在毕业论文（设计）中对侵犯任何方面知识产权的行为，由本人承担相应的法律责任。

作者签名:

\_\_\_\_\_年\_\_\_\_月\_\_\_\_日

# COMMITMENT OF HONESTY

1. I solemnly promise that the paper presented comes from my independent research work under my supervisor's supervision. All statistics and images are real and reliable.
2. Except for the annotated reference, the paper contents no other published work or achievement by person or group. All people making important contributions to the study of the paper have been indicated clearly in the paper.
3. I promise that I did not plagiarize other people's research achievement or forge related data in the process of designing topic and research content.
4. If there is violation of any intellectual property right, I will take legal responsibility myself.

Signature:

Date:

# 利用拓扑卷积神经网络实现手写数字识别

胡子胤

(数学系 指导教师: 朱一飞)

**[摘要]:** 本文基于 LOVE E R, FILIPPENKO B, MAROULAS V, Carlsson G 发表的 Topological convolutional layers for deep learning (2023) 的结果改进了传统的卷积神经网络. 原论文提出了四种新的卷积层, 可以分成两类. 第一类是 Klein One Layer (KOL) 和 Circle One Layer (COL). 第二类是 Klein Filters (KF) layer 和 Circle Filters (CF) layer. 传统的卷积层利用 NOL+NOL 的结构, 论文中作者使用 KF+KOL 的结构. 我们受到原论文的启发, 使用 KF+NOL 以及 KF+KF+Pool 的结构, 我们估计这个结构会对 NOL+NOL 在准确度和运行时间上均有提升. 由于手写数字特征简单容易提取, 对数据作预处理, 然后对最后的线性层进行梯度下降和反向传播. 在我们的结构中, 两层 KF 起到了和预处理相同的作用. 该神经网络在准确率、泛化性、鲁棒性上超过了 NOL+NOL 模型, 在运行时间上缩短到 1 小时以内 (RTX3060 with 8GB).

**[关键词]:** 深度学习, 卷积层, 克莱因瓶, 拓扑

**[ABSTRACT]:** This article bases on the results of Topological Convolutional Layers for Deep Learning and balances the framework of the paper with the traditional framework. The original paper proposed four new convolutional layer structures, which can be divided into two types. The first type is Klein One Layer (KOL) and Circle One Layer (COL). The second type is Klein Filters (KF) layer and Circle Filters (CF) layer. We balance training time and accuracy. Traditional convolutional layers utilize the structure of NOL+NOL, while the authors in the paper use the structure of KF+KOL. We use the structures of KF+KF+Pool. Due to the simplicity and ease of extracting handwritten digit features, the data is preprocessed and then gradient descent and back propagation are applied to the final linear layer. In our structure, the two-layer KF plays the same role as the preprocessing. This neural network surpasses the NOL+NOL model in accuracy, loss rate, generalizability, and robustness, and reduces the running time to half of KF+KOL.

**[Key words]:** Deep Learning, Convolutional Layers, Klein Bottle, Topology

# 目录

<b>1. 引言</b>	<b>1</b>
<b>2. 背景</b>	<b>2</b>
2.1 卷积层的描述	2
2.2 激活函数的描述	3
<b>3. 拓扑卷积层</b>	<b>3</b>
3.1 Klein One Layer (KOL) 与 Circle One layer (COL)	4
3.2 Klein Filters (KF) layer 和 Circle Filters (CF) layer	5
3.3 可视化	6
<b>4. 实现</b>	<b>7</b>
4.1 元参数的选择	7
4.2 普通实验	8
4.3 噪声实验	8
4.4 泛化实验	8
<b>5. 实验结果</b>	<b>8</b>
5.1 普通实验	9
5.2 噪声实验	10
5.3 泛化实验	11
<b>6. 总结</b>	<b>11</b>
<b>参考文献</b>	<b>13</b>
<b>致谢</b>	<b>14</b>

# 1. 引言

我们基于 Topological Convolutional Layers for Deep Learning, 使用了它提出的卷积核和卷积层, 搭建了新的神经网络. 我们通过在卷积层上的附加结构来实现 Klein Bottle 产生图像滤波器.

原论文考虑了一般前馈神经网络的节点图表示, 并且注意到神经网络以及不同层的神经元之间的权重对应于图中的一条边, 如果把不同的神经元从简单的节点图放到结构更复杂的流形上, 如二维圆, Klein Bottle, 有可能直接提取更高级的特征. 具体实现可以对 KleinBottle 进行参数化, 并且给不同的层赋予一个由该参数化产生的特征标. 由此产生的 filters 与 Gabor Filters 的一个子族密切相关. 我们将具有这样的层的神经网络称为拓扑卷积神经网络 (TCNNs). 原论文引入了四种新的神经网络层, 分成两类. 第一类包含 Circle One Layer (COL), Klein One Layer (KOL). 第二类包含 Circle Filters Layer (CF) 和 Klein Filters Layer (KF). 传统的卷积层被称为 Normal One Layer (NOL). 原论文利用两个全新的层搭建了一种新的结构 KF+KOL.

LeCun 等人 (1998) 观察到, 能够帮我们构建卷积神经网络 (CNNs) 的一个启发点是, 它们允许基于图像特征空间的几何形状的稀疏化. TCNN 层的拓扑结构灵感来自对拓扑数据分析 (TDA) 的研究. Carlsson 等人 (2008) 发现, 自然图像数据库中的  $3 \times 3$  像素图聚集在一个与克莱因瓶同态的 2-流形周围, 我们用  $K$  表示.  $K$  有一个简单的代数表示, 原论文利用这种代数表示构造了一个函数  $F_K : KleinBottle \rightarrow (I^2, I), I = [0, 1]$ . 这个函数同时满足了在 Klein 瓶等价关系下取值相同, 并且能够根据点在  $K$  上位置的不同生成不同的 filters, 这些 filters 自身具有对称性, 彼此之间具有相似性, 并且提取不同的特征, 比如与水平成不同角度的直线. 这也是使用这个 filters 的神经网络能够有更好表现的核心原因.

我们着重于实现原论文的卷积核并分析实验结果. 我们使用了原论文里面利用拓扑方法定义的卷积核, 改变了神经网络的结构, 层数不变. 我们主要对训练时间和准确度进行了中和, 即争取实现比传统模型用时更短, 准确率更高, 鲁棒性更强, 泛化性更好的模型. 而且我们发现, 由于 Klein Bottle 拓扑结构比平面复杂, 如果让卷积神经网络的节点分布在 Klein Bottle 上, 我们就可以在更少的数据上, 训练更少的 epoch

就迅速达到和最终训练结果非常接近的水平. 这可能会对某些数据集稀缺的训练任务具有意义.

本文的结构如下: 在第 2 节, 我们回顾了卷积神经网络的数学表示. 在第 3 节, 回顾了原论文定义的 4 个卷积层并对卷积层生成的 filters 做了可视化. 在第 4 节, 我们给出了我们实现实验的元参数和 3 种对比实验的目的和网络结构. 在第 5 节, 我们给出并分析了实验结果.

## 2. 背景

我们在这节使用数学语言描述传统 CNNs 的架构, 并在下一节进一步描述 TCNNs 新定义的几个层. 在图像识别中, 把每个像素看成一个神经元, 如果相邻两层的每个神经元之间都有权重 (同层神经元无权重), 则我们称这个层为一个全连接前向传播层 fully connected, feed-forward (FF) layer, 因此对于一般的神经网络我们可以把它看成一个全连接层的稀疏化. 传统的神经网络探寻潜在图像空间, 即通过卷积核识别图像所蕴含的空间特征. 神经网络必须在局部和整体上做出平衡, 与某个像素之间权重不为 0 的像素数量如果太少, 会造成无法得到很好地拟合结果, 如果数量太多, 会造成运算成本大幅增大, 而且有可能造成由于过滤器 filter 尺寸变大造成的过拟合. 如果能够发现潜在图像空间的拓扑特点, 我们可以定义新的层来实现得到较好的拟合效果, 并且不会增大计算量. 这具体由把每个层利用参数化的 Klein Bottle 进行标记来实现, 只计算距离较近的像素之间的权重 (下文会对这个距离进行定义), 其余权重都冻结为 0.

### 2.1 卷积层的描述

定义 1. 一个前馈传播神经网络 (Feed Forward Neural Network (FFNN)) 是一个有向无环图, 这个有向无环图的节点集  $V(\Gamma)$  满足如下性质: 1.  $V(\Gamma)$  由它的层的无交并组成  $V(\Gamma) = V_0(\Gamma) \sqcup V_1(\Gamma) \dots \sqcup V_r(\Gamma)$  2. 如果  $v \in V_i(\Gamma)$ , 那么  $\Gamma$  的每条边  $(v, w) \in V_{i+1}(\Gamma)$  3. 对于每个非初始节点  $w \in V_i(\Gamma): i > 0$ , 存在至少一个  $v \in V_{i-1}(\Gamma)$  使得  $(v, w)$  是  $\Gamma$  的一条边.

为了描述相邻层之间的关联, 我们把所有在  $V_i$  和  $V_{i+1}$  两个节点之间的边作为集合  $C$  的元素, 那么  $C \subset V_i \times V_{i+1}$ . 对于  $v_0 \in V_i$  和  $w_0 \in V_{i+1}$ , 我们定义子集

$C(v_0) := \{w \in V_{i+1} | (v_0, w) \in C\}, C^{-1}(w_0) := \{v \in V_i | (v, w_0) \in C\}$ .

当  $C = V_i \times V_{i+1}$  时，这是结构最简单的卷积层，即全连接层，我们定义如下：

定义 2. 我们称  $V_{i+1}$  是一个全连接层，如果  $C = V_i \times V_{i+1}$ .

我们继续描述卷积层，我们用  $\mathbb{Z}^2$  来给手写数字图像中的每一个像素赋予一个指标，每个像素颜色可以对应于一个数字  $i \in [0, 1]$ . 等价地，一个图像就是一个映射  $\mathbb{Z}^2 \rightarrow [0, 1]$ . 每个图像经过一层神经网络之后得到新的图像，所以我们可以用  $\chi \times \mathbb{Z}^2$  来表示每一个层，即  $V_i = \chi \times \mathbb{Z}^2$ .

定义 3. 让  $V_{i+1}$  是 FFNN 中的一个层。我们把它称为一个 Normal One Layer (NOL) 如果  $V_i = \chi \times \mathbb{Z}^2$  且  $V_{i+1} = \chi' \times \mathbb{Z}^2$  对于某些有限集合  $\chi$  和  $\chi'$ ，而且对于某个固定的阈值  $s \geq 0$ ,  $C = C_c \times C_{d,N}(s)$ , 其中  $C_c = \chi \times \chi'$  且  $C_{d,N}(s) \subset \mathbb{Z}^2 \times \mathbb{Z}^2$  给出如下  $C_{d,N}(s)^{-1}(x') := \{x \in \mathbb{Z}^2 | d_{\mathbb{Z}^2}(x, x') \leq s\}$ . 其中， $d$  是定义在  $\mathbb{Z}^2$  上的  $L^\infty$ -距离.

一个卷积神经网络 (CNN) 是一个 FFNN 使得前面的层  $V_0, \dots, v_i$  是卷积层并且后面的层  $V_{i+1}, \dots, V_r$  是全连接层.

## 2.2 激活函数的描述

激活函数的作用在于为神经网络提供非线性拟合，由于卷积层都是线性的，多个线性函数的叠加仍然是线性的，如果不在卷积层之间加入非线性层，多个卷积层的效果等价于一个卷积层. 除了非常简单的任务，我们都需要加入激活函数增加神经网络的拟合能力. 输出值是输出节点的概率分布，所以输出是一个非负实数  $x \in \mathbb{R}, 0 \leq x \leq 1$ . 激活函数的种类很多，对于非最终层我们选择最常用的 ReLU 函数  $f(x) = \max(0, x)$ , 对于最后一层的分类任务我们选择 softmax 函数  $\sigma(x_i) = e^{x_i} / (\sum e^{x_j})^{-1}$ . 优化器负责决定如何基于梯度下降进行反向传播，我们选择通常的 Adam 优化器. 在我们的实验中，激活函数和优化器的选择和原论文是一致的，所以不会对实验结果产生这方面的影响.

## 3. 拓扑卷积层

我们在前一节中已经看到，卷积神经网络可以直接从像素网格上的图结构中构造出来. 卷积条件也可以直接从这些图的结构中推导出来的. 这意味着人们可以从其他几何图形或流形中构造类似的层和神经网络. 在这种情况下，层中的节点将对应于

图的顶点，而连接只存在于在距离比较近的节点之间，并且对应于图的边. 我们使用克莱因瓶  $K$  和二维圆  $S^1$ , 其中  $S^1$  可以嵌入  $K$ , 所以在卷积层方面,  $S^1$  是  $K$  的一种特殊情况. 具体来说, 我们先对  $K$  进行二维参数化, 然后令其中一个参数固定, 从而得到  $S^1$ . 在训练过程中, KF 层内部的参数是被冻结的, 可以看作一个固定的滤波器, 起到的对图像进行预处理的作用. KOL 能够消除不相关之间的像素的联系, 在由拓扑得出的图像隐藏特征上进行高效的提取. 相同的论述可以应用到 CF 和 COL.

与 NOL 进行对比, 在 KF 和 CF 层上我们可以明显地发现更强的可解释性和鲁棒性. 从原因上来看, 由于 CNN 学习局部条件, 由方向导数进行梯度下降得到  $S^1$  可以从切方向体现局部信息, 所以神经网络可以直接学习二阶局部信息. 在  $K$  中, 一些二阶局部信息被直接包含在内, 所以神经网络可以学习额外的二阶信息, 甚至是三阶信息. 从结果上来看, 可解释性, 观察得到的 KF 层, 即一个滤波器, 每一个 KF 层的排布是有规律的, 沿着某条轴线对称, 或者是黑白反转的, 这种旋转不变性和黑白反转不变性使得它能够应对数字只有位置发生改变或者是笔迹颜色发生改变而形状不变, 我们要学习的关键特征也就是数字的形状. 鲁棒性, 将 KF 的得到的卷积层作用于手写数字 5 上, 我们发现其忽略了很多细小的特征, 比如由于手写导致的数字边缘抖动, 这些特征不是数字 5 本身的特征, 所以忽略这些特征有利于模型学习, 同时防止过拟合. 而传统的 NOL 由于其神经元之间的联系没有被限制, 所以有时会对这些特征做出反应, 当训练的 epoch 增大, 会产生过拟合.

在 3.1 节中我们介绍 KOL 和 COL. 在 3.2 中介绍 KF 和 CF, KF 和 CF 由 KOL 和 COL 直接产生, 事实上只需要对定义的函数在  $K$  或  $S^1$  参数化后在标准正方形区域上计算二重定积分. 在 3.3 中, 我们可视化由 KF 和 CF 产生滤波器, 以及该滤波器作用在 5 上的结果.

### 3.1 Klein One Layer (KOL) 与 Circle One layer (COL)

定义 4.  $M$  是一个流形且  $\chi, \chi' \subset M$  是  $M$  的两个离散化, 即有限个点组成的集合. 令  $V_i = \chi \times \mathbb{Z}^2$  且  $V_{i+1} = \chi' \times \mathbb{Z}^2$  是 FFNN 中两个相邻的层. 固定一个阈值  $s \geq 0$ .  $d$  是  $M$  上的一个度量. 定义  $C(s) \subset \chi \times \chi', C(s)^{-1}(\kappa') = \{\kappa \in \chi | d(\kappa, \kappa') \leq s\}$  对于所有的  $\kappa' \in \chi'$ . 同时固定另一个阈值  $s' \geq 0$ , 这定义了  $C \subset V_i \times V_{i+1}$  由  $C = C(s) \times C_{d,N}(s')$ .

定义 4 给出了在流形上定义卷积层的一般方法, 但对于特殊的例子我们还需要对其中的度量  $d$  进行定义, 我们接下来考虑一个二维圆  $S^1$ , 它的一个特殊的离散化是  $n$  次单位根  $\chi = \{e^{2\pi ik/n} | 0 \leq k \leq n-1\}$ , 其中  $n \geq 1$ .

定义 5.  $\chi, \chi' \subset S^1$  是一个圆的两个离散化. 令  $V_i = \chi \times \mathbb{Z}^2$  且  $V_{i+1} = \chi' \times \mathbb{Z}^2$  是 FFNN 中两个相邻的层. 固定一个阈值  $s \geq 0$ .  $C_S(s) \subset \chi \times \chi'$  定义如下,  $C_S(s)^{-1}(\kappa') = \{\kappa \in \chi | d_S(\kappa, \kappa') \leq s\}$  对于所有的  $\kappa' \in \chi'$ , 其中度量  $d_S$  给出如下,  $d_S(\kappa, \kappa') = \cos^{-1}(\kappa, \kappa')$  对于  $\kappa, \kappa' \in S^1$ . 我们把  $V_{i+1}$  叫作 Circle One Layer (COL) 如果对于某个其他的阈值  $s' \geq 0$ ,  $C = C_S(s) \times C_{d,2}(s')$ .

现在我们定义 Klein Bottle 上的一个函数. 首先对 Klein Bottle 作参数化. Klein Bottle 可以由  $\mathbb{R}^2$  上的等价关系  $(\theta_1, \theta_2) \sim (\theta_1 + 2k\pi, \theta_2 + 2l\pi)$  对于  $k, l \in \mathbb{Z}$  和  $(\theta_1, \theta_2) \sim (\theta_1 + \pi, -\theta_2)$  得到. 我们定义函数  $F_K(\theta_1, \theta_2)(x, y) = \sin(\theta_2)(\cos(\theta_1)x + \sin(\theta_1)y) + \cos(\theta_2)Q(\cos(\theta_1)x + \sin(\theta_1)y)$ , 其中  $Q(t) = 2t^2 - 1$ . 这个函数满足  $F_K(\theta_1, \theta_2) = F_K(\theta_1 + 2k\pi, \theta_2 + 2l\pi)$  和  $F_K(\theta_1 + \pi, -\theta_2) = F_K(\theta_1, \theta_2)$ .  $F$  是从 Klein Bottle 到  $I^2 \rightarrow I$  的函数, 其中  $I = [0, 1]$ .

定义 6.  $\chi, \chi' \subset K$  是一个 Klein Bottle 的两个离散化. 令  $V_i = \chi \times \mathbb{Z}^2$  且  $V_{i+1} = \chi' \times \mathbb{Z}^2$  是 FFNN 中两个相邻的层. 固定一个阈值  $s \geq 0$ .  $C_K(s) \subset \chi \times \chi'$  定义如下,  $C_K(s)^{-1}(\kappa') = \{\kappa \in \chi | d_S(\kappa, \kappa') \leq s\}$  对于所有的  $\kappa' \in \chi'$ , 其中度量  $d_S$  给出如下,  $d_S(\kappa, \kappa') = (\int_{[-1,1]^2} (F_K(\kappa)(x, y) - F_K(\kappa')(x, y))^2)^{1/2}$  对于  $\kappa, \kappa' \in S^1$ . 我们把  $V_{i+1}$  叫作 Klein One Layer (KOL) 如果对于某个其他的阈值  $s' \geq 0$ ,  $C = C_K(s) \times C_{d,2}(s')$ .

### 3.2 Klein Filters (KF) layer 和 Circle Filters (CF) layer

基于上面的定义, 我们下面定义两个新的层 KF 和 CF. 首先, 我们给出  $S^1$  到  $K$  的嵌入并定义  $F_S^1(\theta)(x, y) := F_K(\theta, \pi/2)(x, y) = \cos(\theta)x + \sin(\theta)y$ . 这里令第二个参数取  $\pi/2$  将  $S^1$  嵌入  $K$ .

定义 7. 令  $M = S^1$  或  $K$  并且让  $\chi \subset M$  是一个有限集. 令  $V_i = \chi \times \mathbb{Z}^2$  且  $V_{i+1} = \chi' \times \mathbb{Z}^2$  是 FFNN 中两个相邻的层. 假设  $V_{i+1}$  是一个阈值  $s \geq 0$  的卷积层 (Definition 3). 则  $V_{i+1}$  被叫做一个 Klein Filters (KF) layer 或一个 Circle Filters (CF) layer, 如果神经元 (像素) 之间的权重  $\lambda_{-, (\kappa, -, -)}$ , 其中  $\kappa$  已经被给出, 由尺寸为  $(2s+1) \times (2s+1)$  的滤波

器给出, 其中值由  $Filter(\kappa)(n, m) = \int_{-1+2m/(2s+1)}^{-1+2(m+1)/(2s+1)} \int_{-1+2n/(2s+1)}^{-1+2(n+1)/(2s+1)} F_M(x, y) dx dy$  给出. 我们在这里可以看出, KF 和 CF 层中的参数由滤波器完全确定, 而滤波器由离散化集合中的  $\kappa$  和该滤波器的尺寸决定. 而且, 当滤波器的 size 变大, 被积分范围并不会改变, 始终是  $[-1, 1]^2$ , 所以 size 增大只是将  $[-1, 1]^2$  细分, 由于函数是连续的, size 变大意味着过渡更加平缓. 为了中和过渡平缓性和计算量, 我们选择神经网络中常用的 size  $5*5$ , 其中  $s = 2$ . 这里选择更大的滤波器可能会使实验结果更好, 但是计算量会倍增, 每一层由  $5*5$  增加到  $7*7$ .

### 3.3 可视化

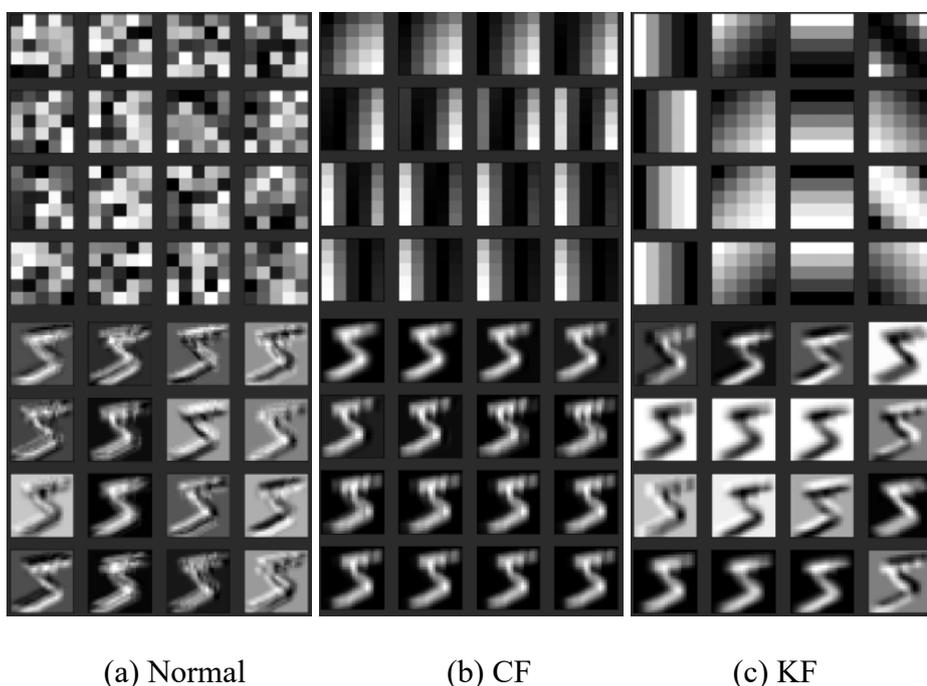


图 1

正如我们在本节开头承诺, 明显观察到 CF 和 KF 产生的 filter 中的图像对称性, 包含旋转不变性, 翻转不变性, 黑白反转不变性. 另一方面, 观察卷积之后的结果, 仍然能够看出是数字 5. KF 和 CF 在棱角处明显比 Normal 要光滑, 这能同时增加鲁棒性和泛化性. 鲁棒性可以由噪声测试体现, 这是由于添加噪声之后可能会对原本非关键特征的特征造成加剧, 所以越平滑越能抵消这种噪声带来的非关键特征加剧, 从而越能抵消其对模型带来的特征误导. 泛化性指模型学习的迁移能力, 这个性质与过拟合紧密相关. 一般来说泛化性强, 表示学习的是重要特征, 从而不容易产生过拟合, 泛化性弱可能是由于学习了数据集所特有的特征, 所以容易造成过拟合, 比如某一个手

写数字数据集的数字可能由一个人写了很多，这会让模型去学习这个人的写字习惯造成的特征，而不是数字本身所具有的特征. 我们的计算结果和原论文计算出的卷积核有些不同，可能是导致实验结果没有原论文显著优越于传统模型的原因之一. 但需要注意的是，这里得到的卷积核与原论文不完全相同，我们得到的 kernel 在整体上是相同的，但是每一个 kernel 并不完全相同. 举例来说，我们的某个 kernel 可能由原论文的两个 kernel 的一半拼接而成，或者旋转一定角度. 这可能与不同的参数化和积分域有关，我们尚未完全解决这个问题.

## 4. 实现

我们主要实现了 KF 和 CF 层，并把其作为我们神经网络的第一层. 我们搭建了如下的神经网络分别进行了三组对比实验：普通实验，噪声实验和泛化性实验. 在训练过程中，按照原论文描述，我们冻结这些层，仅仅把这些层作为数据的预处理. 我们通过反向传播训练的只是后面的 NOL 层.

我们使用 pytorch 框架来实现这些模型. 这些模型由两层带有 ReLU 激活函数的卷积层和后面的平展层组成，平展层仅作为处理数据的必要手段，与常规的平展层相同，我们没有作额外处理. 损失函数我们选择了 cross-entropy loss function，这个函数在传统的手写数字识别任务中表现很好，作为计算梯度下降的工具，这不会对我们的实验造成很大影响. 优化器我们选择了最常用的 Adam 优化器. 所有实验均在 RTX 3060 with 8GB of VRAM 上运行.

### 4.1 元参数的选择

	普通实验	噪声实验	泛化实验
EPOCH	60	15	20
LEARNING_RATE	1.00E-05	1.00E-05	1.00E-05
BATCH_SIZE	500	500	500
KERNEL_SIZE	3	3	3
CONV_SLICES	64	64	64

表 1

batch 的选择会影响训练时间，我们尝试了 batch 的值分别为 100, 300, 500 得到的结果非常近似，我们只讨论 epoch 为 500 时候的情况. 这是由于 epoch 增大，训练速度加快，但对 GPU 性能要求提高，当 batch=500 时，总训练时间不超过 1 个小时，

而且能在性能较好的笔记本电脑上运行。而当  $\text{batch}=100$  时，总训练时间达到 6 小时以上，而训练结果等于甚至可能劣于  $\text{batch}=500$  时。

## 4.2 普通实验

在本实验中，我们使用可视化了三个卷积层（NOL、CF、KF）的权重，并将它们的激活情况应用于 5 的 MNIST 图像。CF 滤波器是通过指定圆周围 16 个均匀间隔的角度来生成的，16 个 KF 滤波器是通过为克莱因瓶上的 2 个角度指定 4 个均匀间隔的值来生成的。不需要对 CF 层和 KF 层进行训练，因为这些层的权重总是固定的。我们训练了以下模型：NOL+NOL、KF+NOL、CF+NOL。

## 4.3 噪声实验

这个实验为了研究 TCNN 在不同数据集上的抗噪声表现（鲁棒性），我们在 MNIST 数据集上添加了服从高斯分布  $N(0.2, 0.04)$  的噪声。并且进行了正反测试，在有噪声上训练在无噪声测试，以及在无噪声上训练在有噪声上测试。由于 CF 和 KF 在对手写数字 5 的处理的优良表现，我们预测 CF 和 KF 能够在噪声处理中同样有优良表现。我们训练了以下模型：NOL+NOL、KF+NOL、CF+NOL。

## 4.4 泛化实验

这个实验为了研究 TCNN 的泛化性，即学习迁移能力的强弱，能否从一个数据集迁移到另一个数据集。我们训练了以下模型：KF+KF、KF+KF+Pool、NOL+NOL、NOL+NOL+Pool。

## 5. 实验结果

每个实验结果都包含两张图，分别是准确率和交叉熵损失函数计算的损失。损失函数用来计算输出值和真实值之间的差异，趋势总是和准确率相反，所以我们在这里只讨论准确率。

5.1 包含了三个模型的一次对比实验。5.2 包含了三个模型的两次对比实验。5.3 包含了三个模型的两次对比实验。

## 5.1 普通实验

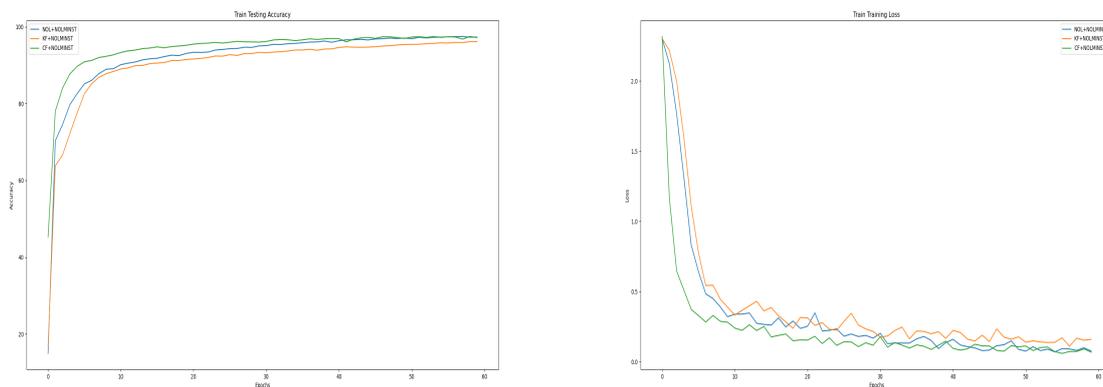


图 2

从左至右分别是：准确率和损失我们的测试结果显示，只关注 50 个 epoch 之后的准确率，CF 层的结果优于 NOL，NOL 明显优于 KF. 在接近准确的速度上，CF 明显由于 NOL，这里和原论文以及我们的预测一致，CF 可以提取快速地提取某些高级特征. 可能对于更复杂的分类任务，这点可以极大减少样本量和训练时间.

## 5.2 噪声实验

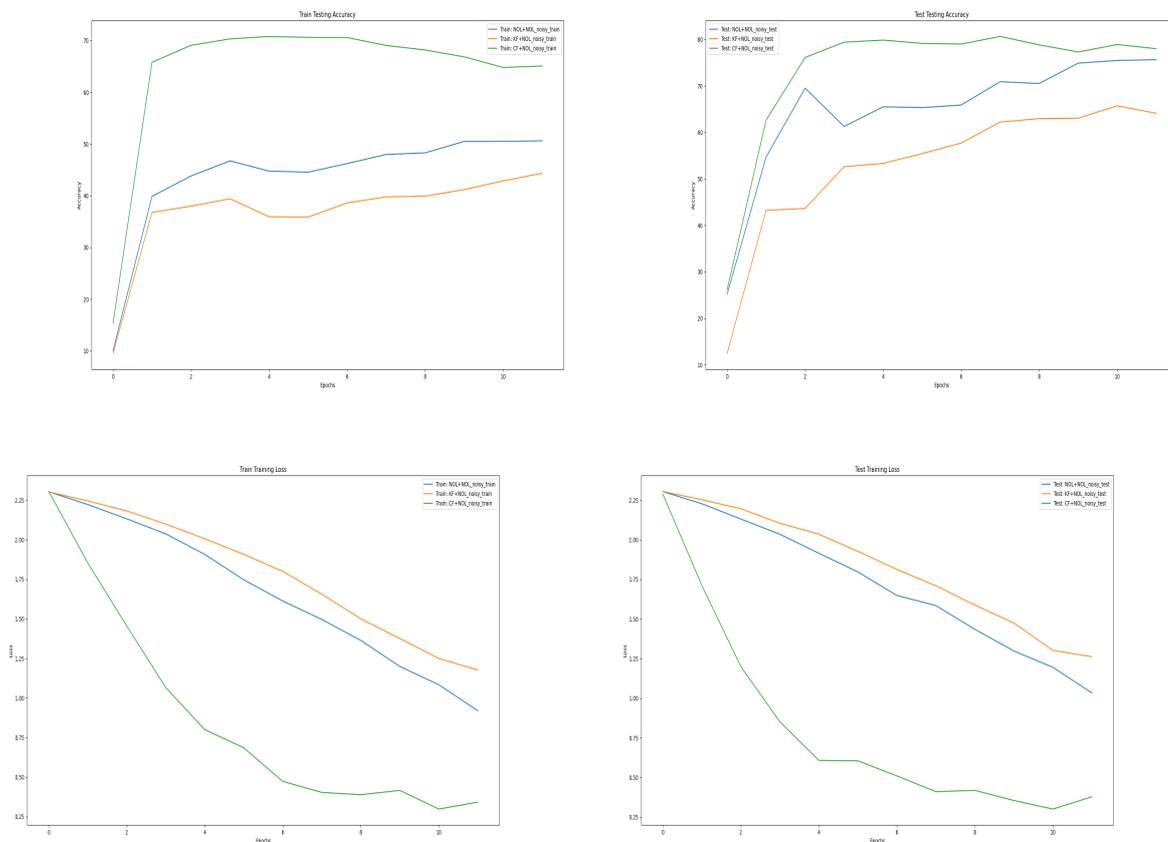


图 3

从左至右，从上至下分别是：干净数据集训练噪声数据集测试的准确率、噪声数据集训练干净数据集测试的准确率、干净数据集训练噪声数据集测试的损失、噪声数据集训练噪声数据集测试的损失。在噪声实验中，CF 层的准确率明显优于 NOL，而且 CF 层也比 NOL 先到达自身峰值，这说明 CF 可以少样本，快速地达到更好地效果。与我们之前看到的 CF 比 NOL 更为平滑的可视化结果相符合。进行了 5 次以上的实验，我们都发现在有噪声的数据集上训练，干净的数据集上测试时（右上），NOL 的准确率都会在 2, 3, 4 其中一个 epoch 出现下跌（文中只展示了一次实验结果，这次下跌出现在第 3 个 epoch），NOL 可能出现了过拟合，提取了过多干扰特征，这点在 5.3 中的泛化实验中更为明显。同时，这说明传统的架构如果只搭建两层，难以应对具有噪声的训练样本。其实同样的下落也出现在干净训练噪声测试（左上），这里的下落没有之前提到的明显。

### 5.3 泛化实验

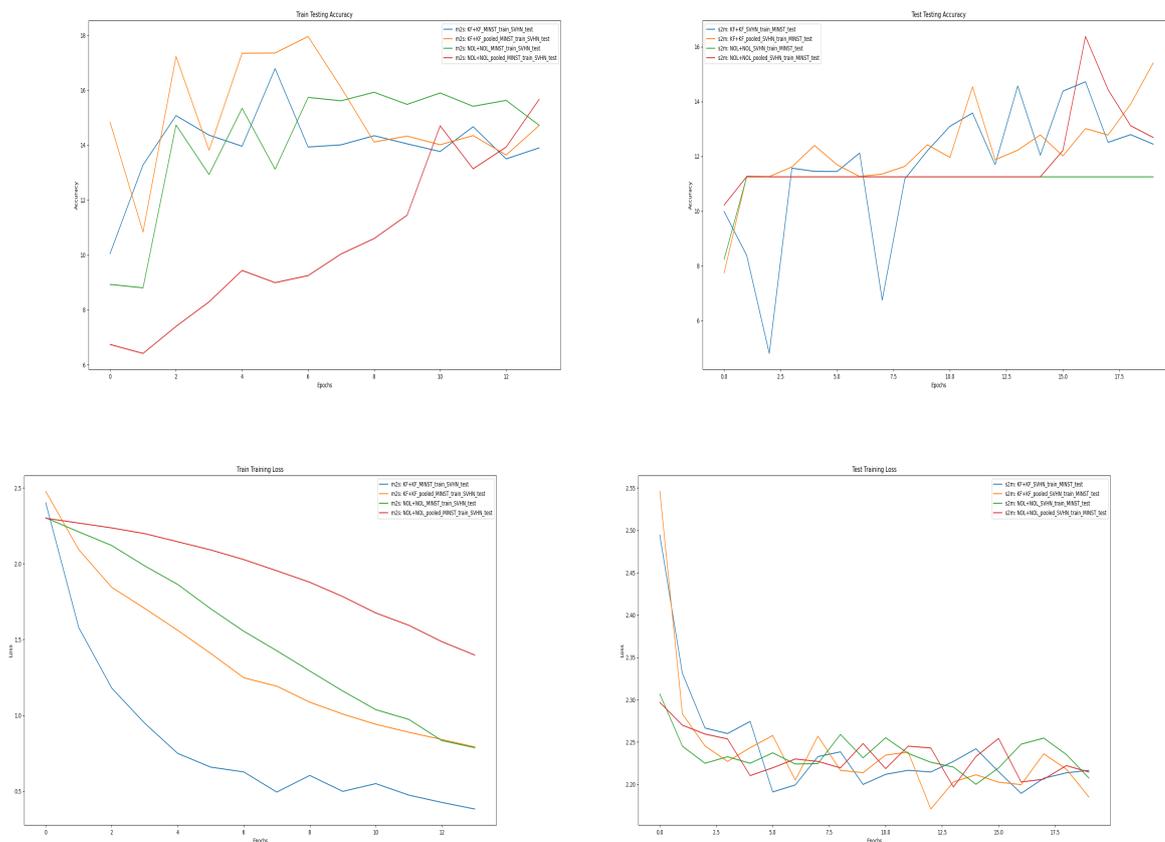


图 4

从左至右，从上至下分别是：MNIST 训练 SVHN 测试的准确率、S 训 M 测的准确率、M 训 S 测的损失、S 训 M 测的损失。

左上图最终四个网络结果相近，含 Pool 的网络效果上好于不含 Pool 的网络。但是含 KF 层均出现了先上升后下降的趋势，这是由于训练的 epoch 太多，如果只训练 5 个 epoch 会让 KF 优势看起来更明显，而 MNIST 中数字比较规范，能够提取的特征较少，KF 由能够迅速提取图像的高级特征，所以容易发生过拟合。同样的情况在右上图明显改善，这是由于 SVHN 更大，其中的数字书写更不规范（如弯钩、擦痕等），所以 KF 面对较为复杂的分类工作会有优势，充分提取数据特征。

## 6. 总结

TCNN 使用了拓扑方法定义新的卷积层，帮助快速提取高级特征，并且使得模型准确率更高，鲁棒性更强，泛化性更强，学习样本减少，学习速度增快。在三组对比实

验中，以上优势均有体现. 准确率方面没有显著提高，在鲁棒性，泛化性，学习样本数量上有显著提高. 我们把这归因于 KF 层和 CF 层能够直接展现图像的高级特征从而对其快速提取. 实验不足之处在于，没有实现原论文的 KF+KOL, CF+COL 层，这两层可能会使得自然图像本具有的拓扑特征更容易被提取，可能是我们的模型在准确率方面没有显著提高的原因. 而且在泛化测试中，模型准确率整体偏低，可能是层数太少，提取特征不够，而且层数太少容易发生过拟合的原因. 最后，原论文提及的自然图像聚在 Klein Bottle 附近，函数的构造方式，以及如何从广泛的 Gabor Filters 中找到 KF 层，我并不是很清晰，所以只能尽可能取实现原论文定义的卷积层.

## 参考文献

- [1] LOVE E R, FILIPPENKO B, MAROULAS V, et al. Topological convolutional layers for deep learning[J]. Journal of Machine Learning Research, 2023, 24(59): 1-35.
- [2] CARLSSON G, GABRIELSSON R B. Topological approaches to deep learning[C]. in: Topological Data Analysis: The Abel Symposium 2018. 2020: 119-146.
- [3] LECUN Y, BOTTOUL L, BENGIO Y, et al. Gradient-based learning applied to document recognition [J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.

## 致谢

感谢各位老师对我的指导，特别是我的学术导师朱一飞老师。感谢各位同学四年的陪伴。