

# 关于深度神经网络在交叉领域的情感分类任务中的可迁移性探究

涂曼姝<sup>1</sup> 潘接林<sup>2</sup>

1. 中国科学院声学研究所 语言声学与内容理解重点实验室 北京 100190;  
2. 中国科学院大学 北京 100049

**摘要** 领域自适应中交叉领域的情感分析研究近几年获得了广泛的关注。传统的做法是寻找源域和目标域相同的特征表示,或将源域和目标域的特征映射在高维空间中,使目标域的特征表达在高维空间中更加靠近源域以达到领域自适应的目的。然而由于这些方法都是直接将源域的全部网络权重迁移至目标域,没有考虑到神经网络中文本特征可能具有的层级性,因此我们提出一种基于超深度卷积神经网络(VDCNN)的层级迁移方法(TVDCNN)以探究如下几个问题:(1)交叉领域情感分类任务中可迁移的网络层有几层(2)哪几层获得了最好的迁移特征(3)不同的领域可迁移层数是否一致。在中英文的两个数据集的实验结果表明,文本的特征提取确实具有层级特征,前三层的迁移性最好,不同领域的可迁移层数基本一致,且在迁移之后对网络进行微调可以进一步提高正确率。

**关键词:** 领域自使用; TVDCNN; 交叉领域的情感分类

**中图分类号:** G35

开放科学(资源服务)标识码(OSID)



## How Features Transferred in Very Deep Neural Networks on Cross Domain Sentiment Classification

TU Manshu<sup>1,2</sup> PAN Jielin<sup>1,2</sup>

1. Key Laboratory of Speech Acoustics and Content Understanding, Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China;  
2. University of Chinese Academy of Sciences, Beijing 100049, China

**基金项目:** 国家自然科学基金(61650202, 11590770-4, U1536117)。

**作者简介:** 涂曼姝(1991-), 博士研究生, 研究方向: 倾向性分析, E-mail: tumanshu@hcccl.ioa.ac.cn; 潘接林(1967-), 博士, 研究员, 研究方向: 语音信号处理、语音识别、大数据分析。

**Abstract** Domain adaptation has raised a lot attention in recent years especially in the field of cross-domain sentiment classification. The traditional approach is to find the same feature representation of the source domain and the target domain, or map the source and target domain in a high dimensional space, so that the target domain features are closed to the source domain to achieve domain adaptation. However, because these methods directly transfer the network weights learned from source domain to the target domain, they do not take into account the hierarchical nature of text feature extraction in the network. Based on these reasons, we proposed a hierarchical transfer function based on the VDCNN to explore these questions: (1) how many layers can be transferred in cross domain sentiment classification. (2) which layer get best transferring features. (3) whether different domain get same transferability. The experiments on two different datasets demonstrated that documents features have hierarchical character. The first three layers have the best transferability. Different domains have same layers for transferring and the accuracy be further improved when fine-tune after the transfer

**Keywords:** Domain adaptation; TVDCNN; cross domain sentiment classification

## 引言

情感分类是自然语言处理中一个重要的任务，它可以用来识别用户倾向，应用在推荐系统或舆情分析中<sup>[1]</sup>。其任务是识别文本的倾向性（正向或负向）。随着深度学习的发展，许多基于神经网络的情感分析方法在公开数据集例如 IMDB 电影评论数据集上已经取得了很好的效果<sup>[2-4]</sup>。但有效的深度学习方法往往依赖于大量的标注数据，而标注数据的获得往往需要大量的人力和时间。且实际的数据有时候是不可得的并且数据量很小，在同样的网络结构上不能达到在公开数据集上所达到的效果。一种有效的解决思路是从已标注数据中学习有关知识迁移到相似领域的未标注数据中<sup>[5-7]</sup>。

迁移学习由其数据分布不同和标签分布不同可以分为域适配问题，多任务学习问题，自学习问题等<sup>[8]</sup>，交叉领域的情感分类任务由于标签的一致性属于领域自适应问题 (DA)。从早期的基于实例，基于模型的迁移，到现在的深度迁移，

诞生了许多经典的 DA 方法。Glorot 提出了利用栈式降噪自编码器 (SDA) 来学习所有领域中具有鲁棒性的特征表示，以实现源域到目标域的自适应<sup>[9]</sup>。在此基础上，Chen 提出了使用线性的降噪自编码器作为堆叠深度网络的基本层<sup>[10]</sup>，解决了 SDA 计算开销大所需时间多的问题。Ganin 首次提出使用对抗的方式使目标域数据在网络中的特征表示更加靠近源域数据，从而对目标域进行分类的方法<sup>[11]</sup>。Zheng 提出了使用 3 层的对抗记忆网络 (AMN) 自动提取源域和目标域的共同情感词以达到对目标域的自适应<sup>[12]</sup>。

虽然以上方法在公开的 Amazon 数据集上都取得了不错的效果，但其都使用了较浅层的神经网络并未考虑文本特征在网络中的层级性。浅层的前馈网络虽然能够表示任何函数，但网络容易出现过拟合<sup>[13]</sup>，尤其在迁移中当目标域的带标签数据不参与训练时，浅层网络会过分拟合源域数据，迁移的效果会受到影响。而深层网络在提取特征时具有层级性<sup>[20]</sup>，迁移层数更加随机。为了探究交叉领域的情感分类在自适

应时，每一层的可迁移能力，我们提出了一种层级迁移的网络的结构。该网络可以将源域训练好的权重按层级取出迁移至目标域网络中。

## 1 相关工作

自 AlexNet<sup>[14]</sup> 以来，最先进的 CNN 架构的层数已经越来越深。AlexNet 只有 5 个卷积层，而之后的 VGG 网络<sup>[15]</sup> 和 GoogleNet<sup>[16]</sup> 分别有 19 层和 22 层，CNN 中的层数深度的增加也使得图像识别率不断提高。Shin 在医学图像上使用 GoogLeNet 等深度 CNN 网络进行迁移学习取得了显著的效果<sup>[17]</sup>。随着超深度神经网络在图像分类上的成功，研究者们也尝试通过增加 CNN 的网络层数来解决 NLP 中的分类问题，Zhang<sup>[18]</sup> 首次提出了使用字符级特征的 6 层 CNN 进行文本分类，Poria 提出使用多层卷积与多卷积核共同作用的深度卷积神经网络<sup>[19]</sup>，在此基础上 Conneau 提出的超深度卷积神经网络（VDCNN）也在情感分类公开数据集上取得了良好的效果<sup>[20]</sup>，本文提出的层级迁移实验是在 Conneau 的 VDCNN 的基础上进行的，同时根据 Yosinski<sup>[21]</sup> 在图像数据集上进行的可迁移性探究，对摘要中提出的几个问题进行探讨。

## 2 TVDCNN

本章我们对交叉领域的情感分类问题进行描述，并详细描述我们提出的 TVDCNN 模型。

### 2.1 问题定义

交叉领域的情感分类是指：将源域数据

$X_s^l = \{x_s^i, y_s^i\}_{i=1}^{N_s^l}$ （其中  $N_s^l$  表示源域数据有标签数据的数量）通过训练所学的知识应用在目标域数据  $X_t^u = \{x_t^i\}_{i=1}^{N_t^u}$  上，（其中  $N_t^u$  表示目标域数据量），从而对目标域数据进行分类。源域和目标域的数据分布不同，但分类目标一致，在本文的所有实验中都是情感极性二分类，标签为正或者负向。

### 2.2 模型说明

我们提出的 TVDCNN 模型是在 VDCNN 模型基础上进行的迁移模型，之所以选择 VDCNN，是因为它实现了深度的层级结构，适合层级迁移并且其在多个公开的分类数据集上都有很好的表现。根据我们的实验结果，当迁移层数超过 9 层时，迁移效果会下降，因此我们选择在 9 层的网络上进行迁移，模型结构如图所示，A 和 B 表示两个不同的数据集，A 表示源域，B 表示目标域。A 和 B 的标签分别表示源域和目标域的分类结果。BaseA 和 BaseB 模型是在 VDCNN 网络上通过标准的后向传播方式训练得到的结果。

如图 1 所示，图中的前两行分别表示的是 BaseA 和 BaseB 网络，是在 A、B 数据集上使用标准的后向传播算法进行训练的。网络的输入是字或字母为单位的矢量矩阵，维度记为  $emb$ ，记句子长度为  $s$ ，网络中滤波器的宽度为  $d$ ，输入矩阵为  $s * emb$ 。图中横条纹矩形（Convolutional）表示的是一个基本的卷积层<sup>[22]</sup>，用于将输入编码成特征，第一层的滤波器数为 16。图中无条纹的矩形表示的是卷积块（Convolutional Block），它由两个卷积层连接而成，对于一个卷积块中的两个卷积层来说，它们的

输出通道数  $c$  相同, 网络的 2-9 层通道数分别是 64,128,256 和 512, 即分别会产生 64,128,256 和 512 个特征矩阵。卷积不改变每一个特征矩阵的维度, 图中的下采样 (pooling) 使每一个特征矩阵维度减半。最后一层的下采样是  $k\_max\_pooling$ , 即采样每个特征矩阵中最大的  $k$  个特征值, 因此最终送入全连接层的特征维度是  $512*k*l$ 。图中竖条纹矩形表示全连接层, 大小分别为  $(k*512)*1024$ ,  $1024*2048$ ,

$2048*num\_class$ ,  $num\_class$  是最终分类的类别数, 在该实验中  $num\_class=2$ 。在矩形上面的标注 ( $W_{A2}&W_{A3}$ ) 表示的是这一层学习到的权重信息, 对于每一个 Convolutional Block 来说就有 2 个权重矩阵, 矩形中的颜色表示其权重的来源, 例如蓝色表示其权重是从 BaseA 中学习得到的, 绿色表示随机初始化。矩形框中的锁图形式表示该层是否参与训练, 锁关闭表示该层不参与训练, 打开表示参与训练。

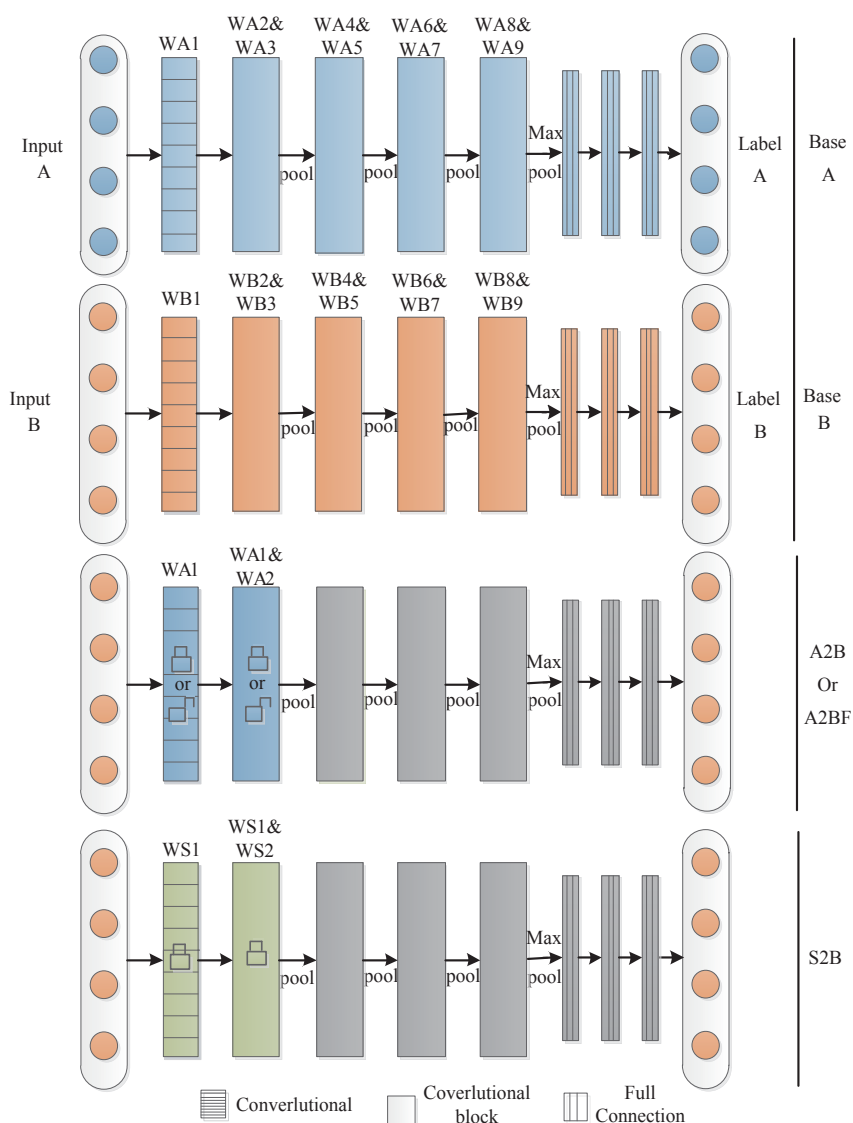


图 1 TVDCNN 结构

● 迁移网络 AnB: 前  $n$  层从 BaseA 迁移过来并保持不变, 剩下的  $9-n$  层随机初始化后在 B 数据集上进行训练。即我们使用 A 数据集训练得到网络的前  $n$  层, 然后在 B 数据集上学习剩下的  $9-n$  层从而对 B 数据集进行分类。若 AnB 在 B 数据集上的表现和 BaseB 一致的话, 有理由说在基于深度卷积网络的情感分类中, 网络的第  $n$  层学到的是比较泛化的特征。反之, 如果效果下降, 则网络的第  $n$  层学到的就是与 A 数据有关的私有的特征。如图 1 第 3 行所示。

● 对照网络 SnB: 网络的前  $n$  层随机初始化并保持不变, 剩下的  $9-n$  层也随机初始化后在 B 数据集上进行训练。若 SnB 的表现和 BaseB 的表现一样, 说明网络的第  $n$  层对分类没有贡献。反之, 则说明前  $n$  层对分类都有贡献。如图 1 第 4 行所示。为了使对照试验有意义, 随机初始化的矩阵将被保存并适用于所有实验。

● 对网络所有的层  $n \in [1, 9]$  都进行同样的迁移实验。上面我们提出的 AnB 和 SnB 中的  $n$  是只迁移但不参与训练的层数。我们也在 AnB 的基础上提出了对前  $n$  层进行微调的网络结构。

迁移网络 AnBF: 与 AnB 一样, 但是所有的层都参与训练。如图 1 第 3 行所示。

### 3 试验和分析

该章节会详细介绍数据, 实验设置参数和对实验结果的详细分析。

#### 3.1 数据说明

实验使用了英文和中文两种数据集。训练集和测试集的比例分配是 9:1, 正向和负向的

比例为 1:1。英文数据集选取的是在交叉领域的情感分类中被广泛使用的亚马逊评论数据集<sup>[23]</sup>, 以下简称 AMZ。中文数据集是中国移动的客服和客户之间的对话语料以下简称 CMCC, 该数据由两部分组成, 一部分是河南省的客服语料, 以下简称 CMCCCH, 另一部分是江苏省的客服语料, 以下简称 CMCCJ。表 1 是详细的数据分布。

表 1 数据分布

	AMZ	CMCCCH	CMCCJ
正向	15451	5742	2501
负向	14242	5742	2501

AMZ: 包括 16 种商品的评论语料, 例如电子产品、婴儿用品等。首先随机将这 16 种语料随机分成两部分, 一部分作为源域数据 A, 另一部分作为目标域数据 B。然后将上述分数据的方式重复 10 次, 并保证每次的 A 和 B 数据集都与之前的每次不一致, 由此得到 10 个数据集对。

CMCC: 将 CMCCCH 作为源域数据, 数据的原始标签是客户的需求分类, 例如查询套餐、变更套餐、办理套餐、投诉等十余个类别。我们仅关注客户的需求是否是投诉, 并将投诉的文本视为负向, 其他类别视为正向。将 CMCCJ 作为目标域数据, 其原始标签是客户对客服的服务态度的评分, 评分分为四个等级(非常满意、满意、一般、不满意), 我们将非常满意视为正向, 其它视为负向。

#### 3.2 模型参数说明

下面的设置在我们所有的实验中都通用且实验结果表明这些参数的效果最好。网络的输入以字符为单位, 是语义理解的最小单位, 类似于图像中的像素<sup>[3]</sup>。AMZ 的词典由 Conneau 提出的英文词典<sup>[19]</sup>加上中文的常用标点“., ! ? “ ””总共 75 个字符组成。CMCC



的字典由 3500 个常用中国汉字<sup>①</sup>，再加上 AMZ 的词典一共 3575 个字符组成。AMZ 的句子长度为 1014，字向量维度为 16，即输入矩阵为  $1014 \times 16$ ，所有的滤波器宽度为 3，下采样的宽度也都为 3。CMCC 的句子长度为 2500，字向量维度为 50，即输入矩阵为  $2500 \times 50$ ，CMCC 中第一层 CNN 的滤波器宽度为 10，剩下的滤波器宽为 3，下采样的步长都为 10。句子长度超过设置的长度会被截断，不足的补零。CMCC 的原始数据是对话形式，我们使用空格将其连接成连续的一个段落。最后一层的最大下采样中  $k=8$ 。所有字向量在网络中随机初始化不经过预训练，但 SnB、AnBF、AnB 与 BaseB 的字向量保持一致。所有实验均采用 5 折交叉验证。

实验中采用 ADAM<sup>[24]</sup> 的优化方式进行优化，学习率设置为 0.001。

### 3.3 实验结果探究

1. 首先纵向对比两个数据集的实验表现，AMZ 的正确率在每一个模型的每一层都比 CMCC 的正确率高。例如在 BaseB 上 AMZ 的正确率是 77.51%，在 CMCC 是 67.35%，AMZ 比 CMCC 的基线正确率要高 10.16%。其原因一方面是 AMZ 数据集是一个对商品的评论数据集，其本身比由对话组成的 CMCC 数据要更清楚的表达正向或者负向情绪，另一方面，CMCC 的数据是由语音转写成文本的，由于方言等一些其它原因识别效果并不是特别理想也没有恰当的断句，例如句子“昨天晚上才充好，今天早上怎么就全都没有了。”识别结果是“昨天晚上才充好今天早上怎么么么么就全都没有了”，这也对文本的分类造成了一定的干扰。

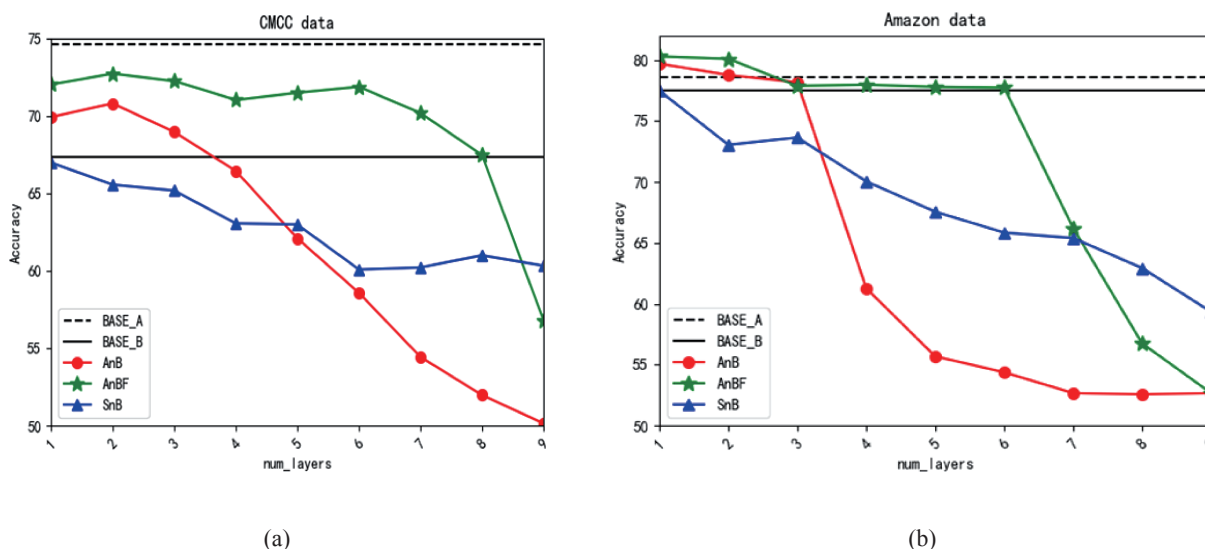


图2 BaseA, BaseB, AnB, AnBF 和 SnB 五种实验的准确率比较。横坐标表示迁移或随机初始化的层数，纵坐标是数据集在多次重复试验下的平均分类正确率。红色有圆点的线表示 AnB 的结果，绿色有五角星的线表示 AnBF 的结果，蓝色有三角的线表示 SnB 的结果。黑色实线是 BaseB 的结果，黑色虚线是 BaseA 的结果，作为基线与各模型在各层做对比。图 (a) 是 AMZ 的实验结果，图 (b) 是 CMCC 的实验结果。

① <https://wenku.baidu.com/view/930365cb650e52ea5418981c.htm>

2.SnB 网络用来验证从 BaseA 网络中得到的权重矩阵迁移到 B 数据上是否真的对分类有效。如图所示,其在 AMZ 的前 2 层,CMCC 的前 4 层都低于 AnB 的正确率,说明通过这几层迁移的权重矩阵得到的特征是对分类有效的。SnB 在两个数据集上每一层的正确率都要低于 BaseB 的正确率,这意味着在训练 VDCNN 的时候,训练得到的特征是对最终分类有用的特征。SnB 在第一层的正确率在两个实验中都与 BaseB 的正确率相差很少,说明通过第一层权重矩阵得到的特征对于分类来说不太重要,这一层提取的特征是泛化特征,类似于图像中的像素块或者由像素块组成的线条,该种特征无法直接用于分类。但这也说明深度神经网络理解语言的过程像理解图像从像素到线条到色块的过程一样有类似的层级性。

3.AnB 模型展示的是每一层的权重矩阵从源域到目标域的迁移能力。在 AMZ 的实验结果中,AnB 前两层的正确率高于 BaseB 2%-3%,CMCC 前三层也有这样的表现。这意味着,对于这两个数据集,深度神经网络不仅仅是在第一层提取了泛化的特征,第二层的特征也是泛化的,更加令人惊喜的是,在 CMCC 的第三层上,特征依然是泛化的。数据集经过上述这几层迁移后的权重所得到的特征在不同程度上提高了最终的正确率,这种提高在 CMCC 上更加明显,一方面是因为 CMCC 的数据量是 CMCCJ 的两倍,另一方面 CMCC 的语料中有相当一部分使用了北方口音,相比起 CMCCJ 中使用的南方口音前方系统识别率相对较好,从而使得迁移促进了 CMCCJ 的分类,当然这也是对 CMCC 本身分类效果(BaseA)比

CMCCJ(BaseB) 要好的原因。在 CMCC 和 AMZ 中,迁移大于四层时,正确率就会大幅度下降,这意味着数据通过这些层的权重矩阵得到的特征是目标域数据的私有特征。CMCC 迁移前四层时正确率要低于 BaseB 但是比 SnB 的正确率高且与迁移前三层正确率相比只是有一个轻微的下降,表明第四层学习得到的特征并不是泛化特征但是也不是私有特征,它学到的特征是一种过渡特征,也说明在语言理解上,神经网络并没有像图像一样有特定的提取线条或颜色块特征的权重层,它是一个模糊的过程,但是这种基础特征例如字或词的提取发生在了神经网络的前 4 层。

4.AnBF 模型表现显著:当迁移后,再对网络进行微调,分类的效果在 AMZ 迁移  $n$  层( $n < 7$ ),CMCC 迁移  $n$  层( $n < 9$ )时都比 BaseA 的结果好,AMZ 的前两层的分类效果甚至超过了 BaseA。在 CMCC 中, $n=2$  时正确率达到峰值 72.71% 比 BaseB 高 5.36%。在 AMZ 中,当  $n=1$  时,正确率达到峰值 80.28% 比 BaseB 高 2.77%。当迁移所有 9 层时分类效果不好的原因是很好解释的,此时,网络不能学习目标域数据的任何私有特征。对于 AMZ,从第七层开始正确率就有了显著的下降,说明第七层第八层的特征是对源域分类私有的特征,迁移后即使网络在学习过程中进行微调也不能使分类的效果变好。为了验证我们的想法,我们在将 AMZ 数据分成源域与目标域的时候人为的将 16 种商品分为接近的两部分,显而易见的是当商品的种类相近时,人们对其的评价用词会比较接近,如表 2 所示,在源域中的 magazines 和目标域中的 books 的数据较为相似,类似的还有 dvd 和

video 等, 表中的第四行为迁移对应层数所得的

正确率, 此时其表现与 CMCC 就比较一致。

表 2 A 是源域, B 是目标域, 第二行是源域和目标域所包括的商品评论的类别, 第三行表示的是迁移的层数, 第四行是其相对应的正确率

	A				B				
	Sports&outdoors,camera&photo,dvd,software,health&personal_care,magazines,music_electronics				kitchen&housewares,books,video,baby,apparel, office_products,toys&games_computer&video_games				
Base B	1	2	3	4	5	6	7	8	9
77.29	80.91	81.42	77.87	78.88	78.23	77.14	77.36	77.80	53.41

5. 同时我们也用 Proxy A-Distance(PAD)<sup>[25]</sup> 计算源域和目标域数据的相似程度,  $e$  为分类错误率

$$PAD=2(1-2e)$$

我们使用下面描述的过程获取 PAD 值:

- 混合源域和目标域的数据, 或者混合从源域数据经过 BaseA 和目标域数据经过 AnB 或 AnBF 后在同一层取出的特征, 源域数据标签设为 1, 目标域数据标签设为 0. 新的数据集记为 U, 将 U 分为训练数据集和测试数据集, 比例为 9:1

- 使用训练数据集训练 VDCNN 分类器, 在测试集上得到分类错误率  $e$

PAD 越小, 表示两个数据的相似度越大。如图 3 所示, 我们选取了两个迁移网络的其中几层进行实验。AB 的 PAD 值为 95.57%, 在一定程度上表示 CMCC 和 CMCCJ 原始数据的距离比较远, 分类器比较好分离这两种数据, 实际由我们观察, 这两个数据在内容表达上是相近的, 都是客服和客户之间的对话, 但是由于转写工具的不同, 以及语言一定的方言特色使得分类器较好的分离了这两个数据。A5BF、A7BF 和 A2B 的 PAD 值都比 AB 的要小, 说明经过 AnBF 或 AnB, CMCCJ 的数据更加靠近 CMCC, 在特征上更不容易将它们分开, 即这

两种网络学习了某些 CMCC 的特征, 并在迁移后使得 CMCCJ 的特征更加靠近 CMCC, CMCC 的分类正确率比 CMCCJ 的正确率高从而提高了 CMCCJ 的分类准确率。而 A5BF 的 PAD 值比 A6BF 高, 说明迁移 6 层并微调后得到的特征比迁移前 5 层微调网络 CMCCJ 得到的特征进一步靠近了 CMCC 的特征, 在正确率上的表现迁移前 6 层也确实比迁移前 5 层高 0.5%, 不过这种靠近不一定对分类是完全有正向作用的, 因为它会舍弃一些 CMCCJ 的私有特征从而导致分类错误, 如何设计网络, 使网络只学到对目标域分类有效的相似特征也是我们之后要做的工作。

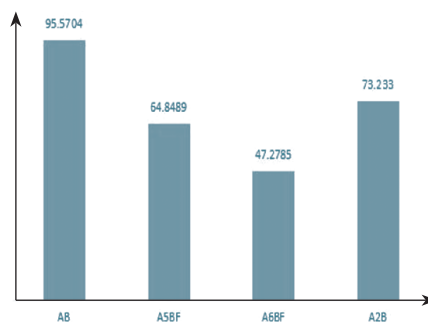


图 3 PAD 的值, 横坐标表示的是两个数据集的来源, AB: 混合 CMCC 和 CMCCJ 的原始数据。A5BF: 混合 CMCC 经过 BaseA 所得到的特征和 CMCCJ 数据经过 A5BF 模型所得到的特征。A7BF: 与 A5BF 类似, CMCCJ 数据是经过 A7BF 得到的。A2B: CMCC 的来源与 A5BF 一样, CMCCJ 数据是经过 A2B 所得到的特征



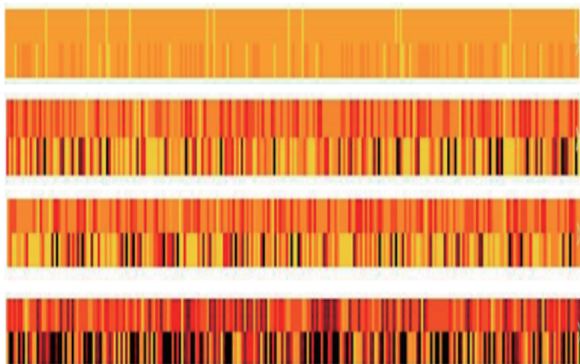


图4 可视化结果，总共四行，从上到下依次表示的是四次采样分别对应到原始文本的结果，每一个采样结果都由上层的 BaseB 的可视化结果与下层表示的 A6BF 的可视化结果做对比。颜色的深浅表示的是原文中该位置的字被采样的次数，采样的次数越多颜色越深，反之越浅

6. 为了探究迁移后与迁移前网络学到的特征有什么不同，我们选取了一个在 BaseB 上分类错误，在 A6BF 上分类正确的例子，具体见附录中给出的实例 A。实例的正确分类应该为负向，在判断极性时，句子“好是确实是你们这里有问题了”，是判断用户的情绪倾向的关

键句子，在采样时，希望这个句子中的字被采样的次数更多，尤其是“有问题”这三个字。

可视化方法如算法 1 所示，结果如图 4 所示，由图 4 中第一行的颜色分布的上一层可以看到，对于 BaseB 第一次采样的时候对话中的每一个词的被采样次数都比较平均，但是 A6BF 已经对一些词和句子有所侧重，这种侧重在第二三四次采样后变得越来越明显。而 BaseB 在二三四次采样中虽然没有平均采样，但是侧重的并不明显，每个字的被采样次数基本没有差别。A6BF 最后一次下采样的结果，可视化为黑色的字有 {3, g, 信, 号, 耐, 心, 等, 待, 问, 题……}，虽然被采样多次可视化为黑色部分的字比较多，但是我们关注的词“问题”这两个字是在这个列表里的，而“问题”两个字并不在 BaseB 的最后一次采样的可视化为黑色的部分中。这也直观的解释了该例子在 BaseB 中分类错误而在 A6BF 中分类正确的原因。

#### Algorithm 1 可视化算法

定义网络的每一层： $P_0, C_1, C_2, C_3, P_4, C_5, C_6, P_7, C_8, C_9, P_{10}, C_{11}, C_{12}, P_{13}$ ， $C$  是卷积层， $P$  是下采样层， $P_0$  是输入的内容，为方便记录，记为第一个下采样层。

得到下采样序列  $P_n$  所对应的卷积层  $C_{n-1}$  的位置  $[p_1, p_2, p_3] \dots$

Repeat

根据该下采样层到前一个下采样层所含的卷积层的滤波器的宽度列表  $[f_1, f_2]$  计算  $[p_1, p_2, p_3] \dots$  所对应的前一个下采样层的位置。

$$p_{n\_begin} = p_{n\_end} = p_n$$

For  $i$  in  $[f_1, f_2]$  do:

$$p_{n\_begin} : p_{n\_end} - \text{int}(i/2)$$

$$p_{n\_end} : p_{n\_end} + \text{int}(i/2)$$

End for

得到一个字典  $p_n : [p_{n\_begin}, p_{n\_end}]$ ，即  $p_n$  所对应到的前一个下采样层的位置。

until 得到对应到第一个下采样层  $p_0$  的字典  $p_0 : [p_{0\_begin}, p_{0\_end}]$

最终得到的字典表明从  $p_{0\_begin}$  到  $p_{0\_end}$  的所有字被采样了一次，将字典中的所有对应的次数相加，得到原文中每一个字被采样的次数。将字对应的采样次数拆成 RGB 值，得到对应的颜色。

## 4 结论

我们使用 TVDCNN 对中英文两个数据集进行了迁移实验, 实验结果在两个数据集上有轻微的不同, 大体上表现一致, 都表明了网络前几层的权重矩阵具有不错的迁移能力, 经过前几层的权值矩阵得到的特征是泛化特征, 经过中间和后面几层权值矩阵的特征是训练数据的私有特征, 不具有泛化能力。值得注意的是, 不论迁移前面的几层还是中间或后面的几层, 网络经过微调, 正确率基本上都能提升 2-6 个百分点。这个发现对于使用迁移在深度神经网络中进一步提高分类正确率很有帮助。当然, 通过与 SnB 的对照试验我们也发现 VDCNN 在分类的过程中学到了对分类有用的特征。同时我们也提出了一种可视化的算法, 该算法对采样后的特征矩阵进行反向推导, 用颜色的深浅表示原始数据中每个字被采样的次数, 颜色越深采样次数越多, 这种可视化方式在一定程度上解释了神经网络在解决 NLP 的问题上是如何工作的。在接下来的工作中, 根据我们探究得到的情感分类在深度卷积网络中迁移的这些性质, 进一步改进网络使得网络更加有效和有目标的学习源域和目标域之间相似的特征, 从而提高目标域的分类正确率。

### 参考文献

- [1] Liu B, Zhang L. A Survey of Opinion Mining and Sentiment Analysis[M]. Berlin: Springer, 2012.
- [2] Yang Z, Yang D, Dyer C, et al. Hierarchical Attention Networks for Document Classification[C]. Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2017:1480-1489.
- [3] Zhang X, Zhao J, Lecun Y. Character-level convolutional networks for text classification[C]. 28<sup>th</sup> Advances in Neural Information Processing Systems. 2015:649-657.
- [4] Catal C, Nangir M. A sentiment classification model based on multiple classifiers[J]. Applied Soft Computing, 2017(50): 135-141.
- [5] Li Z, Wei Y, Zhang Y, et al. Hierarchical Attention Transfer Network for Cross-domain Sentiment Classification. In Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, AAAI 2018, New Orleans, Louisiana, USA, February 2-7, 2018.
- [6] Long M, Cao Y, Wang J, et al. Learning Transferable Features with Deep Adaptation Networks[J]. arXiv preprint arXiv:1502.02791, 2015.
- [7] Melville P, Gryc W, Lawrence R D. Sentiment analysis of blogs by combining lexical knowledge with text classification[C]. In Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2009: 1275-1284.
- [8] Pan S J, Yang Q. A Survey on Transfer Learning[J]. IEEE Transactions on Knowledge & Data Engineering, 2010, 22(10): 1345-1359.
- [9] Glorot X, Bordes A, Bengio Y. Domain adaptation for large-scale sentiment classification: a deep learning approach[C]. International Conference on International Conference on Machine Learning. Omnipress, 2011:513-520.
- [10] Chen M, Xu Z, Weinberger K, et al. Marginalized denoising autoencoders for domain adaptation[J]. arXiv preprint arXiv:1206.4683, 2012.
- [11] Ganin Y, Ustinova E, Ajakan H, et al. Domain-

- adversarial training of neural networks[J]. Journal of Machine Learning Research, 2016, 17(1): 2096-2030.
- [12] Li Z, Zhang Y, Wei Y, et al. End-to-End Adversarial Memory Network for Cross-domain Sentiment Classification[C]. Twenty-Sixth International Joint Conference on Artificial Intelligence. 2017: 2237-2243.
- [13] 李明国, 郁文贤. 神经网络的函数逼近理论[J]. 国防科技大学学报, 1998(4): 70-76.
- [14] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C]. International Conference on Neural Information Processing Systems. Curran Associates Inc. 2012:1097-1105.
- [15] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv: 1409.1556, 2014.
- [16] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[C]. IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2015:1-9.
- [17] Shin H C, Roth H R, Gao M, et al. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning[J]. IEEE Transactions on Medical Imaging, 2016, 35(5): 1285.
- [18] Zhang X, Zhao J, Lecun Y. Character-level convolutional networks for text classification[C]. International Conference on Neural Information Processing Systems. MIT Press, 2015.
- [19] Poria S, Cambria E, Gelbukh A. Deep convolutional neural network textual features and multiple kernel learning for utterance-level multimodal sentiment analysis[C]. Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. 2015: 2539-2544.
- [20] Conneau A, Schwenk H, Barrault L, et al. Very Deep Convolutional Networks for Text Classification[J]. 2017, arXiv:1606.01781.
- [21] Yosinski J, Clune J, Bengio Y, et al. How transferable are features in deep neural networks?[J]. 2014, arXiv:3320-3328.
- [22] Kim Y. Convolutional neural networks for sentence classification[J]. 2014, arXiv: 1408.5882.
- [23] Blitzer J, Dredze M, Pereira F. Biographies, Bollywood, Boom-boxes and Blenders: Domain Adaptation for Sentiment Classification[C]. ACL 2007, Proceedings of the Meeting of the Association for Computational Linguistics, June 23-30, 2007, Prague, Czech Republic. DBLP, 2007:187-205.
- [24] Kingma D P, Ba J. Adam: A method for stochastic optimization[J]. 2014, arXiv:1412.6980.
- [25] Ganin Y, Ustinova E, Ajakan H, et al. Domain-adversarial training of neural networks[J]. Journal of Machine Learning Research, 2017, 17(1): 2096-2030.

## 附录

## 实例A

A: 您好,我是正在断很高兴为您服务。

B: 是这样的,现在我就给我们现在有那个 杨高北路中。

A: 泉路与蒙。

B: 城关镇这里然后都没打几个信号好,那我。

A: 知道了。对。杨高路是吧。

B: 是这个么二是黄冈系统中没有记录。

A: 杨康路靠近 40 么罗先生。

B: 阳到北京的 885 洞。

A: 阳江北路 350。

B: 六。对。

A: 您请稍等一下,我帮您查讯一下。

B: 好的。

A: 好先生是这样子非常感谢您对移动公司 关注您本人高北路 38856 地区信号不好,问题已经去美国分钟数我们已经有一家建议您就是就是发有已经有加镇规划来改三 G 这边信号 覆盖由

于工程的话是个数字然后是这工作请您 耐心等待我们会支持第一时间改善当地的信号 问题。

B: 是这样的吧。

A: 唉,对的对的。

B: 好是确实是你们这里有问题了。

A: 最后,我们这边已经有加急建设计划后 你们现在的话,就是说暂时建议您耐心等待一下。

B: 噢,大概是多久才会把型号报给我好吧,嗯目前我们。

A: 已经领取之后呢,直接发目前已经是进展到那个购机制单一 G 的话是在 2014 年 12 月 31 号可以完工请您耐心等待我们会支持感谢您的配合问题。不好意思啊。非常抱歉这边 听不变的。

B: 是这样对吧。

A: 对对。

B: 好的好的,谢谢啊。

A: 嗯,好的,那先生如果你对我服务满意 请点十分评价祝您生活愉快,再见。