

# 基于梯形网络和改进三训练法的半监督分类

莫建文<sup>1</sup> 贾鹏<sup>1</sup>

**摘要** 为了提高半监督深层生成模型的分类性能, 提出一种基于梯形网络和改进三训练法的半监督分类模型. 该模型在梯形网络框架有噪编码器的最高层添加 3 个分类器, 结合改进的三训练法提高图像分类性能. 首先, 用基于类别抽样的方法将有标记数据分为 3 份, 模型以有标记数据的标签误差和未标记数据的重构误差相结合的方式调整参数, 训练得到 3 个 Large-margin Softmax 分类器; 接着, 用改进的三训练法对未标记数据添加伪标签, 并对新的标记数据分配不同权重, 扩充训练集; 最后, 利用扩充的训练集更新模型. 训练完成后, 对分类器进行加权投票, 得到分类结果. 模型得到的梯形网络的特征有更好的低维流形表示, 可以有效地避免因为样本数据分布不均而导致的分类误差, 增强泛化能力. 模型分别在 MNIST 数据库, SVHN 数据库和 CIFAR10 数据库上进行实验, 并且与其他半监督深层生成模型进行了比较, 结果表明本文所提出的模型得到了更高的分类精度.

**关键词** 梯形网络, 改进的三训练法, 半监督学习, Large-margin Softmax 分类器

**引用格式** 莫建文, 贾鹏. 基于梯形网络和改进三训练法的半监督分类. 自动化学报, 2022, 48(8): 2088–2096

**DOI** 10.16383/j.aas.c190869

## Semi-supervised Classification Model Based on Ladder Network and Improved Tri-training

MO Jian-Wen<sup>1</sup> JIA Peng<sup>1</sup>

**Abstract** In order to improve the classification performance of semi-supervised deep generation models, a new model based on ladder network and improved tri-training algorithm is proposed. This model adds three classifiers to the highest layer of the noisy coding layer of the ladder network, and improves the image classification performance by combining the improved tri-training. First, the labeled data is divided into three parts by using a class-based sampling method. The model adjusts the parameters by combining the labeled error of labeled data and the reconstruction error of unlabeled data, and is trained to get three Large-margin Softmax classifiers. Next, the improved tri-training algorithm is developed to add pseudo-labels to the unlabeled data, and different weights are assigned to the new labeled data to expand the training set. Finally, the expanded training set is applied to update the model. After the above practice, a weighted voting is performed on the classifier to obtain the classification result. The characteristics of the ladder network obtained by this model have better low-dimensional manifold representations, which can effectively avoid classification error caused by uneven distribution of sample data and enhance generalization ability. The model is tested on the MNIST, SVHN and CIFAR10 respectively. In comparison with other semi-supervised deep generation models, test results show that the model proposed in this paper has obtained state-of-the-art classification accuracy over existing semi-supervised learning methods.

**Key words** Ladder network, improved tri-training, semi-supervised learning, Large-margin Softmax classifier

**Citation** Mo Jian-Wen, Jia Peng. Semi-supervised classification model based on ladder network and improved tri-training. *Acta Automatica Sinica*, 2022, 48(8): 2088–2096

在实际工作中, 获取标记数据需要花费大量的

财力物力, 而未标记数据的获取往往更容易, 因此半监督学习应运而生. 利用少量有标记数据和大量无标记数据一同指导训练过程, 提高分类器的性能. 对于传统的半监督分类学习算法如自训练 (Self-training)、TSVM (Transduced support vector machines)、EM (Expectation maximization) 算法等, 由于其简单易实现使得半监督分类学习得到推广, 同时也由于其误差较大, 使得研究学者在提高分类性能的道路上不断探索新的半监督分类学习方法<sup>[1]</sup>. 例如 Chau 等<sup>[2]</sup> 结合自训练提出 Tri-Forest 算法, 解决了课堂级学生分类的问题, Lu 等<sup>[3]</sup> 结合随机森林提出了 Rotation Forest, 提高特征的多样性, 以

收稿日期 2019-12-20 录用日期 2020-07-21

Manuscript received December 20, 2019; accepted July 21, 2020

国家自然科学基金 (61661017, 61967005, U1501252), 广西自然科学基金 (2017GXNSFB198212), 桂林电子科技大学研究生创新基金项目基金 (2019YCX020) 资助

Supported by National Natural Science Foundation of China (61661017, 61967005, U1501252), Natural Science Foundation of Guangxi (2017GXNSFB198212), and Postgraduate Innovation Project of Guilin University of Electronic Science and Technology (2019YCX020)

本文责任编辑 潘泉

Recommended by Associate Editor PAN Quan

1. 桂林电子科技大学信息与通信学院 桂林 541000

1. School of Information and Communications, Guilin University of Electronic Science and Technology, Guilin 541000

及 Cheung 等<sup>[4]</sup> 结合正则化参数的自训练方法提出 STAR-SVM (Self-training with adaptive regularization for SVM), 解决了某些数据集预测性较差的问题.

随着深度学习的发展, 深层生成模型以其强大的拟合能力、表征能力及灵活性等优点, 在半监督分类学习中得到广泛应用. 主流的生成模型有自回归模型<sup>[5]</sup>; 深层卷积神经网络<sup>[6]</sup>; 生成对抗网络 (Generative adversarial network, GAN)<sup>[7]</sup> 等, 对深层生成模型的研究可以提高网络对更深层次图像特征的学习能力, 避免模型陷入局部最优. 近年来, 许多研究学者通过优化网络结构来构建的半监督分类学习框架都能取得不错的分类效果. 如 Springenberg 等<sup>[8]</sup> 从损失函数的角度对网络优化, 提出 CAT-GAN (Categorical GAN), 它改变了网络的训练误差, 通过正则化信息最大化的框架提高了模型的鲁棒性. Salimans 等<sup>[9]</sup> 从判别器角度对网络优化, 提出 Improved-GAN, 它在判别器后接一个分类器, 训练网络后得到的分类判别器可以直接将数据分为原始图像类别和生成图像类别, 提高了生成样本质量和模型的稳定性. Fu 等<sup>[10]</sup> 从生成器的角度对网络优化, 提出 SSE-GAN (Semi-supervised encoder GAN), 它在生成器对应位置添加一个编码器, 直接获得原始数据的特征变量用于分类, 进一步提高模型分类精度. 有研究学者通过将优化的分类算法与深层生成模型相结合, 也能得到很好的分类效果. 如 Pezeshki 等<sup>[11]</sup> 提出的含有两个编码器和一个解码器的梯形网络. 它能够将与有监督分类有关的信息和与无监督重构有关的信息分开, 但是同时又能利用到这些信息. 之后, Rasmus 等<sup>[12]</sup> 在此工作的基础上提出了梯形网络与半监督学习的算法, 在无监督梯形网络编码的最高层添加分类器, 这样将无监督网络和有监督网络结合起来, 通过对有监督和无监督损失函数总和的最小化反向传播优化网络参数, 来实现半监督学习. Zhao 等<sup>[13]</sup> 提出将层次变分自编码器用于梯形网络中, 这种层次化的模型, 可以学习到更丰富的特征, 并用于可视化直观显示. 还有研究学者通过组合多个分类器、编辑伪标签、多样性增强的方法<sup>[14-15]</sup>, 能够得到质量更好的伪标签去训练分类器. 但是, 这些半监督的方法未能够充分讨论未标记数据对分类结果的影响.

为了提高半监督深层生成模型的性能, 本文在梯形网络框架的基础上, 分析了不同编解码器网络结构与提取到特征之间的联系, 并进一步讨论了分类器损失函数对模型分类精度的影响, 提出一种基于梯形网络和改进三训练法的半监督分类模型 (Ladder network and tri-training semi-supervised classification, LN-TT SSC). 该模型在梯形网络框架有噪编码器的最高层添加 3 个分类器, 结合

改进的三训练法提高图像分类性能. 首先, 用基于类别抽样的方法将有标记数据分为 3 份, 结合未标记数据一同传入改进的梯形网络中, 训练得到 3 个 Large-margin Softmax 基础分类器; 然后, 用改进的三训练法对未标记数据预测, 添加伪标签, 对筛选得到的新的标记数据分配不同权重后, 扩充初始训练集, 更新基础分类器; 最后, 对训练完成的分类器进行加权投票, 得到最终结果.

## 1 LN-TT SSC 模型

为了综合利用梯形网络框架的学习能力和三训练算法的分类能力, 提出 LN-TT SSC 模型. 在梯形网络框架有噪编码的最高层添加 3 个分类器, 结合改进的三训练算法提高图像分类的精度. 模型结构如图 1 所示, 主要包括训练过程和决策过程.

LN-TT SSC 模型的训练过程如图 1(a) 所示, 包括改进的梯形网络框架和改进的三训练法. 图 1(a) 中,  $S_i$  是初始有标记训练集;  $X_U$  是未标记数据;  $\mathbf{n}$  是随机高斯噪声;  $\tilde{\mathbf{z}}^{(L)}$ 、 $\hat{\mathbf{z}}^{(L)}$  是每层编码的输出;  $g^{(L)}$  是降噪函数;  $\mathbf{z}^{(L)}$  是每层解码的输出;  $\mathbf{y}_i$  是有噪编码的输出;  $F$  是全连接层或者全局池化层; LM-Softmax 是 Large-margin Softmax 函数;  $\tilde{\mathbf{y}}_i$  是传入数据经  $F$ 、LM-Softmax 处理得到的特征向量;  $h_i$  是基础分类器;  $Li_x$ 、 $Li_y$  是新的标记数据. 改进的梯形网络框架共有 3 个网络支路: 有噪编码器、解码器和无噪编码器, 其中编码器和解码器用 VGG 网络<sup>[16]</sup> 或  $\Pi$  模型<sup>[17]</sup>. 有噪编码器各层得到的特征变量通过跳跃连接 (Skip connection) 映射到对应的解码器上, 而无噪编码器则辅助解码器进行无监督训练达到对有噪数据的最佳映射效果. 分类器改用 Large-margin Softmax 构造有监督损失函数. 模型在训练时首先采用基于类别抽样的方法将有标记数据分为 3 份  $S_1$ 、 $S_2$ 、 $S_3$ , 每一份初始有标记训练数据  $S_i$  分别结合未标记数据  $X_U$  传入梯形网络框架, 以有标记数据的标签损失和未标记数据的重构损失相结合的方式共同调整网络参数, 并训练得到 3 个 Large-margin Softmax 基础分类器  $h_1$ 、 $h_2$ 、 $h_3$ . 然后, 用改进的三训练法预测未标记数据  $X_U$ , 添加伪标签, 并对筛选出的新的标记数据 ( $Li_x$ 、 $Li_y$ ) 分配一个权重  $w$ , 用新的标记数据扩充初始训练集并更新基础分类器.

LN-TT SSC 模型的决策过程如图 1(b) 所示. 图 1(b) 中,  $h'_i$  是更新后的分类器;  $H$  是最终分类器. LN-TT SSC 模型对更新完成的分类器  $h'_i$  加权投票得到分类器  $H$ , 输入测试数据, 传入无噪编码器, 得到最终的分类结果.

### 1.1 LN-TT SSC 模型的网络结构

LN-TT SSC 模型对一般梯形网络的编解码器

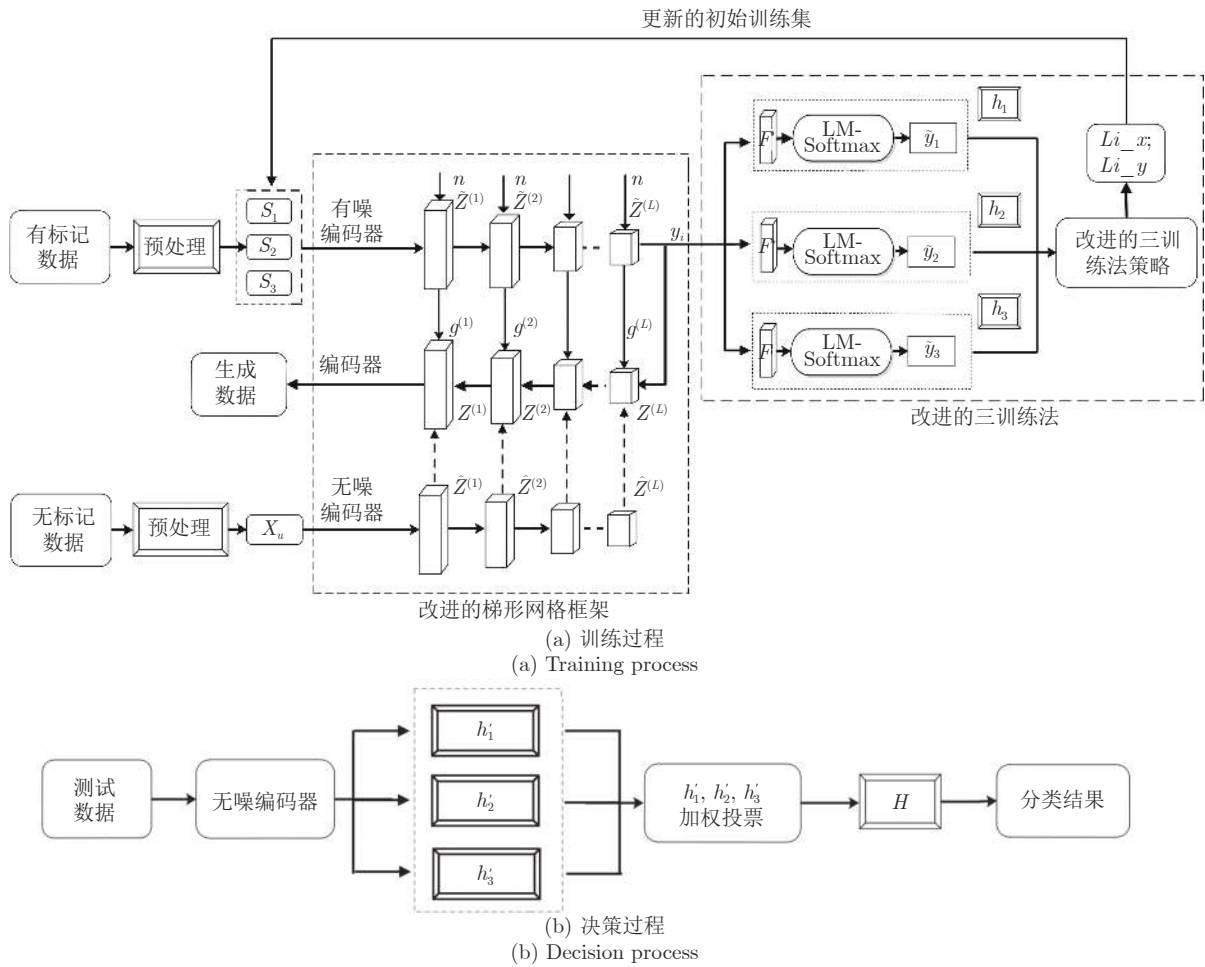


图 1 LN-TT SSC 模型  
Fig.1 LN-TT SSC model

结构和构造分类器的损失函数做了改进，着重提升网络对输入图片提取特征的能力。借用梯形网络框架的优势，对解码器每一层构造损失函数，叠加起来构造总的无监督重构误差函数，采用 Large-margin Softmax<sup>[18]</sup> 构造有监督分类损失函数。

1.1.1 重构损失

针对梯形网络框架能够与有监督算法很好的兼容这一特性，在处理复杂数据时，LN-TT SSC 模型采用了在图像特征提取方面有着卓越表现力的 VGG 网络模型<sup>[16]</sup>。相比传统的卷积神经网络，VGG 使用了连续的几个较小的卷积核来代替较大的卷积核，这样做可以在保证具有相同感受野的情况下，通过不断加深网络的深度达到增强网络提取特征的能力。除此之外，梯形网络框架中编码器的每一层与解码层之间都有一个跳跃连接，它可以避免梯度消失的问题，使得 VGG 等深层神经网络适用于梯形网络框架。

有噪编码器的每一层施加随机高斯噪声  $\mathbf{n}_l$  ( $l = 0, 1, 2, 3, \dots$ )，得到有噪声输入为 ( $\mathbf{x}$  为未添加噪声

数据)

$$\tilde{\mathbf{z}}^{(0)} = \mathbf{x} + \mathbf{n}_0 \tag{1}$$

为了提高网络的收敛速度并且防止因为网络深度的增加出现过拟合的问题，对有噪编码器的第  $l$  层网络输出的特征变量  $\tilde{\mathbf{z}}^{(l)}$  进行批归一化处理：

$$\tilde{\mathbf{z}}^{(l)} = f \left( f_{BN} \left( \tilde{\mathbf{z}}^{(l-1)} \right) \tilde{\mathbf{w}}^{(l)} + \mathbf{n}_l + \tilde{\mathbf{b}}^{(l)} \right) \tag{2}$$

其中， $\tilde{\mathbf{w}}^{(l)}$  和  $\tilde{\mathbf{b}}^{(l)}$  是第  $l$  层神经元的权重和偏置项， $f(x)$  是激活函数， $f_{BN}(x)$  是批正归一化函数，记为

$$\begin{cases} f_{BN}(x) = \frac{x - \mu}{\sqrt{\sigma}} \\ \mu = \frac{1}{m} \sum_{i=1}^m x_i \\ \sigma = \frac{1}{m} \sum_{i=1}^m (x_i - \mu)^2 \end{cases} \tag{3}$$

接着，有噪编码各层的特征变量通过跳跃连接

映射到对应的解码层, 将无噪编码器每层的特征变量作为目标值, 通过无监督训练尽可能多地恢复出未添加噪声数据的信息特征. 按照有噪编码器的思路, 对解码器的每层都进行批归一化处理, 不同的是在此之前, 需要进行降噪. 根据 Pezeshki 等<sup>[11]</sup>的研究成果, 已知有噪数据的特征变量  $\mu$  和无噪数据的先验分布  $\varepsilon$ , 可以得到最优降噪函数  $g^{(l)}$  为

$$g^{(l)}(\mathbf{x}) = \varepsilon \mathbf{x} + (1 - \varepsilon)\mu = (\mathbf{x} - \mu)\varepsilon + \mu \quad (4)$$

解码层每层的特征变量表示为

$$\mathbf{z}^{(l)} = g^{(l)}(\hat{\mathbf{z}}^{(l)}) \quad (5)$$

同样地, 对无噪编码的每层进行批归一化处理

$$\hat{\mathbf{z}}^{(l)} = f(f_{BN}(\hat{\mathbf{z}}^{(l-1)}) \times \hat{\mathbf{w}}^{(l)} + \hat{\mathbf{b}}^{(l)}) \quad (6)$$

这样将式 (5) 得到解码器每一层的特征变量  $\mathbf{z}^{(l)}$ , 结合无噪编码器每层的特征变量  $\hat{\mathbf{z}}^{(l)}$ , 可以得到无监督重构损失函数为

$$R_{\text{loss}} = \sum_{i=m+1}^{m+n} \sum_{l=1}^L \lambda_l (\hat{\mathbf{z}}^{(l)}(i) - \mathbf{z}^{(l)}(i))^2 \quad (7)$$

$\lambda_l$  超参数是第  $l$  层占的比重. 通过最小化重构损失函数, 使输出的重构样本尽可能多的恢复原有数据信息.

### 1.1.2 分类损失

LN-TT SSC 模型采用 Liu 等<sup>[18]</sup>提出的 Large-margin Softmax 替代传统梯形网络中的 Softmax, 来构建有监督分类损失. 它能够引导网络学习到类别之间区分性更大的特征, 从而提高分类精度. 其函数形式为

$$LM_{\text{softmax}}(w, \theta) = -\log \left( \frac{e^{\|\mathbf{w}_{y_i}\| \|\mathbf{x}_i\| \psi(\theta_{y_i})}}{e^{\|\mathbf{w}_j\| \|\mathbf{x}_i\| \psi(\theta_{y_i})} + \sum_{j \neq y_i} e^{\|\mathbf{w}_j\| \|\mathbf{x}_i\| \cos(\theta_j)}} \right) \quad (8)$$

其中,  $\psi(\theta)$  定义为

$$\psi(\theta) = \begin{cases} \cos(m\theta), & 0 \leq \theta \leq \frac{\pi}{m} \\ D(\theta), & \frac{\pi}{m} < \theta \leq \pi \end{cases} \quad (9)$$

式中,  $D(\theta)$  是一个单调减函数且  $D(\pi/m) = \cos(\pi/m)$ .

对于一个多分类问题, 要满足  $j$  的概率大于  $k$  的概率, 必须保证

$$\|\mathbf{w}_j\| \|\mathbf{x}\| \cos(\theta_j) > \|\mathbf{w}_j\| \|\mathbf{x}\| \cos(m\theta_j) > \|\mathbf{w}_k\| \|\mathbf{x}\| \cos(\theta_k) \quad (10)$$

其中,  $0 \leq \theta_1 \leq \pi/m$ , 由于  $\cos(\theta)$  在区间  $[0, \pi]$  之间是递减函数,  $m$  是一个正整数, 当  $m$  越大时, 分类的边界越大, 对  $\mathbf{w}_j$  和  $\mathbf{w}_k$  的学习就要更加困难, 这样就使得类别  $j$  与类别  $k$  之间有了区分度更高的差

异. 有噪编码器的输出  $\mathbf{y}_i$  经  $F$  层得到  $F(\mathbf{y}_i)$ , 接着经 Large-margin Softmax 处理得到  $p$  个  $1 \times q$  维的向量  $\tilde{\mathbf{y}}$ ,  $p$  是样本个数,  $q$  是样本类别数:

$$\tilde{\mathbf{y}}_i = LM_{\text{softmax}}(F(\mathbf{y}_i)) \quad (11)$$

假设一组数据, 有  $m$  个有标记样本  $X_L = \{(x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_m, y_m)\}$ ,  $n$  个未标记样本  $X_U = \{(x_{m+1}), (x_{m+2}), (x_{m+3}), \dots, (x_{m+n})\}$ , 将有标记样本分为 3 份  $m_1, m_2, m_3$ , 分别训练 3 个基础分类器, 记训练第  $i$  个基础分类器的有监督分类损失为

$$S_{\text{loss}_i} = -\sum_{j=1}^{m_i} [\mathbf{y}_j \log_2(\tilde{\mathbf{y}}_{ij}) + (1 - \mathbf{y}_j) \log_2(1 - \tilde{\mathbf{y}}_{ij})], \quad i = 1, 2, 3 \quad (12)$$

$\tilde{\mathbf{y}}_{ij}$  是第  $i$  个分类器对第  $j$  个样本的预测值, 通过梯度下降法使  $S_{\text{loss}_i}$  最小化, 迫使网络输出值去拟合真实标签.

根据式 (7) 和式 (12), 可以得到训练第  $i$  个基础分类器总的损失函数为

$$\text{loss}_i = S_{\text{loss}_i} + R_{\text{loss}} \quad (13)$$

有监督损失和无监督损失都可以通过梯度下降法来达到最小化, 因此采用将二者相结合的方式共同调整网络参数, 进一步地增强网络提取特征的能力.

## 1.2 改进的三训练法

给定的  $m$  个有标记样本  $X_L = \{(x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_m, y_m)\}$ ,  $n$  个未标记样本  $X_U = \{(x_{m+1}), (x_{m+2}), (x_{m+3}), \dots, (x_{m+n})\}$ , 考虑到能够充分利用有标记样本, 避免资源的浪费, 模型采用基于类别抽样的方法将有标记数据  $X_L$  分为 3 个初始训练集  $S_1, S_2, S_3$ , 依次结合未标记数据  $X_U$ , 训练得到 3 个基础分类器  $h_1, h_2, h_3$ . 这样既能保证训练集之间的差异性又能覆盖到所有的有标记数据.

传统三训练算法, 采用两两分类器结合为联合分类器的方式对未标记数据预测, 通过判断联合分类器预测是否一致挑选未标记数据, 但是如果基础分类器性能较差会导致这种标记置信度的方式不够准确. 因此, LN-TT SSC 模型提出一种改进的三训练法对未标记数据标注, 添加伪标签.

经过第 1.1 节得到  $p$  个  $1 \times q$  维的向量  $\tilde{\mathbf{y}}$ , 首先根据向量衡量第  $i$  个分类器对第  $k$  个未标记样本预测的好坏:

$$e_{ik} = -\tilde{\mathbf{y}}_{ik} \ln \tilde{\mathbf{y}}_{ik} \quad (14)$$

熵值越小表明分类器对未标记数据的预测结果

越好. 接着计算任意两个分类器  $a, b$  对同一个未标记数据  $r$  预测值的欧氏距离

$$d_{ab} = \sqrt{(\tilde{\mathbf{y}}_{ar} - \tilde{\mathbf{y}}_{br})^2} \quad (15)$$

改进的三训练法用式 (16) 衡量  $a, b$  的分类差异为

$$D_{ab} = \eta \left( \frac{e_{ak} + e_{bk}}{2} \right) + (1 - \eta)d_{ab} \quad (16)$$

当  $D_{ab} < \varphi$  时, 表明分类器  $a$  和分类器  $b$  对同一个未标记数据标注相同, 则将该样本连同伪标签作为新的标记数据加入到第 3 个分类器的初始训练集中.  $\eta$  超参数用来确定熵值和欧氏距离对  $D_{ab}$  值大小的影响,  $\varphi$  是很小的正数, 根据具体分类问题确定  $\varphi$  的值.

为了解决每个样本具有不同确定度的问题, 使用熵值法为每一个新的标记数据分配一个权重  $w$ , 用以反映不同分类器的分类性能. 根据式 (14), 计算第  $k$  个样本的信息熵, 假设  $e_{ak} < e_{bk}$ , 定义权重  $w$  的计算式为

$$w_k = 1 - \frac{e_{ak}}{\log_2(N)} \quad (17)$$

$N$  是样本类别数, 某个样本的信息熵  $e$  越小, 权重  $w$  越大, 表明它提供的信息也就越多, 所起到的作用也就越大.

LN-TT SSC 模型采用改进的三训练法对未标记数据预测, 标注伪标签并筛选分别得到  $t_1, t_2, t_3$  个新的标记样本, 用新的标记样本扩充训练集更新基础分类器. 此时, 更新第  $i$  个基础分类器的有监督损失可以表示为

$$S_{\text{loss}_i} = - \sum_{j=1}^{m_i} [\mathbf{y}_j \log(\tilde{\mathbf{y}}_{ij}) + (1 - \mathbf{y}_j) \log(1 - \tilde{\mathbf{y}}_{ij})] - \sum_{k=m_i+1}^{m_i+t_i} w_k [\mathbf{y}_k \log(\tilde{\mathbf{y}}_{ik}) + (1 - \mathbf{y}_k) \log(1 - \tilde{\mathbf{y}}_{ik})], \quad i = 1, 2, 3 \quad (18)$$

模型在训练时, 网络传入 4 组数据, 其中包括 3 个有标记数据的训练集和一个无标记数据的训练集, 定义整个模型的损失函数为

$$\text{Loss} = - \sum_{i=1}^3 S_{\text{loss}_i} + R_{\text{loss}} \quad (19)$$

前半部分为有标记数据的交叉熵损失, 后半部分为无标记数据每层解码器重构损失之和. 式 (19) 表示 LN-TT SSC 模型采用 3 个有标记数据的标记损失和未标记数据的重构损失相结合的方式反向传

播, 调整整体模型网络参数. 后续实验证明, 该种方式能够进一步提升网络的学习能力.

### 1.3 LN-TT SSC 模型的训练过程

LN-TT SSC 模型训练时主要有两个步骤: 1) 改进的梯形网络框架做特征提取并训练基础分类器; 2) 改进的三训练法对未标记数据预测, 标注伪标签并筛选新的标记数据, 扩充初始训练集更新基础分类器. LN-TT SSC 模型具体的训练过程如下:

**步骤 1.** 数据集划分, 对有标记数据  $X_L$  采用基于类别抽样的方法分为 3 份  $S_1, S_2, S_3$ ;

**步骤 2.** 训练基础分类器, 分别结合  $S_1$  与  $X_U, S_2$  与  $X_U, S_3$  与  $X_U$  传入改进的梯形网络框架, 并按照式 (13) 训练, 得到 3 个 Large-margin Softmax 的基础分类器  $h_1, h_2, h_3$ ;

**步骤 3.** 改进的三训练法对未标记数据  $X_U$  预测, 标注伪标签, 通过式 (16) 筛选得到新的标记数据  $(Li_x; Li_y)$ , 按照式 (17) 对新的标记数据分配权重  $w$ , 扩充初始训练集  $S_i$ ;

**步骤 4.** 根据扩充后得到的训练集  $S'_i$ , 更新基础分类器  $h_i$ ;

**步骤 5.** 按照式 (19) 计算模型损失函数, 采用标记数据损失和未标记数据损失相结合的方式反向传播, 调整模型参数.

### 1.4 LN-TT SSC 模型的决策过程

由于每个分类器性能不一样, LN-TT SSC 模型依据每个分类器准确率所占的比重对更新完全的分类器用加权投票法进行集成,  $p_i$  是第  $i$  个更新过的分类器对初始已标记数据  $S_i$  预测的准确率

$$H(x) = \sum_{i=1}^3 \frac{p_i}{p_1 + p_2 + p_3} \times h_i(x) \quad (20)$$

$H(x)$  对输入的测试样本  $X_{\text{test}}$  预测, 得到  $1 \times q$  维的向量, 取  $q$  维中概率最大的维数所属类别为该样本最终的预测结果. 通过对更新完全的 3 个分类器进行加权投票, 既可以提高分类精度, 又能增强模型整体的泛化能力.

## 2 实验仿真及结果分析

### 2.1 数据集及实验配置

MNSIT 手写字符数据集: MNIST 有 10 个类别, 包括 60000 个训练样本, 10000 个测试样本. 图像均为单通道黑白图像, 大小为  $28 \times 28$  的手写字符. 参考文献 [8-11, 19], 从训练样本中分别选取

$N_1 = 100$ ,  $N_2 = 1\ 000$  个有标记数据作有监督训练, 其余为无标记数据.

SVHN 街牌号数据集: SVHN 包括 73 257 个训练样本和 26 032 个测试样本. 图像为彩色图像, 大小为  $32 \times 32$ , 每张图片上有一个或者多个数字, 且图像类别以识别的正中数据为准. 参考文献 [8–11, 19], 从训练样本中分别选取  $N_1 = 100$ ,  $N_2 = 1\ 000$  个有标记数据作有监督训练, 其余为无标记数据.

CIFAR10 自然图像数据集: CIFAR10 有 10 个类别, 包括 50 000 个训练样本和 10 000 个测试样本. 图像为彩色图像, 大小为  $32 \times 32$ . CIFAR10 数据集中的图片大都来自真实物体, 图片之间差异性较大, 识别比较困难, 所以参考文献 [8, 10–11, 17, 19–21], 从训练样本中分别选取  $N_1 = 2\ 000$ ,  $N_2 = 4\ 000$  个有标记数据作有监督训练, 其余为无标记数据.

实验硬件采用 Intel Xeon E5-2687W CPU, 32 GB 内存和 GTX 1080Ti GPU 平台; 软件采用 Windows 系统, python 语言, tensorflow 深度学习框架.

对较简单的 MNIST 数据集, 网络结构采用  $\Pi$  模型<sup>[7]</sup>, 主要由 9 个卷积层、3 个池化层和 1 个全连接层组成. 对较复杂的 SVHN、CIFAR10 数据集, 网络结构采用 VGG-19<sup>[6]</sup>, 由 16 个卷积层、5 个池化层和 3 个全连接层组成. 模型为减少梯度对参数大小的依赖性, 对每层编、解码结构都采用了批归一化处理. 且设置批次大小为 100. 另外, 实验中将遍历次数设置为总样本数除以批次大小, 以增加多样性. 为保证模型的收敛速度, 采用指数衰减的方式更新学习率, 并且定义初始学习率为 0.02 (MNIST), 0.003 (SVHN), 0.003 (CIFAR10). 实验设置迭代 120 个 epoch (MNIST), 150 个 epoch (SVHN, CIFAR10).

## 2.2 实验结果

为了进一步分析不同编解码器网络结构和不同分类器损失函数对模型分类精度的影响, 对较复杂数据集 SVHN (训练集中采样 1 000 个有标记样本) 和 CIFAR10 (训练集中采样 4 000 个有标记样本), 以基础模型框架 ( $\Pi$  模型与 Softmax 损失的结合方式) 为基准, 本文提出两种混合方案. 方案 A: 不使用 VGG-19 网络, 编解码器结构采用  $\Pi$  模型; 方案 B: 不使用 Large-margin Softmax 损失, 使用基础 Softmax 损失. 实验结果如表 1 所示.

表 1 显示, LN-TT SSC 模型显著提高了这两个数据集的分类精度. 观察可以发现 VGG-19 网络和 Large-margin Softmax 损失都是提升模型性能

表 1 LN-TT SSC 模型及混合方案在 SVHN 和 CIFAR10 数据集上的分类精度 (%)

Table 1 Classification accuracy of LN-TT SSC model and hybrid scheme on SVHN and CIFAR10 (%)

模型	SVHN	CIFAR10
基准实验 ( $\Pi$ + softmax)	88.75	83.56
方案 A ( $\Pi$ + LM-Softmax)	91.35	86.11
方案 B (VGG19 + softmax)	89.72	84.73
LN-TT SSC	93.05	88.64

的重要因素. 在 SVHN 和 CIFAR10 数据集上, VGG-19 网络的改进率分别为 0.97% 和 1.17%, Large-margin Softmax 损失的改进率分别为 2.6% 和 2.55%, 通过结合这两个因素, 改进率为 4.3% 和 5.08%. 实验结果显示, 利用 VGG-19 网络和 Large-margin Softmax 损失结合的模式使得网络具有更强的学习能力, 同时也表明 LN-TT SSC 模型更有效性.

## 2.3 实验分析

为了验证模型的训练效果, 对 MNIST 数据集随机选取 2 000 个有标记数据用 t-SNE<sup>[20]</sup> 对经过 Large-margin Softmax 函数之后的图像特征进行降维, 得到如图 2 所示的流形图. 观察可以发现, LN-TT SSC 模型训练后得到的梯形网络的特征分布较为集中, 表明类别之间的区分性较强, 可以有效地避免因为样本数据分布不均而导致的分类误差, 提高模型分类精度.

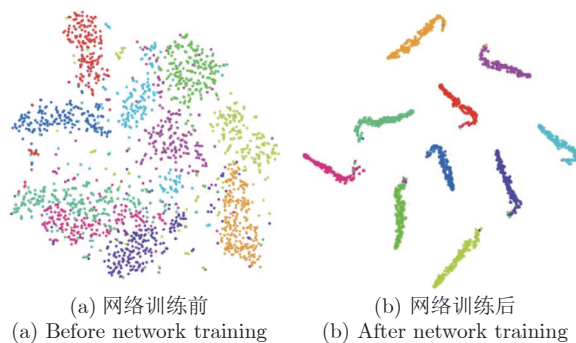
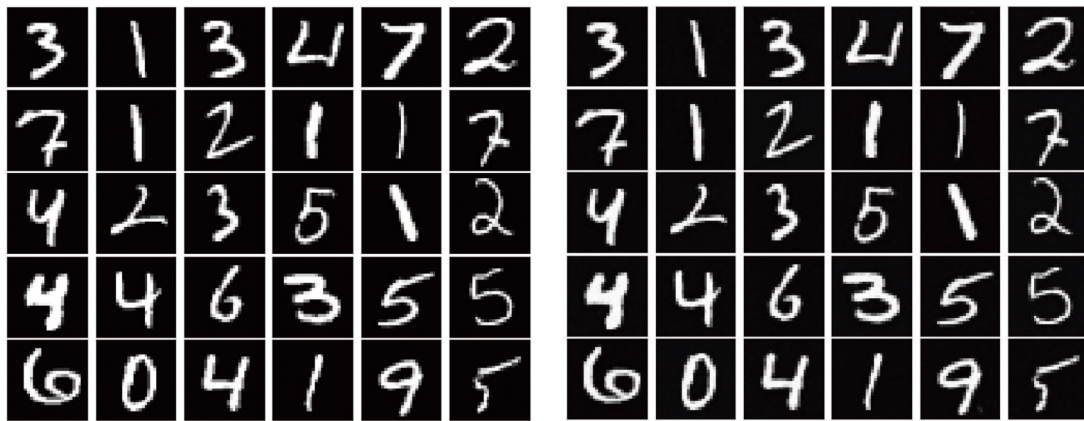


图 2 MNIST 数据的流行表示图

Fig. 2 A popular representation of MNIST

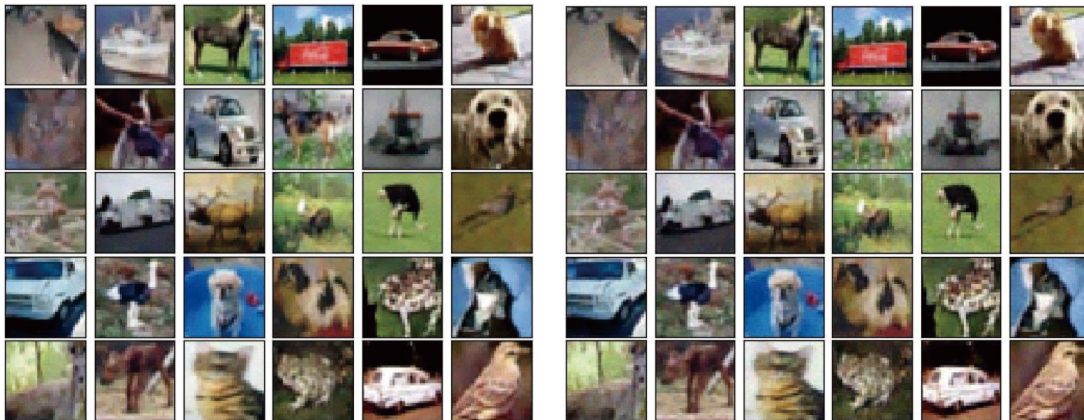
梯形网络架构中, 由于解码器是编码器逆运算的一个过程, 所以也可以通过观察经解码之后得到的图片的质量评估模型的收敛程度. 图 3 显示了 LN-TT SSC 模型趋于收敛时的生成图片 (右) 与原始输入图片 (左) 的对比情况. 图 3(a), 对 MNIST 数据集选取 1 000 个有标记数据, 迭代训练 120 次



(a) MNIST 数据集图片对比  
(a) Image comparison of MNIST



(b) SVHN 数据集图片对比  
(b) Image comparison of SVHN



(c) CIFAR10 数据集图片对比  
(c) Image comparison of CIFAR10

图 3 LN-TT SSC 模型趋于收敛时的图片对比

Fig.3 Comparison of image when the LN-TT SSC model tends to converge

趋于收敛,可以看出已基本恢复出原始输入图片的信息,生成图片质量较高.图 3(b),对 SVHN 数据集选取 1000 个有标记数据,迭代训练 150 次趋于收敛,生成的个别图像比较粗糙,但大多都能还原出原始图片的对应信息.观察图 3 可以发现, LN-

TT SSC 模型不仅对数据学习能力较强,而且在处理不同数据集时,该模型都有很好的鲁棒性.

### 2.4 实验结果对比

为了验证 LN-TT SSC 模型的优势,按照第 2.1 节

的实验配置和采样不同有标记样本数目, 以分类精度为评价标准, 与当前主要半监督生成模型进行对比, 实验结果如表 2 ~ 4 所示.

表 2 MNIST 数据集上的分类精度 (%)

Table 2 Classification accuracy on MNIST dataset (%)

模型	$N_1 = 100$	$N_2 = 1000$
Ladder network <sup>[11]</sup>	98.88	99.06
CAT-GAN <sup>[8]</sup>	99.09	99.11
Improved-GAN <sup>[9]</sup>	98.58	99.15
GAR <sup>[19]</sup>	98.92	99.21
SSE-GAN <sup>[10]</sup>	99.10	92.23
LN-TT SSC	99.14	99.30

表 3 SVHN 数据集上的分类精度 (%)

Table 3 Classification accuracy on SVHN dataset (%)

模型	$N_1 = 100$	$N_2 = 1000$
Ladder network <sup>[11]</sup>	75.50	87.06
CAT-GAN <sup>[8]</sup>	77.68	88.90
Improved-GAN <sup>[9]</sup>	81.57	90.78
GAR <sup>[19]</sup>	80.87	92.08
SSE-GAN <sup>[10]</sup>	81.08	92.92
LN-TT SSC	82.13	93.05

表 4 CIFAR10 数据集上的分类精度 (%)

Table 4 Classification accuracy on CIFAR10 dataset (%)

模型	$N_1 = 2000$	$N_2 = 4000$
Ladder network <sup>[11]</sup>	76.56	79.68
CAT-GAN <sup>[8]</sup>	78.83	80.42
Temporal-Ensembling <sup>[17]</sup>	—	87.84
GAR <sup>[19]</sup>	82.10	83.35
SSE-GAN <sup>[10]</sup>	83.66	85.14
VAT <sup>[21]</sup>	82.88	85.61
LN-TT SSC	84.14	88.64

结果显示, 本文提出的 LN-TT SSC 模型有更高的分类精度. 表明模型对数据的分布拟合更加准确, 对复杂数据仍具有较强的学习能力, 同时也表明模型具有很强的泛化能力.

### 3 结束语

为了进一步提高半监督分类模型的精度, 本文将梯形网络框架与三训练算法相结合, 提出了基于梯形网络和三训练法的半监督分类模型. LN-TT SSC 模型相对其他方法有以下几点优势: 1) 利用 VGG 等深层神经网络和 Large-Margin Softmax 损失结合的模式, 使网络具有更强的学习能力, 并且

对特征有更好的低维流形表示, 直观上更容易分析数据分布规律; 2) 利用了梯形网络的优势, 通过对有监督和无监督损失函数总和的最小化来反向传播优化网络参数, 得到基础分类器的分类性能更好; 3) 采用改进的三训练法对未标记数据预测, 标注伪标签、筛选得到的新的标记数据更具代表性. 实验表明针对不同复杂程度的图像数据, LN-TT SSC 模型均可利用少量的有标记数据训练得到更高的分类精度. 同时 LN-TT SSC 模型也有一定的不足, 在用改进三训练法对未标记数据预测时, 超参数的取值需要人为调试, 对最终的结果有一定的影响, 在以后的工作中需要继续研究. 与其他半监督生成模型一样, LN-TT SSC 模型也存在参数过多的问题, 这导致训练过程较为耗时, 在之后的工作中也会继续研究如何能够保证分类精度的同时又能有效地减少训练时间.

### References

- Liu Jian-Wei, Liu Yuan, Luo Xiong-Lin. Semi-supervised learning. *Journal of Computer Science*, 2015, **38**(8): 1592-1617 (刘建伟, 刘媛, 罗雄麟. 半监督学习方法. 计算机学报, 2015, **38**(8): 1592-1617)
- Chau V T N, Phung N H. Combining self-training and tri-training for course-level student classification. In: Proceedings of the 2018 International Conference on Engineering, Applied Sciences, and Technology, Phuket, Thailand: IEEE, 2018. 1-4
- Lu X C, Zhang J P, Li T, Zhang Y. Hyperspectral image classification based on semi-supervised rotation forest. *Remote Sensing*, 2017, **9**(9): 924-938
- Cheung E, Li Y. Self-training with adaptive regularization for S3VM. In: Proceedings of the 2017 International Joint Conference on Neural Networks, Anchorage, AR, USA: IEEE, 2017. 3633-3640
- Kvan Y, Wadysaw S. Convolutional and recurrent neural networks for face image analysis. *Foundations of Computing and Decision Sciences*, 2019, **44**(3): 331-347
- Luo Jian-Hao, Wu Jian-Xin. A review of fine-grained image-classification based on depth convolution features. *Acta Automatica Sinica*, 2017, **43**(8): 1306-1318 (罗建豪, 吴建鑫. 基于深度卷积特征的细粒度图像分类研究综述. 自动化学报, 2017, **43**(8): 1306-1318)
- Lin Yi-Lun, Dai Xing-Yuan, Li Li, Wang Xiao, Wang Fei-Yue. A new frontier in artificial intelligence research: Generative adversarial networks. *Acta Automatica Sinica*, 2018, **44**(5): 775-792 (林懿伦, 戴星原, 李力, 王晓, 王飞跃. 人工智能研究的新前线: 生成对抗网络. 自动化学报, 2018, **44**(5): 775-792)
- Springenberg J T. Unsupervised and semi-supervised learning with categorical generative adversarial networks. In: Proceedings of the 2005 International Conference on Learning Representations, 2015. 1-20
- Salimans T, Goodfellow I, Zaremba W, Cheung V, Radford A, Chen X. Improved techniques for training GANs. In: Proceedings of the 30th International Conference on Neural Information Processing Systems, Barcelona, Spain: ACM, 2016. 2234-2242

- 10 Fu Xiao, Shen Yuan-Tong, Li Hong-Wei, Cheng Xiao-Mei. Image classification model based on semi-supervised coding generative adversarial networks. *Acta Automatica Sinica*, 2020, **46**(3): 531–539  
(付晓, 沈远彤, 李宏伟, 程晓梅. 基于半监督编码生成对抗网络的图像分类模型. *自动化学报*, 2020, **46**(3): 531–539)
- 11 Pezeshki M, Fan L X, Brakel P, Courville A, Bengio Y. Deconstructing the ladder network architecture. In: Proceedings of the 33rd International Conference on Machine Learning, New York, NY, USA: ACM, 2016. 2368–2376
- 12 Rasmus A, Valpola H, Honkala M, Berglund M, Raiko T. Semi-supervised learning with ladder networks. *Computer Science*, 2015, **9**(1): 1–9
- 13 Zhao S, Song J, Ermon S. Learning hierarchical features from generative models. In Proceedings of the 34th International Conference on Machine Learning, Sydney, NSW, Australia: ACM, 2017. 4091–4099
- 14 Saito K, Ushiku Y, Harada T. Asymmetric tri-training for unsupervised domain adaptation. *International Conference on Machine Learning*, 2017: 2988–2997
- 15 Chen D D, Wang W, Gao W, Zhou Z H. Tri-net for semi-supervised deep learning. In: Proceedings of the 27th International Joint Conference on Artificial Intelligence, Stockholm, Sweden: ACM, 2018. 2014–2020
- 16 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *Computer Science and Information Systems*, 2014.
- 17 Samuli L, Timo A. Temporal ensembling for semi-supervised learning. In: Proceedings of the 2017 International Conference on Learning Representation, 2017. 1–13
- 18 Liu W Y, Wen Y D, Yu Z D, Yang M. Large-margin softmax loss for convolutional neural networks. In: Proceedings of the 33rd International Conference on Machine Learning, New York, NY, USA: ACM, 2016. 507–516
- 19 Kilinc O, Uysal I. GAR: An efficient and scalable graph-based activity regularization for semi-supervised learning. *Neuro-computing*, 2018, **296**(28): 46–54
- 20 Maaten L, Hinton G. Visualizing data using t-SNE. *Journal of Machine Learning Research*, 2008, **9**: 2579–2605
- 21 Miyato T, Maeda S, Koyama M, Lshii S. Virtual adversarial training: A regularization method for supervised and semi-supervised learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, **7**(23): 99–112



**莫建文** 桂林电子科技大学信息与通信学院副教授。主要研究方向为机器视觉, 智能图像处理。本文通信作者。

E-mail: mo\_jianwen@126.com

**(MO Jian-Wen** Associate professor at the School of Information and Communication, Guilin University

of Electronic Science and Technology. His research interest covers machine vision and intelligent image processing. Corresponding author of this paper.)



**贾鹏** 桂林电子科技大学信息与通信学院硕士研究生。主要研究方向为智能图像处理。

E-mail: jiapeng\_jay@163.com

**(JIA Peng** Master student at the School of Information and Communication, Guilin University of Elec-

tronic Science and Technology. His main research interest is intelligent image processing.)