

目录

1. CSI fingerprinting with SVM regression to achieve device-free passive localization.....	1
2. Device-free crowd counting with WiFi channel state information and deep neural networks.....	1
3. Device-Free Localization Based on CSI Fingerprints and Deep Neural Networks.....	1
4. Device-Free Presence Detection and Localization With SVM and CSI Fingerprinting.....	1
5. Adaptive Device-free Localization in Dynamic Environments through Adaptive Neural Networks.....	2
6. Deep Spatial – Temporal Model Based Cross-Scene Action Recognition Using Commodity WiFi.....	3
7. Bi-Directional Generation for Unsupervised Domain Adaptation.....	4
8. Learning Domain Adaptive Features with Unlabeled Domain Bridges..	5
9. Deep Subdomain Adaptation Network for Image Classification.....	6
10. Knowledge Distillation for BERT Unsupervised Domain Adaptation....	7
11. Side-Tuning: A Baseline for Network Adaptation via Additive Side Networks....	8
12. Detection of Suspicious Objects Concealed by Walking Pedestrians Using WiFi....	9
13. Wi-Metal: Detecting Metal by Using Wireless Networks....	10

14. Weakly Supervised Object Detection Using Complementary Learning and Instance Clustering.....	11
15. Detecting Suspicious Objects With a Humanoid Robot Having a Metal Detector.....	13
16. Image-to-Image Translation with Conditional Adversarial Networks	14
17. A Survey of Clustering With Deep Learning: From the Perspective of Network Architecture...	15
18. Adversarial Autoencoders...	17
19. Adversarial Latent Autoencoders....	18
20. Learning Sleep Stages from Radio Signals:A Conditional Adversarial Architecture....	21
21. One-Class Convolutional Neural Network...	23
22. Towards K-means-friendly Spaces: Simultaneous Deep Learning and Clustering....	24
23. One-class Classification with Deep Autoencoder....	25
24. Online Deep Clustering for Unsupervised Representation Learning..	26
25. Unsupervised Domain Adaptation via Structurally Regularized Deep Clustering..	28
26. Feature Sampling Based Unsupervised Semantic Clustering for Real Web Multi-View Content..	30
27. A Bottom-Up Clustering Approach to Unsupervised Person Re-Identification...	31

28. Invariant Information Clustering for Unsupervised Image Classification and Segmentation..33
29. Towards inbaggage suspicious object detection using commodity WiFi...35
30. Non-obtrusive detection of concealed metallic objects using commodity WiFi radios...37
31. FiDo Ubiquitous Fine-Grained WiFi-based Localization for Unlabelled Users via Domain Adaptation...39
32. 2021-CCNC-Congestion-Aware Suspicious Object Detection...41
33. 2020-AAAI-Adversarial Training Based Multi-Source Unsupervised Domain Adaptation for Sentiment Analysis...43

CSI fingerprinting with SVM regression to achieve device-free passive localization

Device-free crowd counting with WiFi channel state information and deep neural networks

Device-Free Localization Based on CSI Fingerprints and Deep Neural Networks

Device-Free Presence Detection and Localization With SVM and CSI Fingerprinting

Adaptive Device-free Localization in Dynamic Environments through Adaptive Neural Networks

主要技术:

(1) iForest: 预处理

(2) 多层 CNN 接入 DNN, 当作为定位模型, 利用来自源域的训练样本, 对自适应模型进行训练, 当作为域适应模型

(3) Semantic alignment: 对相同位置 and 不同位置的源和目标域 CSI 指纹数据进行对齐。(Euclidean distance 带入训练)

存在问题:

目标域变化与源域差别太小 (是否开窗), 需要目标域标签

参考:

https://en.wikipedia.org/w/index.php?title=Domain_adaptation&action=edit§ion=2

<https://levelup.gitconnected.com/understanding-domain-adaptation-63b3bb89436f>

将源域的 CSI 指纹和目标域中的少量指纹共享到一个共享空间, 我们希望从源数据集和目标数据集中提取的特征相似。using the Euclidean distance as the measurement to minimize the distribution mismatch. 用 Domain Adaptation (DA), Semantic Alignment (SA) 技术重新建立新的模型来适应目标域的定位。

AdapLoc consists of a localization model based on onedimensional Convolutional Neural Network (1D-CNN) and a domain adaptation model with Semantic Alignment (SA). AdapLoc adapts the localization model of the source domain using the CSI fingerprints from the source domain and a small number of CSI fingerprints from the target domain, by means of mapping the CSI fingerprints from the source and the target domains into a shared space, with the aim of minimizing the distribution divergence between the two domains. Meanwhile, the source and the target domains are semantically aligned, by minimizing the distance of the mapped fingerprints at the same locations and maximizing the distance of the mapped fingerprints at different locations from the source and the target domains.

Its structure is illustrated in Fig. 5. Depending on the input, the adaptive model acts as a localization model or a domain adaptation model. When acting as the localization model, denoted as $f_R(\cdot)$, the adaptive model is trained to minimize the localization loss LR in equation (1), using the training samples from the source domain. When acting as the domain adaptation model, denoted as $f_M(\cdot)$, the adaptive model without the output layer is trained to minimize the domain loss LD in equation (2), using the training samples from both the source and the target domains (一个模型实现两种函数)

Deep Spatial – Temporal Model Based Cross-Scene Action Recognition Using Commodity WiFi

让志愿者做同样动作十次，每次间隔一段时间，收集 CSI 数据并做预处理，用滑动窗口处理 CSI 数据，将数据带入卷积层和全连接层结合的模型 m1 进行行为识别分类并进行训练(监督学习)，观察结果准确率。取出训练好的 m1 模型中的全连接层作为输入数据带入 Bi-LSTM 模型 m2 进行训练，同样做行为识别分类训练，观察准确率。

微调: (适应新环境)

cnn 和 Bi-LSTM 层的参数不做变化，带入新环境的 CSI 数据微调全链接层的参数进行训练

First, in order to capture the spatially local dependencies among CSI channels and adjacent sequences, we divide CSI streams of a sample into multiple segments, from which spatial features are extracted by a CNN model. Then, the original time sequences are converted into CNN feature arrays, each of which denotes spatial features for the corresponding CSI snippet. Furthermore, a Bi-LSTM model which includes

The problem can be solved if less new data and computation are required when the scene changes. To address the challenge and improve the robustness to dynamics, we consider a transfer learning algorithm which fine-tunes the new model based

on the parameters learned on another similar task instead of training from scratch. In detail, we train our proposed model on the activity recognition task T1

$$\theta(T1) = \arg \min_{\theta \in \text{Loss}(\text{Input}(T1), \theta)} (10)$$

where , Input(T1), and Loss, respectively, denote the parameter domain, input data of T1, and the model loss function.

Then, we try to finish another similar task T2 in which the data collection scene and subjects are different from that in T1.

In this article, we adopt the transfer learning approach that fine-tunes the parameters $\theta(T1)$ trained on T1 for the T2 task

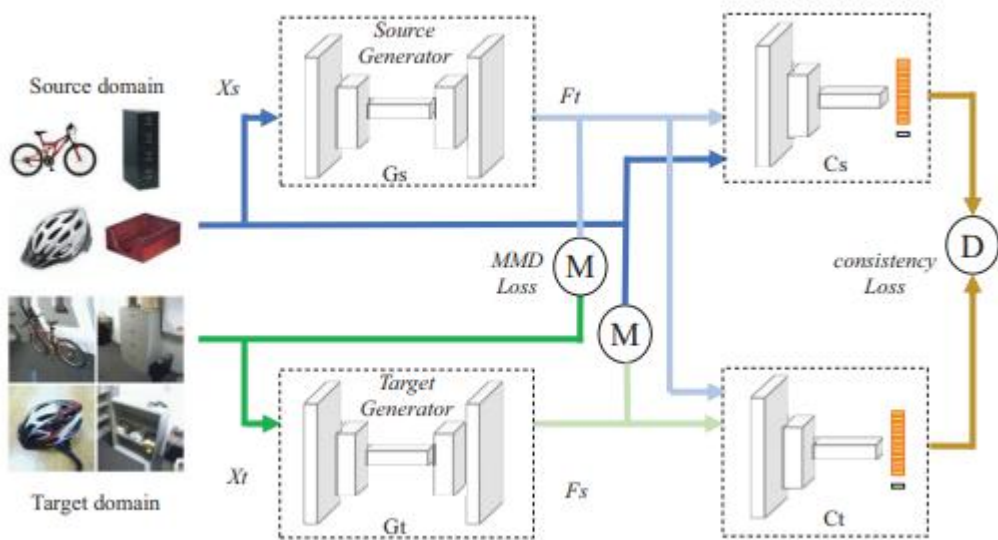
$$\theta(T2) = \arg \min_{\theta \in \text{Loss}(\text{Input}(T2), \theta(T1) + \theta)} (11)$$

where Input(T2) is the subset of the data collected in T2. In our approach, the parameters of convolutional layers in CNN and Bi-LSTM layers are fixed with only the fully connected layer parameters fine-tuned by a part of data in T2. Consequently, the data demand and computation consumption can be reduced when the scene is transferred.

Bi-Directional Generation for Unsupervised Domain Adaptation

The traditional methods like cycle loss and identity loss in cycleGAN are too restricted for domain adaption. Meanwhile, the previous process in domain adaption only focuses on the global transform. Although it can reduce the distribution difference as cross domain, it destroys the class semantic feature in each sample. It leads to that the classifier trained in this way cannot get the ideal result in the target domain.

Differently, we propose a dual generative cross-domain generation framework by interpolating two intermediate domains to bridge the domain gap. Our proposed method leverages bi-directional cross-domain generators to make two intermediate domains and use additional target data with pseudo labels for learning two task-specific classifiers. Our work is on the exploitation of building bi-directional generation network with two classifiers, which has not been fully explored in the literature.



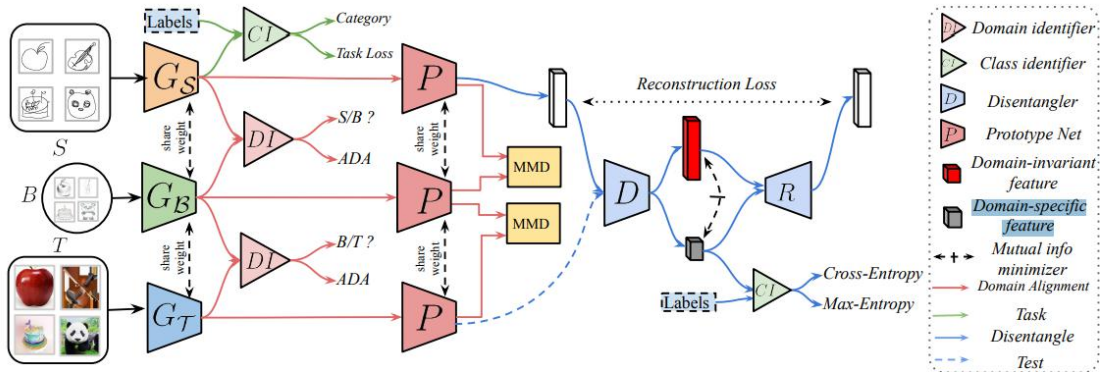
页码:P6617. (有计算公式和解析)

As illustrated in Figure 1, X_s , X_t are source and target samples, respectively. We propose the cross-domain generators G_s , G_t to transfer one domain input to the other domain distribution. Specifically, two generators are defined as $G_s : X_s \rightarrow X_t$ and $G_t : X_t \rightarrow X_s$, respectively. Given the source samples X_s , G_s tries to generate F_t that looks similar to target samples X_t . Similarly, With X_t , G_t aims to generate F_s which looks similar to X_s .

Learning Domain Adaptive Features with Unlabeled Domain Bridges

In this paper, we propose a novel approach to learn domain adaptive features between the largely-gapped source and target domains with unlabeled domain bridges. Firstly, we introduce the framework of Cycle-consistency Flow Generative Adversarial Networks (CFGAN) that utilizes domain bridges to perform image-to-image translation between two distantly distributed domains. Secondly, we propose the Prototypical Adversarial Domain Adaptation (PADA) model which utilizes unlabeled bridge domains to align feature distribution between source and target with a large discrepancy

(1) our approach is devised specifically to tackle the significantly large domain shift, (2) instead of directly synthesizing bridge domains using the source and target domains, we leverage an existing third domain to bridge two distant source and target domains. (是否可以实现跨房间?) CycleGAN [47] introduces a cycle-consistency loss to recover the original images using a cycle of translation and reverse translation. However, these methods assume that the domain gap between the source and target is relatively small (CycleGAN 用于变化较小的域适应)

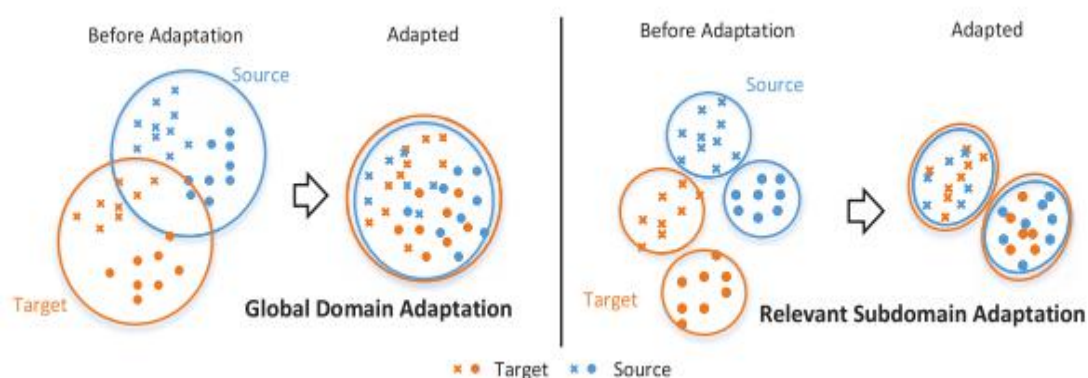


Deep Subdomain Adaptation Network for Image Classification

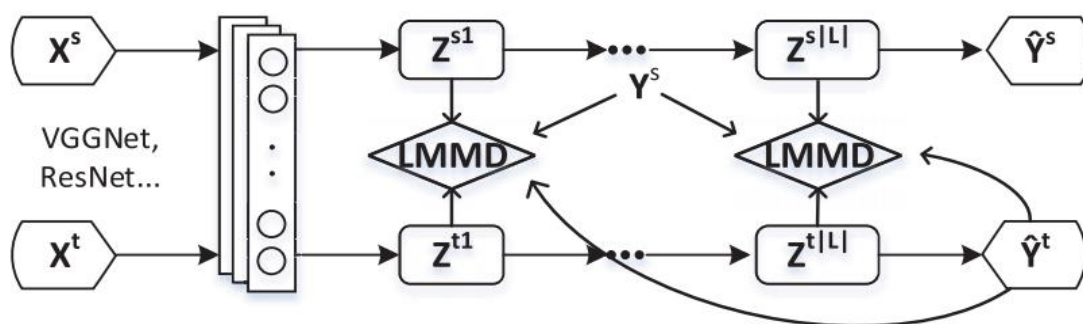
Recently, more and more researchers pay attention to subdomain adaptation that focuses on accurately aligning the distributions of the relevant subdomains. However, most of them are adversarial methods that contain several loss functions and converge slowly. Based on this, we present a deep subdomain adaptation network (DSAN) that learns a transfer network by aligning the relevant subdomain distributions of domain-specific layer activations across different domains based on a local maximum mean discrepancy (LMMD). 对抗犹如用一个对抗网络区分两个不同域，从而使两个域分布接近，而 MMD 直接比较生成器生成的域分布差异来训练。gan 训练完成后，可以通过输入不同的数据生成图片，生成的图片不一定要与区分的图片相同，而是某些特征相同，MMD 则需要对齐两个分布，尤其是 LMMD 分布，对齐两个分布相同标签的域特征

Our code will be available at

<https://github.com/easezyc/deep-transfer-learning>.

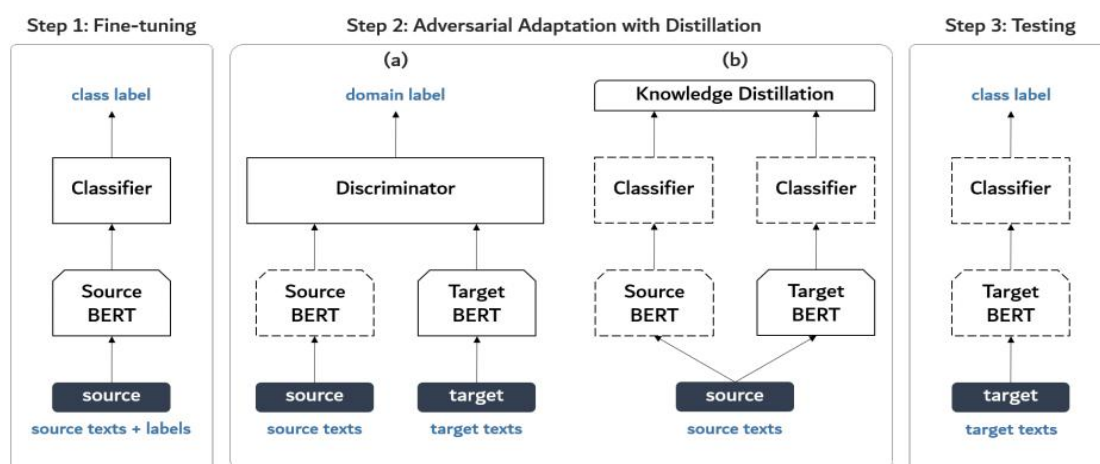


Left: global domain adaptation might lose some fine-grained information. Right: relevant subdomain adaptation can exploit the local affinity to capture the fine-grained information for each category. (更加关注子域的对齐，及相同类别和不同类别的域分布，类似于 Adaptive Device-free Localization in Dynamic Environments through Adaptive Neural Networks 中 Semantic Alignment (SA) 方法) 算法参考论文 p3, p4



Knowledge Distillation for BERT Unsupervised Domain Adaptation

we propose a simple but effective unsupervised domain adaptation method, adversarial adaptation with distillation (AAD), which combines the adversarial discriminative domain adaptation (ADDA) framework with knowledge distillation. We evaluate our approach in the task of cross-domain sentiment classification on 30 domain pairs (情感分类)



$$\mathcal{L}_{KD}(\mathbf{X}_S) = t^2 \times \mathbb{E}_{\mathbf{x}_s \sim \mathbb{X}_S} \sum_{k=1}^K -\text{softmax}(\mathbf{z}_k^S/t) \times \log(\text{softmax}(\mathbf{z}_k^T/t))$$

源码: <https://github.com/bzantium/bert-AAD>

主要方法: distillation

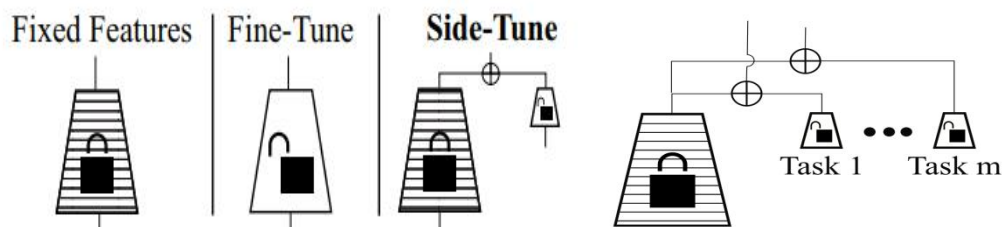
Side-Tuning: A Baseline for Network Adaptation via Additive Side Networks

The goal of side-tuning (and generally network adaptation) is to capitalize on a pretrained model to better learn one or more novel tasks. The side-tuning approach is straightforward: it assumes access to a given (base) model $B : X \rightarrow Y$ that maps the input x onto some representation y . Side-tuning then learns a side model $S : X \rightarrow Y$, so that the curated representations for the target task are

$$R(x) = B(x) \oplus S(x)$$

for some combining operation \oplus . For example, choosing $B(x) \oplus S(x) = \alpha B(x) + (1 - \alpha)S(x)$ (commonly called α -blending) reduces the side-tuning approach to: fine-tuning, feature extraction, and stage-wise training, depending on α (Fig. 2, right). Hence those can be viewed as special cases of the side-tuning approach (Figure 1).

Base Model. The base model $B(x)$ provides some core cognition or perception, and we put no restrictions on how $B(x)$ is computed. We never update $B(x)$, and in our approach it has zero learnable parameters.



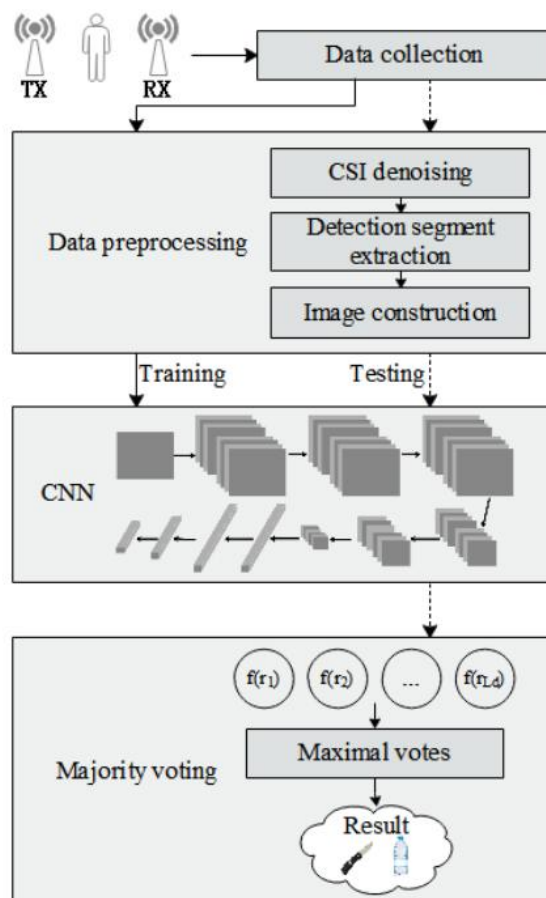
将现有的 fine-tune 机制进行扩展

代码: <http://sidetuning.berkeley.edu>

$$L(x_t, y_t) = \|D_t(\alpha_t B(x_t) + (1 - \alpha_t)S_t(x_t)) - y_t\|$$

Detection of Suspicious Objects Concealed by Walking Pedestrians Using WiFi

Abstract—Security is of vital importance in public places. Detection of suspicious objects such as metal and liquid often requires dedicated and expensive equipment, preventing its wide deployment. This paper proposes a pervasive device-free method to detect suspicious objects concealed by walking pedestrians using WiFi Channel State Information (CSI). By analyzing the different variations of subcarrier amplitude caused by different materials, the proposed method is able to detect suspicious objects such as metal and liquid concealed by pedestrians, when they walk through the transmission link of the WiFi transmitter and receiver. The proposed method employs Convolutional Neural Network (CNN) to classify suspicious objects, on which majority voting is applied to vote for the final result, in order to improve the detection accuracy for walking pedestrians. Evaluations show that the proposed method with majority voting achieve the detection accuracy of 93.3% for metal and liquid concealed by walking pedestrians, 95.6% for exposed metal and liquid carried by walking pedestrians, and 100% for metal and liquid carried by standing pedestrians.



Wi-Metal: Detecting Metal by Using Wireless Networks 2016 IEEE
International Conference on Communications (ICC)

介绍：利用 CSI 数据进行建模实现对不同金属物体的检测，适用于在公共场所的危险物品检测通道。

主要技术：K-Means

Euclidean metric formula (Euclidean distance) as the similarity measure formula and use the sum of squared error formula (SSE) as a measure of clustering quality formula

Euclidean distance: 利用欧氏距离公式计算所有数据到聚类中心的距离

SSE: 计算的是拟合 数据 和原始数据对应点的误差的平方和

Euclidean distance Formula:

$$L_2[(x_1, \dots, x_n), (y_1, \dots, y_n)] = \sqrt{\sum_{i=1}^n |x_i - y_i|^2}$$

Sum of the Squared Error Formula:

$$SSE = \sum_{i=1}^k (x_i - \hat{x}_i)^2$$

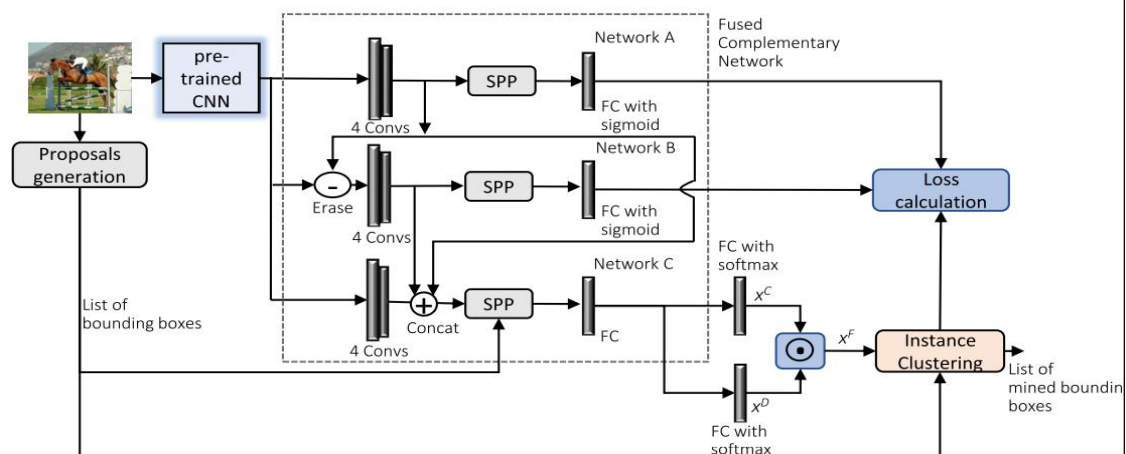
Weakly Supervised Object Detection Using Complementary Learning and Instance Clustering

IEEE Access

Year: 2020 | Volume: 8 | Journal Article | Publisher: IEEE

介绍：监督对象检测方案使用完全注释的训练数据，这是相当昂贵的。而弱监督对象检测（WSOD）仅使用图像级别的注释进行训练，这种注释更容易获取。WSOD是一项具有挑战性的任务，因为它旨在学习使用图像级标签进行对象定位和检测。根据这一主张，在本文中，我们提出了基于判别特征学习的 WSOD 端到端框架。我们使用客观技术从图像中获取初始建议。然后，并行训练两个互补网络以获得判别式图像特征，这些特征图像将与第三网络的特征在通道上进行级联。我们将此专为区分特征学习而设计的分类网络称为**融合互补网络**。该网络学习通过补充特征将整个对象实例包围起来的建议，这些特征最终将学会预测整个对象比仅包含对象部分的建议具有更高的概率。然后，对区域提议进行分层聚类。我们的聚类方法称为实例聚类，首先执行类间聚类，然后使用“交集-联合”度量执行迭代类内聚类，以获得与每个对象实例相对应的空间相邻聚类成员。在每个类内群集迭代中，将高分建议设置为每个类内群集的质心。在 PASCAL VOC2007 和 PASCAL VOC2012 数据集上进行了实验，定性和定量结果均显示在这些基准上 WSOD 性能得到了改善。

互补学习&实例聚类



spp（金字塔池化）的原理：

spp 的做法是使用多个不同大小的 sliding window pooling 对卷积输出的 feature map 进行池化，然后将这多个结果进行 concat 合并得到固定长度的输出。

XC: 计算每个区域分类概率

XD: 建议待检测区域，利用交叉熵损失

通过分别从分类和检测分支中获取两个得分矩阵 x^C 和 x^D 的 Hadamard 乘积（元素乘积）来计算最终得分矩阵（ x^F ）

通过以特定于类的方式对区域进行最大池化来从分类分支计算图像级分类分数。然后将**聚类**应用于提案列表，以获取最终的边界框。

整个网络的损耗函数定义如下：

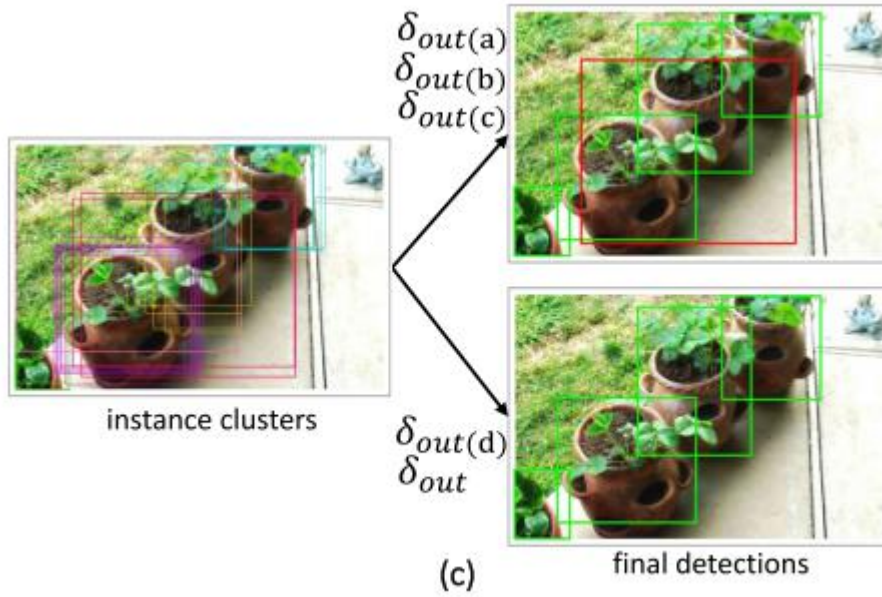
$$L = L_A + L_B + L_C \quad (2)$$

$$L_A = L_B = - \sum_{j=1}^C (y_{ij} \cdot \log(p_{ij}) + (1 - y_{ij}) \cdot \log(p_{ij})) \quad (3)$$

$$L_C = - \sum_{j=1}^C (y_{ij} \cdot \log(p_{ij}) + (1 - y_{ij}) \cdot \log(p_{ij})) - \sum_{s=1}^S (p_s \cdot \log(p_s)) \quad (4)$$

其中，L 是所建议的 WSOD 网络的损耗函数，L_A，L_B 和 L_C 分别是网络 A，B 和 C 的损耗函数。S 是概率分布中离散状态的数量（s 是单个状态）

检测结果：



定性结果（颜色相同的建议表示实例簇）：通过 IC 推断出的中间和最终区域具有不同的异常阈值。在最终检测中，绿色边界框表示真实检测，红色边界框表示错误检测。这些检测结果针对单个对象类别进行了演示。

Detecting Suspicious Objects With a Humanoid Robot Having a Metal Detector

介绍：保安机器人需要在机场，仓库，购物中心等场所巡逻，并找到并应对可疑人员。在这项研究中，我们旨在通过类人机器人实现巡逻任务。操作员使用具有金属探测器的人形机器人进行遥控操作，并将探测器摆动到与犯罪嫌疑人衣服表面短距离的位置，并检查衣服下面是否有隐藏的金属物体。为了开发该人体扫描系统，实验确定了取决于摆动速度的金属探测器的测量范围。基于此知识，在 3 种条件下进行了金属物体检测实验：人类，带有/不带有平衡控制的遥控人形机器人。结果表明，仅通过简单的右手的远程控制就不能充分地执行使用金属探测器的类人机器人代替身体搜索工作。因此，有必要集成视觉系统以掌握工作中机器人手的状况以及扩大工作范围的全身配合动作和脚步动作。

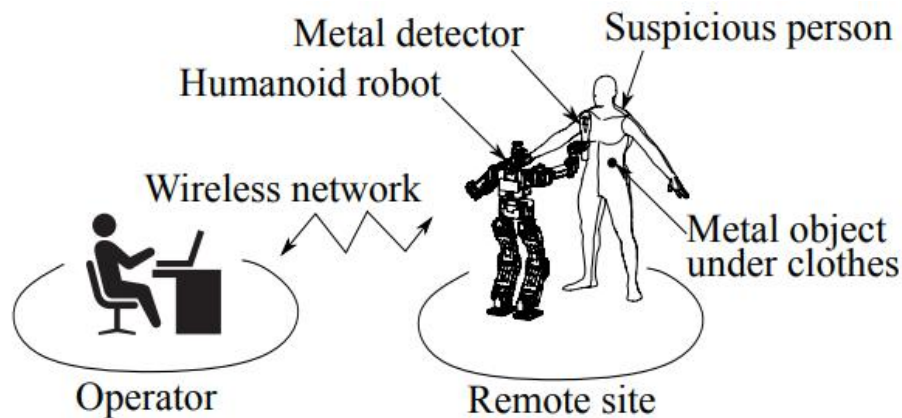


Fig. 1: An overview of body search with a teleoperated humanoid robot.

金属探测器 GC-101H(Doradus Corp.) is used to detect suspicious objects (电流电阻) (动量, 角度)

Image-to-Image Translation with Conditional Adversarial Networks

介绍：图像处理、图形学和视觉中的许多问题都涉及到将输入图像转换为相应的输出图像。这些问题通常使用特定于应用程序的算法来处理，尽管设置总是相同的：将像素映射到像素。条件对抗性网(cGAN)是一种通用的解决方案，它似乎能很好地解决各种各样的此类问题。本文介绍基于 cGAN 的 pix2pix 模型，针对不同的图片生成任务进行测试。同时介绍 cycleGAN 技术实现无标签的图片生成任务。

条件 GAN 的目标函数, 一般的 cGAN 的目标函数如下, 生成器 G 不断的尝试 minimize 下面的目标函数, 而 D 则通过不断的迭代去 maximize 这个目标函数, 即 $G^* = \arg \min_G \max_D \mathcal{L}_{cGAN}(G, D)$:

$$\mathcal{L}_{cGAN}(G, D) = \mathbb{E}_{x,y}[\log D(x, y)] + \mathbb{E}_{x,z}[\log(1 - D(x, G(x, z)))],$$

同时也比较一个无条件变量, 其中判别器不观察 x :

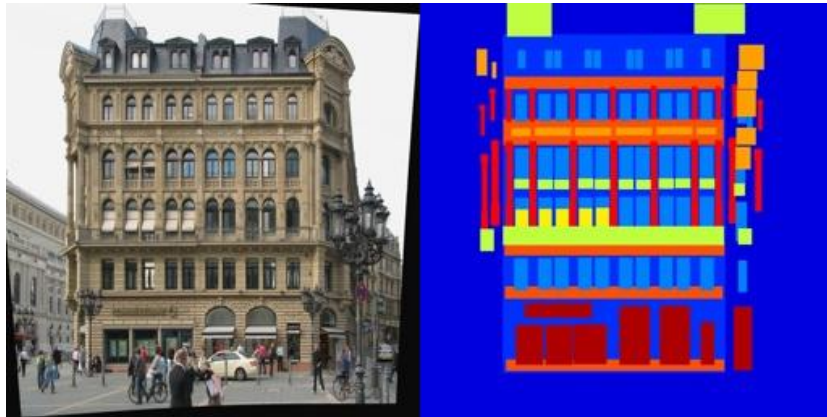
$$\mathcal{L}_{GAN}(G, D) = \mathbb{E}_y[\log D(y)] + \mathbb{E}_{x,z}[\log(1 - D(G(x, z)))].$$

将 GAN 目标与更传统的损失混合是有益的。鉴别器的工作保持不变, 但生成器的任务不仅欺骗鉴别器, 而且在 L2 意义上接近地面真值输出。我在本测试实验中选择使用 L1 距离而不是 L2, 因为 L1 可以减少模糊:

$$\mathcal{L}_{L1}(G) = \mathbb{E}_{x,y,z}[\|y - G(x, z)\|_1].$$

最终目标函数为:

$$G^* = \arg \min_G \max_D \mathcal{L}_{cGAN}(G, D) + \lambda \mathcal{L}_{L1}(G).$$



A Survey of Clustering With Deep Learning: From the Perspective of Network Architecture

介绍：集群是许多数据驱动应用程序领域中的一个基本问题，集群的性能在很大程度上取决于数据表示的质量。因此，线性或非线性特征变换被广泛用于学习更好的聚类数据表示。近年来，大量的研究都集中在利用深度神经网络学习一种对聚类友好的表示，聚类性能得到了显著的提高。在本文中，我们从建筑学的角度对深度学习聚类进行了系统的研究。具体来说，为了更好的理解这一领域，我们首先介绍了初步的知识。在此基础上，提出了一种基于深度学习的聚类方法，并介绍了几种具有代表性的聚类方法。最后，我们提出了一些有趣的机会聚类与深度学习，并给出了一些结论。

Comparison of algorithms based on network architecture and loss function.

Categories	Algorithms	Network Architecture	Network loss	Clustering loss	
				Principal	Auxiliary
AE	DCN	AE	reconstruction loss	k-means loss	N
	DEN	AE	reconstruction loss	N	1) locality-preserving constraint 2) group sparsity constraint
	DSC-Nets	CAE	reconstruction loss	N	self-expressiveness term
	DMC	AE	reconstruction loss	proximity penalty term	locality-preserving loss
	DEPICT	CAE (Denoising)	reconstruction loss	unsupervised cross entropy loss	N
	DCC	AE/CAE	reconstruction loss	robust continuous clustering loss	N
CDNN	DNC	RBM	N	nonparametric maximum margin clustering loss	N
	DEC	FCN	N	cluster assignment hardening loss	N
	DBC	CNN	N	cluster assignment hardening loss	N
	CCNN	CNN	N	k-means	N
	IMSAT	FCN	N	1) regularized information maximization, 2) self-augmented training loss	N
	JULE	CNN	N	agglomerative clustering	N
	DAC ¹	CNN	N	pairwise-classification loss	N
VAE	VaDE	VAE	variational lower bound on the marginal likelihood, with a GMM priori		
	GMVAE	VAE	variational lower bound on the marginal likelihood, with a GMM priori		
GAN	DAC ²	Adversarial autoencoder	reconstruction loss	1) GMM likelihood, 2) adversarial objective	N
	CatGAN	GAN	adversarial objective with a multi-classes priori		
	InfoGAN	GAN	adversarial objective with a multi-classes priori		

¹ Deep Adaptive Clustering

² Deep Adversarial Clustering

Main contributions of the representative algorithms.

Categories	Algorithms	Main contributions to clustering
AE	DCN	perform k-means clustering and feature learning simultaneously, simple but effective
	DEN	learn a clustering-friendly representation
	DSC-Nets	improve the classical subspace clustering by AE
	DMC	improve the classical multi-manifold clustering by AE
	DEPICT	computational efficient, robust, perform well on image datasets
	DCC	avoid alternative optimization, require no prior knowledge of cluster number
CDNN	DNC	improve the classical NMMC clustering by DBN
	DEC	the first well-known deep clustering method, making this field popular
	DBC	improve DEC using CNN
	CCNN	computational efficient, deal with large-scale image datasets
	IMSAT	introduce self-augment training to deep clustering
	JULE	perform well on image datasets, but have high computational and memory cost
VAE	DAC	well-designed clustering loss, achieve the-state-of-art performance on several datasets
GAN	VaDE	combine VAE with clustering
	GMVAE	combine VAE with clustering
GAN	DAC	combine AAE with clustering
	CatGAN	combine GAN with clustering
GAN	InfoGAN	learn disentangled representations

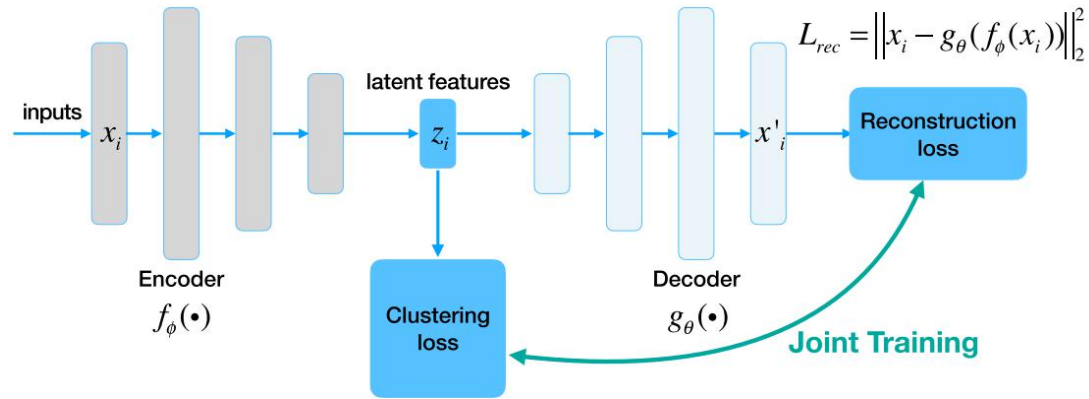


FIGURE 1. Architecture of clustering based on autoencoder. The network is trained by both clustering loss and reconstruction loss.

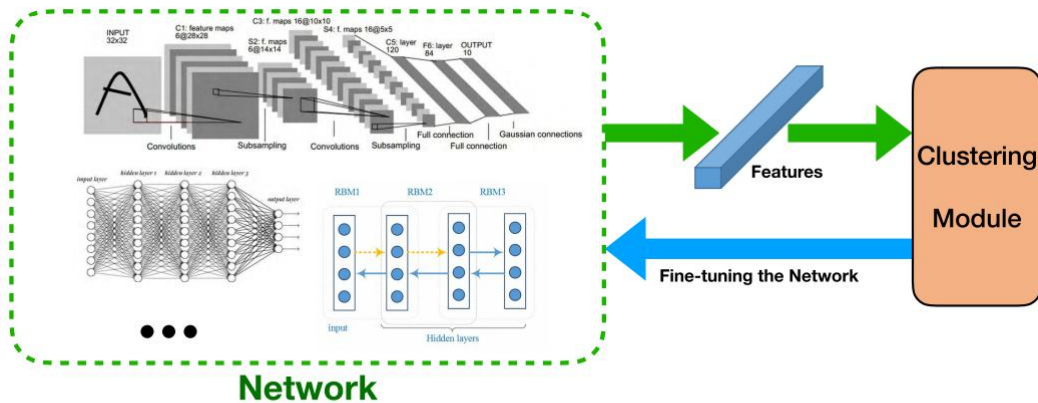


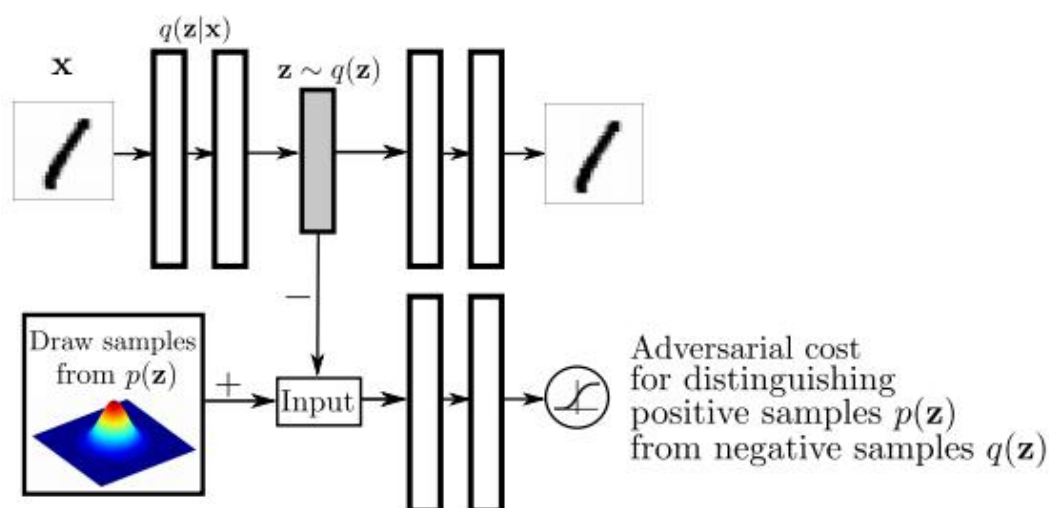
FIGURE 2. Architecture of CDNN-based deep clustering algorithms. The network is only adjusted by the clustering loss. The network architecture can be FCN, CNN, DBN and so on.

Adversarial Autoencoders

A. Makhzani, J. Shlens, N. Jaitly, I. Goodfellow, and B. Frey. (2015).
‘ ‘Adversarial autoencoders.’ ’ [Online]

该文提出“对抗式自动编码器”(AAE)，这是一种概率式的自动编码器，它利用新提出的生成式对抗网络(GAN)，以任意先验分布匹配自编码器隐藏码向量的后验集合来进行变分推理。将聚集的后验与先验进行匹配，可确保从先验空间的任何部分生成有意义的样本。因此，对抗性自动编码器的解码器学习了一个深度生成模型，该模型在数据分布之前映射施加的数据。我们展示了对抗式自动编码器如何在半监督分类、图像解缠风格和内容、无监督聚类、降维和数据可视化等应用中使用。我们在 MNIST、街景房屋号码和多伦多人脸数据集上进行了实验，结果表明对抗自动编码器在生成建模和半监督分类任务中取得了具有竞争力的结果。

目的：1 使隐藏向量更具有区分性. 2 使隐藏向量能生成更加完整分布



$$\min_G \max_D \mathbb{E}_{\mathbf{x} \sim p_{\text{data}}} [\log D(\mathbf{x})] + \mathbb{E}_{\mathbf{z} \sim p(\mathbf{z})} [\log(1 - D(G(\mathbf{z})))]$$

生成器 G 的编码器和下方判别器 D 组成的对抗网络，此处 G 和 D 联系起来的不再是图片数据，而是一个一维向量 z，判别器 D 通过不断学习，预测输入的 z 来自于 real data（服从 q(z) 概率分布）还是 fake data（服从预定义的 p(z) 概率分布）。由于这里的 p(z) 可以是任何我们可以生成的一个概率分布，因此整个对抗学习过程实际上可以认为是通过调整 encoder 不断让其产生数据的概率分布 q(z) 接近我们预定义的 p(z)，当模型训练完成后，由于 p(z) 与 q(z) 十分相近，因此可以直接通过 p(z) 产生我们需要的随机 latent variable，然后借助于解码器产生一个新的图像数据。Adversarial Latent Autoencoders

我们设计了两个自动编码器：一个基于 MLP 编码器，另一个基于 StyleGAN 生成器，我们称之为 StyleALAE。我们验证了这两个架构的解纠缠特性。结果表明，StyleALAE 不仅可以生成与 StyleGAN 相当质量的 1024×1024 人脸图像，而且在相同分辨率下，还可以生成基于真实图像的人脸重建和处理。这使得 ALAE 成为第一个能够与之相比的自动编码器，并且超越了仅产生器类型的架构的能力。

生成图片能力与 StyleGAN 相当

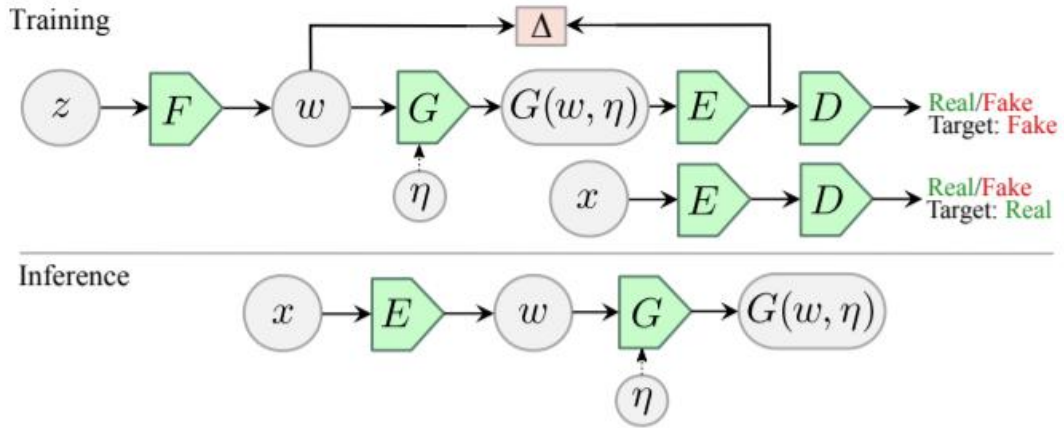


Figure 1: ALAE Architecture. Architecture of an Adversarial Latent Autoencoder.

把原生 GAN 中的 G 分解为 F 与 G 的映射， D 分解为 E 与 D 的映射：

$$G = G \circ F, \quad \text{and} \quad D = D \circ E$$

F 是一个确定性的映射，将噪声 z 编码成隐变量 w 。 E 和 G 是随机的， G 同时取决于隐变量 w 和噪声的输入。 E 将生成的图像进行编码，然后约束由 F 生成的分布与由 E 生成的分布尽可能详尽。这样给定 w 就可以生成图像，给定图像就可以编码 w 。在推理时就可以实现重构。可以看到公式 7 约束的就是隐变量空间 w 的相似度，而非约束原生 AE 中的图像（数据空间）相似度。

$$\min_{F,G} \max_{E,D} V(G \circ F, D \circ E) \quad (6)$$

$$\min_{E,G} \Delta(F \| E \circ G \circ F) \quad (7)$$

上式即为目标函数。相比 BiGAN 重构效果不太受到 label flip 的影响：

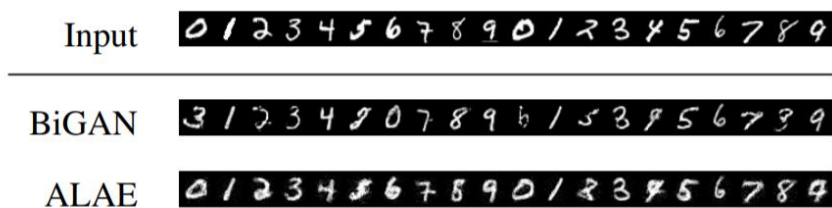


Figure 3: **MNIST reconstruction.** Reconstructions of the permutation-invariant MNIST. Top row: real images. Middle row: BiGAN reconstructions. Bottom row: ALAE reconstructions. The same MLP architecture is used in both methods.

对比在 Z 空间插值和直接在 W 空间插值的结果：后者更平滑，较为分离。



Figure 4: **MNIST traversal.** Reconstructions of the interpolations in the Z space, and the W space, between the same digits. The latter transition appears to be smoother.

基于 StyleGan 的结构：

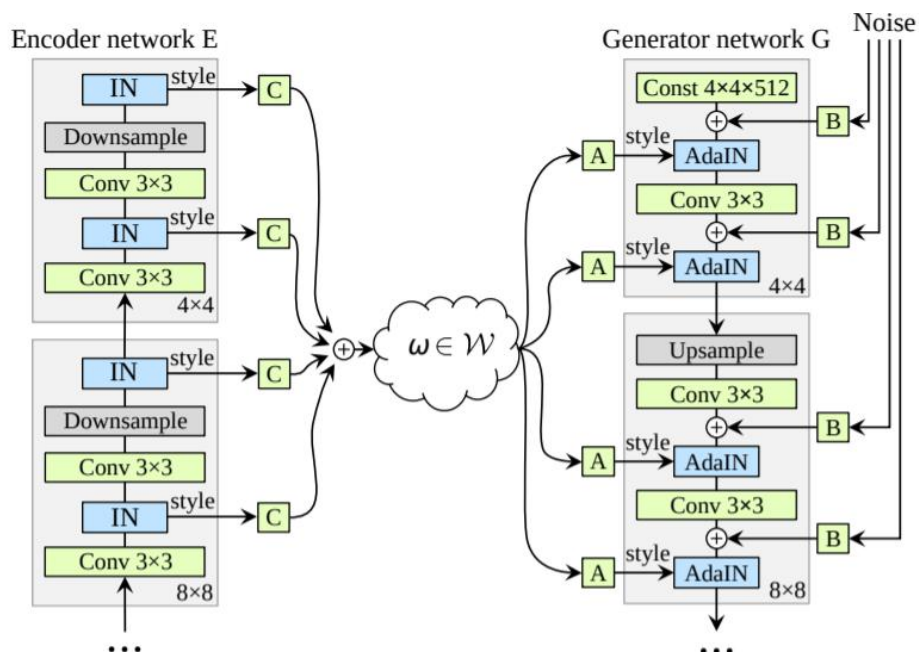


Figure 2: **StyleALAE Architecture.** The StyleALAE encoder has Instance Normalization (IN) layers to extract multiscale style information that is combined into a latent code w via a learnable multilinear map.

各级风格特征（均值方差）经过一个线性层来聚合后放到 GAN 里，此时的 E 就可以编码风格，重建效果：

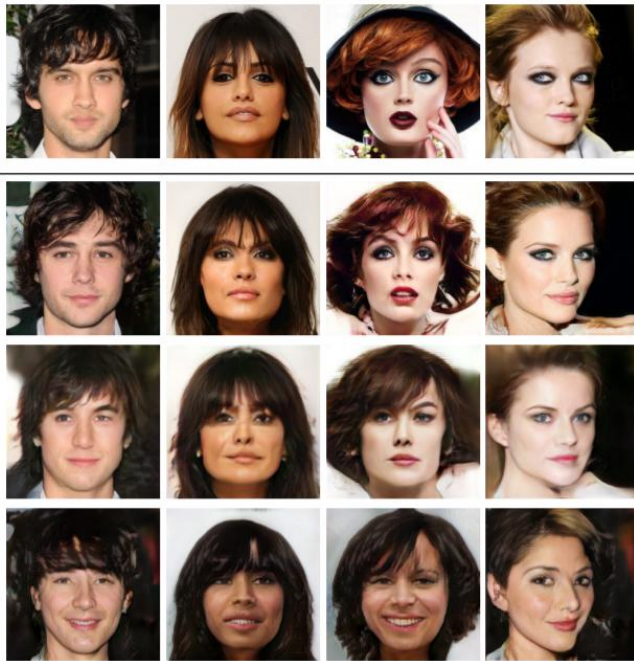


Figure 8: **CelebA-HQ reconstructions.** CelebA-HQ reconstructions of unseen samples at resolution 256×256 . Top row: real images. Second row: StyleALAE. Third row: Balanced PIONEER [17]. Last row: PIONEER [16]. StyleALAE reconstructions look sharper and less distorted.

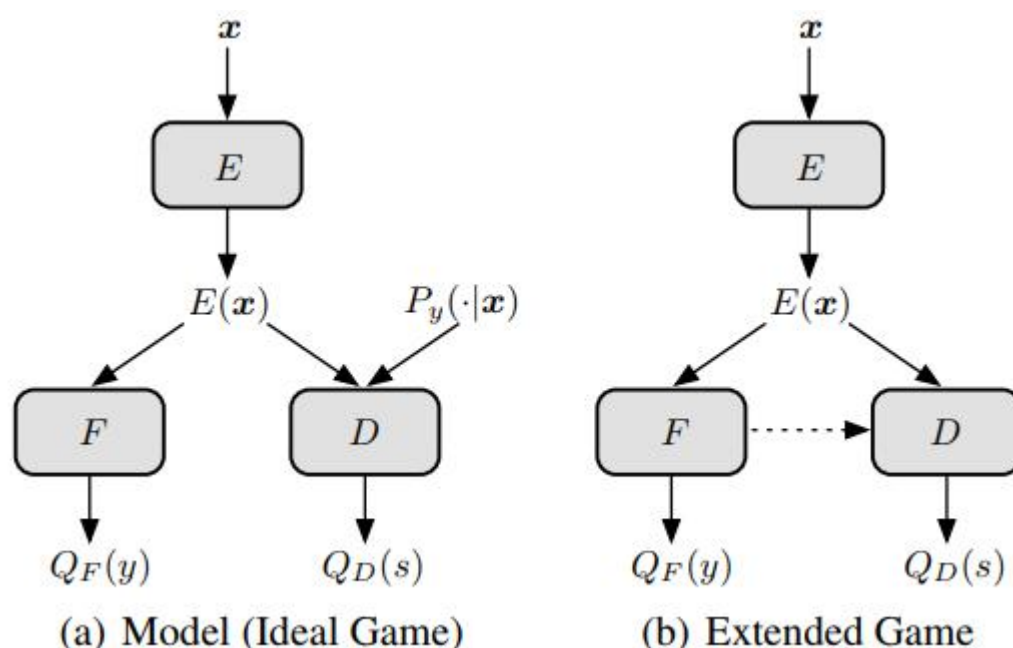
Learning Sleep Stages from Radio Signals:A Conditional Adversarial Architecture ICML 2017

我们专注于通过无线电测量来预测睡眠阶段，而不需要在受试者身上附加任何传感器。我们引入了一种新的预测模型，结合卷积和循环神经网络从 RF 信号中提取睡眠特定的主观不变特征，并捕获睡眠的时间进展。我们方法的一个关键创新是改进的对抗训练机制，它丢弃了特定于个体或测量条件的无关信息，同时保留了与预测任务相关的所有信息。我们分析了我们的博弈论设置，并通过经验证明，我们的模型比最先进的解决方案取得了显著的改进

本文提出了一种新的模型，该模型具有更高的精度，并能很好地推广到新的环境和主题。该模型采用卷积神经网络(CNN)从射频频谱图中提取特定阶段的特征，并与递归神经网络(RNN)耦合，以捕获睡眠阶段的时间动态。

然而，单独的 CNN-RNN 组合仍然容易分散与特定个体或测量条件(例如，源域)相关的特征，因此不能很好地泛化。为了解决这个问题，我们引入了一种新的对抗训练制度，抛弃特定于个体或测量条件的无关信息，同时保留与预测任务相关的所有信息——即。时，对手确保了学习表示和源域之间的条件独立性。

让 $x \in \Omega_x$ 是一个输入样本，和 $y \in \{1, 2, \dots, n_y\}$ 输出标签。设 $s \in \{1, 2, \dots, n_s\}$ 表示一个辅助标签，它指的是特定输入样本的来源。我们定义 $x = [x_1, x_2, \dots, x_t] \in \Omega_x$ 的序列输入样本时间到当前时间 t 的开始。在我们的应用程序的上下文中，上面的符号转化为以下：输入样本 x 是一个 30 秒的射频声谱图，和输出标签 y 是一个睡眠阶段，4 个值之一：醒了，光睡觉，深度睡眠，或快速眼动。向量 x 是指射频频谱图的顺序从晚上开始到当前时间。由于射频信号携带关于主体和测量环境的信息，我们给每个输入 x 分配一个辅助标签，它标识受试者-环境对，以下称为源。



我们的模型如图(a)所示。它有三个组件：一个编码器 E ，标签预测器 F ，和一个源鉴别器 d 。模型设置为一个游戏，其中表示编码器扮演一个合作博弈的标签预测

允许它来预测正确的标签使用编码表示。编码器也会对源判别函数进行极大极小博弈

编码器 E: 编码器 $E(\cdot): \Omega_x \rightarrow \Omega_z$ 是一个函数, 它接受一个输入样本序列 x , 并返回一个向量的摘要 $z = E(x)$ 。标签预测器的损失 F , 给定编码器 E , 定义为标签分布 $P_Y(\cdot | x)$ 和 $Q_F(\cdot | E(x))$ 之间的交叉熵:

$$\mathcal{L}_f(F; E) = \mathbb{E}_{\mathbf{x}, y}[-\log Q_F(y|E(\mathbf{x}))]$$

在训练过程中, 编码器 E 和预测器 F 进行合作博弈, 以最小化标签预测损失。

源鉴别器 D : 我们定义一个源鉴别器 $D(\cdot, \cdot): \Omega_z \times [0, 1] \rightarrow [0, 1]$ 。它以潜在表示 $E(x)$ 和标签分布 $P_Y(\cdot | x)$ 作为输入, 并预测它们从哪个源域(即主题和环境)采样为 $Q_D(\cdot | E(x), P_Y(\cdot | x))$ 。

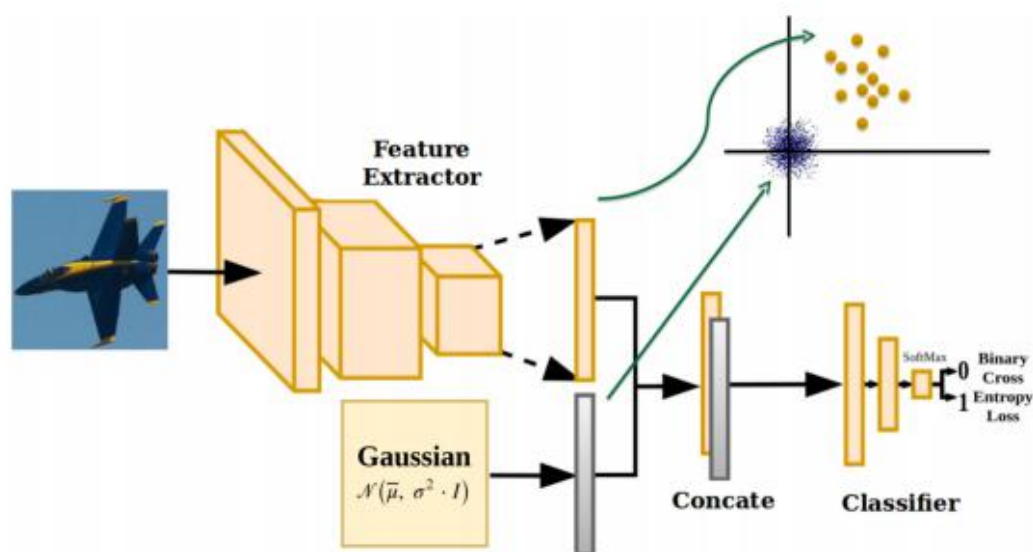
接下来, 我们定义所需的表示不变性。

定义 1(表示不变性)。如果 $E(x)$ 除了 $P_Y(\cdot | x)$ 中已经包含的 s 之外, 不包含 s 的信息, 则表示 E 是不变的; 即 $Q_D(\cdot | E(x), P_Y(\cdot | x)) = Q_D(\cdot | P_Y(\cdot | x))$ 为最优 D 。

在训练过程中, 编码器 E 和鉴别器 D 进行了极大极小博弈: 训练 D 使源预测损失最小, 训练编码器 E 使源预测损失最大, 以达到上述不变性。

One-Class Convolutional Neural Network -CoRR abs(2019)

我们提出了一种基于卷积神经网络 (CNN) 的单类分类方法。该思想是使用潜在空间中的零中心高斯噪声作为伪负类，并使用交叉熵损失训练网络，以学习一个好的表示和给定类的决策边界。该方法的一个重要特点是，任何预先训练好的 CNN 都可以作为单类分类的基础网络。本文提出的一类 CNN 是基于 UMDAA-02 Face、abnormity -1001 和 FounderType200 数据集进行评估的。这些数据集涉及到各种单一类的应用问题，如用户认证、异常检测和新颖性检测。大量的实验表明，该方法比目前最先进的方法有显著的改进。源代码可在 github.com/otkupjnoz/oc-cnn 获得。



Towards K-means-friendly Spaces: Simultaneous Deep Learning and Clustering-ICML 2017

大多数学习方法分别处理降维(DR)和聚类(即顺序),但最近的研究表明,联合优化这两个任务可以显著提高两者的性能。后者的前提是数据样本是通过容易聚类的潜在表示的线性变换获得的;但在实践中,从潜在空间到数据的转换可能会更加复杂。在这项工作中,我们假定这个变换是一个未知的,可能是非线性的函数。为了恢复“聚类友好型”潜在表示并更好地聚类数据,我们提出了一种DR和K-means联合聚类方法,其中DR是通过学习深度神经网络(DNN)来完成的。其动机是保持联合优化两个任务的优势,同时利用深度神经网络近似任何非线性函数的能力。这样一来,所提出的方法可以很好地应用于广泛的生成模型。为此,我们精心设计了DNN结构和相关的联合优化准则,并提出了一种有效的、可扩展的算法来处理制定的优化问题。利用不同的实际数据集进行实验,验证了该方法的有效性。

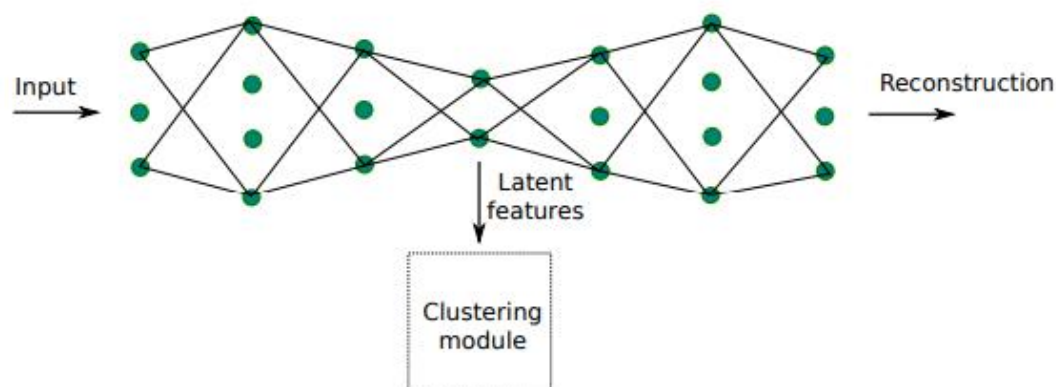


Figure 3. Proposed deep clustering network (DCN).

聚类损失:

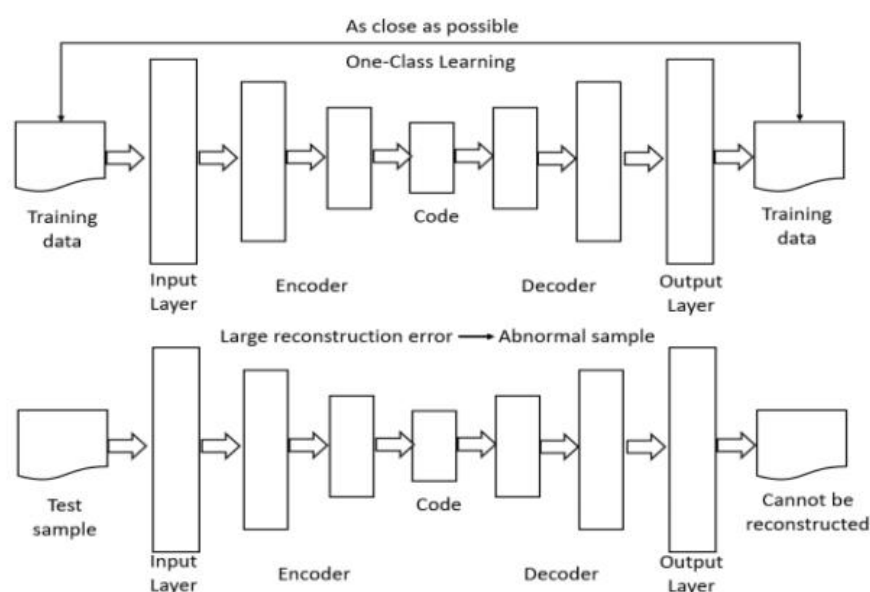
$$\min_{M \in \mathbb{R}^{M \times K}, \{s_i \in \mathbb{R}^K\}} \sum_{i=1}^N \|x_i - Ms_i\|_2^2 \quad (1)$$

$$\text{s.t. } s_{j,i} \in \{0, 1\}, \mathbf{1}^T s_i = 1 \quad \forall i, j,$$

One-class Classification with Deep Autoencoder Neural Networks for Author Verification in Internet Relay Chat-AICCSA 2019

在本文中，我们设计了一个自主的 IRC 监控系统，执行递归深度学习来分类消息的威胁级别，并开发了一种基于深度自编码神经网络的单类分类作者验证方法。实验结果表明，该方法能够成功地对 IRC 用户进行有效的作者验证。
方法：给一组明确的候选人分组，其他类型进行判别直接归为嫌疑人(单分类)

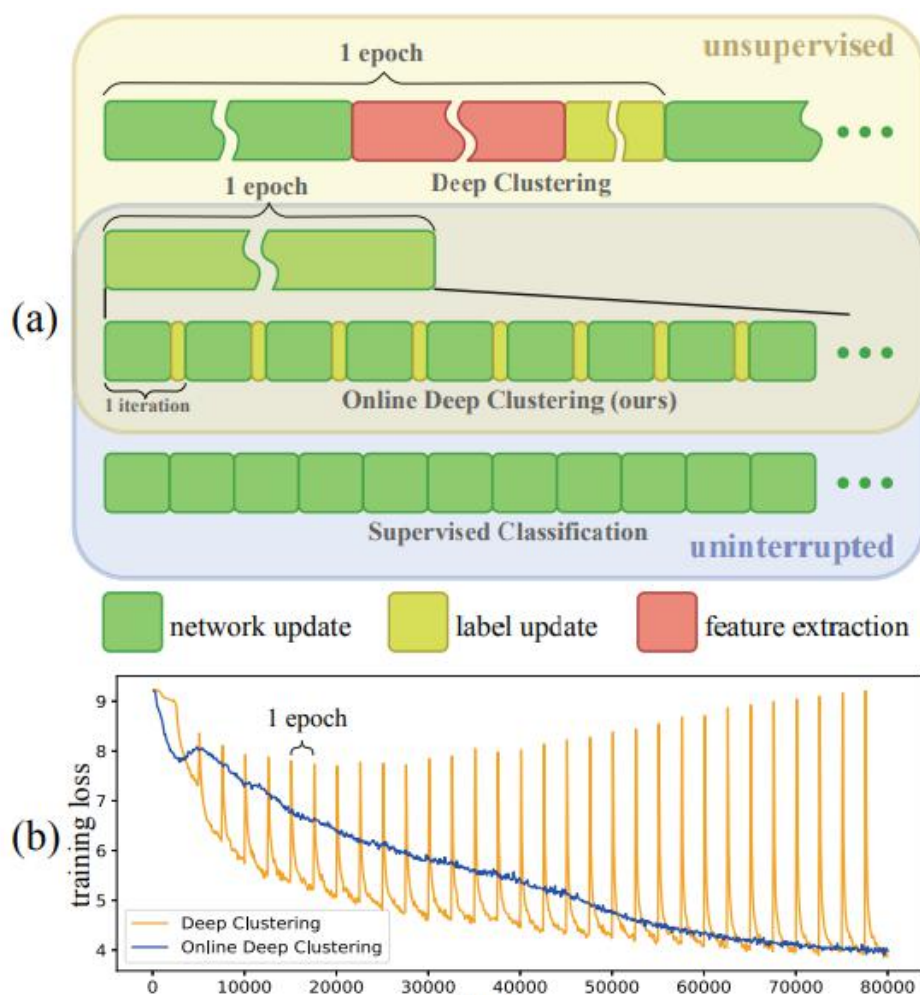
处理 IRC(网络中继聊天)消息，分为正常，警告和高等级



我们的对称形状的深度自动编码器的架构如图 5 所示。输入层有 52 个节点，因为我们提取了 52 个人格洞察特征。因此，由于目的是重构样本，输出层也有 52 个节点。在 autoencoder 中加入 5 个完全连接的层(节点分别为 30、20、10、20、30)作为隐藏层，形成深度自动编码器。中间隐含层，以 10 个节点为代码层，存储输入数据的压缩表示空间。我们的深度自动编码器使用 tanh 激活函数为每个隐藏层和身份函数为输出层。深度自动编码器学习参数使用亚当优化器与小批训练，以最小化均方误差。在完成训练阶段后，我们的深度自动编码器模型可以验证 IRC 消息的作者身份。IRC 消息样本是否属于用户取决于重构错误。在测试阶段，如果一个 IRC 消息样本是由同一作者编写的，那么它的重构错误很低。而另一个作者写的重构误差很大。

联合聚类 and 特征学习方法在无监督表示学习中表现出了显著的性能。然而，训练计划在特征聚类和网络参数更新之间交替，导致视觉表示学习不稳定。为了克服这一挑战，我们提出了在线深度聚类(ODC)，它可以同时进行聚类和网络更新，而不是交替进行。我们的关键见解是，在保持分类器稳定更新的过程中，簇中心应该稳定地演化。具体来说，我们设计并维护了两个动态记忆模块，即用于存储样本标签和特征的样本记忆模块和用于质心进化的质心记忆模块。我们将突然的全局聚类分解为稳定的内存更新和分批次的标签重新分配。这个过程被集成到网络更新迭代中。这样，标签和网络就并肩发展，而不是交替发展。大量的实验表明，ODC 可以稳定训练过程，有效地提高性能

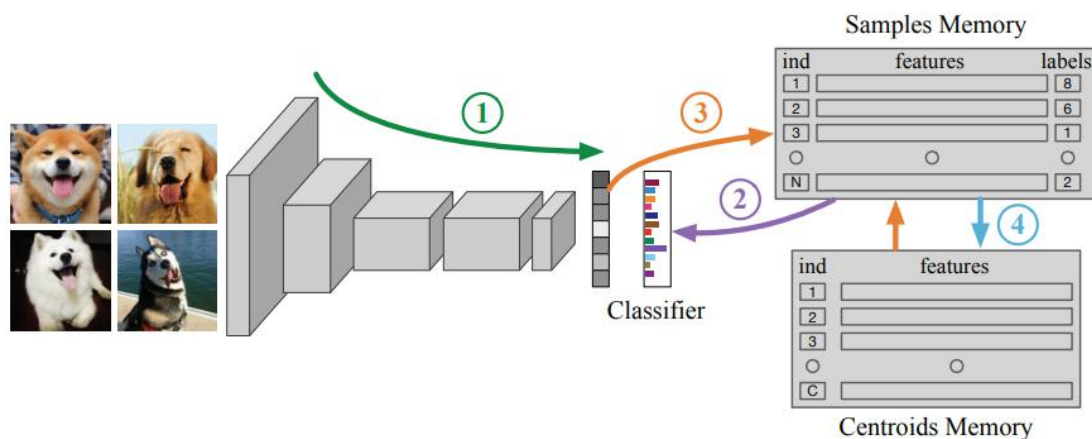
网络结构损失和聚类损失同时进行训练，提高训练稳定性



为了方便在线标签重分配和质心更新，我们设计并维护了两个动态记忆模块，即用于存储样本标签和特征的样本记忆模块和用于质心进化的质心记忆模块。通过

这种方式，ODC 以一种类似于监督分类的不间断方式进行训练，而不需要手工注释。在训练过程中，标签和网络参数并肩进化，而不是交替进化。由于标签在每次迭代中都是连续即时更新的，所以 CNN 中的分类器也进化得更加稳定，从而产生了如图 1 (b) 所示的更加稳定的损失曲线。

图 2。每次 ODC 迭代主要包括四个步骤:前向得到一个紧凑的特征向量;2. 从样本存储器中读取标签，进行反向传播，更新 CNN;3. 通过更新特征和分配新标签来更新样本内存;4. 通过重新计算涉及的质心来更新质心记忆。

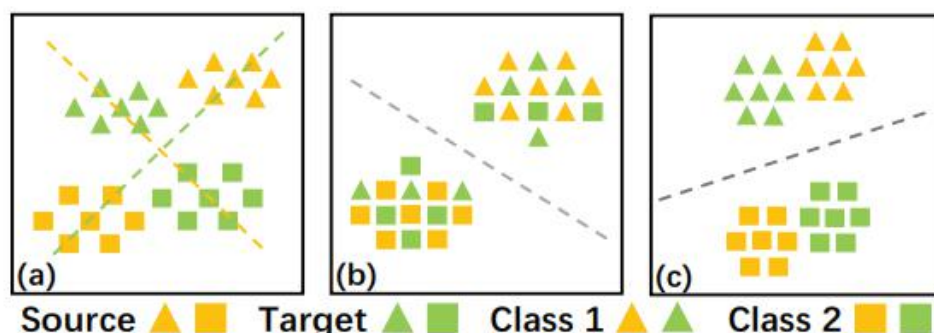


Unsupervised Domain Adaptation via Structurally Regularized Deep Clustering-CVPR2020

无监督域自适应(Unsupervised domain adaptive, UDA)是对目标域上的未标记数据进行预测。主流的 UDA 方法学习两个领域之间的对齐特征,这样根据源特征训练的分类器可以很容易地应用到目标特征上。然而,这种转移策略有可能破坏目标数据的内在判别性。为了减轻这一风险,我们以结构域相似性假设为动机,提出通过对目标数据进行判别聚类,直接揭示目标的内在判别。我们使用结构源正则化来约束聚类解决方案,这取决于我们假定的结构域相似性。技术上,我们采用了一种灵活的基于深度网络的判别聚类框架,使网络预测标签分布与引入的辅助标签分布之间的 KL 差异最小;用源数据地面真实标签形成的辅助分布代替辅助分布,通过联合网络训练的简单策略实现结构化的源正则化。我们将我们提出的方法称为结构正则化深度聚类(SRDC),其中我们还通过中间网络特征聚类增强了目标识别,通过软选择小分歧源实例增强了结构正则化。仔细消融研究表明我们建议的 SRDC 的疗效。值得注意的是,由于没有显式的域对齐,SRDC 在三个 UDA 基准测试上优于所有现有方法。

为了解决通过明确学习域对齐特征而破坏内在数据判别的潜在问题,我们在本研究中提出了一种源正则化的深度判别聚类方法,以直接揭示目标数据之间的内在判别。该方法的动机是我们假设这两个领域的结构相似,为此我们将提出的方法称为结构正则化深度聚类(SRDC)。•为了从技术上实现 SRDC,我们使用了一个灵活的深度聚类框架,该框架首先引入一个辅助分布,然后最小化引入的分布和网络的预测标签分布之间的 KL 差异;用源数据的地真标签代替辅助分布,通过联合网络训练的简单策略实现结构化的源正则化。在 SRDC 中,我们还设计了有用的成分,利用中间网络特征聚类来增强目标识别,并利用低发散源实例的软选择来增强结构正则化。•我们在 UDA 基准数据集上进行仔细的消融研究,以验证 SRDC 中提出的单个成分的疗效。值得注意的是,在没有显式域对齐的情况下,我们提出的 SRDC 在基准数据集上优于所有现有的方法

该框架首先引入一个辅助分布,然后最小化引入的分布和网络的预测标签分布之间的 KL 差异;用源数据的地真标签代替辅助分布



利用对抗和 mmd 方法进行源域和目标域的对齐
源数据和目标域数据满足两个假设:

1. 领域识别假设：在单个领域中存在固有的数据识别结构，即源域或目标域的数据根据共享标签空间进行区分聚类。

2. 类的紧密性假设：对应于同一个类标签的两个域的集群在几何上紧密。

损失函数：KL 散度，网络损失

Qt: 网络预测标签。Pt: 聚类辅助标签

$$\min_{Q^t, \{\theta, \vartheta\}} \mathcal{L}_{f \circ \varphi}^t = \text{KL}(Q^t || P^t) + \sum_{k=1}^K q_k^t \log q_k^t,$$

准确率：Office-31 based on ResNet-50

(1) 源模型，对标记源样本的基础网络进行微调；

(2) SRDC (w/o 结构源正则化)对源预训练模型进行精细调整，即不使用结构源正则化；

(3) SRDC (w/o 特征识别)，表示在特征空间中不进行源和目标识别的训练；

(4) SRDC (w/o 软源样本选择)，表示通过软源样本选择进行不增强训练。

Method	A → W	A → D	D → A	W → A	Avg
Source Model	77.8±0.2	82.1±0.2	64.5±0.2	66.1±0.2	72.6
SRDC (w/o structural source regularization)	87.3±0.0	92.1±0.1	73.9±0.1	75.0±0.1	82.1
SRDC (w/o feature discrimination)	94.2±0.4	94.3±0.4	74.3±0.2	75.5±0.4	84.6
SRDC (w/o soft source sample selection)	94.8±0.2	94.6±0.3	74.6±0.3	75.7±0.3	84.9
SRDC	95.7±0.2	95.8±0.2	76.7±0.3	77.1±0.1	86.3

Method	A → W	D → W	W → D	A → D	D → A	W → A	Avg
Source Model [21]	77.8±0.2	96.9±0.1	99.3±0.1	82.1±0.2	64.5±0.2	66.1±0.2	81.1
DAN [34]	81.3±0.3	97.2±0.0	99.8±0.0	83.1±0.2	66.3±0.0	66.3±0.1	82.3
DANN [16]	81.7±0.2	98.0±0.2	99.8±0.0	83.9±0.7	66.4±0.2	66.0±0.3	82.6
ADDA [53]	86.2±0.5	96.2±0.3	98.4±0.3	77.8±0.3	69.5±0.4	68.9±0.5	82.9
VADA [51]	86.5±0.5	98.2±0.4	99.7±0.2	86.7±0.4	70.1±0.4	70.5±0.4	85.4
SimNet [43]	88.6±0.5	98.2±0.2	99.7±0.2	85.3±0.3	73.4±0.8	71.8±0.6	86.2
MSTN [59]	91.3	98.9	100.0	90.4	72.7	65.6	86.5
GTA [49]	89.5±0.5	97.9±0.3	99.8±0.4	87.7±0.5	72.8±0.3	71.4±0.4	86.5
MCD [48]	88.6±0.2	98.5±0.1	100.0±0.0	92.2±0.2	69.5±0.1	69.7±0.3	86.5
SAFN+ENT [60]	90.1±0.8	98.6±0.2	99.8±0.0	90.7±0.5	73.0±0.2	70.2±0.3	87.1
DAAA [28]	86.8±0.2	99.3±0.1	100.0±0.0	88.8±0.4	74.3±0.2	73.9±0.2	87.2
iCAN [63]	92.5	98.8	100.0	90.1	72.1	69.9	87.2
CDAN+E [35]	94.1±0.1	98.6±0.1	100.0±0.0	92.9±0.2	71.0±0.3	69.3±0.3	87.7
MSTN+DSBN [5]	92.7	99.0	100.0	92.2	71.7	74.4	88.3
TADA [56]	94.3±0.3	98.7±0.1	99.8±0.2	91.6±0.3	72.9±0.2	73.0±0.3	88.4
TAT [33]	92.5±0.3	99.3±0.1	100.0±0.0	93.2±0.2	73.1±0.3	72.1±0.3	88.4
SymNets [68]	90.8±0.1	98.8±0.3	100.0±0.0	93.9±0.5	74.6±0.6	72.5±0.5	88.4
BSP+CDAN [9]	93.3±0.2	98.2±0.2	100.0±0.0	93.0±0.2	73.6±0.3	72.6±0.3	88.5
MDD [66]	94.5±0.3	98.4±0.1	100.0±0.0	93.5±0.2	74.6±0.3	72.2±0.1	88.9
CAN [27]	94.5±0.3	99.1±0.2	99.8±0.2	95.0±0.3	78.0±0.3	77.0±0.3	90.6
SRDC	95.7±0.2	99.2±0.1	100.0±0.0	95.8±0.2	76.7±0.3	77.1±0.1	90.8

Table 3. Results (%) on Office-31 (ResNet-50).

Feature Sampling Based Unsupervised Semantic Clustering for Real Web Multi-View Content-AAAI2019

真正的 web 数据集通常与多个视图相关联,比如长评论和短评论、用户偏好等等。然而,随着用户生成文本的快速增长,数据集的每个视图都有很大的特征空间,这给矩阵分解过程带来了计算上的挑战。在本文中,我们提出了一种基于非负矩阵分解的多视角聚类算法,该算法试图利用特征采样策略来降低迭代过程中的复杂度。特别地,我们的方法在学习过程中利用无监督语义信息,通过图正则化来捕获内在相似性。此外,我们使用 Hilbert Schmidt 独立准则(HSIC)来探索一个网页项目的多视图内容之间的无监督语义多样性信息。总体目标是最小化多视图非负矩阵分解的损失函数,该分解结合了一个语义内相似图正则化器和一个语义间多样性术语。与一些最先进的方法相比,我们在一个大的真实数据集 Doucom 和其他三个较小的数据集上证明了我们提出的方法的有效性。

辅助空间

大规模的多视图 web 内容通常有以下属性:1)一个 web item 的每个视图都有自己的特性集。2)同一条目的不同观点在语义信息上有一定的一致性。就像一个新的故事可能会被不同的新闻资源报道,但其背后的内容不会改变。3)一个 web 项目表示之间的多样性信息存在于多视图数据中,每个表示对应于一个单一的视图内容。以我们的 Doucom 数据集为例,web 社区中的一个电影条目包含多个视图,包括摘要、短评论、长评论和用户组。而电影项目的每次查看主要包含相关内容中的具体信息。

本文提出了一种针对真实 web 多视图内容的多视图无监督语义聚类方法——特征采样无监督语义聚类(F SUSC)。我们使用图正则化来捕获内在的语义相似性信息。我们也使用 Hilbert Schmidt 独立准则(HSIC) (Arthur et al. 2005)作为协正则项来加强共同学习表示的语义多样性。虽然现有的一些方法(Wang et al. 2015;Zhu 等人, 2013;郑等, 2011;杨 et al . 2014 年)图像正则化适用于各种应用程序,他们只观察原始数据信息,激发我们探索语义结构信息和识别的一种常见的低维空间中的数据实例在多个视图我们的工作,从而减少内存消耗如果数据指的是一个大数字。我们将我们的多视图聚类规划为一个联合优化问题,以最小化在多个视图上的重构误差。本文的贡献可以总结为以下几点:

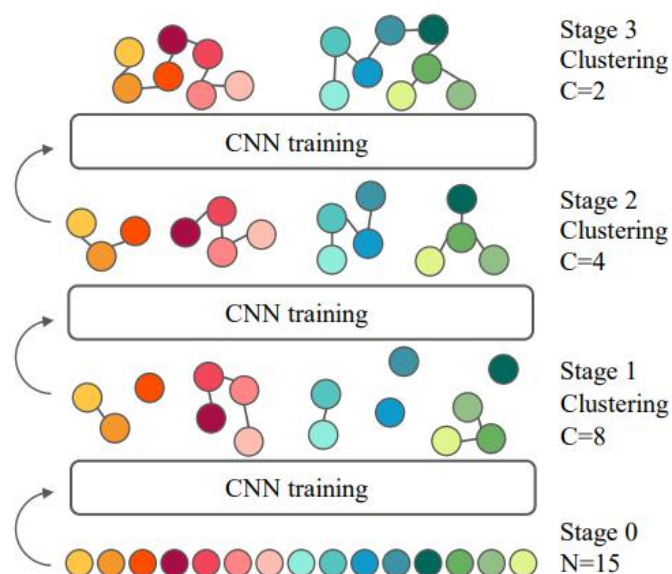
A Bottom-Up Clustering Approach to Unsupervised Person Re-Identification-AAAI2019 未发 csdn

大多数人再识别(re-ID)方法都基于监督学习,这需要对训练数据进行密集的人工标注。然而,获取身份注释不仅是资源密集型的,而且对大规模的真实世界数据进行标记也是不切实际的。为了解决这个问题,我们提出了一种自底向上聚类(BUC)的方法来联合优化卷积神经网络(CNN)和个体样本之间的关系。我们的算法考虑了在重新识别任务中的两个基本事实,即不同身份之间的多样性和同一身份内部的相似性。具体地说,我们的算法从将单个样本作为一个不同的身份开始,这将最大化每个身份的多样性。然后它逐渐将相似的样本分组成一个身份,这增加了每个身份内部的相似性。我们在自底向上聚类过程中使用了一个多样性正则化项来平衡每个聚类的数据量。最后,该模型实现了多样性和相似性之间的有效平衡。我们在大规模的图像和视频 re-ID 数据集上进行了大量的实验,包括 Market-1501、DukeMTMCreID、MARS 和 DukeMTMC-VideoReID。实验结果表明,该算法不仅优于目前最先进的无监督 re-ID 方法,而且优于竞争的迁移学习和半监督学习方法。

我们提出了一个自底向上的聚类框架来解决无监督的 re-ID 问题。通过利用身份之间的内在多样性和每个身份之间的相似性,我们的框架可以学习鲁棒的和有区别的特征。

- 我们采用排斥损失对无标签模型进行优化。排斥损耗直接优化了每个样本/簇之间的余弦距离。它可以促进模型利用每个聚类内的相似性,并最大限度地提高每个身份之间的多样性。
- 我们提出了一个多样性正则化项,以使每个簇中的图像数量平衡。使聚类结果与真实世界的分布保持一致。
- 实验结果表明,我们的方法在基于图像和基于视频 re-ID 数据集上都优于最先进的方法。在 Market-1501 (Zheng et al. 2015) 和 MARS (Zheng et al. 2016) 上,我们的准确率分别达到了 66.2% 和 61.1%。此外,一次性的 re-ID 方法比我们的方法使用了更多的注释,而我们的方法也获得了更高的性能。

通过利用身份之间的内在多样性和每个身份内部的相似性



如图 1 所示，在训练过程中，将单个的样本聚集成簇，并逐渐合并。具体来说，我们的框架 BUC 以一种迭代的方式应用网络训练和自底向上聚类。我们提出的排斥损耗可以在没有任何标签的情况下优化网络，并获得良好的初始精度。在网络训练的初始，我们将单个图像作为范例，即每个图像属于一个独立的集群。然后，我们通过自底向上的聚类，将相似的图像(聚类)合并到一个聚类中，逐渐合并身份内部的相似性。此外，在实践中，不同的身份被相机捕捉到的概率应该是相似的，因此不同簇的图像数量应该是平衡的。为了加强这一性质，我们在合并过程中加入了一个多样性正则化项。最后，在从下到上的聚类过程中，我们的框架利用相似性和多样性来学习判别特征。

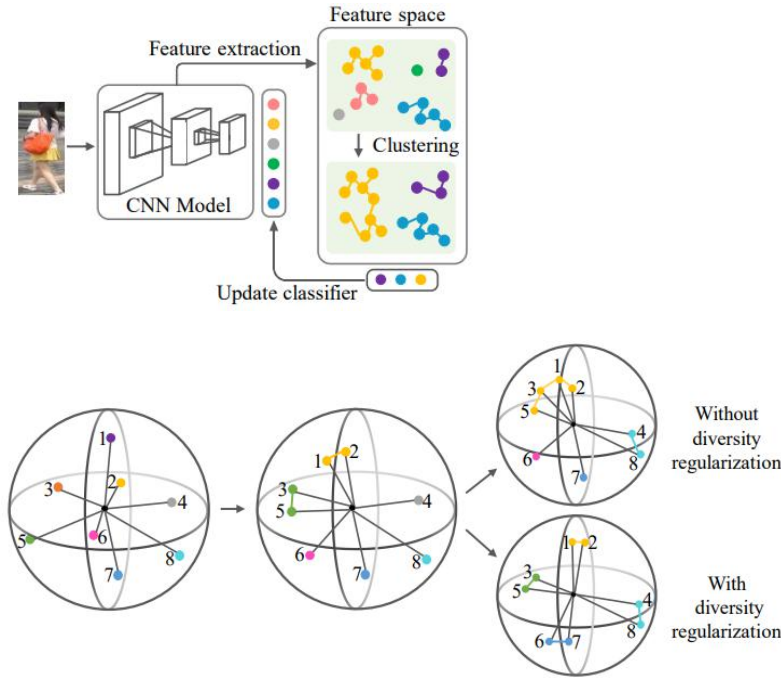
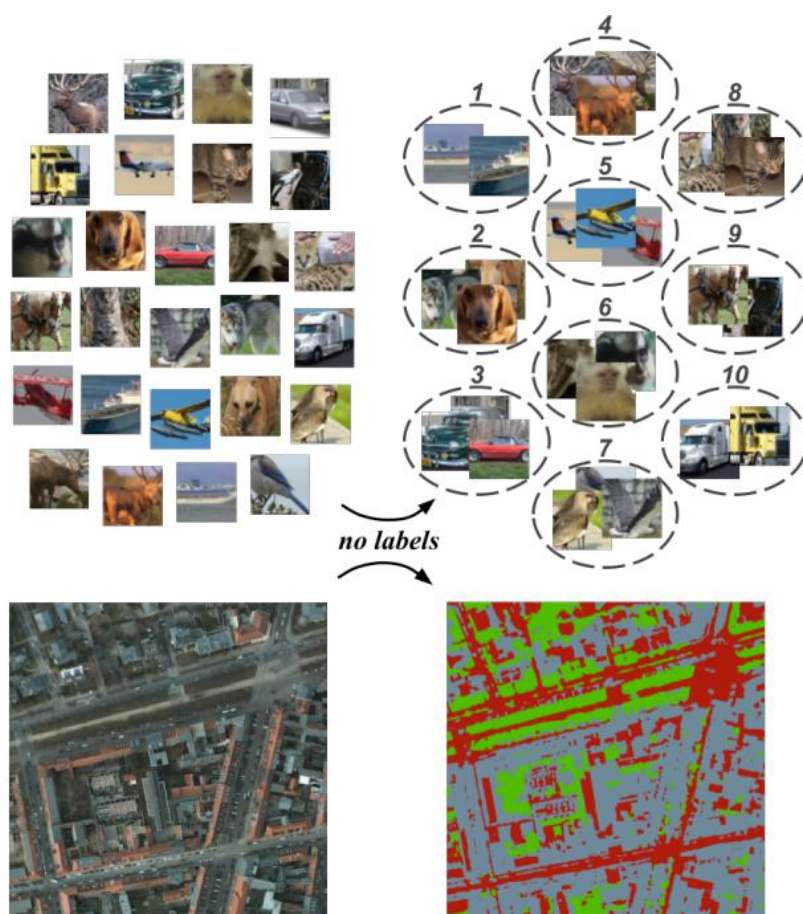


图 2:(a).本文提出的框架以未标记的图像作为输入，训练网络，提取图像特征进行聚类。该框架交替进行三个步骤，即提取每幅图像的 CNN 特征，合并整个训练集的簇，并重新训练 CNN 模型。图(b)-(d)描述了集群合并的过程。在这个例子中，每一步我们合并两个具有最小差异的簇对。相同颜色的实心点代表同一簇中的图像。彩色线表示连接的两个集群已合并为一个。在(b)中，学习到的特征区分跨越一个单元球，该单元球的多样性是最大的。在(c)中，合并簇后，同一簇的特征嵌入距离球体更近。在(d)中，上球面显示了没有分集正则化的聚类合并结果:(点 1, 点 3)和(点 4, 点 8)距离最短，合并为一个聚类。下面的球面显示的是分集正则化后的聚类合并结果:虽然黄色和绿色的聚类之间的距离最短，但这两个聚类太大，不应该合并。点 6 和点 7 被合并。

Methods	Venue	Labels	Market-1501				DukeMTMC-reID			
			rank-1	rank-5	rank-10	mAP	rank-1	rank-5	rank-10	mAP
BOW (Zheng et al. 2015)	ICCV15	None	35.8	52.4	60.3	14.8	17.1	28.8	34.9	8.3
OIM* (Xiao et al. 2017)	CVPR18	None	38.0	58.0	66.3	14.0	24.5	38.8	46.0	11.3
UMDL (Peng et al. 2016)	CVPR16	Transfer	34.5	52.6	59.6	12.4	18.5	31.4	37.6	7.3
PUL (Hehe et al. 2018)	TOMM18	Transfer	44.7	59.1	65.6	20.1	30.4	46.4	50.7	16.4
EUG* (Wu et al. 2018a)	CVPR18	OneEx	49.8	66.4	72.7	22.5	45.2	59.2	63.4	24.5
SPGAN (Deng et al. 2018)	CVPR18	Transfer	58.1	76.0	82.7	26.7	46.9	62.6	68.5	26.4
TJ-AIDL (Wang et al. 2018)	CVPR18	Transfer	58.2	-	-	26.5	44.3	-	-	23.0
BUC without diversity regularizer	AAAI19	None	62.9	77.1	82.7	33.8	41.3	55.8	62.5	22.5
BUC	AAAI19	None	66.2	79.6	84.5	38.3	47.4	62.6	68.4	27.5

Invariant Information Clustering for Unsupervised Image Classification and Segmentation-ICCV2019

训练后的网络直接输出语义标签,而不是需要外部处理才能用于语义聚类的高维表示。目标是使每对学生的课堂作业之间的相互信息最大化。它很容易实现,并且严格地以信息论为基础,这意味着我们毫不费力地避免了其他聚类方法容易受到的退化解。



不变的信息聚类

影响其他方法的两个问题:

第一个是聚类退化,即单个聚类控制预测或聚类消失的趋势。由于互信息中的熵最大化成分,如果所有图像被分配到同一类,损失不会最小化。与此同时,由于条件熵最小化,最理想的模型是对每个图像确定地预测单个类(即一个热点)。

第二个问题是带有未知类或干扰类的噪声数据。IIC 通过使用一个平行于主输出层的辅助输出层来解决这个问题,这个辅助输出层经过训练来产生一个在测试时被忽略的超聚类。

辅助过聚类是一种通用技术,对其他算法也很有用。IIC 的这两个特性使得它成为我们的无监督基线中唯一的方法,该方法足够健壮,可以利用 STL10 的无标记噪声子集,STL10 是 ImageNet [14] 的一个版本,专门设计作为无监督聚类的基准。

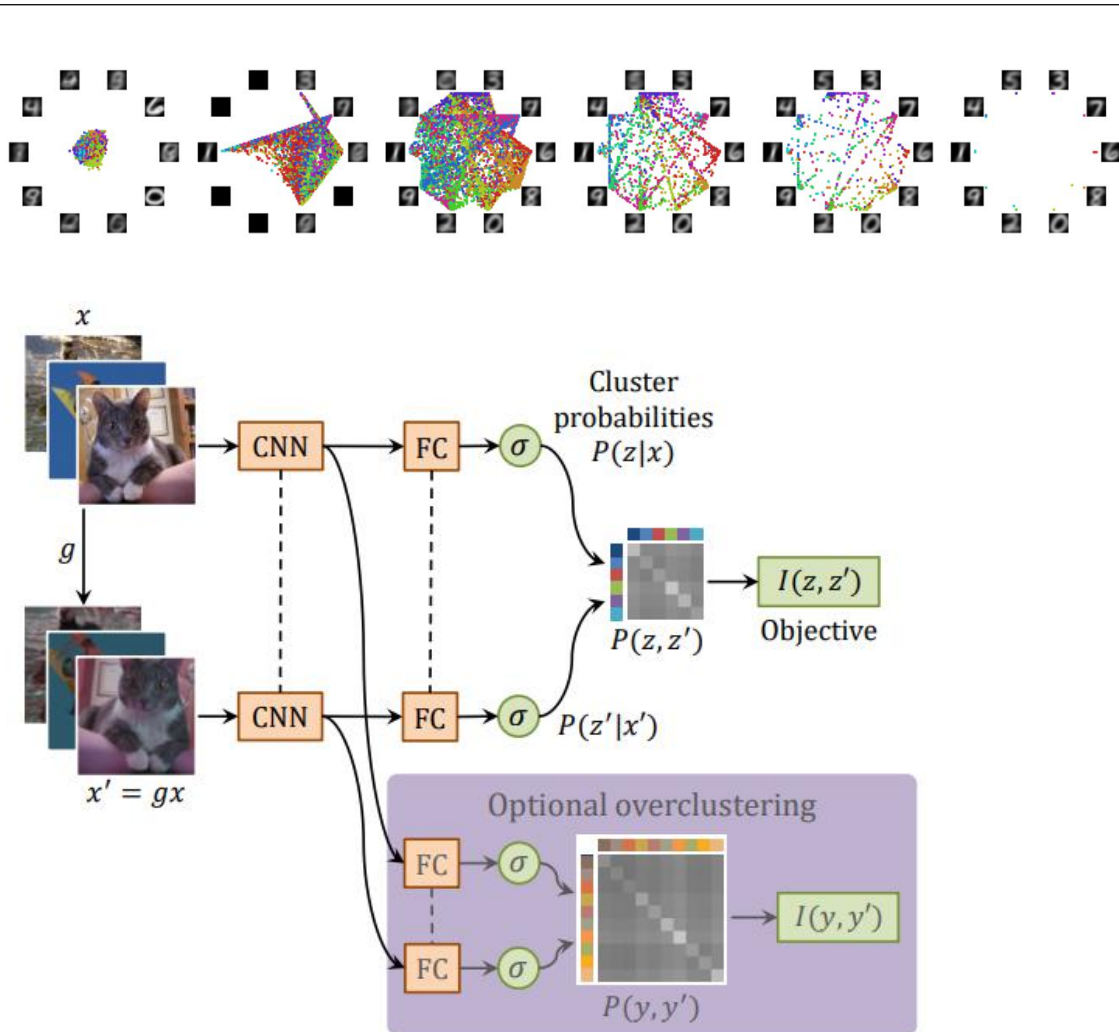


图 2: 图像聚类的 IIC。虚线表示共享参数, g 为随机变换, I 为互信息 (式 (3))。

$\Phi : X \rightarrow Y, \max I(\Phi(x), \Phi(x_0))$,
最大化 I 即丢弃特定于实例的细节

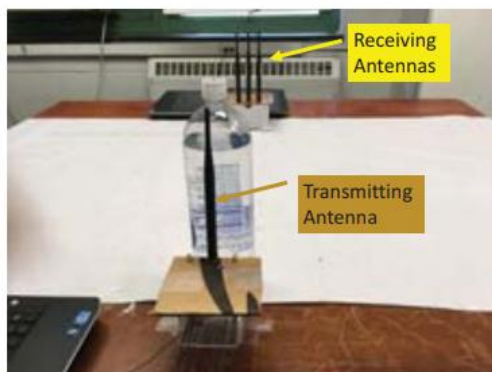
Towards inbaggage suspicious object detection using commodity WiFi-

CNS 2018

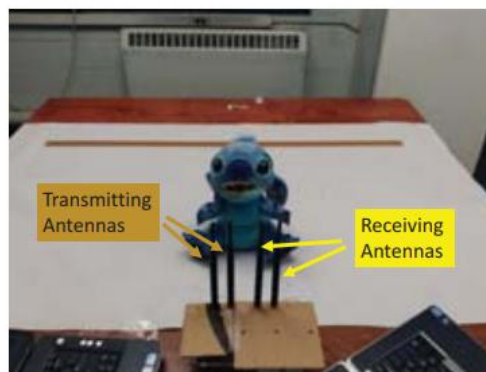
特别是，为了识别不同的材料，我们利用了通过或绕过物体传输的 WiFi 信号，这导致了来自天线的 CSI 复值的不同特性(即吸收、折射和反射)及其差异。此外，我们从 CSI 中提取物体反射的信号，根据发现反射信号的强度与物体的反射面积成比例，来估计其形状(如宽度和高度)或体积。与现有的工作相比，我们的方法独特地分离了由物体的两个影响因素(即，材料和形状)引起的无线干扰，利用不同的信号波束包含在 CSI 复合体中。我们的系统只需要拥有 2 到 3 个天线的 WiFi 设备，并可以以低成本和部署努力集成到现有的 WiFi 网络中，使其比使用专用仪器(例如 x 射线和 60GHz 雷达)的方法更具可扩展性和实用性。要使用现成的 WiFi 实现该系统，需要解决许多挑战。首先，WiFi 信号测得的 CSI 会受到一组物体物理属性(如材料、形状、大小、位置等)的影响，很难区分不同的影响，很难分别识别物体的材料和形状。其次，WiFi 信号相对于目标物体的尺寸，其波长相对较长，不太适合物体成像，造成强烈的衍射，导致成像分辨率低。第三，检测行李中隐藏的物体需要减轻各种类型的行李的影响。为了解决这些挑战，我们开发了两种专门设计的系统方法，用于从 CSI 复合体中分离折射信号和反射信号，并分别识别物体的材料和形状。该系统消除了 CSI 的原始相位噪声，重构了 CSI 复合体，即使可疑物体隐藏在行李中，也能很好地捕获可疑物体材料引起的主要干扰。我们还从 CSI 复合体中获得反射通道，这使我们能够利用长波长 WiFi 信号在更细的水平上估计物体的形状和体积。

- 我们证明了来自低成本设备的 WiFi 信号可以穿透视线受阻的行李，并在没有专用设备或信号的情况下方便可疑物体的检测和识别。
- 我们利用 CSI 复杂系统中丰富的信息来检测可疑的行李物品并识别其类别(即金属和液体)。
- 我们开发了基于反射的风险水平估计方法，根据液体的估计体积和金属的形状成像确定可疑物体的风险水平。我们证明，在不需要大型天线阵列或修改传输的情况下，可以从 WiFi 设备中不完美的 CSI(受未预测位移的影响)中提取物体的纯反射。
- 在为期 6 个月的时间里，用 15 个有代表性的对象，6 种类型的袋子/盒子进行了广泛的实验。结果表明，该系统对可疑物体的识别和材料类型的确定精度分别达到 95%和 90%以上，对液体体积和金属物体形状的估计平均误差分别达到 16ml 和 0.5cm

金属几乎不让 wifi 信号通过，甚至还能反射可以用来估计金属物体维度



(a) Setup1: The Tx and Rx are placed apart to identify material



(b) Setup2: The Tx and Rx are placed closely to capture reflected signal

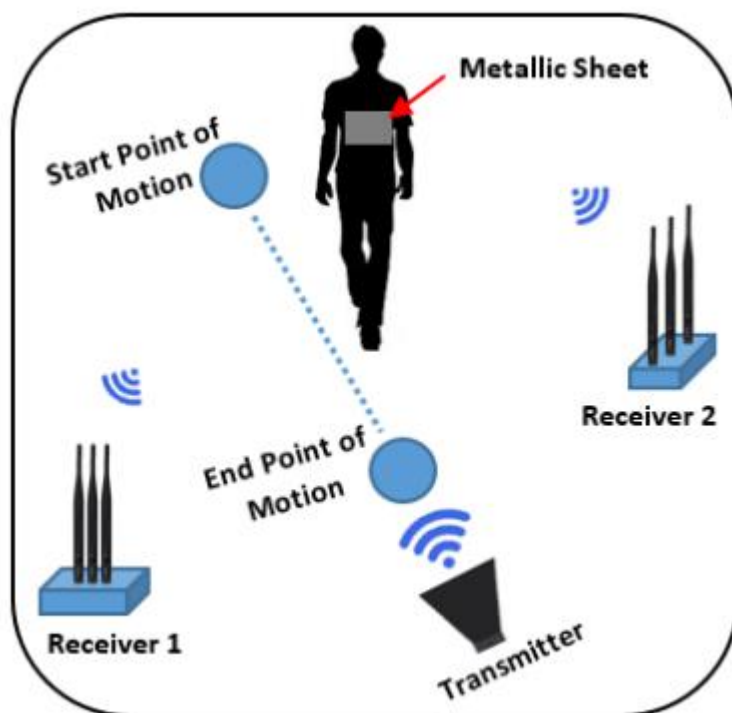
A, 检测材料。B, 检测反射信号

1) 两步法材料分类, 重点分析材料类型, 检测黑匣子内可疑物体, 同时减少物体大小、形状、位置等影响因素; 2) 基于信号反射的目标风险估计, 可以从 CSI 中提取目标的反射信号, 进行形状成像和体积估计, 估计可疑目标的风险水平。

Non-obtrusive detection of concealed metallic objects using commodity

WiFi radios-GLOBECOM2018

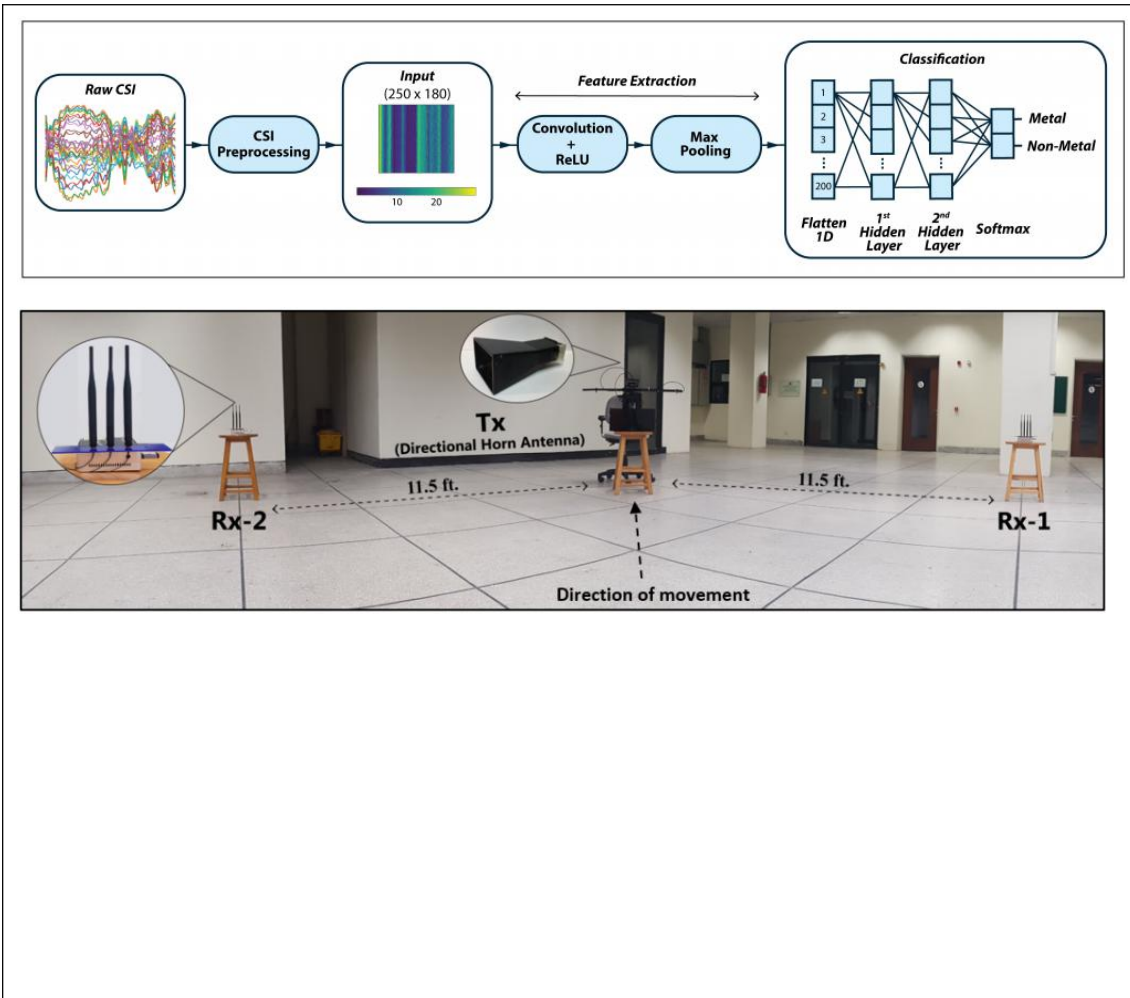
随着人们对隐蔽金属武器检测的日益关注,对具有大覆盖范围、鲁棒性和非突击性的金属检测系统的需求日益迫切。基于电磁感应或 x 射线的传统系统是有效的,但除了需要昂贵的基础设施外,覆盖范围较小。在本文中,我们探索利用无处不在的 WiFi 信号对隐藏的金属物体进行非侵入性检测。为目的,我们构建一个原型系统组成的单天线商品 WiFi 无线发射机,和两个 multi-antenna 无线电接收器放置在一个室内环境约 42 英尺×39 英尺。与我们进行广泛的实验学科走过的设置(或没有)的金属板放置在胸。我们利用从接收机收集到的信道状态信息训练深度卷积神经网络,发现所提出的系统能够区分金属和非金属的情况,平均准确率为 86.44%。



我们进行了大量的实验,让受试者在没有或有一块金属板放置在他们的胸部附近的照明区域行走。使用从两个接收节点收集到的对应 CSI 作为训练数据,我们使用深度卷积神经网络(CNN)来区分这两种情况。通过对来自 459 个实验和 4 个不同对象的数据进行 10 倍交叉验证,发现该方法的平均精度为 86.44%。

使用振幅,移动人物

由于两个接收节点各有 3 个天线,因此共 6 个接收天线。对每个天线对应的 CSI 流进行预处理,仅提取目标运动对应的信息。去掉初始瞬态和最终瞬态后,得到 5000 个 epoch 流,每个 epoch 对应一个单包索引(每个包对应的传输持续时间为 1ms,因此对应的记录窗口为 5 秒)

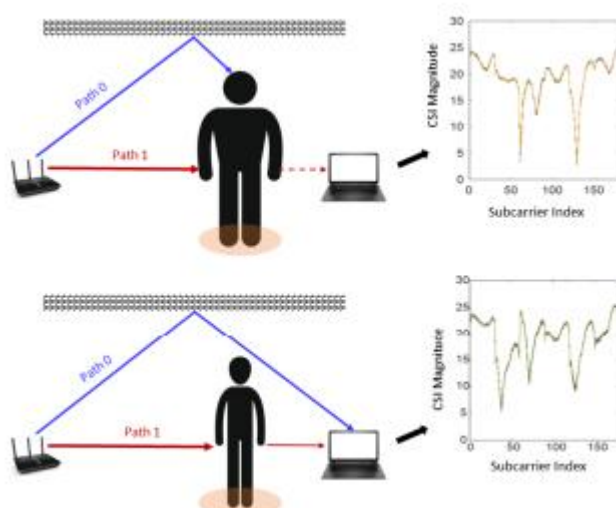


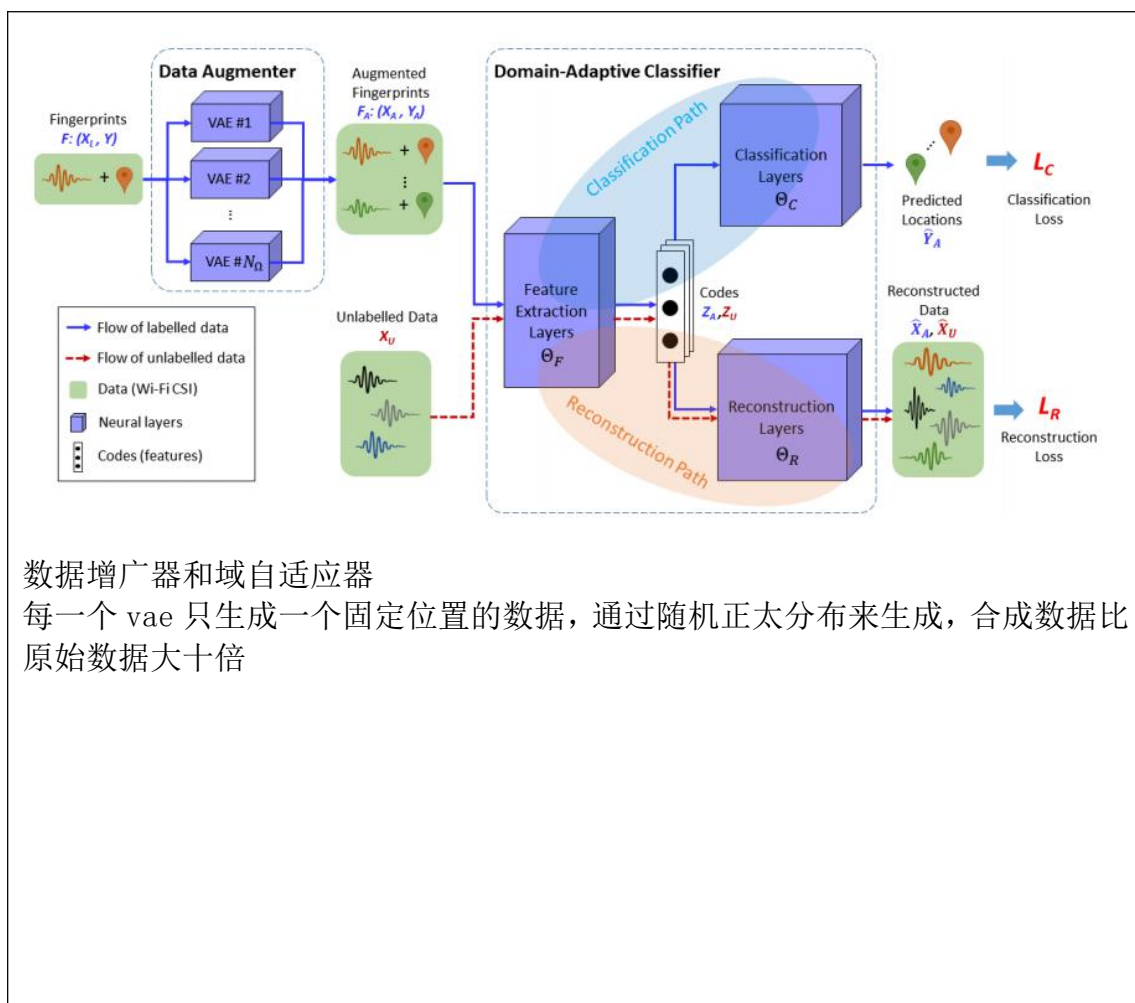
20-WWW-FiDo Ubiquitous Fine-Grained WiFi-based Localization for Unlabelled Users via Domain Adaptation

我们提出了一种基于 WiFi 的域自适应系统 FiDo，该系统能够使用来自一个或两个示例用户的标记数据来本地化许多不同的用户。 FiDo 包含两个模块：1) 一个数据增幅器，它使用变分自动编码器（VAE）引入数据分集； 2) 域自适应分类器，使用联合分类重构结构将自身调整为新收集的未标记数据。 与现有技术相比，FiDo 将平均 F1 得分提高了 11.8%，将最坏情况下的准确性提高了 20.2%。

Challenge 1: It is unrealistic to create a comprehensive fingerprint database that covers all possible users. To address Challenge 1, FiDo utilizes a data augmenter to generate synthetic fingerprints that are similar to yet different from the collected ones. This augmenter uses Variational Autoencoders (VAEs) to summarize the statistical features of the WiFi fingerprints. It then generates synthetic fingerprints, which are samples from the random distribution described by the statistical features. By doing the above steps, FiDo augments the fingerprints collected from a small group of example users (e.g., one or two) to a large group of virtual users, and thus is able to cover more users

基于 WiFi 的位置指纹的主要缺点在于它们在不同人群之间的不一致。 例如，如图 2 所示，一个高个子的人阻止了 WiFi 传播路径 0，而一个矮个子的人则阻止了 WiFi 传播路径 0。 同时，大人物吸收路径 1 上的更多信号功率，而小人物吸收更少的信号功率。 结果，WiFi 传播特性（在此示例中，我们使用 CSI 幅度）在不同人群之间差异很大，即使他们站在同一位置。 根据一个用户的数据进行训练，本地化系统很可能会对其他看不见的用户产生可怕的影响。





数据增广器和域自适应器

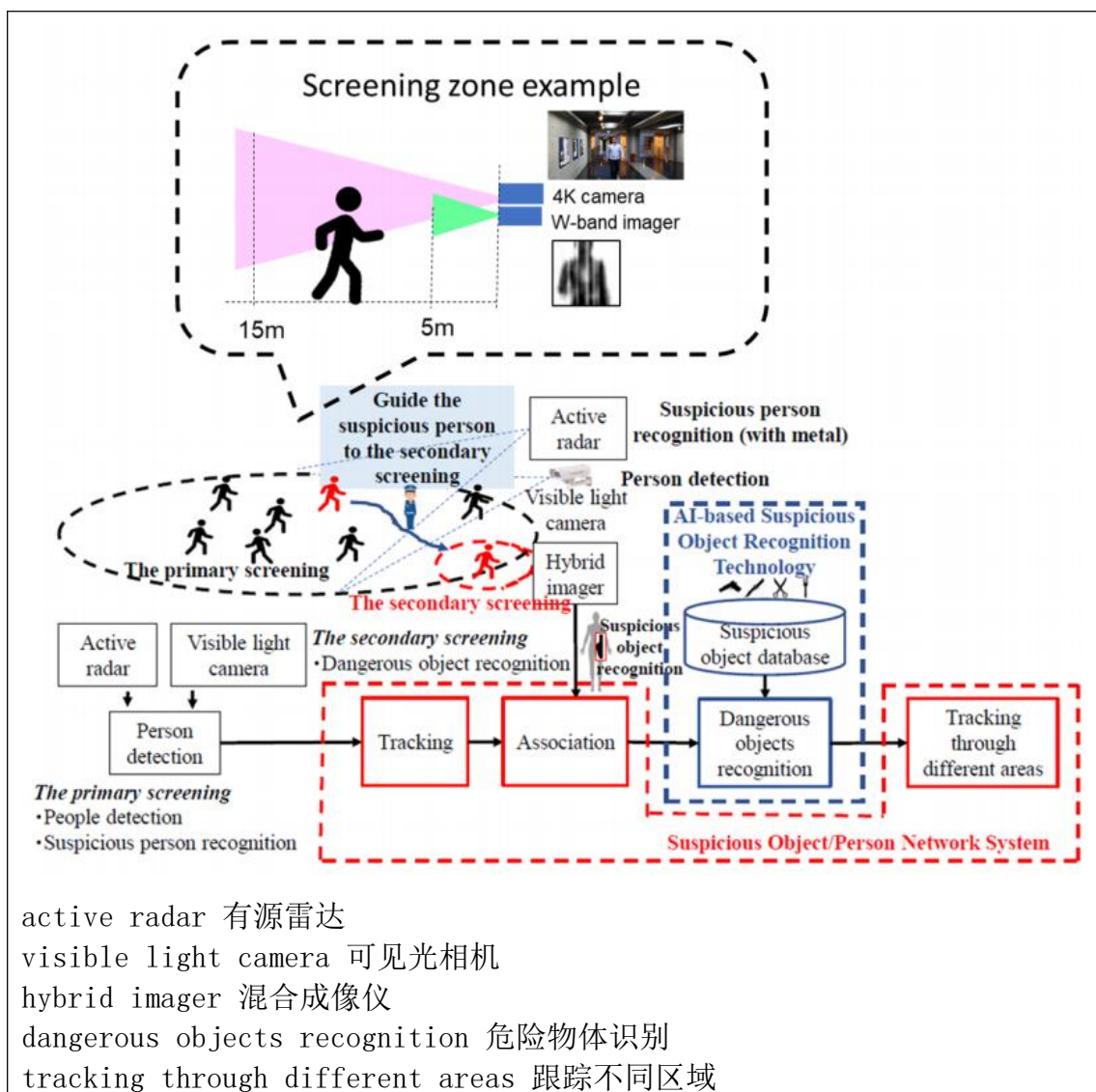
每一个 vae 只生成一个固定位置的数据，通过随机正太分布来生成，合成数据比原始数据大十倍

2021-CCNC-Congestion-Aware Suspicious Object Detection

摘要——致命疾病和恐怖袭击严重威胁着人类安全，对全球安全构成挑战。为了解决这个问题，城市监控系统正在以成熟但效率低下的解决方案在大规模网络中快速应用。当监控网络正在管理从多个边缘节点生成的数据时，由于数据流量集中和数据传输机制效率低下，很容易造成拥堵。同时，5G 技术应对爆发式移动数据流量增长和海量设备连接，可以实现真正的“万物互联”，构建社会经济数字化转型。在本文中，在 5G 技术的背景下，我们基于下一代网络概念设计的可疑对象网络系统（SONS）提出了一个以信息为中心的网络（ICN）监视系统。在该解决方案中，网络中的边缘节点分发计算和数据存储需求。我们首先描述当前的监控问题和我们提出的系统架构。然后我们使用仿真来验证和评估传统的全通集中监控系统和基于 ICN 的分散监控系统之间的系统性能。

- 1) 我们设计了一个基于内容的监控系统命名数据网络架构。
- 2) 我们在监控系统中引入了内容名称数据传递机制。
- 3) 实现了模拟评估比较传输传统集中式网络之间的容量要求和一个去中心化的网络。

可疑对象网络系统本着“保证足够的安全级别而不阻止人流”的原则，我们设计了一个可疑对象检测系统（SONS）来识别人类隐藏的可疑对象[15]。图 2 显示了 SONS 的整体结构。在网络中产生和传输数据之前，有两个筛选程序，初级筛选和二级筛选。在初步筛查中，我们结合使用传感器、多个可见光摄像头和 W 波段有源雷达成像仪，在最远 15 米的范围内识别和识别一个人是否有可疑物体（金属等）。在一个区域内，有一套可见光相机和一个成像仪。相机可检测到 6m 15m 内的人，成像器可分别检测 1m 5m 内的人。当相机检测到一个人并确认该人正在进入成像仪检测区域时，成像仪开始扫描此人。如果成像器检测到粗略归类为金属的可疑物体，则在第一次筛选过程中将此入确定为可疑人员。如果识别出可疑人员，我们的系统会将该人员的信息提供给安全人员。该信息包括传感器图像和其他数据。然后保安可以找到并引导可疑人员到二次筛查地点进行更详细的检查。在二次筛查中，混合成像仪用于进一步识别物体形状和材料，以识别可疑人员携带的可疑物体。在两阶段筛查过程中，我们的系统会记录和跟踪每个可疑人员的面部和其他信息并进行缓存。



2020-AAAI-Adversarial Training Based Multi-Source Unsupervised Domain Adaptation for Sentiment Analysis

用于情感分析 (SA) 的多源无监督域适应 (MS-UDA) 旨在利用多个源域中的有用信息来帮助在没有监督信息的未标记目标域中进行 SA。MS-UDA 的现有算法要么仅利用共享特征, 即域不变信息, 要么基于 NLP 中的一些弱假设, 例如平滑假设。为了避免这些问题, 我们提出了两种基于多源域适应方法的迁移学习框架, 通过结合源假设来推导出良好的目标假设。第一个框架的关键特征是一种新颖的基于加权方案的无监督域自适应框架多源无监督域自适应 (MS-UDA) 用于情感分析 (SA) 旨在利用多个源域中的有用信息来帮助在未标记的目标域中进行 SA 没有监督信息的领域。MS-UDA 的现有算法要么仅利用共享特征, 即域不变信息, 要么基于 NLP 中的一些弱假设, 例如平滑假设。为了避免这些问题, 我们提出了两种基于多源域适应方法的迁移学习框架, 通过结合源假设来推导出良好的目标假设。第一个框架的关键特征是一种新颖的基于加权方案的无监督域适应框架 (WS-UDA), 它结合了源分类器来直接获取目标实例的伪标签。而第二个框架是基于两阶段训练的无监督域适应框架 (2ST-UDA), 它进一步利用这些伪标签来训练目标私有提取器。重要的是, 分配给每个源分类器的权重基于目标实例和源域之间的关系, 这是由判别器通过对抗性训练测量的。此外, 通过相同的鉴别器, 我们还实现了共享特征和私有特征的分离。在两个 SA 数据集上的实验结果证明了我们框架的有前途的性能, 其性能优于无监督的最先进的竞争对手。ework (WS-UDA), 结合源分类器直接获取目标实例的伪标签。而第二个框架是基于两阶段训练的无监督域适应框架 (2ST-UDA), 它进一步利用这些伪标签来训练目标私有提取器。重要的是, 分配给每个源分类器的权重基于目标实例和源域之间的关系, 这是由判别器通过对抗性训练测量的。此外, 通过相同的鉴别器, 我们还实现了共享特征和私有特征的分离。在两个 SA 数据集上的实验结果证明了我们框架的有希望的性能, 优于无监督的最先进的竞争对手。

在本文中, 我们专注于 SA 的 MS-UDA, 并希望结合多个标记源域的假设来为未标记的目标域推导出一个好的假设。为此, 我们引入了两个迁移学习框架。第一个框架是基于加权方案的无监督域适应 (WS-UDA), 其中我们集成源分类器以直接注释目标实例的伪标签。我们的第二个框架是基于两阶段训练的无监督域适应方法 (2ST-UDA), 它进一步利用伪标签来训练特定于目标的提取器。我们框架的主要特点包括: 首先, 我们通过将鉴别器视为概率分布估计器来在我们的模型之前引入数据依赖。具体来说, 我们利用鉴别器来衡量不同源域和目标实例之间的实例到域的关系, 在此基础上我们实现实例级加权方案, 为每个源分类器分配不同的权重; 其次, 我们的框架明确地对域表示的私有和共享组件进行建模, 并鼓励它们分离或独立, 这可以抵抗与底层共享分布相关的噪声污染 (Salzmann et al. 2010) 并对系统有益性能 (Bousmalis 等人, 2016 年)。详细地说, 我们的框架强制共享特征是域不变的, 并且私有特征通过对抗性训练来包含特定于域的信息, 而不是 Bousmalis 等人采用的正交性约束; 刘、邱和黄 (2016 年; 2017 年)。

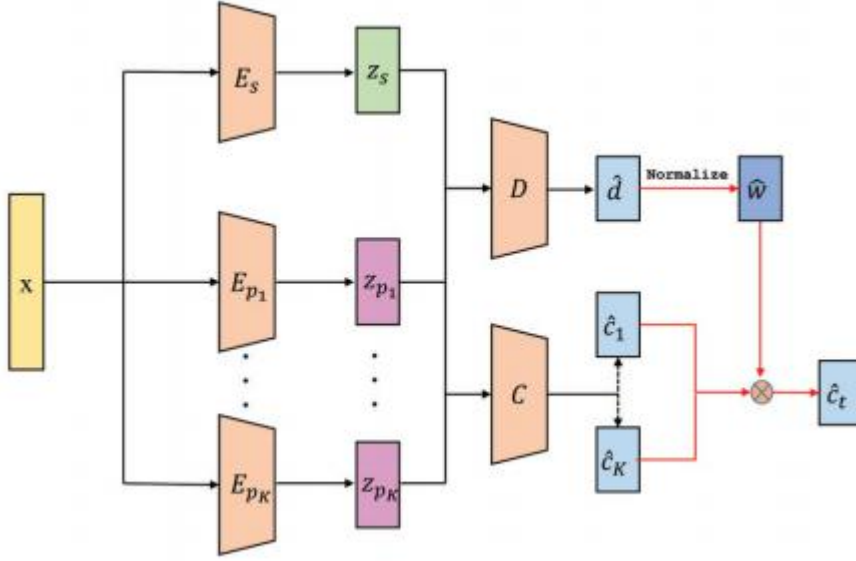


Figure 1: WS-UDA: The shared-wight extractor E_s captures shared features z_s for all domains. Each domain-specific extractor E_{p_j} K $j=1$ captures private features z_{p_j} K $j=1$ for each source domain. The classifier D strives to discriminate which domain the instances coming from and force z_s domain-invariant and z_{p_j} K $j=1$ domain-informative. The C estimates sentiment polarities $\{\hat{c}_j\}$ K $j=1$ from the views of different source domains as a traditional classifier. \hat{d} is normalized to tell what confidence we can give to each $\{\hat{c}_j\}$ K $j=1$. Finally, the sentiment polarities of target domain are assembled by the weighted sum of $\{\hat{c}_j\}$ K $j=1$. tion. We first present the problem definition and notations, followed by an overview of each framework. Then we detail the frameworks with all components successively