

## 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 Networks14

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&section=2](https://en.wikipedia.org/w/index.php?title=Domain_adaptation&action=edit&section=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

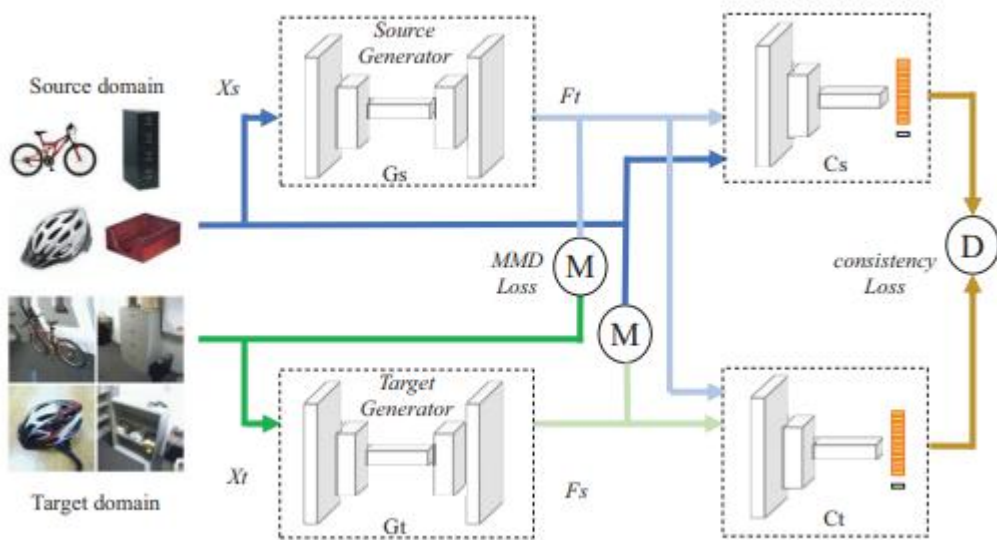
$$\theta(T2) = \arg \min_{\theta \in \text{Loss}(\text{Input}(T2), \theta(T1) + \theta)} (11)$$

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.



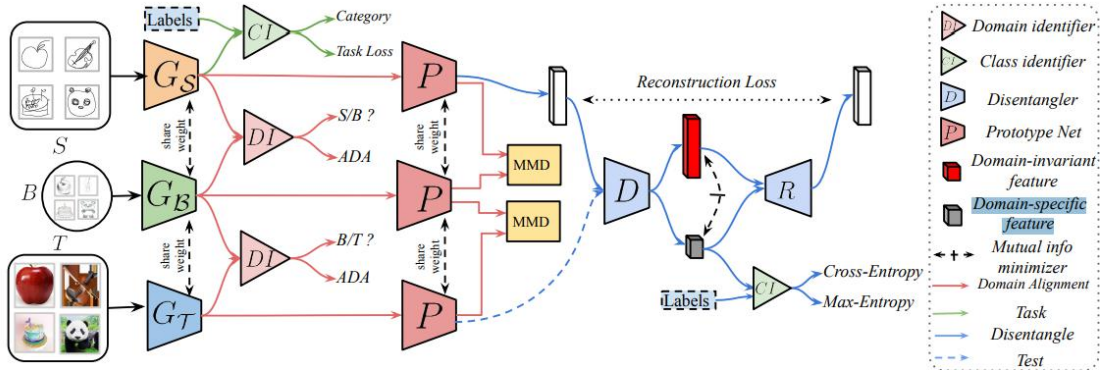
页码:P6617. (有计算公式和解析)

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)

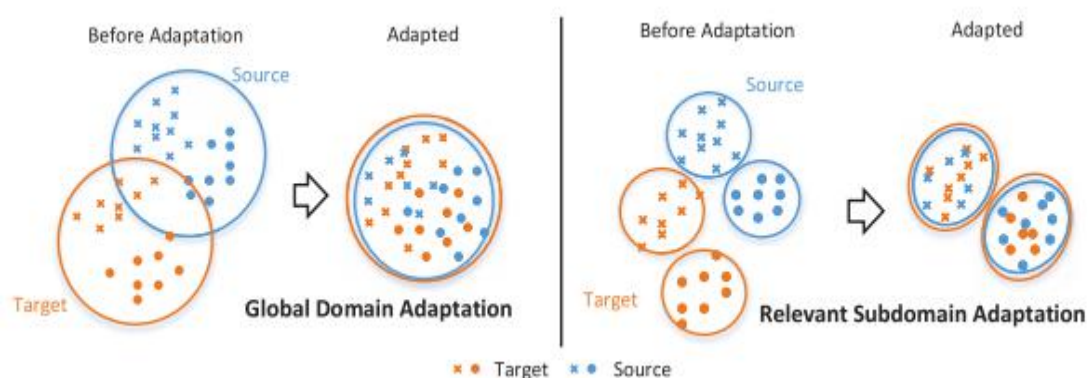


## Deep Subdomain Adaptation Network for Image Classification

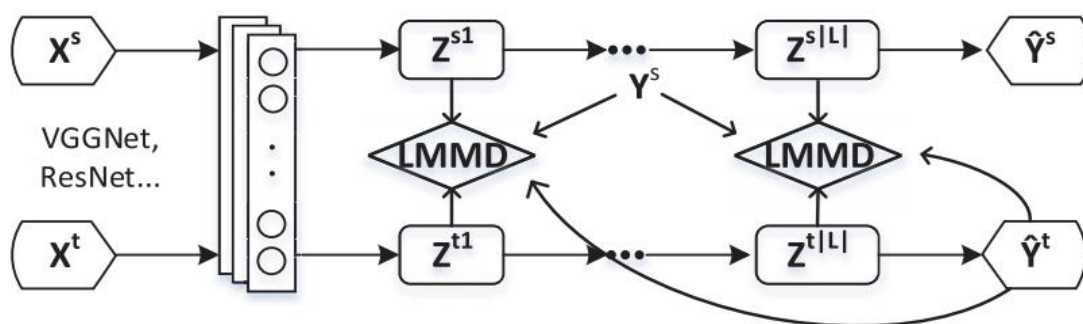
Recently, more and more researchers pay attention to subdomain adaptation that focuses on accurately aligning the distributions of the relevant subdomains. However, most of them are adversarial methods that contain several loss functions and converge slowly. Based on this, we present a deep subdomain 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>.



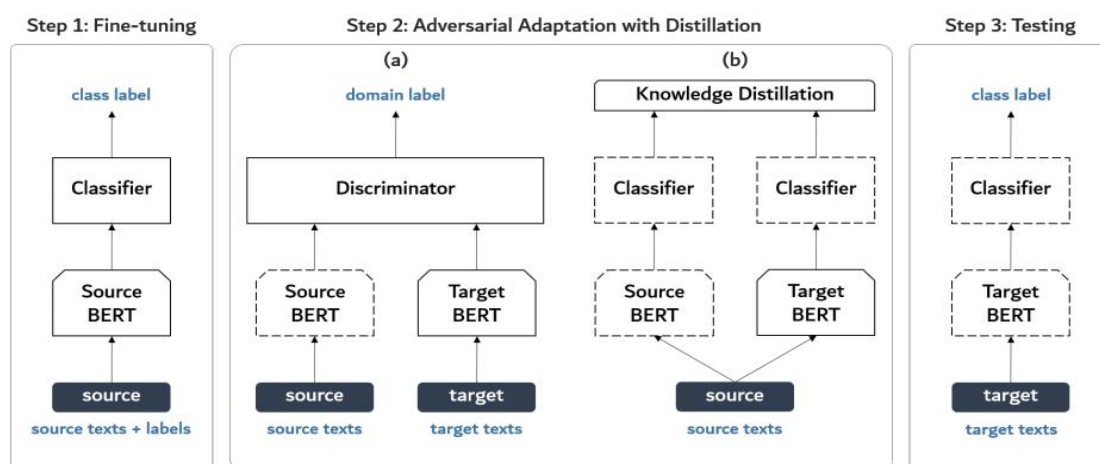
Left: global domain adaptation might lose some fine-grained information. Right: relevant subdomain adaptation can exploit the local affinity to capture the fine-grained information for each category. (更加关注子域的对齐，及相同类别和不同类别的域分布，类似于 Adaptive Device-free Localization in Dynamic Environments through Adaptive 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

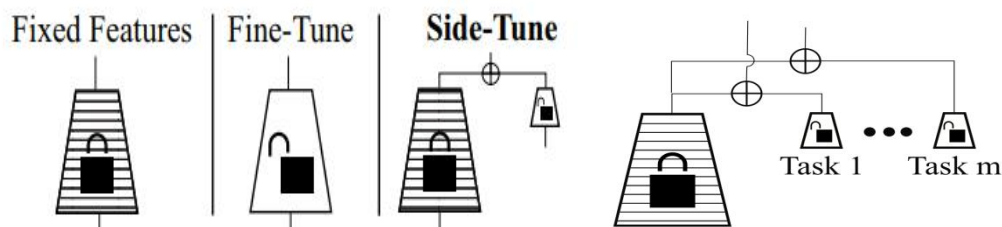
## Side-Tuning: A Baseline for Network Adaptation via Additive Side Networks

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  $\alpha$ -blending) reduces the side-tuning approach to: fine-tuning, feature extraction, and stage-wise training, depending on  $\alpha$  (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.



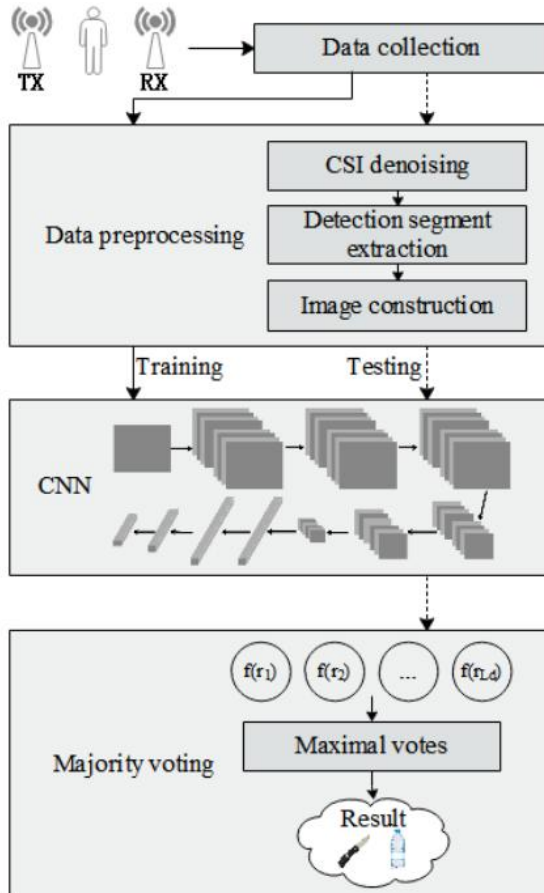
将现有的 fine-tune 机制进行扩展

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:**

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

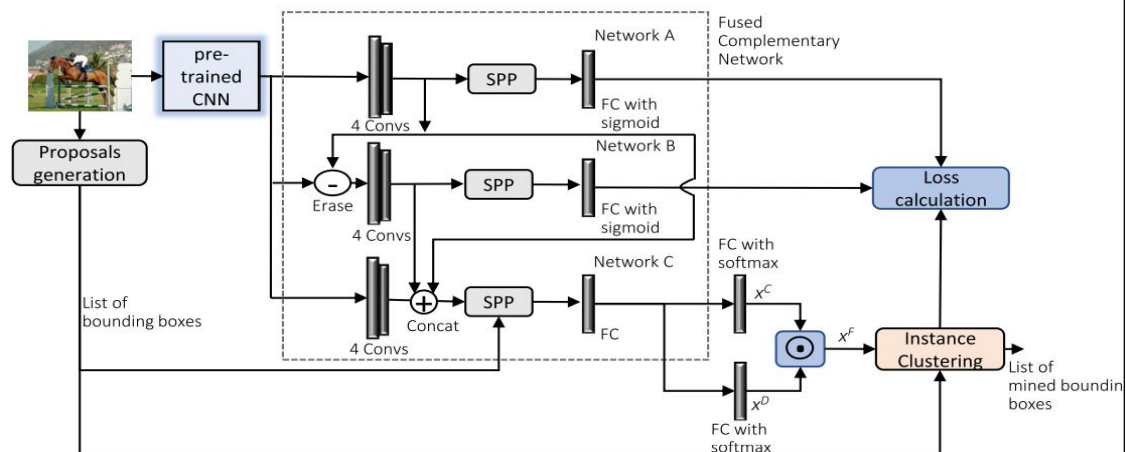
# Weakly Supervised Object Detection Using Complementary Learning and Instance Clustering

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：

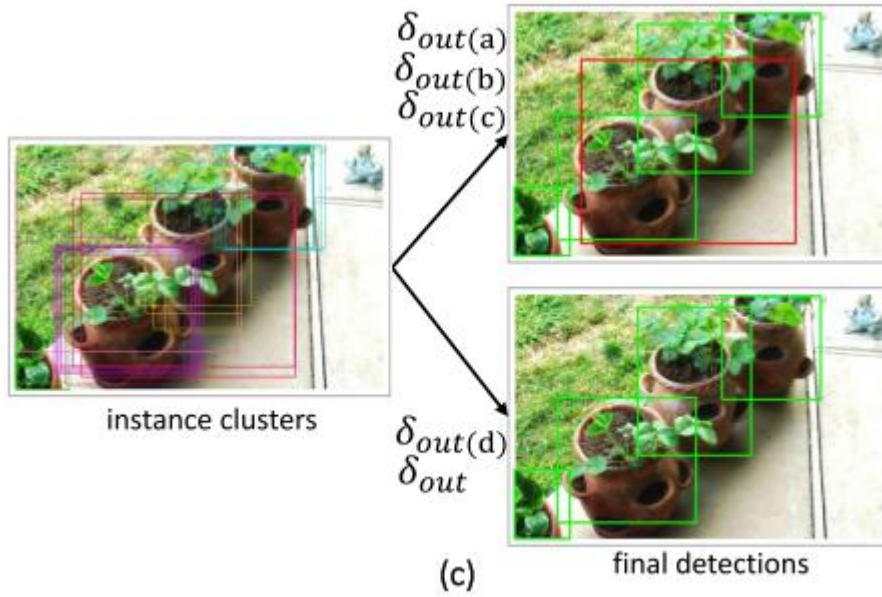
$$L = L_A + L_B + L_C \quad (2)$$

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

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

其中，L 是所建议的 WSOD 网络的损耗函数，L<sub>A</sub>，L<sub>B</sub> 和 L<sub>C</sub> 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。

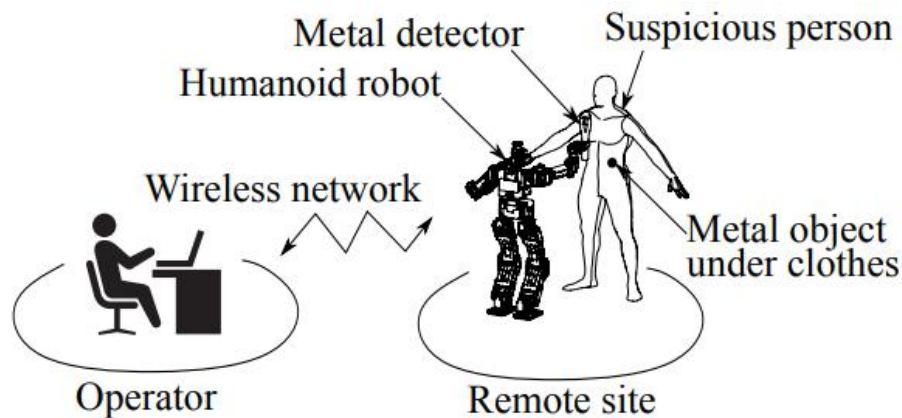


Fig. 1: An overview of body search with a teleoperated humanoid robot.

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:

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

