Active learning & MMD论文调研

Multi-Class Active Learning by Integrating Uncertainty and Diversity (2018)

In this paper, we **integrate uncertainty and diversity** into one formula by multi-class settings. Uncertainty is measured by the **margin minimum** while diversity is measured by the maximum mean discrepancy, which is popular to measure the distribution between two data sets. By minimizing the upper bound for the true risk of the integrating formula, we find the samples that not only uncertainty but also diversity with each other.

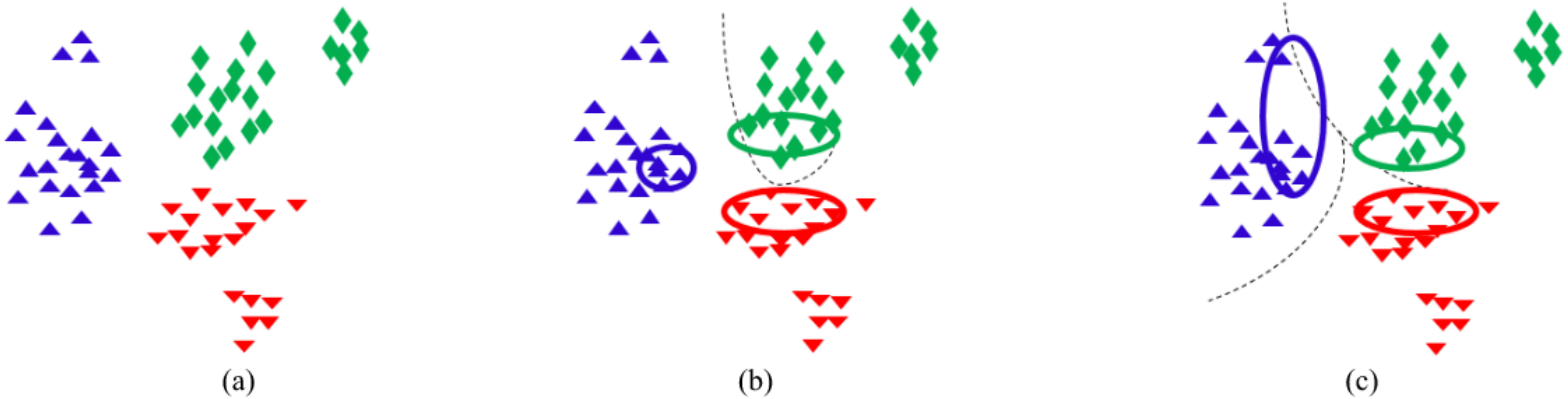


FIGURE 1. An illustrative example for selecting most informative samples with uncertainty and diversity (representativeness) for binary and multiple class settings. (a) Data distribution. (b) Samples queried based on binary class setting with uncertainty and diversity. (c) Samples queried based on multiple class setting with uncertainty and diversity.

Active Transfer Learning (2020)

- Select informative and discriminative subsets that are class balanced and highly similar to the target domain

保证选取的样例满足

- 1）类别平衡
- 2）有较好的数据分布广度(diversity)
- 3）和target domain中的样例相似。

按照传统的最小化Maximum Mean Discrepancy(MMD)方法，加上对数据加权，就可以学习到一个正交投影以及source domain中的样本权值。样本权值越大，越和target domain中的数据相似，那么我们更倾向于选择他们。

第二大目标的1)2)两点需要对权值 α 在优化过程中加上新的限定。一方面引入矩阵 K, 其元素 $k_{i,j} = [[xl_i = xl_j]]$。也就是说如果第i个样例和第j个样例属于同一类，那么对应下标元素取1，否择取0。另一方面引入矩阵 W，与前一个矩阵原理相似但是功能不同。

$$w_{ij} = [[\phi(i,j]]) = \begin{cases} 1, \text{if } x_i \in N_k(x_j) \text{ or } x_j \in N_k(x_i) \\ 0, \text{otherwise} \end{cases}$$

如果第i个元素属于第j个元素的k近邻，或者第j个元素属于第i个元素的k近邻，那么 W 矩阵中对应下标元素取1，否择取0。 这两个矩阵中取1的部分，都是为了最小化这两个元素的权值积 $\alpha_i\alpha_j$，使得他们不同时成为高 α 而被选择。

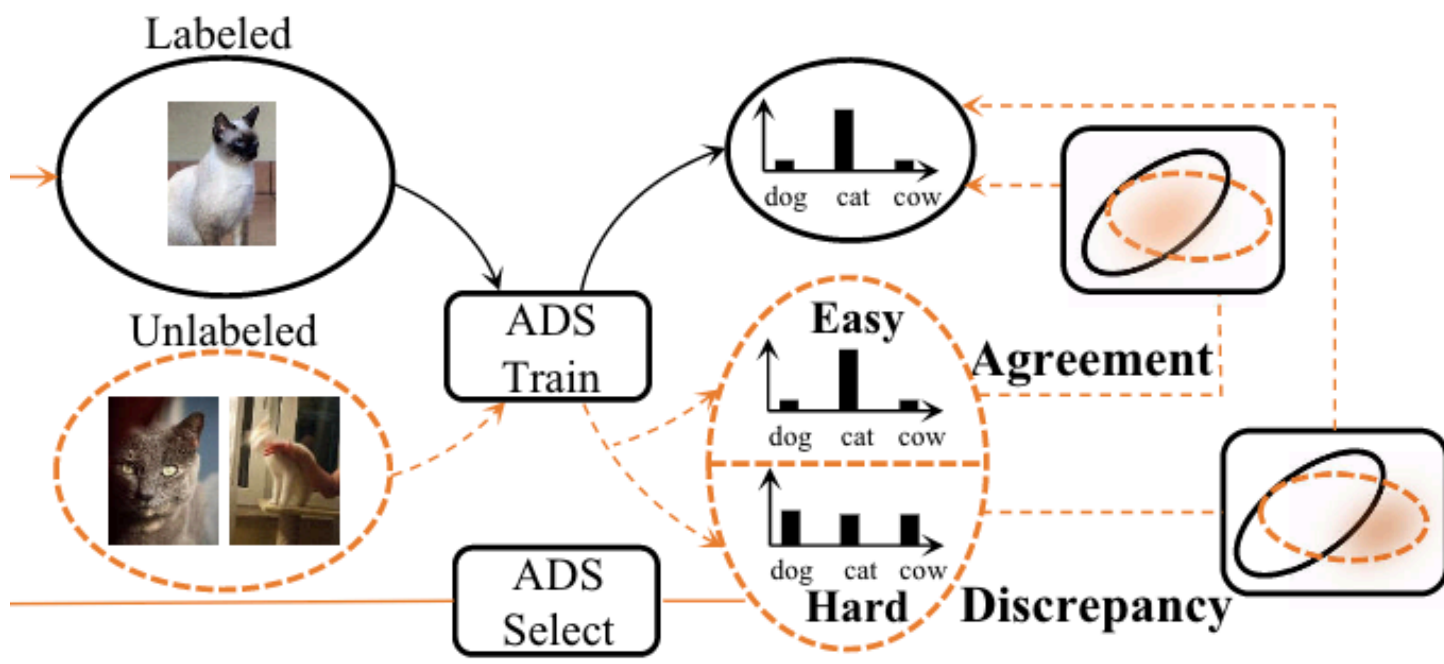
综合以上三个部分，Active Transfer Learning(ATL)的损失函数为

$$\min \mathcal{L}(P,\alpha) = ||\frac{1}{s}\sum_{i=1}^{s} P^T x_i \alpha_i - \frac{1}{t}\sum_{j=1}^{t} P^T y_j||_F^2 + \lambda_1 \sum_{i=1}^{s}\sum_{j=1}^{s} \alpha_i\alpha_j k_{i,j} + \lambda_2 \sum_{i=1}^{s}\sum_{j=1}^{s} \alpha_i\alpha_j w_{i,j}$$

$$\text{s.t.} \quad P^T P = I. \quad \sum_{i=1}^{s} \alpha_1 = 1, \alpha_i \geqslant 0.$$

Agreement-Discrepancy-Selection: Active Learning with Progressive Distribution Alignment (2021)

- In this paper, we propose an agreement-discrepancy-selection (ADS) approach, and target at unifying distribution alignment with sample selection by introducing adversarial classifiers to the convolutional neu-ral network (CNN).



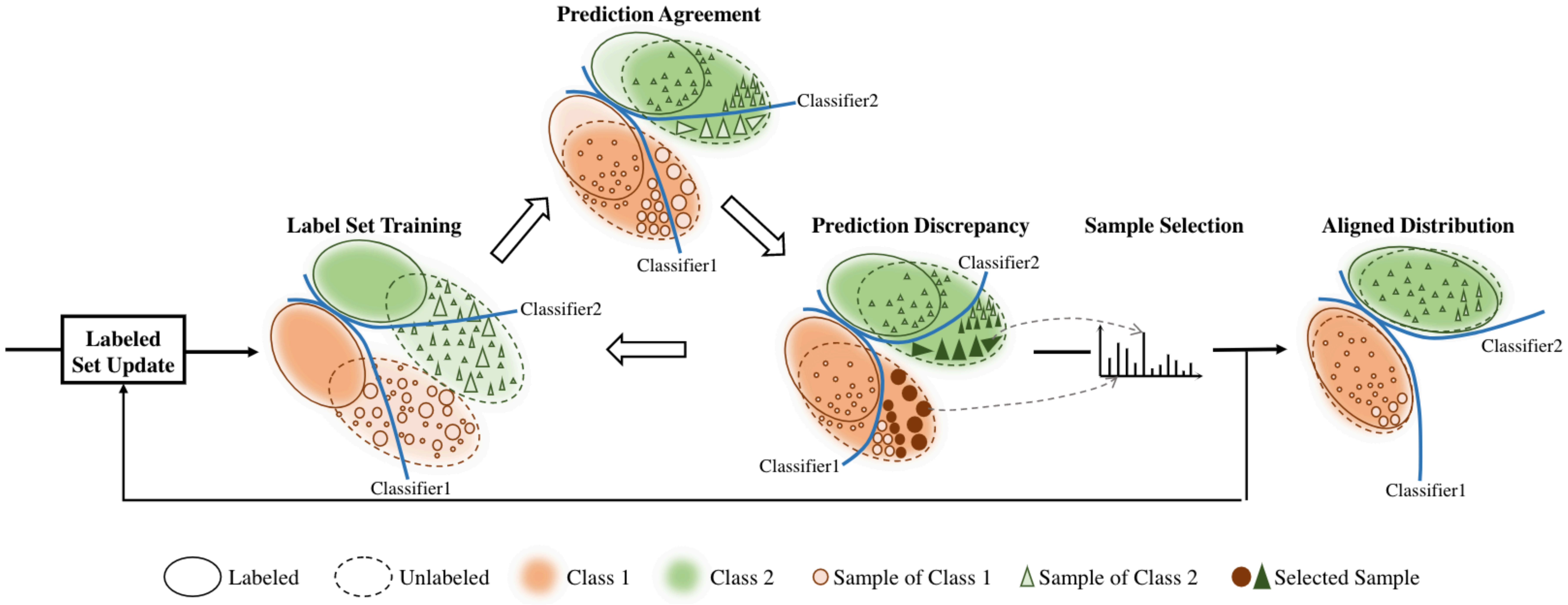


Figure 2: ADS flowchart.The prediction agreement step "pulls" the distributions of labeled and unlabeled samples with low and mid entropy together by updating features while the prediction discrepancy step "push" the distribution of unlabeled samples with high and middle entropy out of the alignment area by updating classifiers. Iterative agreement-discrepancy progressively aligns distributions of unlabeled samples with those of labeled samples. Larger circles/triangles denote more informative samples with larger entropy.

1. Batch Mode Active Learning for Object Detection based on Maximum Mean Discrepancy (2015)

Algorithm 1 MMD-based active learning for object detection

Require: Initial detector: D_1, training image set $IMG = \{img_1, img_2, \cdots, img_n\}$, active learning step N_{si}.

Ensure: Object detector D_{target}.

Create initial windows proposal set WIN_1 by applying D_1 to training image set IMG:

$$WIN_1 = \{w_1, \cdots, w_n\}.$$

$IMG_L = \emptyset$: the labelled images in the last iteration.

$NEG_L = \emptyset$: the negative windows obtained from labelled images IMG_L.

for $Iteration : index = 1 \rightarrow m$ **do**

Image selection:

Select N_{si} images $IMG_S = \{img_i, img_j, \cdots, img_k\}$ from the unlabelled images, based on the proposed MMD-based image selection method according to the distribution of detected windows WIN_{index}.

Negative data creation:

Create negative data NEG_S from IMG_S.

Update IMG_L and NEG_L:

$IMG_L = IMG_L \cup IMG_S$, $NEG_L = NEG_L \cup NEG_S$.

Train object detector $D_{index+1}$ on IMG_L and NEG_L.

Created new windows proposal set $WIN_{index+1}$ by applying $D_{index+1}$ on IMG.

end for

$D_{target} = D_{m+1}$

Incorporating Distribution Matching into Uncertainty for Multiple Kernel Active Learning (2019)

- multiple kernel learning(MKL), Combine the MKL and active learning together

- we propose a multiple kernel active learning algorithm by incorporating the distribution infor-mation into uncertainty with MKL, denoted as MKLAL

- we take the distribution information as a group regularizer and then add the group regularizer into uncertainty. Specifically, we use the minimum margin approach with MKL as the uncertainty criterion.

- By introducing the multi-kernel variant of MMD, we fuse the multiple kernel variant of both minimum margin approach and MMD into one formulation. The multiple kernel variant of MMD is mainly adopted to decrease the distribution difference between the labeled data and the unlabeled data in the learning space. Since the kernel weights are updated by the classification loss and MMD, the uncertainty can be calculated with both the distribution information and the margin information.

2. Active Label Distribution Learning Based on Marginal Probability Distribution Matching （2020）

- 基于边际概率分布匹配的主动标记分布学习(Active Label Distribution Learning Based on Marginal Probability Distribution Matching,ALDL-MMD)

- ALDL-MMD算法训练了一个线性回归模型,在保证其训练误差最小的同时,学习一个反映未标记数据上选点需求的稀疏向量,使选点后的训练集和未标记集的数据分布尽量相似,并对这个向量做松弛化处理,以简计算。

- **Author Keywords**主动学习 标记分布学习 最大平均差异 边际概率分布匹配 线性模型

3. Unsupervised Domain Adaptation for Object Detection Using Distribution Matching in Various Feature Level

- we aim to learn a model to generalize well in target domain of object detection by using MMD in various feature levels. We adjust MMD based on **single shot multibox detector (SSD) model** which is a single stage detector that learns to localize objects with various size using a multi-layer design of bounding box regression and infers object class simultaneously.