

融合深度主动学习的视觉目标检测模型

曹玉东, 廖鑫林, 陈 鑫, 贾 旭

(辽宁工业大学 电子与信息工程学院, 辽宁 锦州 121001)

摘 要: 车辆自动驾驶对周边目标的感知是保障交通安全的重要手段, 基于深度学习的目标检测模型被广泛应用, 但是需要海量的标注数据进行训练。本文提出一种采用高斯混合分布估计未标注图像不确定度的主动视觉目标检测模型, 以减少模型训练对标注数据的依赖。首先, 采用混合密度网络作为检测头, 以深度神经网络提取的图像特征为输入, 估计目标预测框分类和定位的概率分布; 其次, 将目标预测框的分类得分值映射到概率空间, 通过边缘不确定度计算目标的分类不确定度, 用预测框定位方差度量目标的定位不确定度; 最后, 挑选最不稳定的样本进行标注。在 VOC 数据集上的结果表明: 与其他典型的主动学习采样策略相比, 本文模型取得了最优性能, 仅用 54% 的数据标注量就能达到 YOLOX 监督学习 98.8% 的性能, 节省近 45% 的数据标注量。

关键词: 主动学习; 目标检测; 高斯分布; 标注代价; 不确定度估计

中图分类号: TP391 **文献标志码:** A **文章编号:** 1671-5497(2025)11-3697-08

DOI: 10.13229/j.cnki.jdxbgxb.20240223

Vision object detection model with deep active learning

CAO Yu-dong, LIAO Xin-lin, CHEN Xin, JIA Xu

(School of Electronics and Information Engineering, Liaoning University of Technology, Jinzhou 121001, China)

Abstract: The perception of surrounding objects by vehicle autonomous driving is an important means to ensure traffic safety. Object detection model with deep learning is adopted widely, but they requires a large amount of annotated data for training. In this paper, an active vision object detection model is proposed using Gaussian mixture distribution to estimate the uncertainty of unlabeled images, reduces the dependence of model training on labeled data. Firstly, the mixed density network is adopted as the detection head, taking the image feature extracted by the deep neural network as input, estimates the probability distribution of classification and location of the object predicted boxes. Secondly, the classification score of the object predicted boxes is mapped into the probability space, and the classification uncertainty of the object is calculated by edge uncertainty; the location variance of the predicted boxes is used to measure the location uncertainty of object. Finally, the most unstable samples were selected for labeling. The results on the VOC dataset show that compared with other typical active learning sampling strategies, the proposed model achieved the best performance. The proposed model using only 54% of the

收稿日期: 2024-03-06.

基金项目: 国家自然科学基金项目(12371363); 辽宁省应用基础研究计划项目(2022JH2/101300279); 辽宁省教育厅基本科研课题项目(JYTMS20230861).

作者简介: 曹玉东(1971-), 男, 教授, 博士. 研究方向: 计算机视觉与机器学习. E-mail: caoyd@lnut.edu.cn

data annotation volume can achieve the 98.8% performance of YOLOX with supervised learning, which saves up nearly 45% of the data annotation volume.

Key words: active learning; object detection; Gaussian distribution; labeling cost; uncertainty estimation

0 引言

基于深度学习的目标检测是当前计算机视觉领域的研究热点^[1],在行人、车辆等动态目标检测与识别方面应用较多。深度学习模型的训练需要依赖大量的标注数据,但是专业领域的标注成本很高。因此,主动学习^[2]开始受到研究人员的关注。主动学习方法模拟人的学习过程,从大量未标注的数据中选择对模型训练更有效的样本数据加入训练集,然后逐步优化网络模型参数,达到节省数据标注成本的目的^[3]。

不确定度可以度量未标注样本包含的信息量,不确定度越高的样本对模型的优化越有效。目前,多数主动目标检测(Active object detection, AOD)算法仅考虑图像的分类不确定度。例如,文献[4]以 SSD 网络作为主动学习的基础分类器,然后构建分类器委员会(committee of classifiers),将类别输出差异最大的样本判定为信息量丰富的样本;Gal 等^[5]通过估计图像分类网络权值的先验分布,采用 dropout 采样权值并预测变异率,将变异率高的图像视为不确定度高的图像,然后计算未标记图像的信息量得分,该方法的主要缺点是没有考虑所选样本间的相似性,导致样本信息冗余;Elhamifar 等^[6]考虑所选样本在特征空间中的相似性及其信息量评分,在凸优化框架内实现主动学习;Paul 等^[7]通过引入核心集减少样本数据的信息冗余;张新生等^[8]采用基于 Bootstrap 的主动学习方法选择未标注样本数据并训练分类器;Lakshminarayanan 等^[9]尝试将主动学习应用于目标检测的回归任务;Kao 等^[10]利用定位紧密度和定位稳定度对未标注图像进行排序,前者测量边界框的紧密程度,后者评估边界框在原始图像和有噪声图像中的稳定性。上述主动学习方法中,部分仅在主动视觉目标检测中考虑分类不确定度^[5-8,11],部分则专注于单一的回归不确定性^[9]。文献[10]虽然利用预测框(RPN 候选框)和真实框的交并比估计不确定度,但是没有估计交并比的概率分布。

主动学习方法虽然能降低数据的标注成本,但在处理高维数据时其性能并不理想。深度学习与主动学习都是机器学习的重要分支,为了综合二者的优势,有必要探索深度主动学习方法。然而,找到具有不确定性和代表性的未标记样本并非易事^[2]。针对上述文献中所提方法的不足,本文在分类不确定性的基础上,提出了融合定位不确定性采样的深度主动视觉目标检测模型,并将其命名为高斯 YOLOX(Gaussian YOLOX)。本文的主要工作包括:①改进网络结构,使神经网络的输出层可预测概率分布;②改进样本不确定性的度量方法,除分类不确定度外,还采用目标框位置分布的方差信息度量位置不确定度;③在目标检测模型的损失函数中增加定位不确定性损失,采用随机梯度下降法优化目标框位置的定位概率分布;④采用单一的检测网络和单次前向推理估计图像不确定度。与基于查询委员会(Query by committee, QBC)的主动学习方法^[4]相比,本文方法显著降低了计算成本和模型复杂度。

1 主动目标检测模型

基于主动学习的目标检测模型由查询策略、目标检测模型、领域专家、有标记样本集合和无标记样本集合 5 部分组成^[12],本文重点关注查询策略,即如何筛选“高品质”的未标注样本。图 1 描述了 Gaussian YOLOX 模型的主动学习循环过程。首先,基于有标记样本集合,采用划分聚类算法^[13]构造初始训练集,开始训练目标检测模型;未标注样本图像经过目标检测模型的单次前向推理过程,输出深度特征图。其次,从深度特征图中获取若干前景目标的预测框信息,计算各目标区域的分类不确定度和回归不确定度得分的加权和,将所有未标注图像的不确定度从大到小进行排序。最后,取出排序前 p 幅图像,由“领域专家”标注后加入训练集,开始下一轮训练,直到满足某种停止条件。

1.1 网络模型结构

与基于深度学习的 YOLOv3、YOLOv4 和

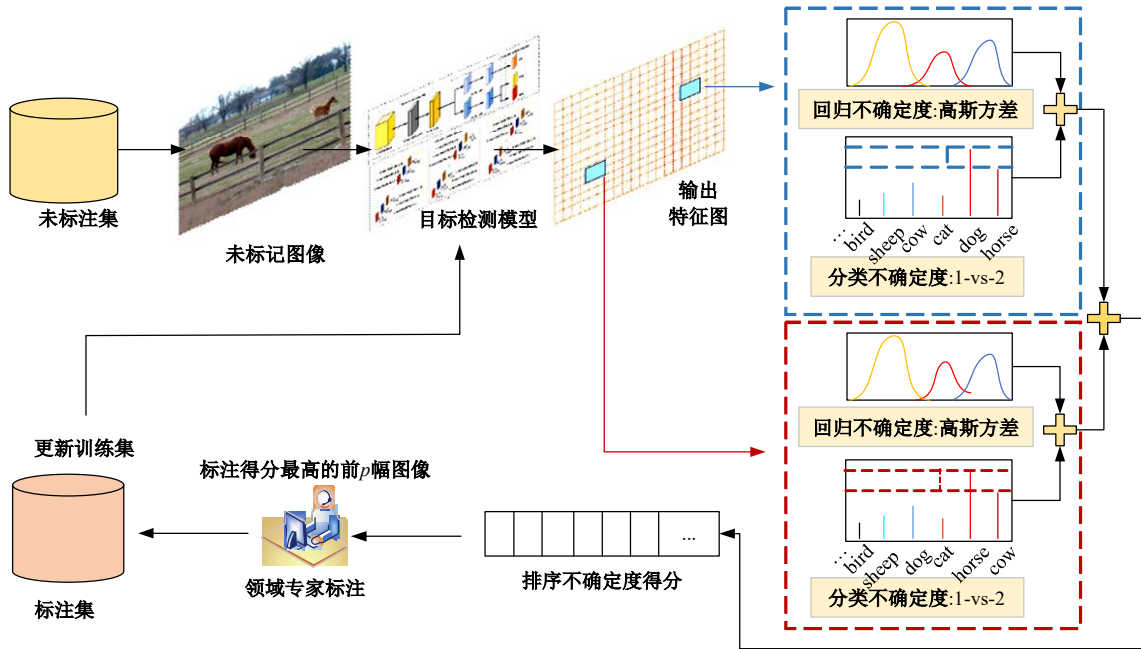


图 1 Gaussian YOLOX 主动学习循环过程

Fig. 1 Active learning cyclic process of Gaussian YOLOX

YOLOv5 版本相比, YOLOX^[14]不再采用锚点方式搜索目标框,既降低了计算量,又缓解了正负样本不平衡的问题。本文在 YOLOX 的基础上,改进了目标检测模型,其网络结构如图 2 所示。采用混合密度网络 (Mixture density network, MDN)^[15] 替换 YOLOX 的检测头 (detection head),网络输出由高斯混合模型的参数组成,即均值 μ 、方差 ν 和权重 π ,利用这些参数估计未标注图像的不确定性。MDN 包括 3 个分支,即定位分支、分类分支和置信度分支,每个分支分别输出服从高斯混合分布的 K 组卷积核参数,即 $\hat{\pi}_{\text{reg}}^k$ 、 $\hat{\mu}_{\text{reg}}^k$ 和 $\hat{\nu}_{\text{reg}}^k$ ($k=1, 2, \dots, K$)。其中, $\hat{\pi}_{\text{reg}}$ 为高斯分布的权重估计值, $\hat{\mu}_{\text{reg}}$ 为预测框中心和宽高参数的估计值, $\hat{\nu}_{\text{reg}}$ 为高斯分布的方差估计值。此外,主干特征提取网络用于获取图像中不同大小感受野的特征图,特征融合金字塔网络用于融合不同大小感受野的特征图。

基于 QBC 的主动学习方法^[4]需要用到多个分类器模型 (构成委员会),而本文提出的模型采用单一网络,其复杂度显著降低。

1.2 损失函数设计

本文的贡献之一是,采用目标框位置概率分布的方差度量图像的定位不确定性,并通过梯度下降算法优化目标框位置的概率分布。与文献^[16]类似,MDN 输出参数的计算公式为:

$$\begin{cases} \pi^k = \frac{\exp(\hat{\pi}^j)}{\sum_{j=1}^K \exp(\hat{\pi}^j)} \\ \mu^k = \hat{\mu}^k \\ \nu^k = \frac{1}{1 + \exp(-\hat{\nu}^k)} \end{cases} \quad (1)$$

式中: K 为高斯混合分布的组分数目,也是回归 MDN 的数量。

基于高斯混合分布拟合 MDN 网络定位分支的输出,并以交并比 (Intersection over union, IOU) 作为高斯混合分布的均值参数 $\hat{\mu}_{\text{iou}}$ 。相应地, MDN 网络的回归分支输出中另外两个输出作为高斯混合分布的权重 $\hat{\pi}_{\text{reg}}$ 和方差参数 $\hat{\nu}_{\text{reg}}$ 。负对数似然损失函数适用于概率建模和多类别分类问题,因此本文采用负对数似然损失函数设计回归损失函数,具体形式为:

$$L_{\text{reg}} = -\frac{1}{M} \log \left(\sum_{k=1}^K \hat{\pi}_{\text{reg}}^{i,k} \frac{1}{\sqrt{2\pi} \hat{\nu}_{\text{reg}}^{i,k}} e^{-\frac{(1 - \hat{\mu}_{\text{iou}}^{i,k})^2}{2(\hat{\nu}_{\text{reg}}^{i,k})^2}} + \epsilon \right) \quad (2)$$

式中: M 为与真实框 (标注框, ground-truth) 中目标类别匹配的正负样本预测框的数量; ϵ 取极小正数,用于保证对数值的稳定性。

图像的分类任务具有离散性,而高斯分布具有连续性,因此不能直接将 MDN 分类网络输出的权重、均值和方差作为高斯分布的参数。首先将分类 MDN 的均值 $\hat{\mu}_{\text{cls}}$ 参数映射到概率空间,然

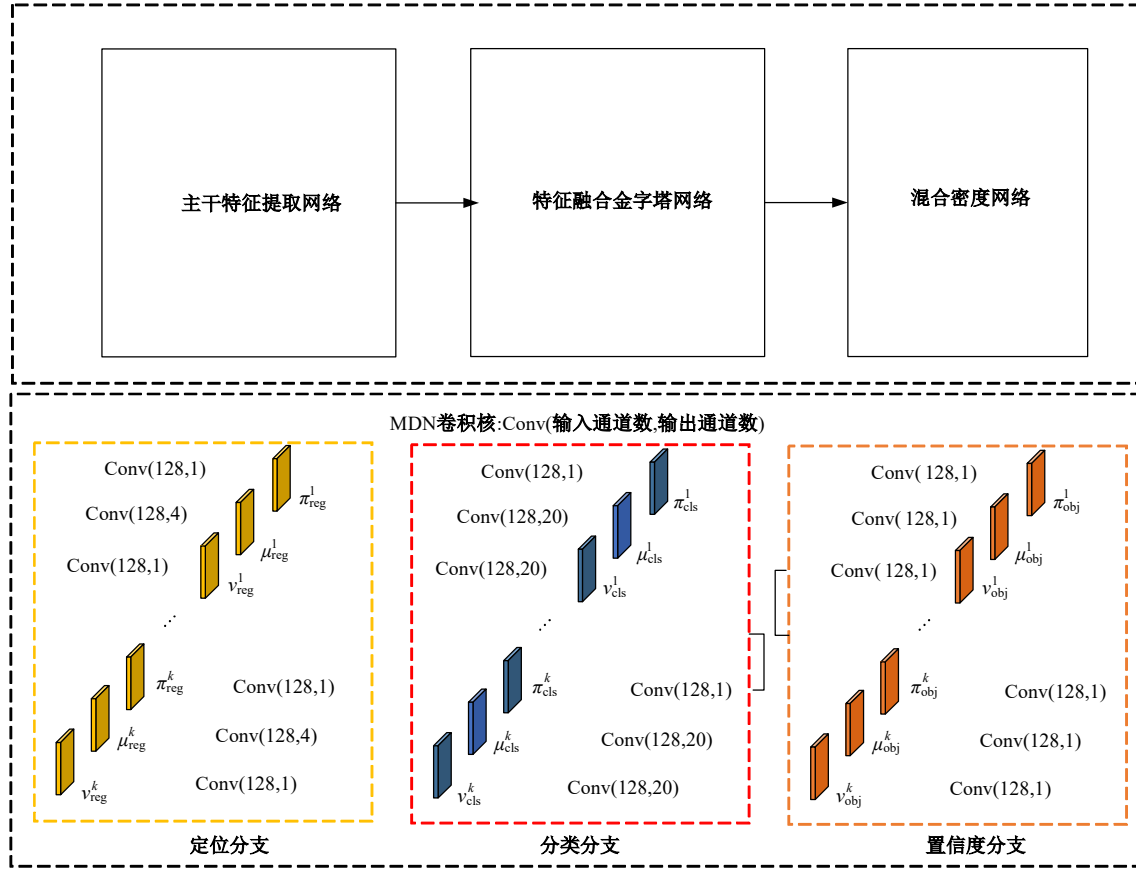


图 2 Gaussian YOLOX 的网络结构

Fig. 2 Network structure of Gaussian YOLOX

后将分类 MDN 参数中的均值和方差组合为 $\hat{c}^k = \hat{\mu}_{\text{cls}}^k + \sqrt{\hat{\nu}_{\text{cls}}^k} \gamma$, 其中 γ 服从标准高斯分布。交叉熵损失函数可衡量预测概率分布和标签值的差异, 通过惩罚错误分类样本优化分类边界^[17]。因此, 本文采用交叉熵设计分类损失函数, 具体形式为:

$$L_{\text{cls}} = -\frac{1}{N} \sum_{i \in \text{pos} + \text{neg}} \sum_{k=1}^K \hat{\pi}_{\text{cls}}^{i,k} \left[y_{\text{cls}}^i \log \left(\frac{1}{1 + \exp(-\hat{c}^{i,k})} \right) + (1 - y_{\text{cls}}^i) \log \left(\frac{\exp(-\hat{c}^{i,k})}{1 + \exp(-\hat{c}^{i,k})} \right) \right] \quad (3)$$

式中: N 为目标框中预测类别的正负样本数量; pos 为正样本; neg 为负样本; y_{cls}^i 为第 i 幅图像的分类预测目标。

同理, 基于交叉熵设计 Gaussian YOLOX 的置信度损失函数为:

$$L_{\text{obj}} = -\frac{1}{Q} \sum_{i \in \text{pos} + \text{neg}} \sum_{k=1}^K \hat{\pi}_{\text{obj}}^{i,k} \left[y_{\text{obj}}^i \log \left(\frac{1}{1 + \exp(-\hat{o}^{i,k})} \right) + (1 - y_{\text{obj}}^i) \log \left(\frac{\exp(-\hat{o}^{i,k})}{1 + \exp(-\hat{o}^{i,k})} \right) \right] \quad (4)$$

式中: Q 为 SimOTA^[14] 在前景目标特征点匹配过程中计算置信度的正负样本数量; y_{obj}^i 为置信度预测目标; $\hat{o}^{i,k}$ 的处理方式与式 (3) 类似, 即 $\hat{o}^{i,k} = \hat{\mu}_{\text{obj}}^k + \sqrt{\hat{\nu}_{\text{obj}}^k} \gamma$ 。

综上, Gaussian YOLOX 的总体损失函数为:

$$L = L_{\text{reg}} + L_{\text{cls}} + L_{\text{obj}} \quad (5)$$

1.3 不确定度计算

本文采用 IOU 的方差 $\hat{\nu}_{\text{reg}}$ 描述图像的定位不确定度, 采用边缘置信度描述图像的分类不确定度。计算图像中目标的定位方差、位置信息、类别概率和置信度的公式分别为:

$$\nu_{\text{reg}} = \sum_{k=1}^K \hat{\pi}_{\text{reg}}^k \sigma(\hat{\nu}_{\text{reg}}^k) \quad (6)$$

$$\text{reg} = \sum_{k=1}^K \hat{\pi}_{\text{reg}}^k \hat{\mu}_{\text{reg}}^k \quad (7)$$

$$\text{cls} = \sum_{k=1}^K \hat{\pi}_{\text{cls}}^k \sigma(\hat{\mu}_{\text{cls}}^k) \quad (8)$$

$$\text{obj} = \sum_{k=1}^K \hat{\pi}_{\text{obj}}^k \sigma(\hat{\mu}_{\text{obj}}^k) \quad (9)$$

式中: σ 为 sigmoid 函数; ν_{reg} 为目标预测框的定位

不确定度; reg 为预测框的中心坐标、宽度及高度; cls 为预测框中目标的预测类别概率; obj 为预测框中存在目标的概率。

由于网络模型无法直接输出分类的概率分布,因此需将分类输出映射到概率空间。对于某个目标框,其对应的类别预测概率可表示为 $\{Prob_n^c | c = c_1, c_2, \dots, c_I\}$,其中 n 为目标预测框的序号, I 为类别的总数量。分类输出映射到概率空间的计算公式为:

$$Prob_n^c = \frac{\sum_{k=1}^K \pi_{cls}^{i,k} \sigma(\mu_{cls}^{i,k})}{\sum_{i=1}^I \sum_{k=1}^K \pi_{cls}^{i,k} \sigma(\mu_{cls}^{i,k})} \quad (10)$$

式中: σ 为sigmoid激活函数; K 为MDN的数量。

目标的分类不确定度可表示为:

$$U_{cls}(X) = \sum_{k=1}^K \sum_{n=1}^{N_b} 1 - (Prob_n^{k,c_1} - Prob_n^{k,c_2}) \quad (11)$$

式中: c_1 和 c_2 为某预测框分类预测概率最大的两个类别; N_b 为目标预测框数量; X 为未标注的输入图像。

相应地,该目标的定位不确定度可表示为:

$$U_{reg}(X) = \sum_{k=1}^K \sum_{n=1}^{N_b} \pi_{reg}^{k,n} \sigma(\nu_{reg}^{k,n}) \quad (12)$$

未标注图像的总不确定度得分为:

$$U(X) = \alpha \cdot U_{cls}(X) + (1 - \alpha) \cdot U_{reg}(X) \quad (13)$$

式中: α 为权值,默认取值为0.5。

若目标预测框的最终得分大于设定的阈值,则判定该特征点对应的区域为目标区域。目标预测框最终得分 s_{score} 的计算公式为:

$$s_{score} = \sum_{k=1}^K \hat{\pi}_{cls}^k \sigma(\hat{\mu}_{cls}^k) * \sum_{k=1}^K \hat{\pi}_{obj}^k \sigma(\hat{\mu}_{obj}^k) \quad (14)$$

式中: σ 为sigmoid激活函数。

主动学习是一个周期循环、迭代采样的过程,为了使视觉目标检测模型达到预期的性能,需要不断选出对模型训练“贡献大”的未标注样本,因此样本选择的优劣将直接影响模型的性能。本文以未标注图像中每个目标的分类不确定度和定位不确定度之和作为图像的总不确定度,在每个主动学习周期结束后,通过前一轮训练得到的模型对剩余的未标注图像进行前向推理,获取未标注图像的总不确定度,然后计算不确定度得分并排序,保留不确定度排序前 p 幅图像,标注后加入训练集。

1.4 算法描述

Gaussian YOLOX 的伪代码描述如下:

输入:未标注样本池 $D_U(X)$;主动学习每周期末未标注图像采样数量 p ;标注专家Oracle;前景目标阈值 th (默认取0.5)。

输出:最终的模型和性能。

初始化:包含少量标注样本的样本池 $D_L(X, Y)$ 。

repeat:

(1)用已标注样本池 $D_L(X, Y)$ 训练模型。

(2)选择步骤(1)中性能最优的模型,在未标注样本池中进行前向推理,得到每个未标注样本的预测框信息 reg 、回归方差 v_{reg} 、分类概率 cls 、置信度 obj 。

(3)对步骤(2)中前向推理的结果进行非极大值抑制处理,将未标注样本中保留的预测框放入正样本集,即 $pos \leftarrow NMS(v_{reg}, cls, reg, obj)$ 。

(4)依据式(8)(9)计算预测框的分类得分和置信度得分,代入式(14)的sigmoid函数,保留大于前景目标阈值 th 的预测框,即 $\sigma(cls[pos]) * \sigma(obj[pos]) \geq th$ 。

(5)计算保留预测框的分类得分,根据式(10)映射到概率空间,再根据式(11)计算分类不确定度 $U_{cls}(X)$ 。

(6)依据式(12)计算保留预测框的回归不确定度 $U_{reg}(X)$ 。

(7)依据式(13)计算图像的总不确定度得分 $U(X)$ 。

(8)将未标注样本池中的未标注样本按对应的不确定度得分从大到小排序,保留不确定度最大的前 p 个未标注样本 $X_1, X_2, \dots, X_p$ 。

(9)对 p 个样本进行专家标注:

$(X_i, Y_i) | i \in (1, 2, \dots, p) \leftarrow Oracle(X_1, X_2, \dots, X_p)$

(10)更新数据集:

$D_L(X, Y) \leftarrow D_L(X, Y) \cup (X_i, Y_i), i \in (1, 2, \dots, p)$

$D_U(X) \leftarrow D_U(X) \setminus (X_i), i \in (1, 2, \dots, p)$

until 模型满足预定的性能或终止条件。

2 实验与分析

在分类分支中,均值和方差卷积核的输出通道均设为20。在置信度分支中,均值和方差卷积核的输出通道均为1。在所有分支中,权重卷积核的输出通道均为1,所有卷积核的步长也均为1。对比实验分为监督学习和主动学习两部分。

2.1 实验环境

操作系统为 Windows10;处理器为 i7-10700F (2.9 GHz),内存容量为 32 G,GPU 型号为 RTX3060 (12 G);深度学习框架为 Pytorch1.9、CUDA11.1,软件编程环境为 PyCharm。

2.2 数据集和评价指标

实验采用 PASCAL VOC 数据集^[18],包含人、猫、狗等 20 个类别。其中,VOC2007 包含 5 011 幅图像,VOC2012 包含 11 540 幅图像,采用 VOC2007 的测试集评价模型的性能。实验使用的性能评价指标为 AP_{50} ,即当 IOU 的阈值取 50% 时,计算平均查准率(mean average precision, mAP)^[19,20]。

2.3 实验

2.3.1 实验参数设置

Gaussian YOLOX 模型的初始学习率(learning rate,lr)设为 0.001,其下降策略与 YOLOX 一致,初始学习率下降比例取 5%,后续根据损失值的变化情况自适应调整,逐渐降低学习率以提高训练的稳定性。训练所选样本数(batch_size)取 8,模型训练轮数(epoch)取 50。Gaussian YOLOX 的分类、回归、置信度分支的 MDN 数量 K 取 4。主动学习的采样周期设为 9,每周期采样图像数量 p 为数据集的 6%,因此 9 个周期的图像累积采样数量占比依次为 6%、12%、18%、24%、30%、36%、42%、48% 和 54%。在第 1 个迭代周期中,将未标注数据中采样不确定度最大的前 p 幅图像加入训练集重新训练,后续周期以此类推。

2.3.2 对比实验

对比实验是在所有超参数设置相同的条件下于 VOC 数据集^[18]上完成的,实验结果为在相同的随机数种子下训练 3 次的平均值,如表 1 所示,可以看出,当模型训练迭代 10 次时,Gaussian YOLOX 比 YOLOX(监督学习)的 AP_{50} 性能提高 1.5 百分点,在整个迭代训练过程中,Gaussian YOLOX 性能始终优于 YOLOX,表明通过估计检测网络输出特征图的概率分布进行分类和定位预测是有效的。

计算不确定度(uncertainty)的方法通常有多种。本文将提出的 Gaussian YOLOX 不确定性采样方法与 5 种典型的主动学习采样方法进行比较,包括随机(Random)不确定性采样^[21]、熵(En-

表 1 Gaussian YOLOX 与 YOLOX 的 AP_{50} 性能比较

迭代次数	10	20	30	40	50
YOLOX	74.0	78.3	79.5	82.5	82.7
Gaussian YOLOX	75.5	79.2	80.3	82.7	83.1

ropy)不确定性采样^[11,22]、最小置信度(L least confidence, LC)采样^[11,22]、边缘置信度(Margin confidence, MC)采样^[11,23]和核心集(Core-set)采样^[7,24]。上述主动学习采样方法在 github 官网均开源。

上述 5 种主动学习采样方法和本文提出的 Gaussian YOLOX 采样方法均通过 YOLOX 目标检测模型实现,结果如图 3 所示。为保证实验的可比性,所有方法均设置相同的阈值。由图 3 可以看出,当数据标注量为 6% 时,Gaussian YOLOX 的采样性能明显优于其他采样方法。在迭代采样过程中,Gaussian YOLOX 的采样性能始终优于其他 5 种不确定性采样方法,其中随机不确定性采样方法的性能最差。

主动学习是一个周期迭代的采样过程,目的是筛选出对当前目标检测模型最有效的无标注样本。在数据标注量逐渐增加的过程中,将上述几种主动学习不确定度采样方法,在不同标注数据量下与 YOLOX 监督学习(100% 标准数据)的性能进行比较,实验结果如图 4 所示,其中纵坐标为主动学习采样方法与 YOLOX 监督学习 AP_{50} 性能的比率。当数据标注量为 30% 时,Random 采样方法的性能达到 YOLOX 监督学习的 89.6%,而 Gaussian YOLOX 达到 YOLOX 监督学习的

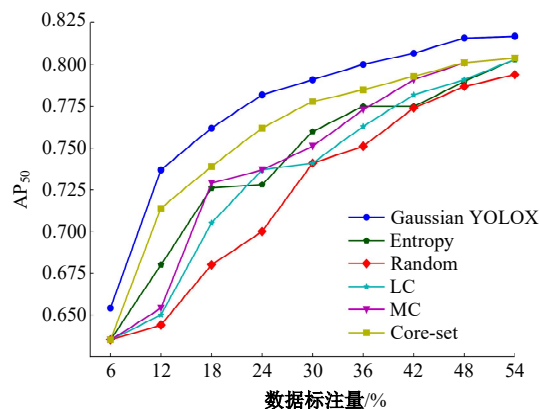


图 3 不同采样方法的性能比较

Fig. 3 Performance comparison of several sampling methods

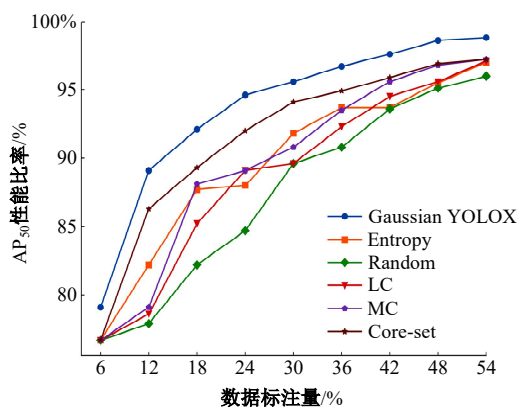


图 4 不同采样方法与监督学习性能比较

Fig. 4 Comparison between under different sampling methods and supervised learning performance

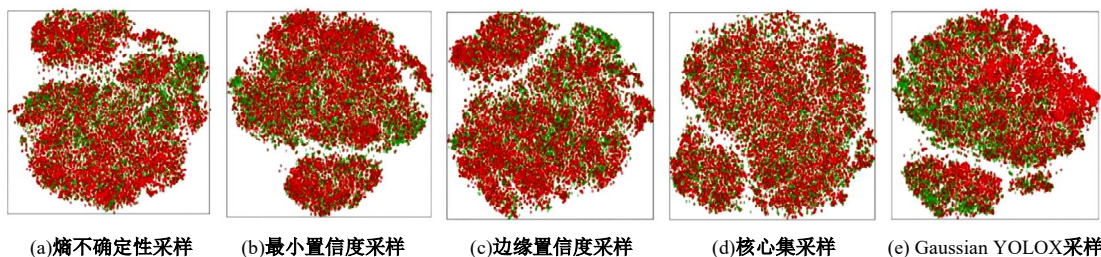


图 5 不同采样方法的数据分布情况对比

Fig. 5 Data distribution comparison under different sampling methods

3 结束语

视觉目标检测是车辆自动驾驶的关键技术之一,主动学习旨在以最小的训练代价有效地提高视觉目标检测模型的性能。在标注样本数量相同的条件下,主动学习采样方法选取效果优于随机选取。本文改进了未标注图像不确定度的估计方法,在此基础上提出了融合深度主动学习的目标检测模型。该模型采用MDN作为检测分支,基于高斯混合分布改进模型的回归、分类和置信度损失函数。在PASCAL VOC数据集上的实验证明了本文方法的有效性。当数据标注量为30%时,Gaussian YOLOX的性能已经达到了YOLOX监督学习的95.6%;使用54%的数据标注量便达到了YOLOX监督学习98.8%的性能。

参考文献:

- [1] 曲优,李文辉. 基于锚框变换的单阶段旋转目标检测方法[J]. 吉林大学学报:工学版, 2022, 52(1): 162-173.
- Qu You, Li Wen-hui. Single-stage rotated object detection network based on anchor transformation[J].

95.6%。最终,Gaussian YOLOX主动目标检测模型仅使用54%的标注数据,便实现了YOLOX监督学习98.8%的性能,说明本文提出的Gaussian YOLOX不确定度采样方法具有有效性。

2.4 可视化分析

采用t-SNE^[25]将样本数据映射到二维空间,以观察被选样本的分布情况。如图5所示,红色点代表数据集中被采样的图像,绿色点代表未被采样的图像。由图5可以看出,除Gaussian YOLOX采样方法外,其他采样方法的样本数据可视化结果更接近均匀分布;而Gaussian YOLOX采样方法有效规避了随机采样的盲目性,能够选取有“价值”的查询样本。

Journal of Jilin University (Engineering and Technology Edition), 2022, 52(1): 162-173.

- [2] 吴伟宁,刘扬,郭茂祖,等. 基于采样策略的主动学习算法研究进展[J]. 计算机研究与发展, 2012, 49(6): 1162-1173.
- Wu Wei-ning, Liu Yang, Guo Mao-zu, et al. Advances in active learning algorithms based on sampling strategy[J]. Journal of Computer Research and Development, 2012, 49(6): 1162-1173.
- [3] 李延超,肖甫,陈志,等. 自适应主动半监督学习方法[J]. 软件学报, 2020, 31(12): 3808-3822.
- Li Yan-chao, Xiao Fu, Chen Zhi, et al. Adaptive active learning for semi-supervised learning[J]. Journal of Software, 2020, 31(12): 3808-3822.
- [4] Roy S, Unmesh A, Namboodiri V P. Deep active learning for object detection[C]//British Machine Vision Conference. Northumbria, UK: BMVC, 2018: 1-12.
- [5] Gal Y, Ghahramani Z. Bayesian convolutional neural networks with bernoulli approximate variational inference[J]. Computer Science, 2015, 12: 1-12.
- [6] Elhamifar E, Sapiro G, Yang A, et al. A convex optimization framework for active learning[C]//Proceed-

- ings of the IEEE International Conference on Computer Vision. Piscataway, NJ: IEEE, 2013: 209-216.
- [7] Paul R, Feldman D, Rus D, et al. Visual precis generation using coresets[C]//Proceedings of the 2014 IEEE International Conference on Robotics and Automation. Piscataway, NJ: IEEE, 2014: 1304-1311.
- [8] 张新生, 高新波, 王颖, 等. 乳腺图像微钙化簇主动学习检测新方法[J]. 西安电子科技大学学报, 2008(5): 871-877.
- Zhang Xin-sheng, Gao Xin-bo, Wang Ying, et al. New method for microcalcification clusters detection using active learning in the mammogram[J]. Journal of Xidian University, 2008(5): 871-877.
- [9] Lakshminarayanan B, Pritzel A, Blundell C. Simple and scalable predictive uncertainty estimation using deep ensembles[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. New York: Curran Associates Inc., 2017: 6405-6416.
- [10] Kao C C, Lee T Y, Sen P, et al. Localization-aware active learning for object detection[C]//Asian Conference on Computer Vision. Heidelberg, Germany: Springer, 2018: 506-522.
- [11] Settles B. Active learning literature survey[R]. Madison: University of Wisconsin-Madison, 2010.
- [12] 吕佳, 傅屈寒. 基于改进主动学习和自训练的联合算法[J]. 北京师范大学学报:自然科学版, 2022, 58(1): 25-32.
- Lyv Jia, Fu Qu-han. A joint algorithm by combined improved active learning and self-training[J]. Journal of Beijing Normal University (Natural Science), 2022, 58(1): 25-32.
- [13] 徐艳. 基于主动学习的图像标注方法研究[D]. 锦州: 辽宁工业大学信息学院, 2014.
- Xu Yan. Research on image annotation method based on active learning[D]. Jinzhou: School of Information, Liaoning University of Technology, 2014.
- [14] Ge Z, Liu S T, Wang F, et al. YOLOX: Exceeding YOLO series in 2021[J/OL]. [2021-07-18]. <https://arxiv.org/abs/2107.08430>
- [15] Bishop C M. Mixture density networks[EB/OL]. [2023-03-01]. https://publications.aston.ac.uk/id/eprint/373/1/NCRG_94_004.pdf
- [16] Choi S, Lee K, Lim S. Uncertainty-aware learning from demonstration using mixture density networks with sampling-free variance modeling[C]// IEEE International Conference on Robotics and Automation. Piscataway, NJ: IEEE, 2018: 6915-6922.
- [17] Choi J, Elezi I, Lee H J. Active learning for deep object detection via probabilistic modeling[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway, NJ: IEEE, 2021: 10264-10273.
- [18] Everingham M, Van G L, Williams C K I. The pascal visual object classes (VOC) challenge[J]. International Journal of Computer Vision, 2010, 88(2): 303-338.
- [19] Lin T Y, Maire M, Belongie S. Microsoft COCO: common objects in context[C]//Proceedings of the European Conference on Computer Vision. Heidelberg, Germany: Springer, 2014: 740-755.
- [20] Yuan T N, Wan F, Fu M Y, et al. Multiple instance active learning for object detection[C]//Proceedings of the Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2021: 5326-5335.
- [21] Liu W, Anguelov D, Erhan C, et al. SSD: Single shot multibox detector[C]//Proceedings of the European Conference on Computer Vision. Heidelberg, Germany: Springer, 2016: 21-37.
- [22] 杨文柱, 田潇潇, 王思乐, 等. 主动学习算法研究进展[J]. 河北大学学报:自然科学版, 2017, 37(2): 216-224.
- Yang Wen-zhu, Tian Xiao-xiao, Wang Si-le, et al. Recent advances in active learning algorithms[J]. Journal of Hebei University (Natural Science Edition), 2017, 37(2): 216-224.
- [23] Buur S. Active Learning[M]. California: Morgan & Claypool Publishers, 2012.
- [24] Sener O, Savarese S. Active learning for convolutional neural networks: A core-set approach[C]//Proceedings of International Conference on Learning Representation. Vancouver, Canada: ICLR Press, 2018: 1-13.
- [25] Van D M L, Hinton G. Visualizing data using t-SNE[J]. Journal of Machine Learning Research, 2009, 9(11): 2579-2605.