

内河航行的船舶视觉轨迹预测

A PREPRINT

亚历山大·普齐查^{*1}, 康斯坦丁·乌斯特费尔德¹, Kathrin Wilms², and 弗兰克·维希特²

¹Chair of Modeling and Simulation, Computer Science, TU Dortmund University, Germany

²Chair of Computer Graphics, Computer Science, TU Dortmund University, Germany

2025年5月4日

ABSTRACT

未来内河航运越来越依赖于自主系统和远程操作，强调了准确预测船舶轨迹的必要性。本研究通过整合先进的目标检测方法、卡尔曼滤波器和基于样条插值的方法来解决基于视频的船舶跟踪与预测挑战。然而，现有的检测系统经常由于复杂的环境而误分类内河水道中的物体。对包括 BoT-SORT、Deep OC-SORT 和 ByeTrack 在内的跟踪算法进行比较评估，突显了卡尔曼滤波器在提供平滑轨迹方面的稳健性。来自不同场景的实验结果表明，在预测船舶运动方面有所改善，这对于避碰和态势感知至关重要。研究发现强调了为内河导航定制数据集和模型的必要性。未来的工作将扩展数据集并纳入船舶分类以完善预测，支持自主系统和人类操作员在复杂环境中运作。

Keywords 卡尔曼滤波器 · 多目标轨迹预测 · 多目标检测 · 多目标跟踪；自主船只 · 远程操作船 · 类激活图

1 介绍

内河船舶的未来导航越来越依赖远程操作员和自主系统。在这种情况下，操作员可以访问视频流，但缺乏关键的空间和环境深度信息，这使得决策变得复杂。准确预测船舶轨迹对于增强态势感知、确保安全操纵和支持自主导航至关重要。

在这项研究中，我们通过整合先进的对象检测和跟踪方法与专门为内河船舶动态行为定制的预测模型来应对这些挑战。我们的重点是利用计算机视觉技术结合卡尔曼滤波器和基于样条插值的方法分析和预测船舶轨迹。这些工具旨在克服传统检测系统的局限性，例如由于传感器配置错误导致自动识别系统（AIS）数据不准确的问题，并提供强大的跟踪和预测能力。

基于现代目标检测和先进跟踪算法，我们研究了水道上内河和海船的多目标跟踪与预测的准确性。

本研究通过对船舶检测和轨迹预测中的空白进行补充，为内河航道导航的研究领域做出了贡献。预期结果将提高人工操作和自主航行船舶的安全性和效率。

*CONTACT Alexander Puzicha. Email: alexander.puzicha@cs.tu-dortmund.de

2 文献回顾

正如近期研究所显示的 [13, 20, 28, 36], 基于摄像头的系统对于未来遥控和全自主船舶在海洋和内陆水道上的发展是必不可少的。它们的任务主要集中在监视上, 例如检测船只和其他物体、追踪以及分类, 以替代或支持船长进行避碰操作。船舶检测与追踪包括每帧的船舶检测、基于对象的分类以区分不同的实例, 以及跨多帧的船舶追踪。存在一些通用的追踪方法 [26] 来解决这些任务, 并且不需要对航运问题和特定情况进行特别适应。在航运领域, 也区分海洋航运和内河航运 [13]。

在目标检测领域, 可以根据训练数据和标签信息识别出各种方法, 这些方法使用不同程度的标签信息: 监督学习 [18]、半监督学习 [32]、少量样本学习 [16]、弱监督学习 [25] 以及无监督学习 [41]。可用的模型架构主要分为两大类: 两阶段和单阶段目标检测器, 依赖于基于 CNN 或变换器的目标检测模型 [37]。一个单阶段目标检测器的示例是 YOLO (You Only Look Once) 架构 [6]。这一系列网络 [27] 遵循了另一种建议边界框的方法范式。而在两阶段方法中, 提出的区域是由几个固定锚区域的分类结果 [23]。YOLO 应用于特征图上的每个网格单元都会提出边界框作为一个回归任务。通过避免多阶段分析, 这样的网络达到了更高的计算效率 [27]。基于 Transformer 的目标检测是基于 Detection Transformer (DETR) 的基本架构 [3]。DETR 网络依赖于编码器-解码器 transformer 架构 [29] 和二分图匹配损失来简化边界框预测, 使得像其他方法中使用的固定锚或单元区域变得不再必要。其中一个优势是避免了在确定锚框时因可能的大小、位置或分辨率的具体规定而产生的隐性偏差, 并且其结果声称与 Faster R-CNN [4] 相当。RT-DETR [40] 是基本 DETR 的继任者, 在大多数情况下保持其检测性能, 同时声称在实时任务中击败了 YOLO 网络。一个已知的缺点是在小物体情况下的表现较差 [4], 这可能导致船只远离相机时的限制。常见的两阶段探测器是 Faster R-CNN (带有 CNN 特征的更快区域) [23]。Faster R-CNN 通过将其划分为专业子网络来保证精确的对象检测, 但尽管它是 R-CNN 家族中最快的, 它在速度上仍不及 YOLO 和 RT-DETR [27, 40]。

多目标跟踪 (MOT) 的发展方式与检测类似, 因为 MOT 算法的标准方法是通过检测来进行跟踪。对于详细概述, 请参阅相应的综述论文 [5, 10]。然而, 针对船舶的技术发展远远落后于其他物体, 如车辆的检测和跟踪 [34] 或交通标志识别 [42]。尽管如此, 船舶检测仍涉及与通用对象识别相当的挑战, 例如光照、遮挡、多种类别以及类内的变化。此外, 在这个领域的发展往往没有遵循系统的模式。因此, 只有少数综述出版物处理使用深度学习方法确定船只运动的问题, 例如 [7, 2, 24] 或特定可用的数据集 [38]。最先进的技术主要以个别解决方案为特征。我们在评估中选择比较的常用跟踪算法示例有 Deep OC-SORT [19]、BoT-SORT [1] 和 ByteTrack [39]。

Deep OC-SORT [19] 特别专注于通过结合重新识别网络 [11] 来改善相似物体相互遮挡的结果。虽然许多跟踪方法首先丢弃置信度低的单帧检测, ByteTrack [39] 也包括那些置信度低的对象以实现更高的召回率, 通过条目之间的更多关联可能性来达成此目标。BoT-SORT [1] 展现了与前述方法的相似之处: 它使用类似 ByteTrack 的高低置信度检测区分, 并显示出类似于 Deep OC-SORT 的双路径结构。

领先的公司 (例如 SeaAI) 用于海洋船舶检测² 采用两阶段方法来检测并随后分类漂浮物体。他们的系统在 inland 水道上显示出较高的误报率, 因为它被训练用来在空旷的海域中检测每一个漂浮物, 而不是从障碍区域中提取船只。远程航行公司用于 inland 水道 (例如 Seafar³) 使用 YOLO 进行船舶检测。先前关于船只检测和追踪的研究, 例如 [14], 使用改进的 YOLOv3 [8]、深度简单在线实时跟踪 (Deep SORT [35]) 算法。其他基于机器学习的船舶检测和追踪方法利用卫星图像进行海事监视 [33], 由于船对船的角度不同, 与所提出的方案有很大差异。在内河航行场景中, 相对较长的船只接近相机, 因此通常只能观察到部分船只。而在卫星图像中, 几乎总是显示整个船只, 并且大多数船只的顶视图形状相似。这种视角对于侧面并不适用。

²Robosys 和 SEA.AI 创建了基于机器视觉的海上自主水面碰撞避障系统, <https://www.maritimeindustries.org/news/robosys-and-seaai-create-machine-vision-based-maritime-autonomous-surface-collision-avoidance-system> [最后引用日期: 2025-01-05]

³航海者, <https://seafar.eu> [最后引用日期: 2025 年 5 月 1 日]



图 1: 用于检测、跟踪和预测的参数可视化，以一艘向前移动的船只为例。边界框 \mathbf{b}_i^k ，框中心 \mathbf{q}_i^k ，用于跟踪的当前锚点 \mathbf{p}_i^k ，过去帧的轨迹（绿色）和预测轨迹（橙色），以及当前帧索引 k 和对象索引 i 对应的采样时间。值 $j_1 > 0$ 和 $j_2 > 0$ 用于说明时间点。

3 方法论

以下部分详细说明了我们的预测船舶轨迹的方法。在第 3.1 节中，我们描述了检测和跟踪船舶的方法，在第 3.2 节中，我们概述了预测方法。作为对以下子部分的概述，图 1 通过一个示例图像说明了在检测、跟踪和预测中使用的参数。对于每个帧 k ，对象 i 由一个边界框 \mathbf{b}_i^k 表征。为每个边界框确定了一个锚点 \mathbf{p}_i^k 。锚点 \mathbf{p}_i^0 到 \mathbf{p}_i^k 用于跟踪（绿色轨迹）。锚点 \mathbf{p}_i^k 到 $\mathbf{p}_i^{k+j_2}$ 使用基于先前值（橙色轨迹）的样条进行预测。

3.1 语义跟踪

预测物体运动需要两个预备步骤：在视频的各个帧上检测任意数量的物体，以及跟踪，即将在连续帧上检测到的物体分配给一个随时间移动的物体。尽管相对于常见的帧率而言，船只位置的变化通常很小，但相机受到外部影响可能导致整个图像出现颤抖运动，特别是在录制相机放置在船上时。因此，在检测过程中应考虑这一点和其他类型的伪影。

3.1.1 多对象检测

我们将图像检测函数定义为

$$\begin{aligned} \mathcal{B}(\mathbf{I}_k) &= \left\{ \left(\mathbf{b}_i^k, c_i^k, e_i^k \right) \mid i \in \{1, \dots, n_k\} \right\} \\ &= \{ \mathcal{D}_i \mid i \in \{1, \dots, n_k\} \} \end{aligned} \quad (1)$$

该函数指定了在输入视频的第 k 帧中的每次 n_k 检测的边界框 \mathbf{b} 、类别 c 和置信度 e 。

在此上下文中，边界框 $\mathbf{b}_i^k = (\mathbf{q}_i^k, w_{k,i}, h_{k,i})$ 标记了中心 \mathbf{q}_i^k 和最小包含区域 i -th 检测的扩展 (w_i^k, h_i^k) ，在 k -th 帧中。虽然目标是用一个检测精确描述每个对象，但并非每个对象都一定分配了一个边界框。根据具体的检

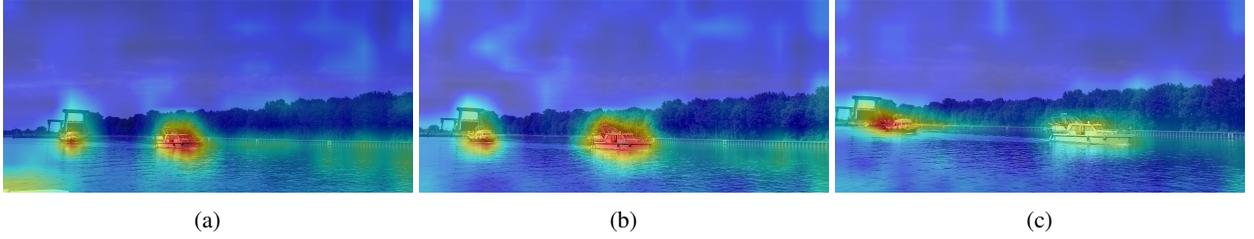


图 2: 图像区域的相关性可视化, 用于将对象分配到船类, 使用在 Microsoft Common Objects in Context (COCO) 数据集上训练的 Faster R-CNN 检测器 [17]。根据 Score-CAM [31] 方法的结果值, 用红色叠加各自的原始图像表示相关性较高, 黄色表示中等相关性, 蓝色表示较低的相关性。这三张图像是来自同一视频的连续帧, 其中船只正在向前移动。请注意轻微的相机运动, 由于视角变化, 这使得跟踪变得更加复杂。

测方法, 可能会有不同的边界框被分配或者根本就没有分配。在显著重叠的边界框这种特殊情况下, 相应的检测可以合并为一个, 例如通过平均来自每次检测的值来融合。在图 2 中, 使用了类激活映射 (CAM) 技术来解释 CNN 检测 [12]。

3.1.2 多目标跟踪

一种视频中多目标跟踪的算法 $V = (I_0, \dots, I_J)$ 可以抽象为算法 1 所描述的一般形式。作为一种简化, 我们假设被追踪的对象, 在这种情况下是船只, 不会分裂或合并。一个具体的多跟踪算法实现并可能扩展这些抽象步骤, 提供专门的函数 h 和 g : $h(D, \mathcal{T})$ 确定检测 D 属于轨道 \mathcal{T} 的可能性, 而 $g(\mathcal{T})$ 是一个删除标准, 如果 \mathcal{T} 被视为无效的跟踪, 则为真。

Algorithm 1 多目标跟踪方法的抽象算法。

Input: 视频 $V = (I_0, \dots, I_J)$, 逐帧检测 \mathcal{B} , 合并过滤器 h , 跟踪删除过滤器 g

```

 $\mathcal{T}_1 = \mathcal{B}(I_1)$ 
for each  $j \in \{2, \dots, J\}$  do
   $\mathcal{T}'_j = \emptyset$ 
  for each  $D \in \mathcal{B}(I_j)$  do
    标志  $\leftarrow 0$ 
    for each  $\mathcal{T} \in \mathcal{T}_{j-1}$  do
      if  $h(D, \mathcal{T}) > h(D', \mathcal{T})$  for all  $D' \in \mathcal{B}(I_j)$  then
        更新  $\mathcal{T}$  通过将其与  $D$  合并
        标志  $\leftarrow 1$ 
      end if
    end for
    if 标志 = 0 then
       $\mathcal{T}'_j \leftarrow \mathcal{T}'_j \cup D$ 
    end if
  end for
   $\mathcal{T}_j \leftarrow \mathcal{T}'_j \setminus \{g(\mathcal{T}) | \mathcal{T} \in \mathcal{T}'_j\}$ 
end for

```

为了获得更有意义的跟踪和可视化位置, 我们从表示船体与水面接触线中心的框 \mathbf{b}_i^k 的坐标计算点 $\mathbf{p}_i^k = \left(\mathbf{q}_i^k - \begin{pmatrix} 0 \\ 0.5 \cdot h_i^k \end{pmatrix} \right)$ 。

3.2 轨迹预测

如引言部分所述（见第 1 节），外部或自主导航需要在早期阶段识别关键情况（例如，碰撞）。因此，本文解决了船舶运动预测问题。这表明，基于轨迹（见第 3.1.2 节） $\mathcal{T}_i = \{\mathbf{p}_i^0, \dots, \mathbf{p}_i^j, \dots, \mathbf{p}_i^k\}$ ， k 当前帧，预期位置 \mathbf{p}_i^j ，对于 $j = k + h$ ，与 $h > 0$ ，需要被预测。由于内河船舶的运动轨迹往往具有明显的平滑性，因此使用三次 Hermite 样条作为插值方法来确定轨迹。这些方法适用是因为有高效的算法可用于计算和分析 [9]。在相关区间 \mathbf{p}_i^l 和 \mathbf{p}_i^r 内的函数通过三次多项式进行插值。一般的三次样条以参数形式给出，根据

$$\mathbf{f}(t) = at^3 + bt^2 + ct + d, \quad (2)$$

其中 $t \in [0, 1]$ [15]。利用 Hermite 样条的约束条件，得到确定系数的如下表示式 [30]：

$$\mathbf{f}(t) = [t^3 \ t^2 \ t \ 1] \begin{bmatrix} 2 & -2 & 1 & 1 \\ -3 & 3 & -2 & -1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{p}_i^l \\ \mathbf{p}_i^r \\ \mathbf{t}_{\mathbf{p}_i^l} \\ \mathbf{t}_{\mathbf{p}_i^r} \end{bmatrix}. \quad (3)$$

这里， $\mathbf{t}_{\mathbf{p}_i^l}$ 和 $\mathbf{t}_{\mathbf{p}_i^r}$ 是点 \mathbf{p}_i^l 和 \mathbf{p}_i^r 处的方向向量，它们由相邻点得出。需要考虑的采样点根据 Chebyshev 条件 [22] 进行。由于内河中的船只通常有比较平滑的轨迹， $l = 0$ 和 $r = k$ 适用。如果情况并非如此， $\mathbf{t}_{\mathbf{p}_i^l}$ 和 $\mathbf{t}_{\mathbf{p}_i^r}$ 之间的角度大于指定误差 ε ，则将轨迹分割成单独的部分，然后对该部分进行插值。样条用于对确定的边界框进行正则化，以便结合检测中的较小变化。此外，样条还用于预测船只未来的移动。如果整个轨迹已经被分成子段，则使用最后一个子段进行预测。这里的一个研究问题是，在预报期间 h 对内河船只和海船的预报质量。

除了样条嵌入外，还集成了卡尔曼滤波器作为第二种鲁棒性机制。这特别适用于船舶检测中的临时问题。此处使用扩展卡尔曼滤波器 (EKF)，其中状态转移和观测模型是非线性的函数，因为无法保证船舶运动（轨迹）的线性。为此，在 EKF 中通过方程 f 对从 \mathbf{p}_i 到 \mathbf{p}_{i+1} 的过渡进行建模，并在预测和校正周期中更新状态变量。EKF 预测步骤的方程如下 [21]：

$$\mathbf{x}_k = f(\mathbf{x}_{k-1}) \quad (4)$$

$$\mathbf{P}_k = \mathbf{A}\mathbf{P}_{k-1}\mathbf{A}^T + \mathbf{Q}. \quad (5)$$

\mathbf{A} 是包含物理模型特性的状态转移矩阵，而 \mathbf{Q} 是状态转移的协方差矩阵，用于建模噪声。在更新步骤中，新值 \mathbf{x}_k 通过计算卡尔曼增益 \mathbf{K}_k 并获取测量值 \mathbf{z}_k 来确定：

$$\mathbf{K}_k = \mathbf{P}_k\mathbf{H}^T (\mathbf{H}\mathbf{P}_k\mathbf{H} + \mathbf{R})^{-1} \quad (6)$$

$$\mathbf{x}_k = \mathbf{x}_k + \mathbf{K}_k(\mathbf{z}_k - h(\mathbf{x}_k)) \quad (7)$$

$$\mathbf{P}_k = \mathbf{P}_k - \mathbf{K}_k\mathbf{H}\mathbf{P}_k. \quad (8)$$

因此，EKF 还可以在检测暂时中断（例如由于遮挡）和误解释（例如由于伪影）的情况下实现跟踪和预测。

4 实验结果

本节展示了我们对不同内陆航行和海洋设置的实验评估，这些设置中待检测、跟踪和预测的船只数量各不相同。

4.1 数据集

我们的数据集由总共九个视频组成。五个视频代表了内河水道，尤其是运河的视角，这些地方的特点是短视距、大部分区域被遮挡以及相对较近的船只。一个视频是从无人机的角度记录的莱茵河的情况。其他四个视频涵盖了北海、大西洋和泰国湾的情景，只展示了船只或小岛。⁴

⁴数据集可在以下位置获取：<https://doi.org/10.5281/zenodo.14243467>

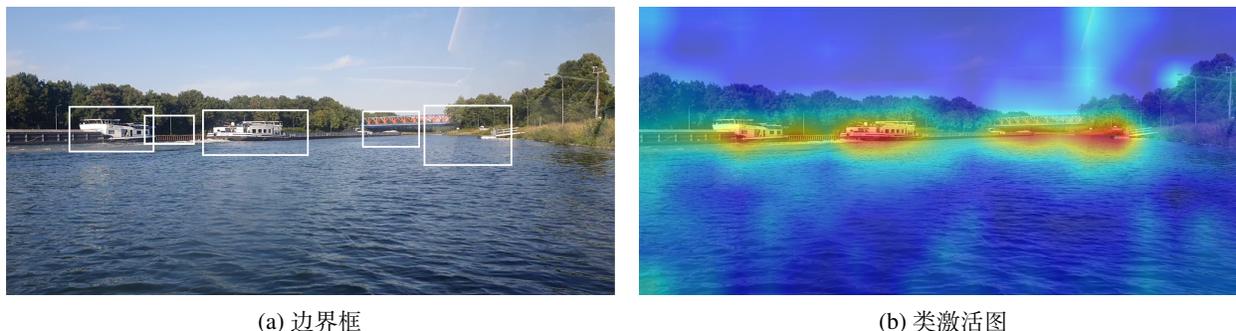


图 3: 类激活图对海岸线上的许多物体显示了高置信度，而对船只的中心部分则显示出低置信度。此外，较长的船只由多个高置信度值表示。

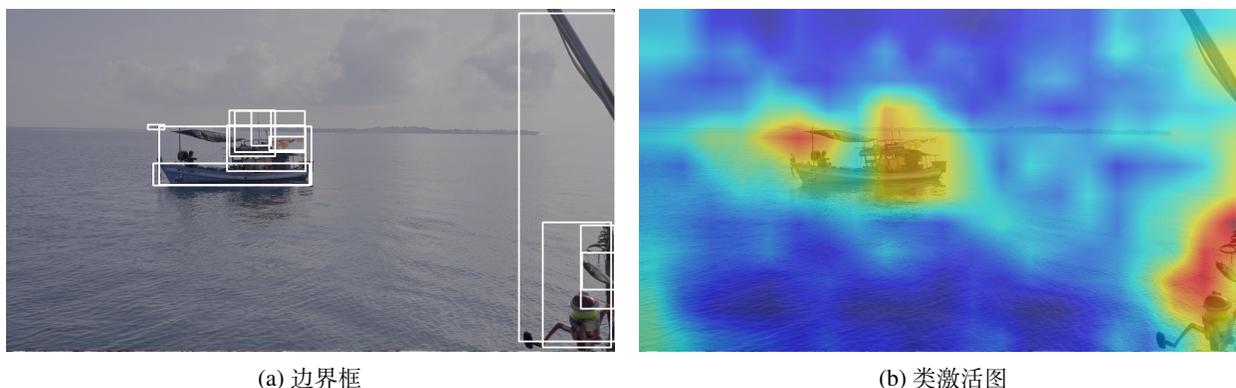


图 4: 类激活图对背景中的钓鱼竿和岛屿呈现出高度的信心。

4.2 分析

标准的 Faster R-CNN 检测模型，该模型是在 Microsoft Common Objects in Context (COCO) 数据集 [17] 上训练的，显示了对内河和海船检测精度较低。然而，内河船只的精度略低。使用 CAM 技术来解释 CNN 检测。CAM 可视化突出显示了许多水面上和岸边的对象，并以较高的置信度进行标注，在图 3 中通过黄色到红色的颜色进行了可视化。另一个海船的例子在图 4 中展示，其中背景中的钓鱼竿和一个小岛获得了最高置信度而不是船只。与精度和 F1 分数相比，召回率的增加表明存在大量的误报检测，这证明了我们关于大量错误分类为船只的检测对象的说法。

图 3 和 4 展示了在内河航行中检测和跟踪船只所面临的部分挑战：前景和背景中的物体，几艘可区分的船只同时出现，较长且不完全可见的船只，以及岸线上大量的细节。这一点也在图 5 中部分表明。对于内河航行，中位数并非中间百分位的平均值。

如图 7、8 和 9 所示，所有三种跟踪算法的跟踪精度几乎相同。因此，作为跟踪结果的预测也几乎是相同的（参见图 10、11 和 12）。在内河船只和海船之间的预测质量上可以观察到显著差异。四个指标，平均距离误差 (ADE)、最终距离误差 (FDE)、平均绝对误差 (MAE) 和均方根误差 (RMSE)，以像素单位表示。视频帧被调整为 1280 x 720，但由于不同摄像机设置，这些像素不能与实际距离相关联。总体而言，错误指标较小且表明预测是有用的。图 6 展示了船只预测轨迹和实际轨迹之间的比较。点表示船只的预测位置，曲线显示实际轨迹。曲线表明卡尔曼滤波器仅在短期内导致偏离预测（例如，见天蓝色曲线）。

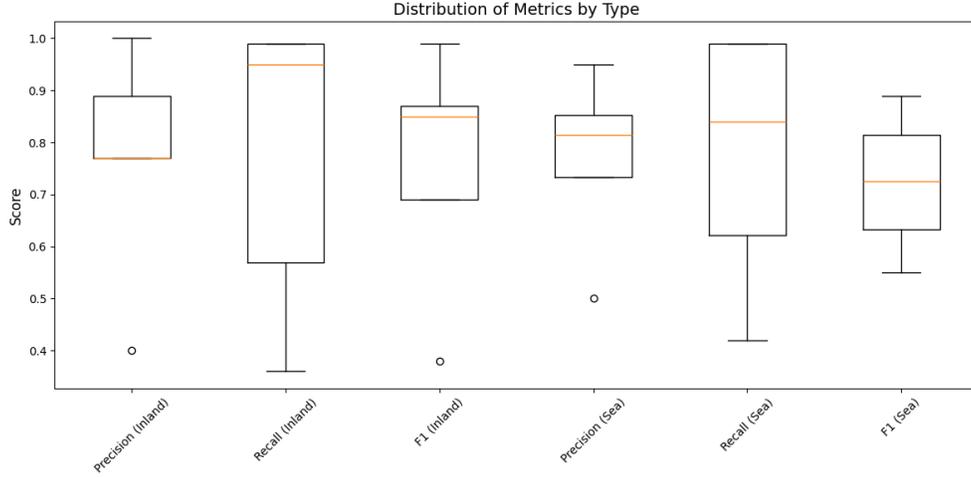


图 5: 内陆船舶与海船检测质量的比较

4.3 讨论

船舶的缓慢移动和短暂的重叠情况导致几乎所有的目标都有良好的跟踪质量。但是，必须考虑到较小船只在被较大船只遮挡时，由于移动缓慢而长时间隐藏的问题。因此，追踪系统需要为每个对象存储更长时间的特征以避免 ID 切换或跟踪丢失。这是与其它车辆跟踪系统的一个主要区别。内河船舶和海船之间预测质量的差异可以基于船只的运动情况来解释。与其他视频主要拍摄在将自身移动传递到视频中的船只上不同，莱茵河上的无人机视频是稳定的。波高和整体条件的不同解释了预测质量的降低。尽管存在这些差异，预测误差指标仍暗示了可能的应用。

检测算法的主要缺点是，它们没有针对船舶特别是内河船舶的检测进行优化。一般来说，所有非极大型水面上的物体，如河流、湖泊或开放海域上的物体，都被分类为船只。这包括固定结构如船队或码头，如图 3 所示。因此，自训练网络对于自主船舶和远程船长辅助系统至关重要。基于卡尔曼滤波器的对象跟踪即使在少量帧缺失检测的情况下也能提供稳定和平滑的轨迹。由于缺乏有关摄像机设置、校准和移动的信息，这些轨迹描述了相对运动，这对于安装固定摄像头的船只上的远程船长或交通控制中心是有用的。为了将信息和预测用于自主系统，例如避碰，基于自身里程计的转换至关重要。此外，这也有助于证明 AIS 数据的正确性。

4.4 结论

我们提出了一种方法，将现有系统结合以追踪和预测内陆水道和公海上的船只。数值结果表明了我们的追踪和预测方法的功能性。然而，该方法主要受低检测精度的影响。因此，揭示出在微软情境中常见对象 (COCO) 数据集 [17] 上训练的最先进的现有检测网络需要调整、训练和微调，尤其是在内陆水道上，因为存在大量的误检船只。所提出的方法及其实施可以与校准好的摄像机设置结合，使船只位置能够精确映射到全球坐标，这是自主船只所需的重要信息。因此，我们将扩展我们的内陆船只数据集以提高检测率，并添加船只分类以突出不同船型对远程船长的潜在风险。这种分类也是识别适用避碰规则所必需的，在仅基于纯 AIS、雷达和激光扫描器数据的情况下是不可能实现这一点的。

参考文献

- [1] Nir Aharon, Roy Orfaig, and Ben-Zion Bobrovsky. Bot-sort: Robust associations multi-pedestrian tracking. *arXiv preprint arXiv:2206.14651*, 2022.

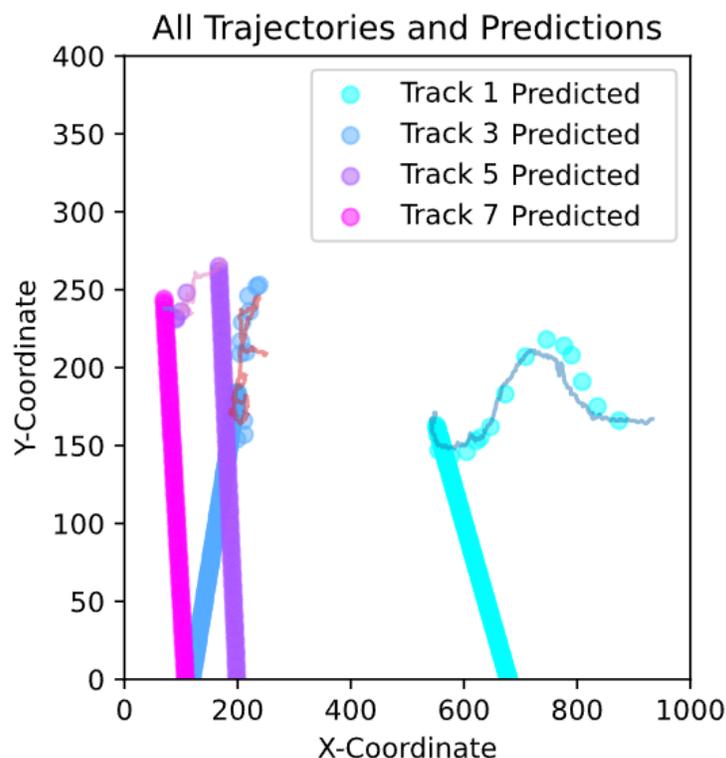


图 6: 比较视频 5 中预测的和实际的船舶轨迹。最初, 预测假设为线性运动, 后来逐渐收敛于实际轨迹。对于底层跟踪, 我们使用了 ByeTrack 跟踪器。

- [2] LI Bo, XIE Xiaoyang, WEI Xingxing, and TANG Wenting. Ship detection and classification from optical remote sensing images: A survey. *Chinese Journal of Aeronautics*, 34(3):145–163, 2021.
- [3] Nicolas Carion, Francisco Massa, Gabriel Synnaeve, Nicolas Usunier, Alexander Kirillov, and Sergey Zagoruyko. End-to-end object detection with transformers. In *European conference on computer vision*, pages 213–229. Springer, 2020.
- [4] Nicolas Carion, Francisco Massa, Gabriel Synnaeve, Nicolas Usunier, Alexander Kirillov, and Sergey Zagoruyko. End-to-end object detection with transformers. In *European conference on computer vision*, pages 213–229. Springer, 2020.
- [5] Gioele Ciaparrone, Francisco Luque Sánchez, Siham Tabik, Luigi Troiano, Roberto Tagliaferri, and Francisco Herrera. Deep learning in video multi-object tracking: A survey. *Neurocomputing*, 381:61–88, 2020.
- [6] Tausif Diwan, G Anirudh, and Jitendra V Tembhurne. Object detection using yolo: Challenges, architectural successors, datasets and applications. *multimedia Tools and Applications*, 82(6):9243–9275, 2023.
- [7] Meng Joo Er, Yani Zhang, Jie Chen, and Wenxiao Gao. Ship detection with deep learning: a survey. *Artificial Intelligence Review*, 56(10):11825–11865, 2023.
- [8] Ali Farhadi and Joseph Redmon. Yolov3: An incremental improvement. In *Computer vision and pattern recognition*, volume 1804, pages 1–6. Springer Berlin/Heidelberg, Germany, 2018.
- [9] Richard Franke and Gregory M Nielson. Scattered data interpolation and applications: A tutorial and survey. *Geometric Modeling: Methods and Applications*, pages 131–160, 1991.

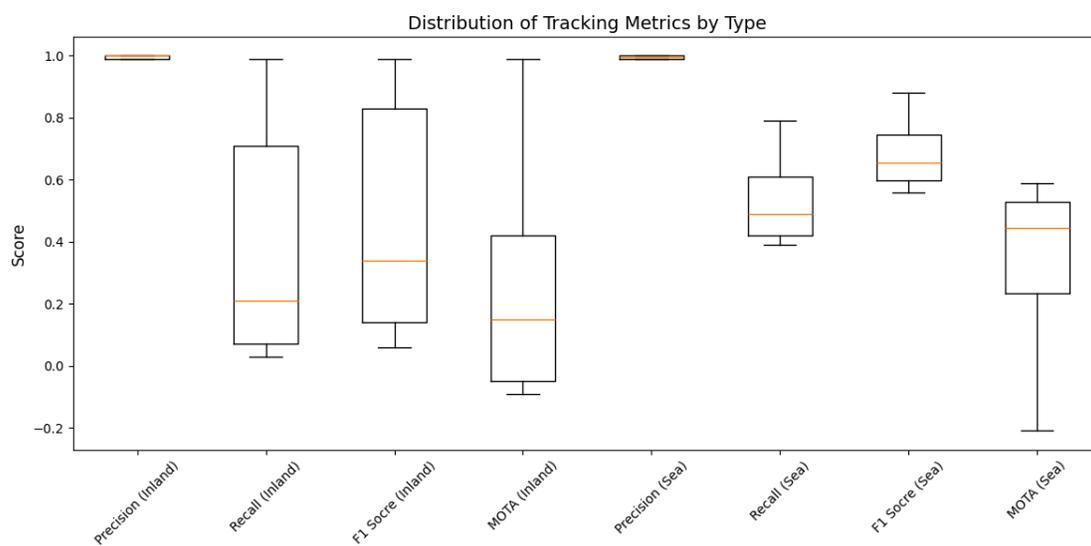


图 7: 内陆与海船之间使用 BoT-SORT 的跟踪质量比较

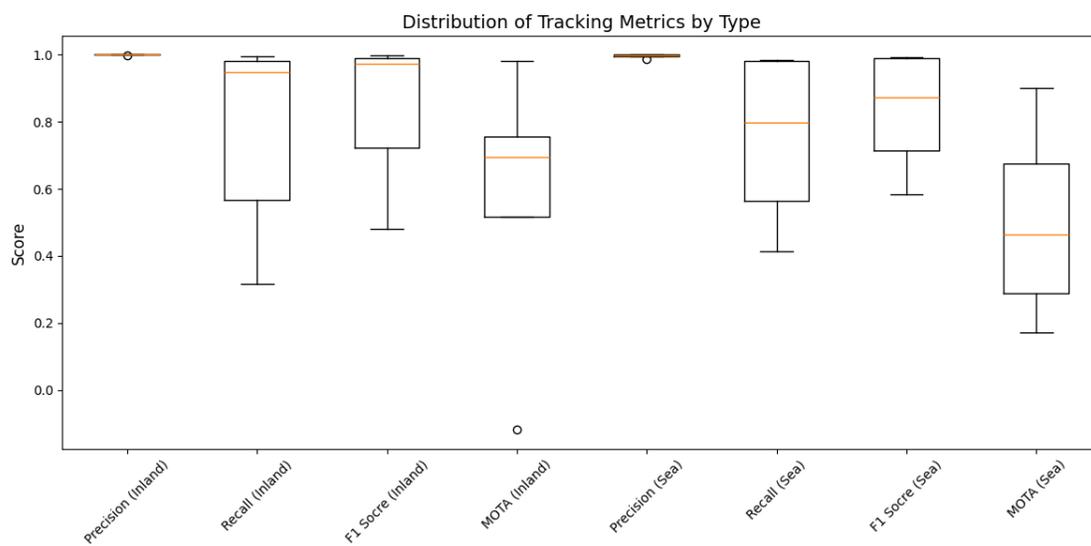


图 8: 内陆与海船追踪质量的对比研究使用 Deep OC-SORT

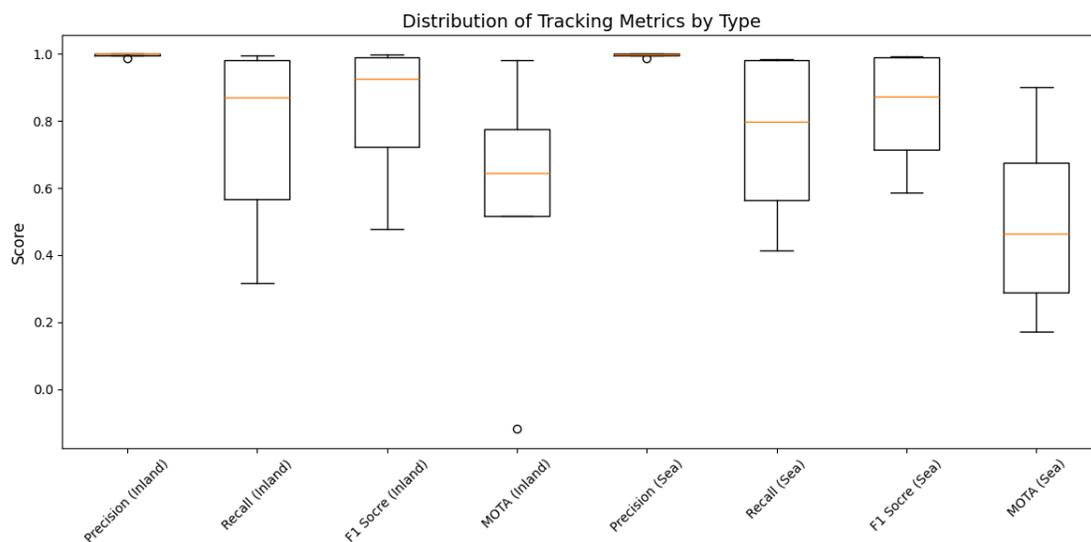


图 9: 内陆与海船追踪质量的对比分析使用 ByeTrack

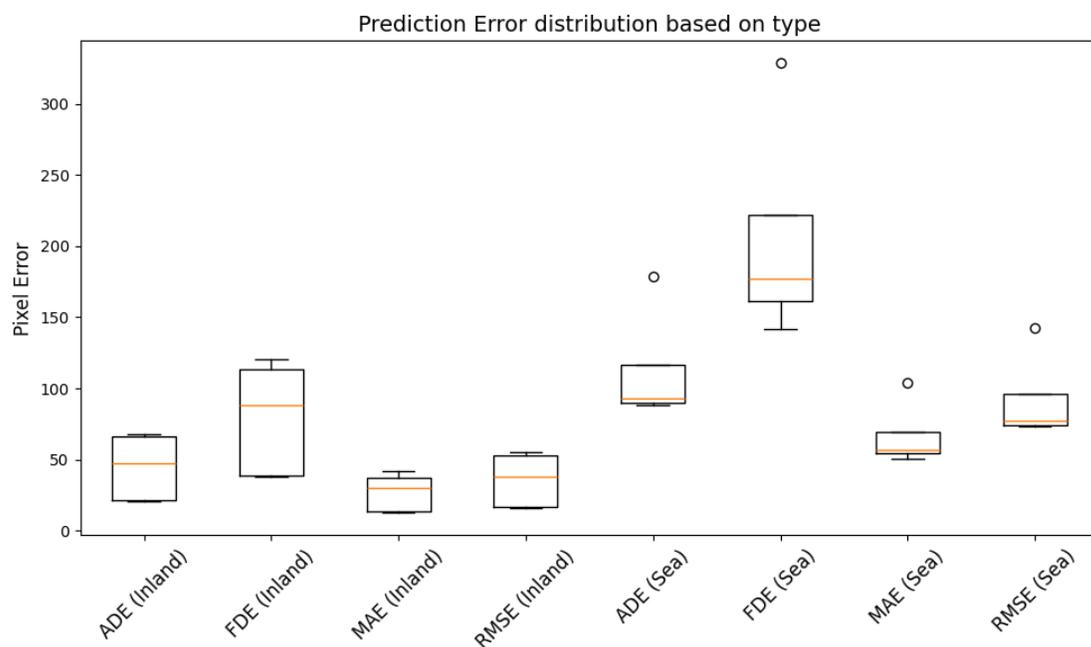


图 10: 内陆船与海船预测质量的比较使用 BoT-SORT。我们展示了视频流类别中的平均距离误差 (ADE)、预测的最终距离误差 (FDE)、平均绝对误差 (MAE) 和均方根误差 (RMSE)。

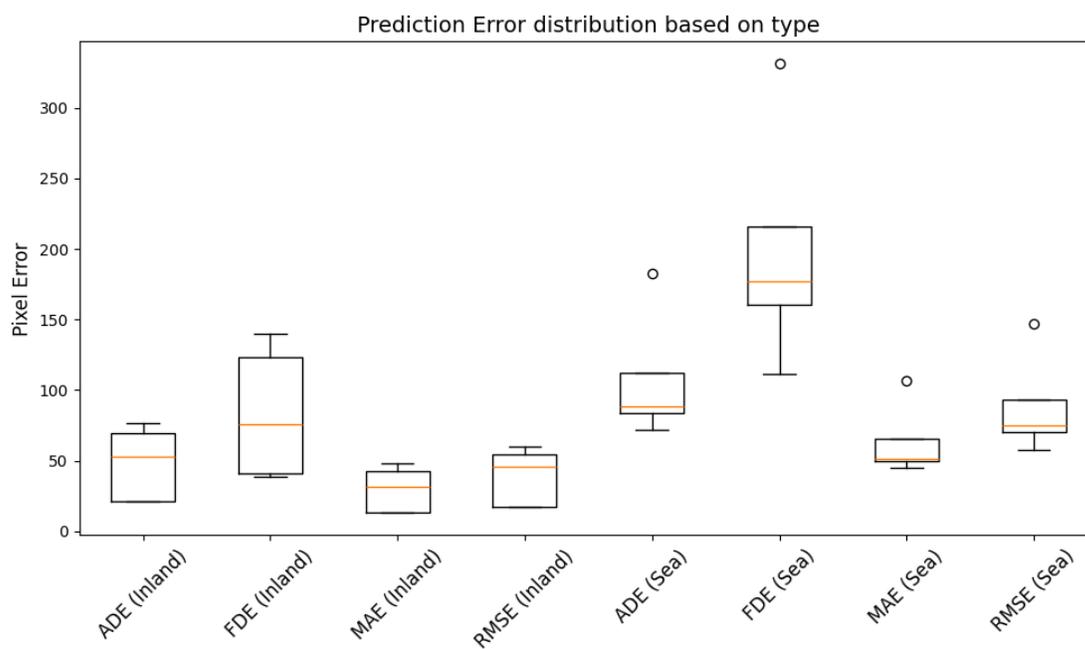


图 11: 内陆船与海船之间 Deep OC-SORT 预测质量的比较。我们给出了视频流各类别的平均距离误差 (ADE)、预测的最终距离误差 (FDE)、平均绝对误差 (MAE) 和均方根误差 (RMSE)。

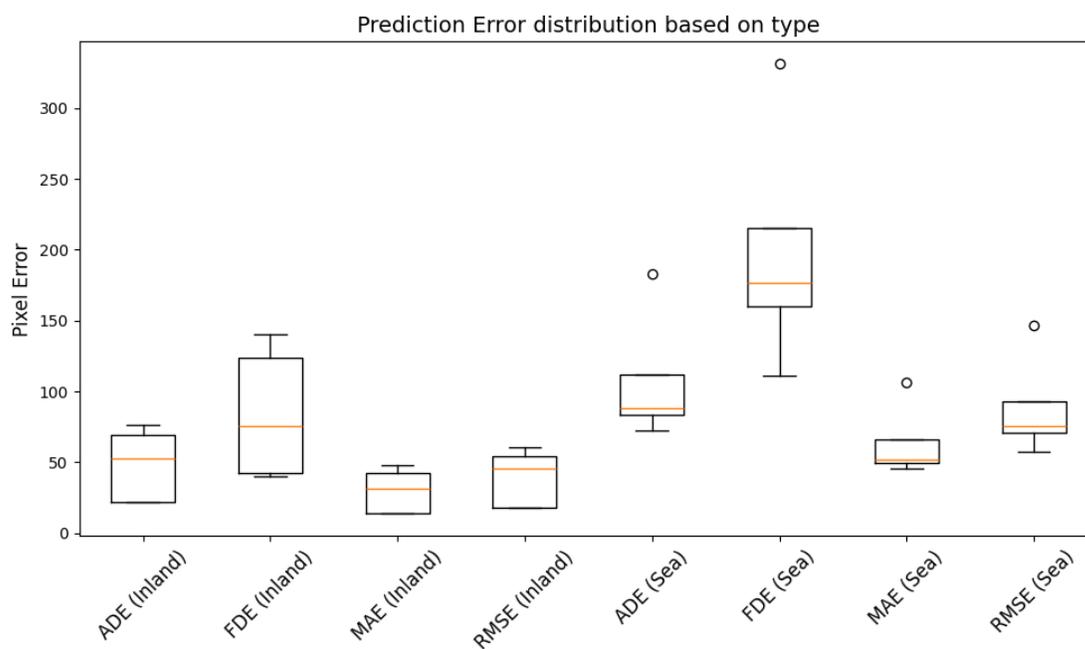


图 12: 内陆船舶与海船预测质量的比较使用 ByeTrack。我们展示了视频流类别中的平均距离误差 (ADE)、预测的最终距离误差 (FDE)、平均绝对误差 (MAE) 和均方根误差 (RMSE)。

- [10] Saif Hassan, Ghulam Mujtaba, Asif Rajput, and Noureen Fatima. Multi-object tracking: a systematic literature review. *Multimedia Tools and Applications*, 83(14):43439–43492, 2024.
- [11] Lingxiao He, Xingyu Liao, Wu Liu, Xinchun Liu, Peng Cheng, and Tao Mei. Fastreid: A pytorch toolbox for general instance re-identification. In *Proceedings of the 31st ACM International Conference on Multimedia*, MM '23, page 9664 – 9667, New York, NY, USA, 2023. Association for Computing Machinery.
- [12] Mingwei He, Bohan Li, and Songlin Sun. A survey of class activation mapping for the interpretability of convolution neural networks. In *International Conference On Signal And Information Processing, Networking And Computers*, pages 399–407. Springer, 2022.
- [13] Rupert Henn and Berthold Holtmann. Autonomes fahren in der binnenschifffahrt: Machbarkeitsstudie für ein testfeld im ruhrgebiet.
- [14] Yang Jie, LilianAsimwe Leonidas, Farhan Mumtaz, and Munsif Ali. Ship detection and tracking in inland waterways using improved yolov3 and deep sort. *Symmetry*, 13(2):308, 2021.
- [15] Arshad Khan, Islam Khan, and Tariq Aziz. A survey on parametric spline function approximation. *Applied mathematics and computation*, 171(2):983–1003, 2005.
- [16] Mona Köhler, Markus Eisenbach, and Horst-Michael Gross. Few-shot object detection: A comprehensive survey. *IEEE Transactions on Neural Networks and Learning Systems*, 2023.
- [17] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C. Lawrence Zitnick. Microsoft coco: Common objects in context. In David Fleet, Tomas Pajdla, Bernt Schiele, and Tinne Tuytelaars, editors, *Computer Vision – ECCV 2014*, pages 740–755, Cham, 2014. Springer International Publishing.
- [18] Li Liu, Wanli Ouyang, Xiaogang Wang, Paul Fieguth, Jie Chen, Xinwang Liu, and Matti Pietikäinen. Deep learning for generic object detection: A survey. *International journal of computer vision*, 128:261–318, 2020.
- [19] Gerard Maggolino, Adnan Ahmad, Jinkun Cao, and Kris Kitani. Deep oc-sort: Multi-pedestrian tracking by adaptive re-identification. *arXiv preprint arXiv:2302.11813*, 2023.
- [20] Michele Martelli, Antonio Virdis, Alberto Gotta, Pietro Cassara, and Maria Di Summa. An outlook on the future marine traffic management system for autonomous ships. *IEEE Access*, 9:157316–157328, 2021.
- [21] Alexandre Schalch Mendes, Pablo Siqueira Meirelles, Janito Vaqueiro Ferreira, and Eduardo Rodrigues de Lima. Extended kalman filter for a monitoring system of the guyed towers. In *Kalman Filter*, chapter 2. IntechOpen, 2022.
- [22] Donatella Occorsio and Woula Themistoclakis. On the filtered polynomial interpolation at chebyshev nodes. *Applied Numerical Mathematics*, 166:272–287, 2021.
- [23] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: towards real-time object detection with region proposal networks. In *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1*, NIPS' 15, page 91 – 99, Cambridge, MA, USA, 2015. MIT Press.
- [24] J Sasikala et al. Ship detection and recognition for offshore and inshore applications: a survey. *International Journal of Intelligent Unmanned Systems*, 7(4):177–188, 2019.
- [25] Feifei Shao, Long Chen, Jian Shao, Wei Ji, Shaoning Xiao, Lu Ye, Yueting Zhuang, and Jun Xiao. Deep learning for weakly-supervised object detection and localization: A survey. *Neurocomputing*, 496:192–207, 2022.
- [26] Zahra Soleimanitaleb, Mohammad Ali Keyvanrad, and Ali Jafari. Object tracking methods:a review. In *2019 9th International Conference on Computer and Knowledge Engineering (ICCKE)*, pages 282–288, 2019.

- [27] Juan Terven, Diana-Margarita Córdova-Esparza, and Julio-Alejandro Romero-González. A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas. *Machine Learning and Knowledge Extraction*, 5(4):1680–1716, 2023.
- [28] Gerasimos Theotokatos, Joao Lucas Dozzi Dantas, Georgia Polychronidi, Georgia Rentifi, and Marco Molica Colella. Autonomous shipping — an analysis of the maritime stakeholder perspectives. *WMU Journal of Maritime Affairs*, 22(1):5–35, 2023.
- [29] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.
- [30] Petr Wagner, Jiri Kotzian, Jan Kordas, and Viktor Michna. Path planning and tracking for robots based on cubic hermite splines in real-time. In *2010 IEEE 15th Conference on Emerging Technologies & Factory Automation (ETFA 2010)*, pages 1–8. IEEE, 2010.
- [31] Haofan Wang, Zifan Wang, Mengnan Du, Fan Yang, Zijian Zhang, Sirui Ding, Piotr Mardziel, and Xia Hu. Scorecam: Score-weighted visual explanations for convolutional neural networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, pages 24–25, 2020.
- [32] Yanyang Wang, Zhaoxiang Liu, and Shiguo Lian. Semi-supervised object detection: A survey on recent research and progress. *arXiv preprint arXiv:2306.14106*, 2023.
- [33] Yu Wang, G Rajesh, X Mercilin Raajini, N Kritika, A Kavinkumar, and Syed Bilal Hussain Shah. Machine learning-based ship detection and tracking using satellite images for maritime surveillance. *Journal of Ambient Intelligence and Smart Environments*, 13(5):361–371, 2021.
- [34] Zhangu Wang, Jun Zhan, Chunguang Duan, Xin Guan, Pingping Lu, and Kai Yang. A review of vehicle detection techniques for intelligent vehicles. *IEEE Transactions on Neural Networks and Learning Systems*, 34(8):3811–3831, 2022.
- [35] Nicolai Wojke, Alex Bewley, and Dietrich Paulus. Simple online and realtime tracking with a deep association metric. In *2017 IEEE International Conference on Image Processing (ICIP)*, pages 3645–3649, 2017.
- [36] Haitong Xu, Lúcia Moreira, and C. Guedes Soares. Maritime autonomous vessels. *Journal of Marine Science and Engineering*, 11(1), 2023.
- [37] Syed Sahil Abbas Zaidi, Mohammad Samar Ansari, Asra Aslam, Nadia Kanwal, Mamoona Asghar, and Brian Lee. A survey of modern deep learning based object detection models. *Digital Signal Processing*, 126:103514, 2022.
- [38] Chi Zhang, Xi Zhang, Gui Gao, Haitao Lang, Genwang Liu, Chenghui Cao, Yuying Song, Yanan Guan, and Yongshou Dai. Development and application of ship detection and classification datasets: A review. *IEEE Geoscience and Remote Sensing Magazine*, 2024.
- [39] Yifu Zhang, Peize Sun, Yi Jiang, Dongdong Yu, Fucheng Weng, Zehuan Yuan, Ping Luo, Wenyu Liu, and Xinggong Wang. Bytetrack: Multi-object tracking by associating every detection box. In *European conference on computer vision*, pages 1–21. Springer, 2022.
- [40] Yian Zhao, Wenyu Lv, Shangliang Xu, Jinman Wei, Guanzhong Wang, Qingqing Dang, Yi Liu, and Jie Chen. Detrs beat yolos on real-time object detection, 2023.
- [41] Zhong-Qiu Zhao, Peng Zheng, Shou-tao Xu, and Xindong Wu. Object detection with deep learning: A review. *IEEE transactions on neural networks and learning systems*, 30(11):3212–3232, 2019.
- [42] Yanzhao Zhu and Wei Qi Yan. Traffic sign recognition based on deep learning. *Multimedia Tools and Applications*, 81(13):17779–17791, 2022.