

CM-Ship-ReID 数据集使用说明

数据主编：徐从安，高龙，刘瑜，宿南

下载说明：CM-Ship-ReID 数据集所有权归海军航空大学、哈尔滨工程大学及合作研究团队所有，由烟台项目组联合采集与标注。《雷达学报》编辑部具有编辑出版权等。数据使用权遵循 Creative Commons Attribution 4.0 国际许可协议（CC BY 4.0），用户可免费下载并用于非商业性学术研究、教学或算法开发，但需在论文、报告等成果中明确引用本数据集及相关论文。如有更多需求，欢迎联系徐从安(xcatougao@163.com)

本数据集引用格式：

[1] 徐从安,高龙,刘瑜,等.跨模态船舶目标重识别数据集(CMShipReID)[OL]. 雷达学报, 2025. <https://radars.ac.cn/web/data/getData?dataType=MaritimeTargets>.

XU Congan, GAO Long, LIU Yu, et al. CMShipReID: A Cross-Modality Ship Dataset for the Re-IDentification Task[OL]. Journal of Radars, 2025. <https://radars.ac.cn/web/data/getData?dataType=MaritimeTargetsEN&pageType=en>.

[2] C. Xu et al., "CMShipReID: A Cross-Modality Ship Dataset for the Reidentification Task," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 18, pp. 10503-10513, 2025, doi: 10.1109/JSTARS.2025.3557509.

数据集简介：CM-Ship-ReID 是面向船舶跨模态重识别任务的公开数据集，旨在推动船舶精细化分析与跨模态匹配算法的研究。该数据集由搭载可见光、近红外和热红外相机的无人机采集，涵盖渔船、油轮、拖船、货船等 10 类船舶，包含 138 个船舶实例，总计 8,337 张图像。其中 VIS 有 3209 张，NIR 有 2668 张、TIR 有 2460 张。

数据集覆盖多种复杂场景，包括白天、夜晚、雨天、雾天等天气条件，并包含尺度变化、遮挡、光照差异、低分辨率等挑战性样本。每艘船舶的视频经 360 度环绕拍摄后分割为 15-30 张图像，具备任意旋转方向、大宽高比以及小目标占比高的特点。

以下为数据集结构：

CM-Ship-ReID/

├── VIS/

| ├── 001/

| | ├── 001_001.jpg

| | ├── 001_002.jpg

| | └── ...

| └── 002/

| └── ...

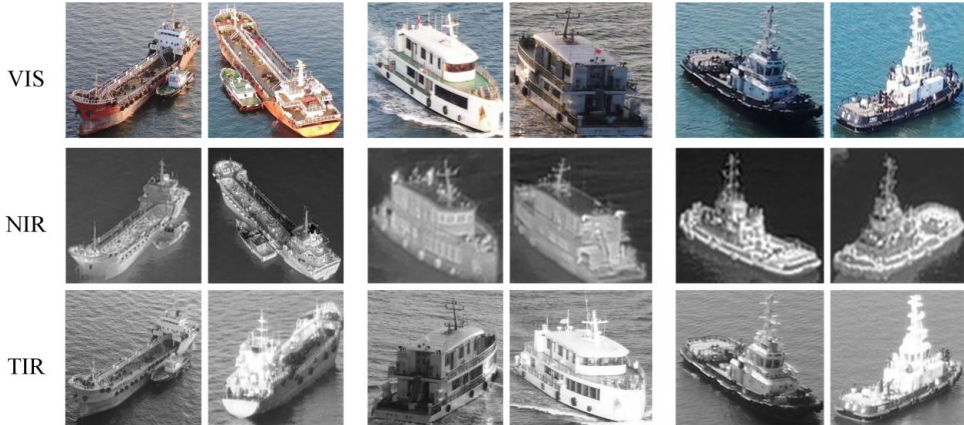
├── NIR/

└── TIR/

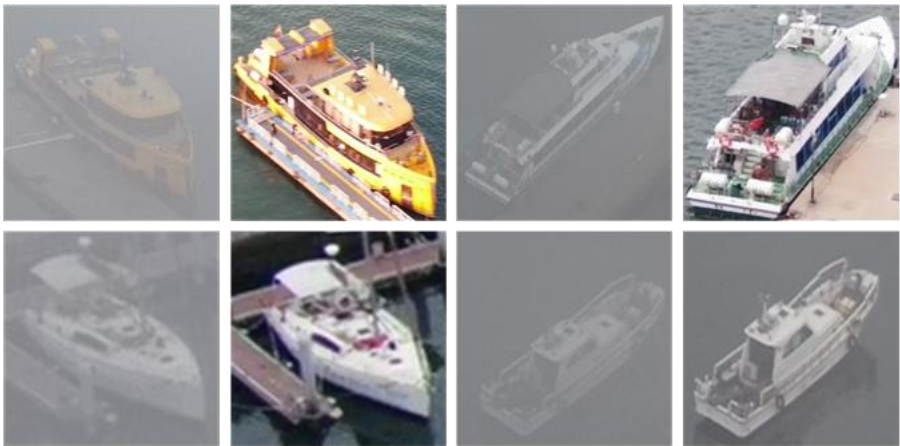
数据集按 7:3 比例划分为训练集与测试集，同时建议采用 6 折交叉验证以提升算法验证可靠性。实验部分对比了 DDAG、DART、DEEN 等主流跨模态 Re-ID 算法，其中 DEEN 在多数任务中表现最优。

CM-Ship-ReID 的发布可有力推动船舶跨模态重识别领域研究，为海事监控、安防与智能分析提供了重要的研究支撑。数据集可通过指定链接公开获取。

数据集的具体内容与样本实例如下：



云遮挡下的样例



不同规模的船



不同光照情况下的船



不同分辨率变化下的船



数据使用说明：

下载数据包含图像数据及标签文件（exp）。exp 文件中包含所有船舶 ID 文件（available.txt）、训练船舶 ID 文件（train.txt）、验证船舶 ID 文件（val.txt）及测试船舶 ID 文件（test.txt），文件中 ID 为随机选取。训练时，train.txt 及 val.txt 两个文件中所有的 ID 均用来进行训练。值得注意的是，只有 134 艘船舶具有近红外数据，其他两模态均 138 艘。

实际使用时，选取任意两个模态数据进行使用，具体结构如下：

```
├── Cam1/
│   ├── 001/
│   │   ├── 001_001.jpg
│   │   ├── 001_002.jpg
│   │   └── ...
│   └── 002/
│       └── ...
├── Cam2/
│   ├── 001/
│   │   ├── 001_001.jpg
│   │   ├── 001_002.jpg
```

| | | ...
| | 002/
| | ...
| exp/

当数据采用模态包含近红外时，100 艘用于训练，34 艘用于测试。其余情况，100 艘用于训练，38 艘用于测试。用于训练的 100 艘船舶 ID 固定。