

基于 BEiT 相对映射网络的单目深度估计

作者姓名 _____ 徐程升 _____

指导教师姓名、职称 _____ 高新波 教授 _____

申请学位类别 _____ 工学硕士 _____

学校代码 10701
分 类 号 TP391

学 号 23021211705
密 级 公开

西安电子科技大学

硕士学位论文

基于 BEiT 相对映射网络的单目深度估计

作者姓名：徐程升

一级学科：控制科学与工程

二级学科（研究方向）：模式识别与智能系统

学位类别：工学硕士

指导教师姓名、职称：高新波 教授

学 院：电子工程学院

提交日期：2026 年 4 月

RGBT Object Tracking Based on Deep Fusion and Attention Mechanism

A thesis submitted to
XIDIAN UNIVERSITY
in partial fulfillment of the requirements
for the degree of Master
in Control Science and Engineering

By
Chengsheng XU
Supervisor: Xinbo GAO Title: Professor
April 2026

西安电子科技大学 学位论文独创性（或创新性）声明

秉承学校严谨的学风和优良的科学道德，本人声明所呈交的论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢中所罗列的内容以外，论文中不包含其他人已经发表或撰写过的研究成果；也不包含为获得西安电子科技大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同事对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

学位论文若有不实之处，本人承担一切法律责任。

本人签名：_____ 日 期：_____

西安电子科技大学 关于论文使用授权的说明

本人完全了解西安电子科技大学有关保留和使用学位论文的规定，即：研究生在校攻读学位期间论文工作的知识产权属于西安电子科技大学。学校有权保留并向国家有关部门或机构送交学位论文的复印件和电子版，以学术交流为目的赠送和交换学位论文，允许学位论文被查阅、借阅和复印，将学位论文的全部或部分内容编入有关数据库进行检索和提供相应阅览服务；允许采用影印、缩印或其它复制手段保存学位论文。同时本人保证，结合学位论文研究成果完成的论文、发明专利等成果，署名单位为西安电子科技大学。

本人保证遵守上述规定。

（保密的论文在解密后遵守此规定）

本人签名：_____ 导师签名：_____

日 期：_____ 日 期：_____

摘要

这是中午摘要。这是中午摘要。这是中午摘要。这是中午摘要。这是中午摘要。
这是中午摘要。这是中午摘要。这是中午摘要。这是中午摘要。这是中午摘要。这是
中午摘要。这是中午摘要。这是中午摘要。这是中午摘要。这是中午摘要。这是中午
摘要。这是中午摘要。这是中午摘要。这是中午摘要。这是中午摘要。这是中午摘要。
这是中午摘要。

这是中午摘要。这是中午摘要。这是中午摘要。这是中午摘要。这是中午摘要。
这是中午摘要。这是中午摘要。这是中午摘要。这是中午摘要。这是中午摘要。这是
中午摘要。这是中午摘要。这是中午摘要。这是中午摘要。这是中午摘要。这是中午
摘要。这是中午摘要。这是中午摘要。这是中午摘要。这是中午摘要。这是中午摘要。
这是中午摘要。这是中午摘要。这是中午摘要。这是中午摘要。这是中午摘要。这是
中午摘要。这是中午摘要。这是中午摘要。这是中午摘要。这是中午摘要。这是中午
摘要。这是中午摘要。这是中午摘要。这是中午摘要。这是中午摘要。这是中午摘要。
这是中午摘要。这是中午摘要。这是中午摘要。这是中午摘要。这是中午摘要。这是
中午摘要。这是中午摘要。这是中午摘要。这是中午摘要。这是中午摘要。这是中午
摘要。这是中午摘要。这是中午摘要。这是中午摘要。这是中午摘要。这是中午摘要。
这是中午摘要。这是中午摘要。

关键词：三维计算机视觉，单目深度估计，绝对深度估计

ABSTRACT

Visual object tracking is a widely studied topic in computer vision and pattern recognition due to its significant theoretical research value and diverse applications in both civilian and military domains, including video surveillance, autonomous driving, and battlefield situation awareness. Object tracking based solely on visible light often face challenges such as smoke interference and varying levels of illumination, leading to frequent failures. Data from both visible and infrared spectra (RGB/Thermal, RGBT) share consistency and provide complementary information about the target, enabling a dual-mode tracker to enhance the robustness and accuracy of visual object tracking. Nonetheless, current RGBT object tracking algorithms suffer from a deficiency in effective feature selection mechanisms for dual-mode feature fusion and the absence of a decision-level fusion algorithm for both modes. This study leverages two frameworks: convolutional neural network and large-scale vision model, incorporating a deep fusion algorithm and attention mechanism to address the identified issues. The main contributions are summarized as follows.

A novel RGBT target tracking algorithm is introduced in this study to address the issues of inadequate network feature representation and variations in the reliability of visible light and thermal infrared for decision fusion, utilizing adaptive attention feature selection and decision fusion techniques. The study employs an adaptive hybrid attention mechanism that integrates channel, spatial, and positional information to improve the network's feature representation, thereby offering more precise evidence for decision-level fusion. The reliability of two modes is modeled using the Dirichlet distribution, the D-S criterion is employed for decision-level evidence fusion, and an online updated multi-mode branch loss adaptive fusion framework is utilized to reinforce the network's robustness in tracking. Extensive experiments conducted on the open datasets GTOT and RGBT234 demonstrate an accuracy and success rate of 90.9%/75.3% and 77.4%/55.6% correspondingly, providing strong evidence for the efficacy of the developed algorithm.

An RGBT target tracking algorithm is introduced in this study to address the issue of limited interaction between template search images and the dynamic changes of the target, utilizing channel space self-attention and template online updating techniques. Leveraging the tracking methods of template and search images in paired networks, a robust benchmark ex-

perimental algorithm is developed by integrating a Transformer-based large-scale vision. The backbone network equipped with dual-branch embedding layers and weights to enhance the feature interaction between visible light and thermal infrared modes. This study introduces a channel space self-attention mechanism based on the correlation between template and search images to improve the interaction and extract diverse complementary features across the modes. Lastly, the study introduces the template online update module to address the issue of model drift due to target time changes, incorporating online template updating and fractional head design mitigate the drift. Extensive experiments conducted on public datasets GTOT and RGBT234 reveal accuracy and success rates of 93.3%/75.6% and 87.2%/63.8%, validating the effectiveness of the proposed algorithm.

Keywords: 3D Computer Vision, Monocular Depth Estimation, Absolute Depth Estimation

插图索引

图 1.1 无人机	2
图 2.1 单目深度估计任务解构与特征提取框架	6

表格索引

符号对照表

符号	符号名称
σ	Sigmoid 函数
\cap	交集
\cup	并集
\exp	e 指数
\log	e 对数
\int	积分
\sum	累加
\prod	累积
\odot	逐点相乘

缩略语对照表

缩略语	英文全称	中英文对照
MDE	Metric Depth Estimation	绝对深度估计
MRF	Markov Random Field	马尔可夫随机场

目 录

摘要.....	I
ABSTRACT	III
插图索引.....	V
表格索引.....	VII
符号对照表	IX
缩略语对照表	XI
第一章 绪论.....	1
1.1 研究意义与背景	1
1.2 国内外研究现状	2
1.2.1 相对深度估计研究现状	2
1.2.2 绝对深度估计研究现状	3
第二章 基于相对深度引导的映射模块.....	5
2.1 引言	5
2.2 模型总体架构设计	6
2.2.1 任务解构与特征提取.....	6
2.2.2 相对几何结构感知分支	7
2.2.3 绝对深度对齐分支	7
第三章 第三章	9
第四章 第四章	11
第五章 第五章	13
参考文献	15
致谢	17
作者简介	19

第一章 绪论

1.1 研究意义与背景

深度估计作为三维计算机视觉领域的核心基础问题，旨在从二维图像中重建场景的几何特征，在过去十余年中经历了从手工特征建模到深度学习驱动的范式变迁^[1-2]。随着传感器技术与计算能力的提升，深度估计已成为智能无人系统实现环境感知与定位导航的关键技术：在自动驾驶中，它为障碍物检测与路径规划提供必要的距离信息；在无人机技术中，它是实现自主避障与三维测绘的基础；在具身智能领域，深度估计则赋予了智能体感知空间结构并进行物理交互的能力。

从技术演进的维度看，单目深度估计的研究历程主要经历了三个阶段。初期阶段主要依赖手工设计的几何特征与先验假设^[3]，通过概率图模型整合图像的底层信息，但在复杂场景下的建模鲁棒性较差。中期阶段随着卷积神经网络（CNN）的兴起，研究重心转向端到端的监督学习^[4]，通过多尺度网络架构显著提升了像素级的预测精度。现阶段则迈入了以大模型和多任务迁移为特征的新时期，Vision Transformer (ViT) 等架构的应用^[2] 极大增强了模型对全局几何上下文的理解。

然而，现有的深度估计方法在非约束场景下的表现仍面临严峻挑战。由于单目深度估计本质上是一个不适定问题，存在固有的比例模糊性^[1,5]，神经网络往往倾向于通过学习训练集中的统计偏见（如物体位置与深度的相关性）来“走捷径”，而非真正理解场景的物理几何。这种策略导致模型过度拟合了特定数据集（如 KITTI 或 NYU Depth V2）的成像特性，使其对相机内参及拍摄视角具有极强的依赖性。一旦应用于光照剧变、极端天气或异质场景，由于领域鸿沟的存在，模型的预测精度往往会出现断崖式下降，限制了其在跨平台部署时的零样本迁移能力。

针对上述挑战，学术界开始探索一种新的研究范式：利用强泛化性的相对深度信息辅助绝对深度的估计^[6]。相对深度虽然不具备物理单位，但其能通过海量异质数据的预训练，捕捉到稳健的几何拓扑关系与遮挡先验，展现出极佳的场景鲁棒性。

本文认为，融合相对深度的几何先验优势与绝对深度的尺度特性，是实现跨场景稳定感知的关键路径。通过设计一种能够解耦几何结构与物理尺度的预测框架，利用大模型提取的全局几何一致性来约束局部尺度恢复，可显著降低模型对特定相机内参的耦合。这种方法旨在打破单目深度估计在未知场景下的精度瓶颈，为无人系统在复杂、全天候环境下的高精度感知提供新的理论支撑与技术方案。



图 1.1 无人机

1.2 国内外研究现状

1.2.1 相对深度估计研究现状

单目深度估计根据输出表征的不同，通常可分为绝对深度估计与相对深度估计。绝对深度估计旨在恢复具有明确物理尺度的度量距离，但在非约束场景下常面临严峻的尺度模糊性挑战；而相对深度估计则侧重于建模场景内物体的几何序关系，在复杂环境及跨领域迁移中展现出更强的鲁棒性。

相对深度特征的建模最早可追溯至 Saxena 等人^[3]的研究，该工作首次利用马尔可夫随机场（MRF）建模像素块间的空间相关性，奠定了几何建模基础。2016 年，Chen 等人^[7]真正将“相对深度估计”确立为一个独立的研究命题，证明了通过人类标注的像素对排序先验（Ranking Perception）即可实现复杂场景下的深度感知。随后，Ranftl 等人^[6]正式提出了跨数据集混合训练方案，通过尺度与平移不变损失函数（Scale-and-shift-invariant Loss），极大提升了模型的零样本（Zero-shot）泛化能力。

进入 2023 年后，相对深度估计在“视觉基础模型”的推动下取得了质的飞跃。以 DINOv2^[8] 为代表的大规模自监督预训练模型证明，强力的语义特征与场景的几何拓扑存在高度一致性。在此基础上，Depth Anything^[9] 系列工作通过海量无标签数据的判别式训练，实现了在任意“野外”场景下极具鲁棒性的深度结构恢复。这些工作不仅提供了高精度的深度图输出，更重要的是构建了一个包含丰富几何纹理与空间上

下文的特征表示空间（Feature Representation Space）。

尽管上述工作在几何结构的稳健性方面取得了显著突破，但由于相对深度输出缺乏真实的物理尺度信息（Metric Scale），导致其在自动驾驶、无人机避障等绝对距离敏感任务中应用受限。因此，如何有效提取并利用相对深度模型中蕴含的强泛化特征块（Feature Blocks），将其作为先验引导来实现高精度的尺度恢复，已成为当前打通“几何结构”与“物理度量”逻辑鸿沟的关键研究方向。

1.2.2 绝对深度估计研究现状

绝对深度估计（Metric Depth Estimation）旨在建立图像特征与真实物理距离之间的回归映射。早期研究如 Eigen 等人^[10]利用多尺度 CNN 实现了端到端的深度回归，奠定了深度学习在该领域的基础。随后，为了解决连续值回归收敛困难的问题，Fu 等人^[4]提出了深度序数回归网络（DORN），通过离散化策略将回归任务转化为有序分类，显著提升了特定数据集上的绝对数值精度。在真值获取方面，针对激光雷达数据稀疏的挑战，Godard 等人^[11]提出了基于光度一致性约束的自监督学习范式，极大拓宽了应用边界。

进入“基础模型”时代以来，单目深度估计正经历从特定场景拟合向零样本泛化（Zero-shot Generalization）的范式转变。以 Depth Anything^[9] 和 Metric3D^[12] 为代表的工作，通过大规模弱监督预训练与数据蒸馏技术，证明了利用海量无标签数据学习稳健几何表示的可能性。然而，即便在基础模型的支撑下，现有的绝对深度估计工作仍面临以下严峻挑战：

首先，模型对相机内参及场景分布存在深层的“尺度-语义耦合”误区。现有的绝对深度预测框架往往将场景几何结构与物理尺度混合建模，导致模型倾向于通过“背诵”训练集中特定相机的成像特性（如焦距、安装高度）以及常见物体的先验尺寸（如车辆、人高）来推断距离，而非基于纯粹的投影几何关系。这种统计过拟合（Statistical Overfitting）导致模型极易受到单目歧义性的干扰：一旦测试环境的焦距或传感器尺寸发生变化，预测结果便会产生剧烈的尺度漂移，无法实现真正的物理一致性。

其次，现有工作对相对深度所蕴含的稳健几何先验利用不足。从数学本质上看，绝对深度预测是一个受限于相机内参的“病态问题”，而相对深度由于具备尺度不变性（Scale-invariant），在捕捉物体边缘、空间拓扑及遮挡关系方面表现出极强的泛化潜能。遗憾的是，多数绝对深度研究仍倾向于构建从图像到绝对数值的直接映射（Direct Mapping）。这种“一步到位”的回归范式往往试图让模型同时学习“复杂的几何感知”与“脆弱的绝对回归”，导致模型在面对异质场景时，容易为了拟合数值精度而丧失基本的几何结构约束，产生结构畸变或物体边缘模糊。

目前，如何在高精度度量需求下实现几何结构与绝对尺度的深度解耦与耦合，已成为学术界关注的焦点。尽管 ZoeDepth^[13] 等工作尝试引入相对深度特征，但其复杂的端到端训练策略仍难以完全摆脱尺度耦合的影响。鉴于此，探讨如何将具有强泛化力的相对深度作为“几何基座”，并通过轻量化的映射机制恢复物理尺度，对于构建全天候、高可靠的无人系统具有重要的学术意义与应用价值。

第二章 基于相对深度引导的映射模块

2.1 引言

随着具身智能技术的普及，单目深度估计（MDE）已成为自动驾驶与机器人感知的核心环节。按照表征形式，MDE 可分为相对深度估计与绝对深度（Metric Depth）估计，后者因能恢复具备真实物理尺度的像素级距离，成为系统从“视觉感知”迈向“物理交互”的关键桥梁。在避障规划、目标抓取及视觉 SLAM 尺度恢复等任务中，绝对尺度信息提供了不可或缺的几何度量约束。相比激光雷达等昂贵设备，单目绝对深度估计凭借低成本与高灵活性，已成为构建全场景无人系统感知的核心基石。

如前文所述，从单幅图像推算绝对深度本质上是一个典型的“不适定问题”（Ill-posed Problem）。目前主流的单目绝对深度估计多采用端到端的直接回归范式，这种“一步到位”的映射机制使得模型不可避免地陷入尺度信息与几何结构的强耦合困境。在这种机制下，模型往往过度依赖训练集中的语义先验（如物体的平均尺寸）来推断距离，而非理解真正的投影几何关系。这导致模型在面对异质场景或相机内参发生波动时，极易产生剧烈的尺度漂移（Scale Drift），并导致物体边缘与空间布局的一致性受损。

值得注意的是，现有的绝对深度估计工作大多倾向于构建从图像特征到物理数值的直接线性或非线性映射，这种粗放的回归方式忽略了相对深度中蕴含的稳健几何拓扑约束。由于绝对深度真值（Ground Truth）通常由稀疏的激光雷达获取，且覆盖场景有限，模型在拟合有限的数值分布时，容易丧失对物体间细粒度几何序关系的把握。这导致模型在处理复杂场景时，虽然能在统计数值上逼近真值，但其三维空间结构常发生畸变，无法满足高精度导航对环境结构化的严苛要求。

然而，现有的绝对深度估计工作大多倾向于构建从图像到物理数值的直接映射，这种“一步到位”的粗放回归方式往往忽略了相对深度中蕴含的稳健几何拓扑约束。这导致模型在处理复杂场景时，虽然能在数值上逼近真值，但其空间结构常发生畸变，且极易受相机参数波动的干扰。

针对上述挑战，本章提出了一种融合相对深度特征先验的二阶段解构式绝对深度估计方法。该方法的核心逻辑在于将绝对深度估计任务解构为“结构感知”与“尺度适配”两个独立且互补的环节。第一阶段利用具有强泛化力的预训练基础模型提取尺度不变的相对几何特征；第二阶段则通过一个轻量化映射模块，利用多尺度特征块（Feature Blocks）中丰富的上下文信息引导模型完成从相对空间到物理尺度的精准转换。这种解耦设计不仅能最大限度保留基础模型在“野外”环境下的结构泛化力，更

通过特征驱动的映射机制实现了物理尺度的精确对齐，为构建高鲁棒性的绝对深度感知系统提供了新思路。

2.2 模型总体架构设计

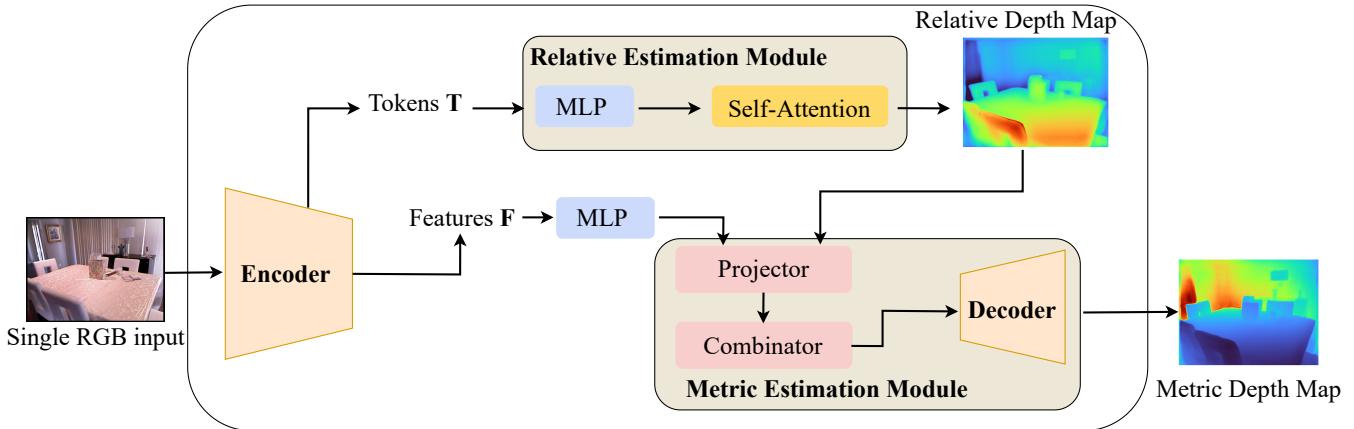


图 2.1 单目深度估计任务解构与特征提取框架

本章所设计的绝对深度估计模型采用了任务解构的二阶段架构，旨在充分利用预训练基础模型的几何结构泛化能力，并通过轻量化的度量映射分支实现物理尺度的精准恢复。模型总体架构如图2.1所示，主要由特征提取编码器（Encoder）、相对深度估计模块（Relative Estimation Module）以及度量深度估计模块（Metric Estimation Module）三部分组成。

2.2.1 任务解构与特征提取

为了从根本上解决单目绝对深度估计中由于投影歧义性导致的尺度与结构强耦合问题，本文将深度恢复过程解构为“结构感知”与“尺度适配”两个解耦的子任务。该设计旨在利用大规模预训练模型（如 DPT-BEiT 或 MiDaS 系列）的几何泛化能力，为绝对尺度的恢复提供稳健的拓扑约束。

在特征提取阶段，系统采用统一的共享编码器（Shared Encoder）作为底层基座。对于输入的单幅 RGB 图像 $I \in \mathbb{R}^{3 \times H \times W}$ ，编码器通过分层表征学习提取出具有异构物理意义的特征。如公式 2-1 所示：

$$\{T, F\} = \text{Encoder}(I) \quad (2-1)$$

其中， T 为蕴含全局几何上下文的标记向量（Tokens）； $F = \{f_1, f_2, f_3, f_4\}$ 为多层级空间特征块（Feature Blocks），对应于代码实现中不同层级的瓶颈层特征与解码块输出，承载了丰富的局部空间细节。

这种多路输出设计实现了任务驱动的特征分配逻辑。第一阶段相对估计分支利用标记向量 T 构建场景的宏观拓扑，确保深度图在几何序关系与物体边缘上的一致性；第二阶段度量映射分支则通过级联的投影器（Projector）处理特征块 F ，并利用局部特征块引导深度分桶（Bins）的自适应动态调整。

相较于传统的端到端直接回归范式，这种解构设计能够有效缓解单目深度估计的不适定性（Ill-posedness）。通过在特征提取阶段将语义先验与尺度特征进行分离，模型能够有效抑制由于相机内参波动或异质场景切换带来的尺度漂移现象。这种解耦机制确保了模型在恢复物理尺度的同时，能够最大限度保留基础模型所提供的强泛化几何结构，从而兼顾了预测结果的数值准确性与空间合理性。

2.2.2 相对几何结构感知分支

在单目深度估计任务中，建立稳健的几何拓扑关系是恢复绝对度量维度的前提。本章所设计的相对几何结构感知分支旨在不考虑绝对物理尺寸的情况下，利用大规模预训练模型捕获的深层语义特征，重构场景内物体的相对位置关系。该分支的输入为共享编码器输出的全局标记向量（Tokens） \mathbf{T} ，如公式 2-2 所示：

$$D_{rel} = \Psi_{rel}(\mathbf{T}) \quad (2-2)$$

其中， Ψ_{rel} 表示相对深度估计映射函数。

标记向量 \mathbf{T} 进入相对估计模块后，首先通过多层感知机（MLP）进行维度对齐与特征压缩，随后进入自注意力机制（Self-Attention）层进行长程依赖建模。自注意力机制的引入使得模型能够摆脱局部感受野的限制，通过计算全图范围内的特征相关性，实现对场景宏观布局的深度理解。这种全局建模能力对于识别复杂的物体间拓扑关系、处理大面积平坦区域以及跨越遮挡边界获取一致的几何先验至关重要。

该分支最终输出一张尺度不变的相对深度图 D_{rel} 。该图能够精确捕获物体的轮廓细节、场景平面的斜率以及物体间的相互遮挡序关系。尽管其数值不具备物理含义，但 D_{rel} 提供的几何稳定性为下一阶段的尺度适配任务奠定了坚实基础。通过将深度估计任务的第一阶段聚焦于几何拓扑的构建，模型能够显著增强在非约束场景下的泛化能力，为后续实现精确的“结构-度量”映射提供可靠的几何度量锚点。

2.2.3 绝对深度对齐分支

第三章 第三章

第四章 第四章

第五章 第五章

参考文献

- [1] EIGEN D, PUHRSCH C, FERGUS R. Depth map prediction from a single image using a multi-scale deep network[J]. Advances in neural information processing systems (NeurIPS), 2014, 27.
- [2] RANFTL R, BOCHKOVSKIY A, KOLTUN V. Vision transformers for dense prediction[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). 2021: 12173-12183.
- [3] SAXENA A, SUN M, NG A Y. Make3d: learning 3d scene structure from a single still image[J]. IEEE transactions on pattern analysis and machine intelligence, 2008, 31(5): 824-840.
- [4] FU H, GONG M, WANG C, et al. Deep ordinal regression network for monocular depth estimation [C]//Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR). 2018: 2002-2011.
- [5] KENDALL A, GAL Y. What uncertainties do we need in bayesian deep learning for computer vision?[C]//Advances in neural information processing systems (NeurIPS). 2017.
- [6] RANFTL R, LASINGER K, HAFNER D, et al. Towards robust monocular depth estimation: mixing datasets for zero-shot cross-dataset transfer[J]. IEEE transactions on pattern analysis and machine intelligence, 2020.
- [7] CHEN W, FU Z, YANG D, et al. Single-image depth perception in the wild[C]//Proceedings of the 30th International Conference on Neural Information Processing Systems (NIPS). 2016: 730-738.
- [8] OQUAB M, DAR CET T, MOUTAKANNI T, et al. DINOV2: learning robust visual features without supervision[A]. 2023.
- [9] YANG L, KANG B, HUANG Z, et al. Depth anything: unleashing the power of large-scale unlabeled data[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2024.
- [10] EIGEN D, PUHRSCH C, FERGUS R. Depth map prediction from a single image using a multi-scale deep network[C]//Advances in neural information processing systems (NeurIPS). 2014.
- [11] GODARD C, MAC AODHA O, FIRMAN M, et al. Digging into self-supervised monocular depth estimation[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). 2019.
- [12] YIN W, ZHANG C, CHEN H, et al. Metric3D: towards zero-shot metric depth prediction via large-scale multi-dataset training[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). 2023.
- [13] BHAT S F, BIRKEL R, BIENECK D, et al. ZoeDepth: zero-shot transfer by combining relative and

metric depth[A]. 2023.

致 谢

我他妈写都写不完我还致谢个 der

作者简介

1. 基本情况

徐程升，男，河南郑州人，1999年2月出生，西安电子科技大学电子工程学院控制科学与工程专业2023级硕士研究生。

2. 教育背景

2018.09~2022.06，中原工学院，本科，专业：自动化

2023.09~，西安电子科技大学，硕士研究生，专业：控制科学与工程

3. 攻读硕士学位期间的研究成果

3.1 发表学术论文

[1]Zhong A, Chen S, Wang T, et al. A Mutual-Attention Guided Feature Extraction and Adaptative Decision Fusion Framework for Fine-Grained Dual-Band Radar Target Classification[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2024.

3.2 参与科研项目及获奖

[1] 国家自然科学基金面上项目，复杂环境下小样本高分辨雷达目标识别方法（62173265），2021.01-2025.12

[2] 中电科38所，开放环境下注意力选择技术，2021.06-2022.05

[3] 国防相关项目，红外可见光****技术研究，2021.12-2022.12

[4] 中电科38所，基于元学习****技术，2023.02-2024.03

