

哈尔滨工业大学深圳校区

毕业论文开题报告

题 目 基于深度学习的无线电信号

时频位置检测及制式识别方法研究

姓 名 黄文浩

学 号 220210809

学 院 信息科学与技术学院

专 业 通信工程

指 导 教 师 张行健副教授

日 期 2025 年 10 月 13 日

目 录

1 课题背景及研究的目的与意义	1
1.1 课题背景	1
1.2 研究的目的与意义	1
2 研究现状及分析	3
2.1 国内外研究现状	3
2.2 国内外文献综述及简析	3
2.2.1 传统信号定位与识别方法研究现状	3
2.2.2 基于深度学习的目标检测方法研究现状	5
2.2.3 基于深度学习的信号检测方法研究现状	6
3 主要研究内容及研究方案	8
3.1 研究内容	8
3.1.1 针对不同尺度信号数据的处理方法研究	8
3.1.2 基于 Transformer 检测头的信号检测模型研究	8
3.1.3 针对窄带短时信号的模型优化策略研究	9
3.2 研究方案	9
4 进度安排及预期目标	12
4.1 进度安排	12
4.2 预期目标	12
5 已具备和所需的条件和经费	14
5.1 实验室条件和经费保障	14
5.1.1 实验室已具备条件	14
5.1.2 实验室经费保障	14
5.2 所需条件和经费	14
6 预计困难及解决方案说明	15
6.1 技术难点与预计困难	15
6.2 解决方案	15
参考文献	17

1 课题背景及研究的目的与意义

1.1 课题背景

随着物联网（IoT）的迅速发展，各种无线通信设备数量呈现爆炸式增长。从智能家居、工业自动化到智慧医疗、无人驾驶，各种无线设备和通信技术共同占用有限的频谱资源，形成了日益拥挤且动态多变的复杂电磁环境。为了提高频谱共享的利用率、减少信号之间的互相干扰，以及监控是否存在未授权的非法信号以保障无线安全，频谱感知的重要性日益凸显。

频谱感知（Spectrum Sensing）作为认知无线电（Cognitive Radio）的核心技术，旨在智能地检测、定位和识别电磁环境中的无线信号，是实现动态频谱接入（DSA）、保障通信可靠性、提升频谱利用率的关键前提。传统的频谱感知方法，如能量检测、匹配滤波和循环平稳特征检测，虽在特定条件下有效，但分别存在对噪声功率敏感、需已知信号先验信息、计算复杂度高等局限性，难以适应当前非协作、高动态的复杂电磁环境。

近年来，深度学习（Deep Learning）的巨大成功为解决该问题提供了全新范式。通过将一维时域信号变换为二维时频图，信号检测任务可以被巧妙地转化为计算机视觉领域中成熟的目标检测（Object Detection）任务。通过将目标检测领域的技术迁移至信号检测领域，可以在无需复杂的信号或信道的先验知识的情况下进行端到端处理，自动学习信号的深层特征，在对信号的时频定位和制式分类方面展现出巨大潜力。

1.2 研究的目的与意义

虽然计算机视觉的目标检测领域的研究已经较为广泛，已经提出了许多先进的方案，但是当前信号检测领域的研究仍处于初步探索阶段，研究成果仍然较少，对目标检测技术的迁移应用尚不成熟，且在迁移应用中存在若干问题：

1. 实验中选用的数据集较为简单：目前，将目标检测应用于信号检测的研究中，往往使用的是计算机模拟生成的信号或模拟信号与真实采集信号混合的数据集，且存在信号制式和调制种类数量较少、信号密度较低、未考虑信号之间的重叠干扰等对实际情景的简化。同时，数据集中的样本均假定以固定采样率、固定采样时长的采集得到，与实际情景中多种采样率、多种采样时长的情况有所出入。这是因为目前信号检测领域尚处于可行性验证通过的研究初期，情景的简化有利于快速验证方案的可行性，但是难以仿真得到实际电磁环境中的模型性能。

2. 对目标检测领域的模型迁移较为僵硬：目前，将目标检测应用于信号检测的研究较为有限。早期的研究往往直接将目标检测领域中的经典模型（如 Faster R-CNN）迁移至信号检测任务中，并没有根据信号检测的实际情景、信号形态与图片中物体形态的差异进行模型结构的针对性优化。最近的研究已经提出针对信号形态和特征对模型结构进行针对改进，但是仍然局限于经典的卷积神经网络模型，并没有尝试引入近年来在计算机视觉领域取得突破性进展的 Transformer 架构（如 DETR 等）及其改进模型（如 Deformable DETR 等）。

本课题旨在针对以上问题进行深入研究，在技术迁移的同时考虑更多更复杂的信号检测的具体情景，针对其中特征进行数据处理、模型结构设计和模型训练等方面的针对性优化，从而使目标检测中的原技术更加适配地融入信号检测的任务中，同时提高信号检测的准确率和鲁棒性，推动深度学习技术在智能频谱感知领域的深化应用与落地实践。

2 研究现状及分析

2.1 国内外研究现状

在无线通信信号处理领域，实现对宽带频谱中各类信号的精确、快速检测与识别，是频谱管理、认知无线电、电子对抗等应用的核心前提。针对这一目标，国内外的研究历程大致可划分为两个阶段：传统信号处理方法阶段和基于深度学习的智能检测阶段。

在传统方法中，信号的“定位”与“分类”是两条相互割裂的技术路线。定位技术从能量检测、匹配滤波等一维分析，发展到基于短时傅里叶变换（STFT）和数学形态学处理的二维时频谱图分析，但普遍存在泛化性差、抗噪能力弱且无法识别信号类型的问题。而信号分类技术，无论是基于专家特征的传统机器学习还是早期的卷积神经网络（CNN）应用，虽能识别调制方式，却无法提供精确的时频位置信息，导致两条技术路径始终未能融合。

深度学习目标检测技术的发展为实现定位与分类一体化提供了全新范式。其技术从经典的 Faster R-CNN、YOLO 等卷积网络模型，演进至以 DETR 及其改进型为代表的、更为先进的端到端 Transformer 架构，为信号检测提供了强大的现成工具集。

受此启发，研究者们开始将目标检测框架应用于信号时频谱图，将信号检测重构为视觉识别任务，并成功验证了该思路的可行性。然而，当前研究仍处于初级阶段，存在三大核心不足：一是数据集局限性，多数研究采用仿真数据或过于简化的场景，与信号密集、重叠、信噪比多变的真实复杂电磁环境严重脱节；二是模型适配性不足，多为直接套用现有视觉模型，缺乏对信号时频谱图物理特性的针对性结构优化；三是工程实用性欠缺，普遍忽略了真实测量中因采样参数变化导致数据尺寸不一的关键问题。

因此，设计能够应对真实复杂电磁环境，并有效解决上述挑战的先进深度学习模型，是当前该领域的研究热点与核心方向。

2.2 国内外文献综述及简析

2.2.1 传统信号定位与识别方法研究现状

2.2.1.1 传统信号定位方式

在信号定位的早期阶段，主要的经典方法是对信号的一维时间序列数据进行能量检测（Energy Detection）和匹配滤波（Matched Filtering）等操作。

能量检测是一种基础的非相干检测方法，其核心思想是在特定频段内对信号能量进行积分，并与预设的门槛进行比较，从而判断信号是否存在，其一般公式为：【公式】。Quan 等在 2008 年的研究^[1] 中为认知无线电介绍了一种多频段联合检测技术，该技术通过联合检测多个子频带的信号能量水平来识别频谱空洞。然而，这类将宽带频谱划分为多个窄带进行独立检测的方法会导致频谱分析的颗粒度较大，频谱利用效率较低；Bkassiny 等在 2012 年的研究^[2] 中提出了采用盲能量检测进行信道感知，然后利用循环平稳性特征来检测周期性信号，并结合无监督聚类进行分类。

匹配滤波（Matched Filtering）则是一种相干检测方法，其核心思想是通过将接收信号与一个已知的信号模板（例如主用户的导频信号或同步序列）进行互相关运算，当存在目标信号时，滤波器输出会达到峰值，从而实现信号的检测，其一般公式如下：【公式】。Bao 等在 2013 年的研究^[3] 中利用了信号与噪声在统计分布上存在差异这一先验信息，通过分析直方图来自动寻找区分信号与噪声基底的阈值，从而实现对占用频段的定位。然而，匹配滤波方法最大的局限性在于其对先验知识的强依赖性。检测方必须预先精确知道待检测信号的波形、调制方式等完整信息，这使得它在非合作通信或盲信号检测等场景下难以适用。此外，当需要检测多种不同类型的信号时，必须设计相应的多个匹配滤波器，这导致系统实现非常复杂，缺乏灵活性。

为了克服一维分析的局限性，研究人员引入了时频分析工具，其中短时傅里叶变换（Short-Time Fourier Transform, STFT）是应用最广泛的一种。通过 STFT，一维的时序信号可以被转换为二维的时频谱图（Spectrogram），直观地展示了信号频率随时间的变化情况。这一转变使得研究者可以借鉴计算机视觉领域中成熟的图像处理技术来分析和检测信号。

在对于二维时频谱图的处理中，基于数学形态学（Mathematical Morphology）的处理方法成为一个研究热点。该方法利用腐蚀、膨胀、开运算和闭运算等一系列形态学操作，对时频谱图进行去噪、边缘增强和目标提取。例如，Mankun 等在 2007 年的研究^[4] 中，研究者利用图像增强和数学形态学技术对跳频（FH）信号的时频谱图进行分析，有效地提取了信号的轮廓和跳变模式；Phonsri 等在 2015 年的研究^[5] 中进一步结合了计算机视觉与双向神经网络，用于从含噪的时频谱图中提取通信信号。这些方法将信号检测问题转化为图像分割或目标提取问题，实现了信号在时频二维平面上的定位。然而，其缺点也十分明显：首先，形态学操作对噪声和干扰非常敏感，且其参数（如结构元素的尺寸和形状）需要根据信号特征进行人工设计，泛化能力较差；其次，这些方法主要集中于“定位”，即找出信号在

时频谱图中的位置和形状，但无法在此基础上进一步提供信号的语义信息，如信号的调制类型或通信制式等。

随着深度学习技术的发展，一些研究开始尝试利用神经网络来处理频域数据。例如，Huang 等在 2020 年的研究^[6] 中采用了全卷积网络（FCN）对宽带功率谱进行分析，以实现载波信号的检测。这种方法虽然在信号的二维时频结构信息基础上利用了深度学习强大的特征提取能力，提升了频率定位的准确性，但其本质上同样无法完成时频联合定位的任务。

2.2.1.2 传统信号分类方式

在深度学习普及之前，基于传统机器学习的信号分类方法是主流。这类方法通常遵循“特征工程 + 分类器”的范式。研究者首先从信号中手动提取一系列专家特征，如高阶累积量、循环谱特征、小波变换特征等，然后将这些特征输入到支持向量机（SVM）、决策树等分类器中进行训练和分类。例如，在 2008 年^[7] 和 2016 年^[8] 的两项研究中，学者们分别利用小波特征和新颖的统计特征，结合 SVM 分类器，对数字信号进行了自动调制识别。这类方法在特定数据集上取得了不错的分类效果，但其核心缺陷在于严重依赖专家知识进行特征设计，导致特征的泛化能力和鲁棒性有限。更重要的是，这些方法的设计初衷是“分类”，输入的是信号片段，输出的是类别标签，完全无法提供信号在时频域中的精确位置信息，无法做到定位与分类一体化。

深度学习的引入为信号分类带来了突破。卷积神经网络（CNN）能够自动从原始数据中学习和提取特征，避免了复杂的手动特征工程。Bitar 等在 2017 年的研究^[9] 中将信号的时频信息转换为基于功率谱密度的频谱图，并将其作为图像输入到一个五层 CNN 中，成功地对 WiFi、蓝牙和 ZigBee 等多种无线技术进行了分类。该研究还探讨了重叠信号的分类问题，展示了深度学习在复杂电磁环境下的应用潜力。然而，这种方法依旧未能解决信号在时间和频率上的定位问题。

综上所述，传统信号处理方法中的定位和分类在技术路线上是割裂的。直到计算机视觉的目标检测任务得到充分的研究，才为实现端到端的时频联合定位与分类提供了可能。

2.2.2 基于深度学习的目标检测方法研究现状

深度学习的浪潮极大地推动了计算机视觉领域的发展，其中核心任务之一即为目标检测，其旨在从图像中同时预测出目标的边界框（Bounding Box）并对目标进行分类。

最初，以 Faster R-CNN^[10] 为代表的两阶段（two-stage）检测器取得了巨大成功。

该方法首先通过区域提议网络（RPN）生成一系列可能包含目标的候选框（anchor-based），然后再对这些候选框进行二次分类和位置精修，精度较高但速度较慢。与之相对的是以 YOLO^[11] 和 SSD^[12] 为代表的单阶段（one-stage）检测器，它们取消了区域提议步骤，直接在特征图上预测目标的类别和位置，实现了更快的检测速度。其中，YOLO 系列是典型的无锚框（anchor-free）算法，而 SSD 则沿用了锚框（anchor-based）机制。

一个关键的里程碑是特征金字塔网络（Feature Pyramid Network, FPN）^[13] 的提出。FPN 通过构建多尺度的特征图，并融合高层语义信息和底层位置信息，极大地增强了模型对不同尺度，尤其是小目标的检测能力。此后，无论是两阶段还是单阶段的检测器，引入 FPN 或其变体思想已成为标准配置。

在 FPN 的基础上，后续工作继续演进。例如，RetinaNet^[14] 通过引入 Focal Loss 解决了单阶段检测器中正负样本极度不平衡的问题，在保持高速的同时达到了媲美两阶段检测器的精度。FCOS^[15] 则是一种更为彻底的 anchor-free 单阶段检测器，它将目标检测视为一种逐像素的预测任务，进一步简化了模型设计。

近年来，Transformer^[16] 架构在自然语言处理领域取得成功后，也被引入到计算机视觉中。DETR^[17] 是首个将 Transformer 完全应用于目标检测的工作，它将目标检测视为一个集合预测问题，无需 NMS（非极大值抑制）等后处理操作，实现了端到端的检测。但早期的 DETR 存在收敛速度慢、对小目标检测能力不足等问题，且未有效利用 FPN 的多尺度思想。为了解决这些问题，Deformable DETR^[18] 应运而生，它通过引入可变形注意力机制，只关注特征图上少数关键的采样点，并结合了多尺度特征图，显著提升了模型的性能和收敛速度。

由于计算机视觉中的目标检测的任务与信号检测的任务十分接近，所以其技术演进为解决信号时频联合检测问题提供了直接的思路和工具。

2.2.3 基于深度学习的信号检测方法研究现状

借鉴于深度学习目标检测的成功，研究者们开始将这类框架应用于二维时频谱图，以实现信号的端到端的时频定位检测与分类识别的一体化处理。

早期的探索性工作验证了这一思路的可行性。例如，Ke 等人在 2019 年的研究^[19] 中将卷积神经网络（CNN）与长短期记忆网络（LSTM）相结合，以分别提取信号的频率特征和时间（空间）特征，对非合作通信信号进行盲检测。然而，该研究并未直接采用成熟的目标检测框架。随后，主流的目标检测模型被陆续引入。Zha 等人在 2019 年的研究^[20] 中采用了 SSD 框架进行信号检测与调制分类，Prasad 等人在 2020 年的研究^[21] 中将 Faster R-CNN 框架用于宽带系统中的盲时频定位，

并在同年的另一项研究^[22]中对其进行了轻量化裁剪；Li 等人在 2021 年的研究^[23]中采用了改进的 YOLOv3 模型进行信号的盲检测。这些工作证明了将经典目标检测模型迁移至信号领域的有效性，但也存在共同的不足：首先，所使用的数据集较为简单，大多是通过算法模拟生成的数字信号，而非真实环境中采集的信号，且信号种类较少（如数据集仅包含 WiFi 信号或仅将信号分类为 Continuous 和 Burst 两类），并且没有考虑感兴趣信号之间相互重叠的复杂情况，未能反映真实电磁环境的复杂性；其次，这些研究更侧重于验证方法的可行性，在网络结构创新和性能提升方面着力不多。

为了解决传统模型在特定信号形态上的检测难题，一些研究开始设计更具针对性的网络结构。Li 等人在 2022 年的研究^[24]中针对高频（HF）宽带频谱图，提出了一种无锚框的框架。该框架利用每个频率点下的所有时间点特征来联合预测信号的中心频率和形状属性。这项工作在数据集的复杂度上有所提升，单个样本中包含了更多的信号实例。然而，其设计的网络结构较为简单，完全基于卷积，缺乏对长距离依赖的建模能力；其预测方式是按频率维度逐点进行，这使得它天生无法处理频率维度上相互重叠的信号。

最新的研究则开始关注更具挑战性的场景和更先进的网络结构。Peng 等人在 2024 年的研究^[25]中提出了一种结合残差扩张网络和水平移位注意力的模型。该研究的数据集包含了拥挤的信号环境，但信号类别和调制方式依然有限（仅分为两类和三种调制方式），且没有考虑在真实测量场景中，由于采样率和采样时间不同导致 STFT 时频谱图尺寸不统一的实际问题。在网络结构方面，该模型仍未引入表现出强大建模能力的 Transformer 架构。

综上所述，将深度学习目标检测应用于信号时频联合检测已成为该领域的主流趋势，并展现出巨大潜力。然而，现有研究仍存在以下主要不足：首先，大多数研究使用的数据集或为仿真生成，或在信号类型、密度、信噪比范围以及重叠复杂性上与真实电磁环境存在较大差距，限制了模型的泛化能力和实用价值；其次，多数研究直接套用计算机视觉领域的成熟模型，缺乏对信号时频谱图特有物理属性（如信号的连续性、时变性、长宽比悬殊等）的针对性设计和优化，尤其在长距离依赖关系建模和重叠信号处理上能力欠缺；最后，没有考虑在真实测量场景中，由于采样率和采样时间不同导致 STFT 时频谱图尺寸不统一的实际问题。因此，本课题拟围绕上述不足展开研究，旨在设计一种更先进的新型深度学习模型，以应对复杂真实电磁环境下的信号时频联合检测与分类挑战。

3 主要研究内容及研究方案

3.1 研究内容

3.1.1 针对不同尺度信号数据的处理方法研究

在真实的电磁监测场景中，信号样本的采集参数（如采样率、采样时长）往往是多样化的，这就带来了一个关键的技术挑战。为了保证信号特征在频域和时域上具有一致的可比性，以便于神经网络进行有效的模式学习，在进行短时傅里叶变换（STFT）时，必须设定固定的频率分辨率和时间分辨率。然而，这一约束导致了不同采样参数的样本在经过 STFT 处理后，会生成尺寸大小不一的二维时频谱图。

传统的计算机视觉模型，特别是包含全连接层的网络，通常要求输入图像具有固定的尺寸。虽然可以通过插值缩放或暴力填充（Padding）的方式统一尺寸，但这两种方法均不适用于本课题。首先，对时频谱图进行插值或缩放会严重破坏信号固有的时频结构和物理意义，导致特征失真；其次，若将所有样本统一填充至数据集中最大样本的尺寸，当尺寸差异悬殊时，会引入大量的零值填充区域，这不仅造成了训练过程中内存与显存资源的巨大浪费，也可能干扰模型的特征学习。

针对此问题，本研究拟借鉴并改造目标检测领域的多尺度训练（Multiscale Training）策略。与目标检测中将多尺度作为一种数据增强、提升模型鲁棒性的可选策略不同，在本课题中，它是一种处理异构数据的必要策略。具体而言，本研究将利用所选模型 DETR（由 CNN 骨干网络、Transformer 及线性预测层构成）对输入尺寸无严格固定要求的结构优势，在训练时将尺寸相同的 STFT 数据组成批次（batch），使得网络既可以输入不同尺寸的二维数据，也可以进行以批次为基本单位的训练，使得模型能够稳健而高效地处理来源多样化的真实信号数据。

3.1.2 基于 Transformer 检测头的信号检测模型研究

模型的检测头，即负责从深度特征图中解码出目标位置和类别的网络末端部分，是决定检测器性能的关键。依据其结构特性，主流的检测头可分为全连接（Fully Connected）、全卷积（Fully Convolutional）和基于 Transformer 三种范式，它们各有优劣：

1. 全连接检测头：其结构要求输入特征图必须为固定尺寸，这与前述的多尺度输入策略相悖，缺乏灵活性。但其优点在于可以人为设定固定数量的预测输出，无需复杂的后处理。

2. 全卷积检测头：该结构能够灵活地适应不同尺寸的输入，但其输出的预测框数量不固定，与输入尺寸相关，因此需要依赖非极大值抑制（NMS）等复杂的后处理算法来滤除冗余的预测框。

3. Transformer 检测头：该结构结合了前两者的优势：既能处理可变尺寸的输入，又能通过其内部的对象查询（Object Queries）机制输出固定数量的预测结果，从而省去了复杂的后处理环节，构建了真正意义上的端到端检测流程。

更重要的是，Transformer 架构的核心——自注意力机制（Self-Attention），使其具备强大的远距离甚至全局依赖建模能力。这对于捕捉信号间的谐波关系、跳频模式等全局性特征具有潜在优势。然而，从时频谱图的视觉先验来看，信号的形态多为局部性较强的矩形或条状。全局注意力机制是否会引入不必要的计算冗余，以及其对于信号局部特征的捕捉是否优于卷积网络，是一个需要通过实验深入探究的关键问题。因此，本研究将重点构建并评估基于 Transformer 的检测头，验证其在信号检测任务中的实际效能。

3.1.3 针对窄带短时信号的模型优化策略研究

在时频谱图中，窄带短时信号（如猝发信号、短时通信等）表现为尺寸微小、持续时间短的“小目标”，这对检测模型提出了极高的要求。标准的检测模型在处理这类小目标时，由于特征图分辨率较低、感受野过大等问题，常常出现漏检。为了解决这一难题，本研究拟引入 Deformable DETR 模型进行优化：首先，借鉴特征金字塔网络（FPN）的思想，Deformable DETR 融合了骨干网络在不同阶段输出的多尺度特征图。高层特征图包含丰富的语义信息，有助于信号分类，底层特征图具有更高的空间分辨率，保留了精确的位置信息。通过将二者有效结合，模型能够同时兼顾对大、小不同尺度信号的检测能力；其次，Deformable DETR 采用了可变形注意力机制。标准的 Transformer 注意力机制会对特征图上的所有像素点进行计算，对于仅占图像一小部分的稀疏信号而言，这会带来巨大的计算浪费。可变形注意力机制不再对整个特征图进行全局计算，而是让网络根据输入动态地学习少数关键采样点的位置，并将注意力集中在这些与目标最相关的区域。这不仅大幅降低了模型的计算复杂度和内存消耗，还使得注意力能够更精准地聚焦于目标区域，从而显著提升对窄带短时等小目标的检测精度。

3.2 研究方案

为实现上述研究内容，本课题计划遵循以下技术路线和研究方案，分阶段进行：

1. 数据集分析与预处理：本研究将基于已有的真实电磁环境信号数据集展开。首先，对数据集进行全面的统计性分析，探究其中信号类型与调制方式的类别均衡性、不同信号的时宽与带宽分布、以及单个样本中的信号密度等关键特征。随后，核心的预处理步骤是将一维时域采样数据通过短时傅里叶变换（STFT）转换为二维时频谱图。在此过程中，将仔细选择窗函数、窗长及重叠率等参数，以在时间和频率分辨率之间取得最佳平衡。完成变换后，将对部分样本进行可视化观察，以直观检验 STFT 参数的合理性，并确认是否需要对生成的时频谱图进行幅度归一化（如 Min-Max Scaling 或 Z-Score Normalization）处理，以利于神经网络模型的稳定训练与快速收敛。

2. 基线模型搭建与实验：为了科学、定量地评估本课题所提方法的有效性，将首先搭建一个基线（Baseline）模型。考虑到 YOLO 系列在目标检测领域中速度与精度的良好平衡性，拟选择 YOLOv5 作为基线模型。将在预处理完成的数据集上对其进行训练与测试，并详细记录其在信号检测任务中的各项性能指标，主要包括平均精度均值（mAP）、精确率（Precision）、召回率（Recall）以及检测速度（FPS）。该基线模型的性能将作为后续所有改进模型进行比较和评估的参照标准，以凸显本研究方法的先进性。

3. 核心模型的设计与实现：本阶段是研究的核心，将重点围绕 Deformable DETR 模型展开，将其迁移至信号时频检测这一特定场景并进行适配与优化。具体实现步骤如下：首先，搭建基于 Deformable DETR 的信号检测框架，该框架由 CNN 骨干网络（如 ResNet）、集成了多尺度特征融合与可变形注意力的 Transformer 编解码器，以及用于目标分类和边界框回归的线性预测层构成。其次，针对性地解决多尺度输入问题，在训练流程中设计数据加载策略，将尺寸相同的时频谱图组织在同一批次（batch）中送入网络，充分利用模型对可变尺寸输入的兼容性。最后，将根据信号数据的特性对模型进行微调，例如调整 Transformer 中的对象查询（Object Queries）数量以匹配数据集中信号的典型密度，或优化损失函数中分类损失与回归损失的权重，使模型更专注于本任务的特定挑战。

4. 综合实验与性能评估：在完成核心模型的设计与实现后，将进行一系列综合性实验来全面评估其性能。（1）整体性能对比：将所提出的 Deformable DETR 模型与基线模型（YOLOv5）在同一测试集上进行比较，分析两者在 mAP、精确率、召回率等关键指标上的差异，验证核心模型的总体优越性。（2）消融实验：为验证模型关键组件的有效性，将设计消融研究。例如，通过对比 Deformable DETR 与使用单尺度特征的基础版 DETR 的性能，来量化多尺度特征融合（FPN 思想）与可变形注意力机制对于检测窄带短时“小目标”信号的具体提升效果。（3）定性

分析：除了定量的指标对比，还将对检测结果进行可视化分析。重点选取一些具有挑战性的样本，如包含低信噪比信号、密集重叠信号或极端长宽比信号的频谱图，直观对比不同模型在这些复杂场景下的检测效果，以展示本研究模型在鲁棒性与精确性上的优势。

5. 总结结论与论文撰写：最后，将系统性地整理和总结所有实验数据与分析结果，凝练出本研究的核心结论，明确所提出的信号检测模型相较于传统方法和基线模型的优势所在，并客观分析其可能存在的局限性。在此基础上，将对未来可能的研究方向进行展望，并最终依据学位论文的规范要求，完成毕业设计的撰写工作。

4 进度安排及预期目标

4.1 进度安排

本课题的研究时间为 2023 年 9 月开始至 2024 年 5 月份进行答辩，根据之前的项目经验，结合自身学习，科研攻关能力，针对课题研究计划，制定出下面的研究进度：

1. 2025 年 9 月至 2025 年 10 月：深入研读相关领域的最新文献，完善并最终确定详细的技术实现方案；对信号数据集进行全面的统计性分析，完成 STFT 转换、数据清洗、归一化及标注格式统一等所有预处理工作；完成深度学习实验环境的配置与调试，并用一个简单的模型跑通数据加载和训练流程，确保环境稳定可用。
2. 2025 年 11 月至 2026 年 3 月：实现 YOLOv5 基线模型以及基于 Deformable DETR 的核心检测模型，并解决多尺度时频谱图的输入与批处理问题，在数据集子集上对核心模型进行初步训练，验证其设计的正确性与可行性，之后在完整数据集上完成训练与评估，获取作为参照标准的各项性能指标，进行全面的性能对比实验与分析。
3. 2026 年 4 月至 2026 年 5 月：根据已有的实验结果，撰写毕业论文，并准备毕业设计答辩。

4.2 预期目标

本课题旨在通过上述研究，达成以下具体目标：

1. 对数据集进行详细且深入的分析，构建科学合理的实验验证环境。深入分析已有的真实电磁环境信号数据集，包括信号类别分布、重叠程度等关键特征。通过短时傅里叶变换（STFT）完成数据的预处理，并针对不同采样参数产生的异构尺寸时频谱图，设计并构建一套能够支撑多尺度输入训练的、科学且规范的实验数据集和验证集。
2. 复现 YOLOv5 的基线模型并将其转化为针对本任务的强力基准。成功复现并实现一个高性能的 YOLOv5 模型作为基线（Baseline）。将该模型适配到信号时频谱图检测任务中，对其检测结果进行全面评估，并获取可靠的性能指标（mAP、精确率、召回率等），为后续核心模型的性能对比提供一个强大的、具有说服力的参照标准。

3. 实现基于 Deformable DETR 的模型，并验证其在信号检测中的优越性。成功将先进的端到端目标检测框架 Deformable DETR 迁移并适配到信号时频联合检测任务中，并对模型结构（如骨干网络、编码器、解码器）进行优化，以更好地捕捉信号的局部与全局特征。通过实验验证该模型相较于传统 CNN 架构检测器的优越性，特别是在复杂电磁环境下对密集、微小信号的鲁棒检测能力。
4. 实现针对本任务的多尺度训练策略，解决异构数据输入的工程难题。针对真实场景中 STFT 时频谱图尺寸不一的问题，创新性地实现一种多尺度训练（**Multiscale Training**）策略。该策略应能够高效、稳定地处理不同尺寸的时频谱图数据批次，避免传统插值或暴力填充导致的特征失真和资源浪费，从根本上解决真实场景中异构数据输入的工程性难题。

5 已具备和所需的条件和经费

5.1 实验室条件和经费保障

5.1.1 实验室已具备条件

实验室已具备高性能计算服务器，其安装有 Ubuntu 系统、Python 环境以及 PyTorch 等必需的配套库，同时配备了 128G 内存和 NVIDIA 4090 GPU，具有充足的内存和显存供数据集的加载和模型的训练。

5.1.2 实验室经费保障

本课题主要通过编写代码完成，无需硬件实现，同时训练所需的硬件设备均已具备，故无需经费支持。

5.2 所需条件和经费

所需条件和经费已经在 5.1 中详细列出并已具备，故无需另外条件和经费。

6 预计困难及解决方案说明

6.1 技术难点与预计困难

1. 数据的异构性与复杂性本研究使用的现成数据集包含因不同采样参数而产生的尺寸不一的时频谱图。如何高效地将这些异构数据送入模型进行训练，是一个关键的工程难题。此外，数据集中包含了种类多样、密度高、且时频重叠的复杂信号场景，这对模型的特征提取与区分能力提出了巨大挑战。

2. 模型的高复杂性与超参数调优本课题选用的 Deformable DETR 模型，其结构（尤其是 Transformer 模块）远比传统 CNN 复杂。这导致了超参数调优的复杂度急剧升高，涉及模型结构参数（如模块层数）、优化器与学习率调度器参数（如学习率、Warmup 轮数）、以及损失函数各部分权重系数等众多变量。寻找最优超参数组合将非常耗时，而调优不当极易导致模型不收敛、欠拟合或过拟合。

3. 微小信号检测的瓶颈在时频谱图中，大量关键信号（如猝发信号）表现为尺寸极小的“小目标”（即时宽窄、带宽窄）。这类目标经过骨干网络的多层下采样后，其特征信息极易丢失，是所有检测模型的公认难题。如何确保模型在信号密集的复杂背景下，依然能够精准地定位和识别这些微小信号，将是衡量模型性能的关键瓶颈。

4. 训练开销与显存瓶颈引入 Transformer 模块后，模型的计算量和显存占用会显著增长。在有限的硬件条件下，若显存溢出，常用的解决方法是降低批大小（Batch Size）。然而，过小的批大小会导致批数量剧增，不仅严重拖慢训练速度，还可能影响模型的收敛效果，使训练变得异常困难。

6.2 解决方案

针对上述可能遇到的困难，本课题已预先制定了如下应对策略和解决方案：

1. 针对数据层面挑战的解决方案：(1) 针对数据集中尺寸不一的固有特性，将通过编写定制化的数据整理函数（Collate Function）来实现一个智能的数据加载流程。该流程可以在构建每个训练批次时，自动将尺寸相同或相近的样本组织在一起，从而在不改变原始数据结构的前提下，高效地进行模型训练；(2) 在现有的训练集上，采用时频遮挡（Time-Frequency Masking）、随机平移等数据增强策略。这并非改变数据集本身，而是通过在训练过程中对输入数据进行随机变换，以提升模型对信号位置变化或部分遮挡的鲁棒性，有效防止模型在固定数据集上产生过拟合。

2. 针对模型层面挑战的解决方案：（1）充分利用迁移学习，加载在大型公开数据集上预训练好的模型权重作为坚实的调整起点。同时，在文献调研中学习经典论文的训练策略，优先测试其超参数设置。（2）采用控制变量法，在调优初期优先调整学习率等关键超参数，积累经验后逐步寻找最优组合。训练中密切监控损失曲线，并使用早停（Early Stopping）策略，避免无效实验造成时间浪费。（3）将验证集上的模型输出结果进行可视化，直观地分析漏检、误检等问题，从而对模型结构或损失函数进行针对性的调整。
3. 针对微小信号检测瓶颈的解决方案：（1）重点发挥 Deformable DETR 的多尺度特征融合能力和可变形注意力机制，确保底层高分辨率特征被有效利用，使注意力能够精准聚焦于微小信号区域。（2）适当调整损失函数中分类损失与边界框回归损失的权重系数，引导模型更加关注对小目标的定位精度。
4. 针对资源层面挑战的解决方案：（1）采用先进训练技术：优先尝试混合精度训练（AMP），理论上可以在几乎不损失精度的情况下大幅度减少显存占用并提高训练速度。同时，可以尝试梯度累积技术，在不增加显存消耗的情况下，实现等效的大批次训练效果。（2）利用分布式训练：如果条件允许，利用多 GPU 进行数据并行或模型并行训练，直接扩大可用的显存上限，从根本上解决显存瓶颈问题。

参考文献

- [1] QUAN Z, CUI S, SAYED A H, et al. Wideband Spectrum Sensing in Cognitive Radio Networks [C/OL] // 2008 IEEE International Conference on Communications, [S.I.], 2008 : 901-906. <http://dx.doi.org/10.1109/ICC.2008.177>.
- [2] BKASSINY M, JAYAWEERA S K, LI Y, et al. Wideband Spectrum Sensing and Non-Parametric Signal Classification for Autonomous Self-Learning Cognitive Radios [J/OL] . IEEE Transactions on Wireless Communications, 2012, 11 (7) : 2596-2605. <http://dx.doi.org/10.1109/TWC.2012.051512.111504>.
- [3] BAO D, DE VITO L, RAPUANO S. A Histogram-Based Segmentation Method for Wideband Spectrum Sensing in Cognitive Radios [J/OL] . IEEE Transactions on Instrumentation and Measurement, 2013, 62 (7) : 1900-1908. <http://dx.doi.org/10.1109/TIM.2013.2251821>.
- [4] MANKUN X, XIJIAN P, TIANYUN L, et al. A New Time-Frequency Spectrogram Analysis of FH Signals by Image Enhancement and Mathematical Morphology [C/OL] // Fourth International Conference on Image and Graphics (ICIG 2007), [S.I.], 2007 : 610-615. <http://dx.doi.org/10.1109/ICIG.2007.154>.
- [5] PHONSRI S, MUKHERJEE S S, SELLATHURAI M. Computer vision and bi-directional neural network for extraction of communications signal from noisy spectrogram [C/OL] // 2015 IEEE Conference on Antenna Measurements & Applications (CAMA), [S.I.], 2015 : 1-4. <http://dx.doi.org/10.1109/CAMA.2015.7428185>.
- [6] HUANG H, LI J-Q, WANG J, et al. FCN-Based Carrier Signal Detection in Broadband Power Spectrum [J/OL] . IEEE Access, 2020, 8 : 113042-113051. <http://dx.doi.org/10.1109/ACCESS.2020.3003683>.
- [7] PARK C-S, CHOI J-H, NAH S-P, et al. Automatic Modulation Recognition of Digital Signals using Wavelet Features and SVM [C/OL] // 2008 10th International Conference on Advanced Communication Technology : Vol 1, [S.I.], 2008 : 387-390. <http://dx.doi.org/10.1109/ICACT.2008.4493784>.
- [8] HASSANPOUR S, PEZESHK A M, BEHNIA F. Automatic Digital Modulation Recognition Based on Novel Features and Support Vector Machine [C/OL] // 2016 12th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS), [S.I.] , 2016 : 172-177. <http://dx.doi.org/10.1109/SITIS.2016.35>.

- [9] BITAR N, MUHAMMAD S, REFAI H H. Wireless technology identification using deep Convolutional Neural Networks [C/OL] //2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), [S.I.], 2017 : 1-6. <http://dx.doi.org/10.1109/PIMRC.2017.8292183>.
- [10] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks [J/OL] . IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39 (6) : 1137-1149. <http://dx.doi.org/10.1109/TPAMI.2016.2577031>.
- [11] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [C] //Proceedings of the IEEE conference on computer vision and pattern recognition, [S.I.], 2016 : 779-788.
- [12] LIU W, ANGUELOV D, ERHAN D, et al. Ssd: Single shot multibox detector [C] //European conference on computer vision, [S.I.], 2016 : 21-37.
- [13] LIN T-Y, DOLLÁR P, GIRSHICK R, et al. Feature Pyramid Networks for Object Detection [C/OL] //2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), [S.I.], 2017 : 936-944. <http://dx.doi.org/10.1109/CVPR.2017.106>.
- [14] LIN T-Y, GOYAL P, GIRSHICK R, et al. Focal Loss for Dense Object Detection [C/OL] //2017 IEEE International Conference on Computer Vision (ICCV), [S.I.], 2017 : 2999-3007. <http://dx.doi.org/10.1109/ICCV.2017.324>.
- [15] TIAN Z, SHEN C, CHEN H, et al. FCOS: Fully Convolutional One-Stage Object Detection [C/OL] //2019 IEEE/CVF International Conference on Computer Vision (ICCV), [S.I.] , 2019 : 9626-9635. <http://dx.doi.org/10.1109/ICCV.2019.00972>.
- [16] VASWANI A, SHAZER N, PARMAR N, et al. Attention is all you need [J] . Advances in neural information processing systems, 2017, 30.
- [17] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with transformers [C] //European conference on computer vision, [S.I.], 2020 : 213-229.
- [18] ZHU X, SU W, LU L, et al. Deformable detr: Deformable transformers for end-to-end object detection [J] . arXiv preprint arXiv:2010.04159, 2020.

- [19] KE D, HUANG Z, WANG X, et al. Blind Detection Techniques for Non-Cooperative Communication Signals Based on Deep Learning [J/OL] . IEEE Access, 2019, 7 : 89218-89225. <http://dx.doi.org/10.1109/ACCESS.2019.2926296>.
- [20] ZHA X, PENG H, QIN X, et al. A deep learning framework for signal detection and modulation classification [J] . Sensors, 2019, 19 (18) : 4042.
- [21] PRASAD K N R S V, DSOUZA K B, BHARGAVA V K, et al. A Deep Learning Framework for Blind Time-Frequency Localization in Wideband Systems [C/OL] //2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring), [S.I.], 2020 : 1-6. <http://dx.doi.org/10.1109/VTC2020-Spring48590>. 2020.9128779.
- [22] PRASAD K N R S V, D' SOUZA K B, BHARGAVA V K. A Downscaled Faster-RCNN Framework for Signal Detection and Time-Frequency Localization in Wideband RF Systems [J/OL] . IEEE Transactions on Wireless Communications, 2020, 19 (7) : 4847-4862. <http://dx.doi.org/10.1109/TWC.2020.2987990>.
- [23] LI R, HU J, LI S, et al. Blind Detection of Communication Signals Based on Improved YOLO3 [C/OL] //2021 6th International Conference on Intelligent Computing and Signal Processing (ICSP), [S.I.], 2021 : 424-429. <http://dx.doi.org/10.1109/ICSP51882.2021.9408998>.
- [24] LI W, WANG K, YOU L, et al. A New Deep Learning Framework for HF Signal Detection in Wideband Spectrogram [J/OL] . IEEE Signal Processing Letters, 2022, 29 : 1342-1346. <http://dx.doi.org/10.1109/LSP.2022.3179958>.
- [25] PENG T, YANG S, FENG Z, et al. Spectrum Sensing via Residual Dilated Network and Horizontal Shift Attention for Cognitive IoT [J/OL] . IEEE Internet of Things Journal, 2024, 11 (22) : 36817-36828. <http://dx.doi.org/10.1109/JIOT.2024.3429200>.