

本科毕业论文

基于深度学习的水下声纳图像去噪

**RESEARCH ON UNDERWATER SONAR
IMAGE DENOISING BASED ON DEEP
LEARNING**

胡章睿

哈尔滨工业大学

2025 年 10 月

密级：公开

本科毕业论文

基于深度学习的水下声纳图像去噪

本 科 生：胡章睿

学 号：220320619

指 导 教 师：胡靓教授

专 业：自动化

学 院：机器人与先进制造学院

答 辩 日 期：2025 年 10 月 30 日

学 校：哈尔滨工业大学

摘 要

前视声纳因其在浑浊水域、低能见度及光照不足等极端条件下仍能稳定成像,已经成为水下探测与水下机器人感知中不可或缺的关键传感方式。然而,声纳图像普遍受到散斑噪声、混响噪声、旁瓣噪声以及结构性噪声等多源干扰的影响,导致图像对比度下降、目标边界模糊及特征提取困难。从而严重制约了后续目标识别、建图与自主导航等任务的性能。传统的声纳图像去噪方法虽然能够降低部分噪声,但常出现过度平滑的现象,难以在复杂水下环境中保持结构与边缘信息。近年来,基于深度学习的监督式方法通过学习噪声分布取得了一定成效,但其依赖大量标注数据,需要大量成对(含噪——干净)声纳数据集,获取成本高且跨设备、跨场景泛化能力不足。

针对上述问题,本文提出了一种基于生成对抗网络的渐进式声纳图像去噪框架 SonarGAN,实现多类型噪声的协同抑制与目标结构的自适应保持。该框架由三个阶段组成:

(1) 非配对初步去噪阶段,利用循环一致性生成对抗网络在非配对的真实噪声图与仿真干净图之间学习随机噪声分布,并完成初步去噪;

(2) 成对结构抑制阶段,在条件生成对抗网络的框架下基于合成的噪声——干净图像对进行训练,显示去除结构性噪声;

(3) 联合约束优化阶段,将前两阶段的结果进行融合与约束优化,实现全局一致的综合去噪。

在次基础上,本文在所有生成器中引入自注意力模块以捕捉声纳图像的长程依赖关系,并引入 VGG 感知损失与图像级监督损失对模型进行精细化调优,从而在噪声抑制与结构保真之间取得平衡。本文主要的研究工作与创新点包括以下四方面:

(1) 提出了一种无需昂贵配对数据的声纳图像去噪框架 SonarGAN,可在一次处理中同时抑制多种类型噪声,显著提升声纳图像质量与下游任务鲁棒性。

(2) 设计了多阶段渐进式生成对抗结构与联合优化机制,通过非配对与成对数据的协同学习,实现从随机噪声到结构性噪声的逐步抑制。

(3) 引入感知损失与自注意力机制以保持结构细节与全局一致性,有效避免了传统 GAN 模型中常见的伪影问题和边缘退化现象。

(4) 在多类型声纳的仿真与真实数据集上开展了系统性实验与对比分析,结果表明 SonarGAN 在图像清晰度、结构保持度及泛化性能方面均优于现有方法,验

证了其在真实水下环境中的可行性与稳定性。此外，本文进一步将去噪后的声纳图像用于三维场景重建任务，并于多种主流去噪算法的重建结果进行对比。实验结果显示，使用 SonarGAN 处理后的声纳图像能够显著提升重建模型的集合精度与结构连续性，生成的三维场景更加完整，细节更清晰，从而充分证明了所提方法在下游重建任务中的有效性和优越性。

综上所述，SonarGAN 为多源噪声环境下的声纳图像去噪提供了一种高效、泛化性强的解决方案，并在三维重建等下游任务中展现出良好的适应性与应用潜力，为水下感知与重建研究提供了新的思路。

关键词： $\text{T}_{\text{E}}\text{X}$ ； $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ ；CJK；论文模板；毕业论文

Abstract

Forward-looking sonar (FLS) has become an indispensable sensing modality in underwater exploration and robotic perception due to its ability to provide reliable imaging in turbid waters, low-visibility conditions, and poor illumination environments. However, sonar images are often corrupted by multiple types of noise, including speckle noise, reverberation, sidelobe interference, and structural noise, which significantly degrade image contrast, blur object boundaries, and hinder robust feature extraction. These degradations in turn limit the performance of downstream tasks such as target detection, mapping, and autonomous navigation. Conventional sonar image denoising methods can reduce part of the noise but often lead to over-smoothing, making it difficult to preserve structural and edge details in complex underwater environments. Recently, deep supervised learning-based approaches have achieved promising results by learning noise distributions; however, they rely heavily on large labeled datasets, which require costly paired (noisy-clean) sonar images, and thus suffer from limited generalization across different sonar types and underwater conditions.

To address these challenges, this paper proposes SonarGAN, a progressive generative adversarial network (GAN)-based framework for sonar image denoising that achieves collaborative suppression of multiple noise types while adaptively preserving structural details. The framework consists of three stages:

(1) Unpaired Preliminary Denoising, which employs a Cycle-Consistent Generative Adversarial Network (CycleGAN) to learn random noise distributions between unpaired real noisy and simulated clean images, achieving initial denoising;

(2) Paired Structural Suppression, which uses a Conditional Generative Adversarial Network (pix2pix) trained on synthetic noisy-clean image pairs to explicitly remove structural noise;

(3) Constrained Joint Refinement, which integrates and optimizes the outputs of the first two stages for globally consistent and comprehensive denoising.

Furthermore, a self-attention module is embedded into all generators to capture long-range dependencies in sonar imagery, and both VGG perceptual loss and image-level supervised loss are introduced to fine-tune the generators, effectively balancing noise suppression and structural fidelity. The main contributions of this work are summarized as follows:

(1) We propose SonarGAN, a sonar image denoising framework that requires no expensive paired datasets and can simultaneously suppress multiple types of noise, significantly improving image quality and the robustness of downstream tasks.

(2) We design a multi-stage progressive GAN architecture with joint optimization, which achieves stepwise suppression from random to structural noise through collaborative learning of unpaired and paired data.

(3) We incorporate perceptual loss and self-attention mechanisms to maintain structural details and global consistency, effectively avoiding artifacts and edge degradation common in traditional GAN-based models.

(4) We conduct systematic experiments and comparative analyses on both simulated and real sonar datasets. Results show that SonarGAN outperforms existing methods in terms of image clarity, structural preservation, and generalization ability, verifying its effectiveness and robustness in real underwater environments. Moreover, we further apply the denoised sonar images to 3D scene reconstruction and compare the results with those obtained using several mainstream denoising algorithms. The experimental results demonstrate that sonar images processed by SonarGAN significantly improve the geometric accuracy and structural continuity of the reconstructed 3D scenes, producing more complete and detailed reconstructions. This validates the effectiveness and superiority of the proposed method in downstream reconstruction tasks.

In summary, SonarGAN provides an efficient and highly generalizable solution for sonar image denoising under multi-source noise conditions, and demonstrates excellent adaptability and application potential in downstream tasks such as 3D reconstruction, offering new insights for underwater perception and reconstruction research.

Keywords: \TeX , \LaTeX , CJK, hitszthesis, thesis

目 录

摘 要	I
Abstract	III
第 1 章 绪 论	1
1.1 课题背景及研究的目的和意义	1
1.1.1 课题背景	1
1.1.2 研究的目的和意义	2
1.2 国内外研究现状	4
1.2.1 水下机器人领域研究现状	4
1.2.2 声纳图像去噪研究现状	5
1.3 本文的主要研究内容	6
第 2 章 基于无配对数据的初步声纳去噪	9
2.1 引言	9
2.2 声纳图像成像原理	9
2.3 循环一致性对抗网络	12
2.4 循环一致性对抗网络	14
2.5 声纳图像无配对去噪网络的构建与优化	16
第 3 章 排版公式	17
3.1 引言	17
3.2 对物理量符号进行注释的情况	17
3.3 子公式	18
3.4 本章小结	18
第 4 章 排版表格	19
4.1 引言	19
4.2 普通表格的绘制方法	19
4.3 长表格的绘制方法	19
4.4 列宽可调表格的绘制方法	21
4.4.1 表格内某单元格内容过长的情况	21
4.4.2 排版横版表格的举例	22
4.5 本章小结	22

第 5 章 引用参考文献	24
5.1 引言	24
5.2 参考文献引用方法	24
5.3 本章小结	24
第 6 章 补充说明	26
6.1 引言	26
6.2 关于数字	26
6.3 索引示例	27
6.4 术语排版举例	27
6.5 定理和定义等	27
6.6 其他杂项	28
6.6.1 右翻页	28
6.6.2 算法	28
6.6.3 脚注	28
6.6.4 源码	28
6.6.5 思源宋体	29
6.6.6 术语词汇管理	29
6.7 本章小结	29
结 论	30
参考文献	31
攻读学士学位期间取得创新性成果	34
哈尔滨工业大学本科毕业论文（设计）原创性声明和使用权限	35
致 谢	36

第 1 章 绪 论

1.1 课题背景及研究的目的和意义

1.1.1 课题背景

随着人类近年来对河流、湖泊及海洋资源的深入探索，海底地图的精确测绘与海洋生物多样性的科学研究变得愈发关键。对海洋生物与海底资源的探索是海底研究活动的重要课题，但是鉴于海底环境的极端复杂性和潜在危险性，探测手段正逐步从依赖潜水员直接作业转向利用水下机器人进行远程探测。相较于光学探测设备在陆地上的广泛应用，在水下探测中光学设备面临着严峻的挑战。光线在水中传播时会被吸收和散射，尤其在浑浊或深水区域，光线衰减迅速；并且由于水对不同波长光的吸收程度不同，水下图像通常存在色彩失真。光线在水中传播受限，易受水体中的浑浊物质和微生物吸收和散射影响，导致成像质量显著下降，因此在水下部署可用的基于视觉的状态估计并不简单，特别是，悬浮颗粒、模糊以及光线和颜色衰减会导致图像特征并不像水面上那样清晰定义。因此，来自不同基于视觉的状态估计包的结果显示大量异常值，从而导致不准确定位，甚至完全丢失跟踪^[1]，使得机器人在水下环境的感知能力减弱，进而影响机器人的定位和自主导航能力。如图1-1所示，海洋中水下机器人视觉退化严重。



图 1-1: 水下视觉图像

相比之下，声纳技术在水下探测领域展现出了显著的优势。其核心在于声波能够在水中实现远距离的有效传播，这一特性使得声纳系统能够覆盖更广阔的水域范围，进行更为全面的探测；尤为重要的是，声纳受水体浑浊度的干扰较小，即便是在水质较差、能见度较低的环境中，依然能保持较高的精度和稳定性。因此，声纳技术成为了水下探测领域的优选方案，为水下作业和科研探索提供了强有力

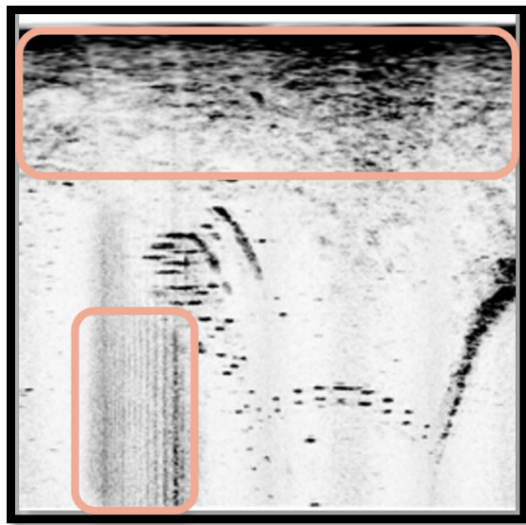
的支持,大多数水下机器人均装备有声纳等声学传感器,可以更好地完成海底基础设施检查、船体检查^[2]和海底测量和测深^[3]等任务。需要注意的是,在本文中,为了视觉的连续性,我们对于所有声纳图像进行了黑白像素翻转,即以白色为声纳图像背景,黑色为声纳探测到的物体部分。

1.1.2 研究的目的是和意义

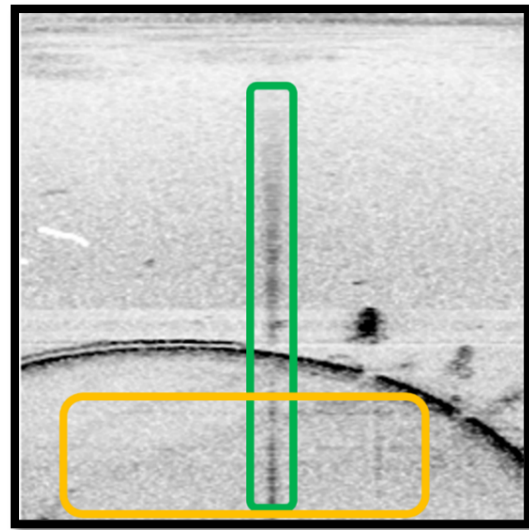
水下声学环境的复杂性源于多种因素,使得声波在传播过程中面临诸多挑战。首先,多路径效应是声波在水中传播时常遇到的现象,声波会因为水下物体的反射和折射而产生多条传播路径,这些路径在抵达接收器时可能会相互干扰,导致信号失真。其次,近场干扰进一步增加了这一复杂性。在近距离内,声波的传播受周围环境的影响显著,声波可能会因为水流、温度梯度和盐度变化等因素而产生波动,影响传感器的准确性。除了这些物理现象,水下声学环境中还存在着多种来源和类型的噪声干扰。来自船只、海洋生物、波浪以及人类活动等的噪声,都会对传感器收集到的原始数据产生显著影响。这些噪声不仅会掩盖目标信号,还可能与目标信号混合,进一步复杂化信号处理过程。因此,传感器常常难以提取有效信息,导致数据质量下降。同时由于受到声波多路径反射的影响,导致真实声纳信号周围产生伪影^[4],从而极大地增加了从这些数据中准确提取特征、识别对象或界定边界的难度。如图1-2所示,不同类型的前视声纳图像中均存在显著的噪声干扰。图1-2 (a)中的含噪声极坐标声纳图像,位于图像顶部区域的噪声是结构性噪声与随机分布的散斑噪声的叠加;位于图像左下方区域的竖线噪声是前视声纳的结构性噪声。图1-2 (b)中的含噪声极坐标图像,位于图像中央的绿色框中的竖线,代表着 Oculus M1200d 声纳的结构性噪声,对于黄色框内的噪声图像,代表着旁瓣噪声的干扰。

现有声纳图像去噪方法主要分为传统滤波类与深度学习类两大类。关于传统的去噪方法,在面对水下声纳图像特有的噪声模式时,往往效果有限,无法有效恢复出清晰高质量的图像信息,如图1-3所示,经过传统滤波处理后,仍存在大量噪声,可见传统滤波处理算法的鲁棒性较差。而深度学习的方法依赖大规模配对数据进行监督训练,而真实水下环境中难以获得与噪声图像严格对应的“干净”样本,导致模型泛化性不足,难以适用于不同声纳设备与环境条件。

针对上述问题,本研究提出一种基于生成对抗网络的渐进式声纳图像去噪框架 SonarGAN,通过结合非配对学习与成对约束,实现多源噪声的联合抑制与结构信息的自适应保持。该方法不仅突破了传统监督方法对配对数据的依赖,还在多种声纳类型及真实水下场景中表现出优异的泛化性能。此外,本文进一步探索了

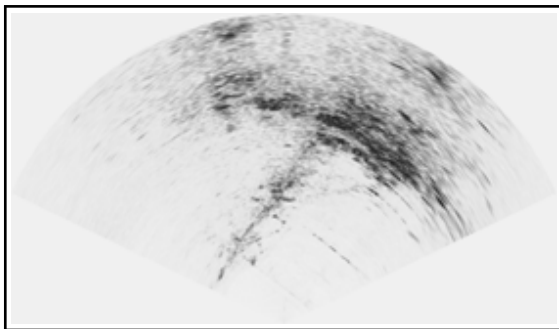


(a) BlueView P900 声纳图像

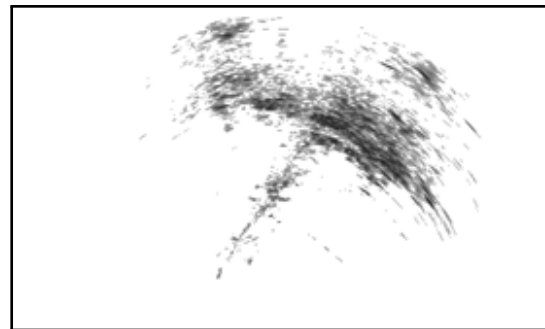


(b) Oculus M1200d 声纳图像

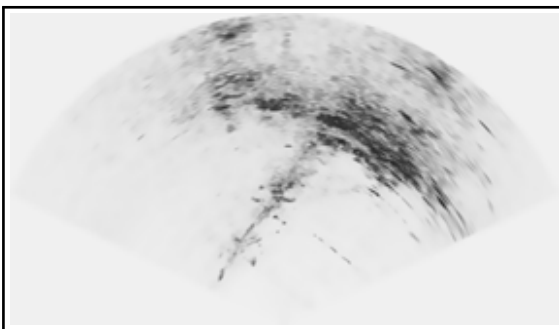
图 1-2: 前视声纳图像噪声示意图



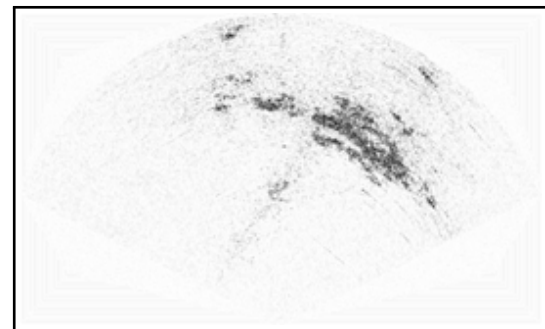
(a) 声纳原始图像



(b) 小波变换去噪后图像



(c) Lee 滤波器去噪后图像



(d) 各向异性去噪后图像

图 1-3: 传统去噪方法效果图

去噪图像在三维场景重建中的应用价值，验证了声纳图像质量提升对下游任务性能提升的直接促进作用。

本研究为前视声纳图像提供了一种高效的去噪方案，在应用上提升了声纳图像在三维重建与水下感知任务中的可用性与鲁棒性，对智能水下机器人、自主探索与海洋测绘等领域具有一定的工程应用价值与推广意义。

1.2 国内外研究现状

水下机器人是在水下作业的一种机器人，随着各国对海洋探索的竞争，水下机器人也在快速发展，加之海洋中油田、天然气资源勘探的发现，使得研发智能化水下机器人变得更加重要，更加迫切。自主水下航行器（AUV）实现了水下作业平台的小巧化、智能化，通过线缆与陆地进行控制和通信，提高了水下作业的灵活度，也更加适用于未知海域的探索，也可以搭配不同的组件，如机械臂，声纳等，完成不同的特定功能。

1.2.1 水下机器人领域研究现状

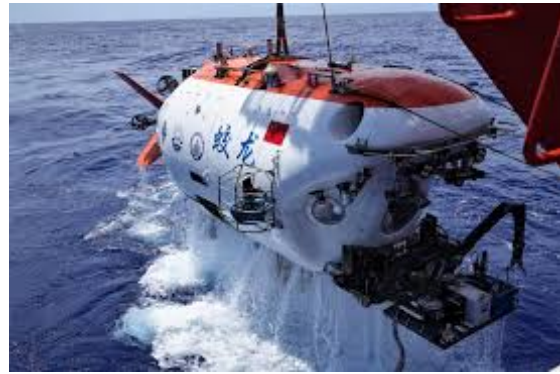
通常，水下机器人可分为自主水下机器人（Autonomous Underwater Vehicle, AUV）和遥控水下机器人（Remotely Operated Vehicle, ROV）。AUV 自带能源，可以自主航行，能够执行大范围探测任务，但作业时间、数据实时性、作业能力有限。ROV 依靠脐带电缆提供动力，水下作业时间长，能够实现数据实时传输，作业能力较强，但作业范围有限。近年来发展的混合式水下机器人——自主遥控水下机器人（Autonomous & Remotely Operated Vehicle, ARV）结合了 AUV 和 ROV 的优点，自带能源，通过光纤微缆实现数据实时传输，既可实现较大范围探测，又可实现水下定点精细观测及轻作业。相关水下机器人如图1-4所示。

国外方面，国外水下机器人研究已有近 70 年的历史。以美国为代表的西方发达国家，先后研发了 ROV、AUV、ARV 以及水下滑翔机等多种不同类型的水下机器人，主要用于深海资源勘查、海洋科学考察和军事应用等领域。20 世纪 50 年代，遥控水下机器人首次用于回收鱼雷。之后在寻求降低成本和减轻操作员风险的过程中，自主遥控水下机器人得到了发展。基于 BlueRov2 ROV 开发的 AUV 节省了大量时间，并且简化了部署，从而降低了运营成本。Nessie-V 是一款 6 自由度研究型 AUV，由赫瑞瓦特大学海洋系统实验室开发，用于重点检查和自主任务的研究。Girona500，这是一种具有可重新配置有效负载和推进系统的自主遥控水下机器人。

国内方面，2008 年，中国科学院沈阳自动化研究所完成了“北极”ARV 的研制，2009 年，中国船舶重工集团公司第 702 研究所完成了“海笋 I 型”ARV 原理



(a) BlueRov2



(b) 蛟龙号

图 1-4: BlueROV2 与蛟龙号水下机器人

样机的研制，并后续投入使用。2013 年，上海海事大学完成了“海事一号”ARV 的研制，可以应用于浅海渔业监测，水下搜救等作业任务。2020 年，由中国科学院沈阳自动化研究所牵头，联合中国 10 余家优势单位共同研制的“海斗一号”全海深自主潜水器完成了万米海底试验验证。

1.2.2 声纳图像去噪研究现状

早期的声纳图像去噪研究主要针对于散斑噪声（speckle noise）展开，经典的滤波算法如 Lee 滤波器、Kuan 滤波器和 Frost 滤波器等最初是为合成孔径声纳（Synthetic Aperture Sonar, SAS）图像设计的^[5]。这些方法基于乘性散斑噪声模型（multiplicative speckle model）来估计局部区域的统计特性，从而实现噪声抑制，但模型的不精确性经常导致角点等精细结构的损失。为改善这一问题，Lopes 等人提出了一种基于双阈值变异系数（dual-threshold coefficient of variation）的自适应方法，从而在去噪的同时更好地保留边缘信息^[6]。随后，一些通用的图像去噪方法，如小波变换（wavelet transform）^[7]、非局部均值法（non-local means）^[8]以及 BM3D 算法^[9]也被应用于声纳图像去噪中，在保留细节的同时提供了更强的噪声抑制效果。尽管这些方法取得了一定进展，但是传统方法在处理前视声纳图像中的结构性噪声（structural noise）时仍显不足，例如条纹状伪影与固定点伪影等问题依然难以有效解决。

随着深度学习的发展，卷积神经网络已经被广泛应用于声纳图像去噪任务中。Ji 等人提出了一种基于 Noise2Void 的自监督去噪方法，能够有效去除高斯噪声和散斑噪声^[10]；Si 等人提出 WTCRNet，将小波变换与对比正则化相结合，用于前视声纳图像去噪^[11]。此外，结合卷积神经网络（CNN）与自注意力变换网络（Transformer）的盲去噪模型 SCUNet 也被引入到声纳图像处理中^[12]。最近 Vishwakarma 提出一

种基于卷积稀疏表示 (convolutional sparse representation) 的去噪与修复方法, 可有效保留边缘与纹理细节^[13]。

生成对抗网络 (Generative Adversarial Networks, GANs) 在声纳去噪领域同样展现出巨大潜力。典型的非配对学习框架包括 UIDNet^[14]、UNIT^[15]、CUT^[16]、Still-GAN^[17] 与 DRGAN^[18], 它们均通过对抗训练提升了图像去噪的有效性。在声纳场景中, Zhao 等人提出了 NCD-GAN, 通过联合对比学习实现了非配对去噪^[19]; Lin 等人基于条件生成对抗网络从前视声纳图像中生成噪声掩膜以滤除噪声^[20]; Zhou 等人提出的 STGAN 将生成对抗网络与自注意力变换网络 (Transformer) 结合, 用于抑制散斑噪声^[21]。这些方法在复杂噪声环境下的表现优于传统算法, 但仍面临训练不稳定与模式坍塌 (mode collapse) 等问题。

近年来, 扩散模型 (Diffusion Models) 也被广泛应用于声纳图像去噪任务。Ho 等人提出了 DDPM (Denoising Diffusion Probabilistic Model), 为扩散模型奠定了理论基础^[22]。Smith 等人将扩散模型用于声纳超分辨率任务中, 有效缓解了条纹噪声问题^[23]; Wang 等人将 DDPM 应用于侧扫声纳 (Side-scan Sonar) 去噪, 在真实数据集上获得了显著性能提升^[24]。然而, 尽管扩散模型的去噪效果优异, 其训练和推理成本较高, 因而不适用于实时声纳图像去噪任务中。

1.3 本文的主要研究内容

近年来, 随着水下机器人技术的快速发展, 环境感知与重建已成为自主作业系统的关键环节。前视声纳 (Forward-Looking Sonar, FLS) 作为一种重要的水下成像传感器, 能够在低能见度、强浑浊等光学图像无法工作的环境下提供高分辨率图像信息, 因此在水下探测、定位和导航任务中得到了广泛应用。然而, 由于声波在传播过程中受到多路径反射、散射以及硬件噪声等影响, 声纳图像常常受到散斑噪声、旁瓣噪声和结构性噪声等多种干扰。这些噪声不仅严重降低图像对比度与分辨率, 还导致目标边界模糊, 显著影响后续的目标识别、三维重建与自主导航等任务。针对上述问题, 本文提出了一种基于生成对抗网络的多阶段声纳图像去噪框架 SonarGAN, 旨在实现多类型噪声的联合抑制与结构信息的有效保持。本文的主要研究内容包括以下几个部分:

1. 基于无配对数据的初步声纳去噪。针对水下声纳图像难以获取真实配对 (噪声——干净) 样本的问题, 本文首先设计了一个无监督的初步去噪框架。该阶段以循环一致性对抗网络 (Cycle-Consistent Adversarial Networks, CycleGAN)^[25] 为基础, 通过构建噪声域与干净域之间的循环一致性约束, 实现不同域之间的特征迁移与映射。本阶段主要针对随机噪声 (如散斑噪声、旁瓣噪声) 进行初步抑制, 在

不依赖配对样本的前提下获得具有较高结构一致性的去噪结果。本阶段中引入感知损失，以在深层特征空间上约束生成器保持目标的边缘与纹理语义。为了避免传统 CycleGAN 容易出现的伪影与纹理漂移问题，本文在生成器中嵌入了自注意力模块（Self-Attention Module），以增强模型的全局感知能力，从而在早期阶段实现噪声的有效压制与纹理信息的保持，为后续阶段提供高质量的特征先验。

2. 基于配对数据的结构性噪声抑制。为了进一步去除结构性噪声并提升图像细节质量，本文在第二阶段引入了基于条件生成对抗网络（Conditional Generative Adversarial Networks, cGAN）^[26] 的精细去噪模块。该阶段利用显式的配对训练集：通过将实测的纯噪声图像叠加到仿真得到的无噪图像上，得到带结构噪声的合成样本对，进行训练。通过显式学习声纳成像机制中结构性噪声的分布特征，增强网络对目标边界及结构性噪声的敏感性。在本阶段得到的生成器能有效识别并剔除稳定存在的结构性干扰，为第三阶段提供结构先验。

3. 联合约束去噪。在第三阶段，本文提出了一个多阶段融合与联合约束机制，对前两个阶段的输出结果进行统一优化与精修。该阶段通过设计显著性掩模约束，在像素层面实现噪声抑制与结构保持的动态平衡。掩模由网络自适应生成，用于区分目标区域与背景噪声区域，使模型在噪声区域增强抑制能力，而在目标区域重点保留纹理细节。同时，本文在损失函数中引入对抗约束与特征一致性约束，实现了从全局结构到局部细节的多层次优化。经本阶段融合后，SonarGAN 能够在多类型噪声共存的复杂水下环境中稳定实现高质量去噪与结构保真，显著提升了模型的鲁棒性与泛化能力。

4. 实验设计与性能验证。为验证所提出方法的有效性与通用性，本文设计并开展了两类实验：去噪性能评估实验与三维重建验证实验。

在去噪性能评估部分，本文分别在仿真数据集与真实声纳数据集上对比了 SonarGAN 与多种主流去噪方法的性能表现，包括传统滤波方法（Lee、BM3D 等）、深度学习方法（DnCNN、S CUNet 等）以及多种生成对抗网络框架（CycleGAN、UNIT、CUT、StillGAN 等）。仿真实验采用虚拟仿真环境生成的噪声自由图像，并结合多种声纳的实测噪声进行合成含噪声图像，以获得多样化的测试样本。真实实验使用同类声纳在实验水池、河流环境以及港口码头中采集的数据，综合评估模型在不同噪声分布下的表现。本文除了采用 PSNR、SSIM、LPIPS 等常规指标，还提出了一种新的基于 Kullback-Leibler（KL）散度^[27] 的统计分布差异指标对去噪性能进行量化评估，并设计了基于平均残差图像的空间一致性度量方法，以进一步衡量模型的噪声去除和目标物体保留能力。实验结果表明，SonarGAN 在多种噪声类型与不同声纳设备下均取得优异的去噪效果，能够有效去除结构性噪声并

保持目标细节。同时，本文通过跨域泛化实验验证了模型的稳健性，即直接将河流数据上训练的模型应用于实验水池数据，仍能获得理想的去噪结果。

在三维重建验证部分，本文进一步评估了去噪质量对水下重建任务的促进作用。通过将去噪前后的声纳图像输入至基于声波传播建模的三维重建算法中，比较重建点云的稠密度、表面平滑性与几何一致性等指标，结果显示经 SonarGAN 去噪后的声纳图像在纹理一致性与深度恢复方面均显著优于原始图像和多种对比算法的去噪后声纳图像，证明了所提方法在水下任务中的实用价值。

第2章 基于无配对数据的初步声纳去噪

2.1 引言

在水下环境中，由于声波传播过程复杂多变，前视声纳图像往往受到多种噪声的共同影响，包括散斑噪声、旁瓣噪声以及由多路径反射产生的伪影等。这些噪声显著降低了图像的清晰度与结构一致性，给后续的目标识别、定位与三维重建任务带来了极大困难。近年来，随着深度学习技术的发展，基于卷积神经网络的图像去噪方法在自然图像领域取得了显著进展。然而，与光学图像不同，声纳图像中“干净图像”的获取几乎不可能。由于声波的非线性传播及声反射的不确定性，同一场景下的“带噪图像”和“无噪图像”无法通过实测获得精确配对样本，这使得传统的监督式去噪方法难以直接应用于声纳场景。

针对这一问题，近年来出现了多种无监督学习框架，用于在无配对样本的条件下实现跨域映射。其中，循环一致性对抗网络（Cycle-Consistent Adversarial Network, CycleGAN）^[25] 通过在两个域之间引入正向与反向映射的循环约束，在保持内容一致性的同时实现了域间风格迁移，为无配对声纳去噪提供了新的思路。受此启发，本文在本阶段中构建了一个基于无配对数据的初步去噪框架，通过在噪声域与干净域之间建立循环一致的映射关系，实现对随机噪声的有效抑制。该阶段的目标是利用未配对的真实噪声图像与模拟干净图像进行对抗式训练，从而获得具有较好结构保持能力的去噪结果，为后续的结构性噪声抑制阶段提供可靠的先验。

2.2 声纳图像成像原理

水下声纳成像原理主要基于声波在水中传播的特性。声纳系统通过发射声波信号，声波在水中传播并与周围物体相互作用，产生反射。当声波遇到物体时，部分声波会被反射回来，声纳接收器接收这些回波信号。声纳成像的关键步骤包括发射、传播、接收和处理。首先，声纳发射器发出短脉冲声波，这些声波在水中传播时会受到环境因素如温度、盐度和水流的影响。回波信号的接收器记录这些反射声波，并根据其到达时间和强度生成数据。随后，通过信号处理技术，将接收到的回波信号转化为图像。常用的方法包括时域分析、频域分析以及图像重建算法。这些处理步骤可以提高图像的清晰度和分辨率，最终生成水下环境的声纳图像。

考虑成像声纳视野中的点 $P(\theta, r, \Phi)$ ，在局部球坐标中参数化，如图图2-1所示，其中 θ, r 和 Φ 分别表示 P 点的方位角、范围和仰角。 P 点到笛卡尔坐标系的坐标 $C(x, y, z)$ 的坐标变换与反变换如下：

$$\mathbf{C} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} r \cos \phi \cos \theta \\ r \cos \phi \sin \theta \\ r \sin \phi \end{bmatrix} \quad (2-1)$$

$$\mathbf{P} = \begin{bmatrix} \theta \\ r \\ \Phi \end{bmatrix} = \begin{bmatrix} \arctan 2(y, x) \\ \sqrt{x^2 + y^2 + z^2} \\ \arctan 2\left(z, \sqrt{x^2 + y^2}\right) \end{bmatrix} \quad (2-2)$$

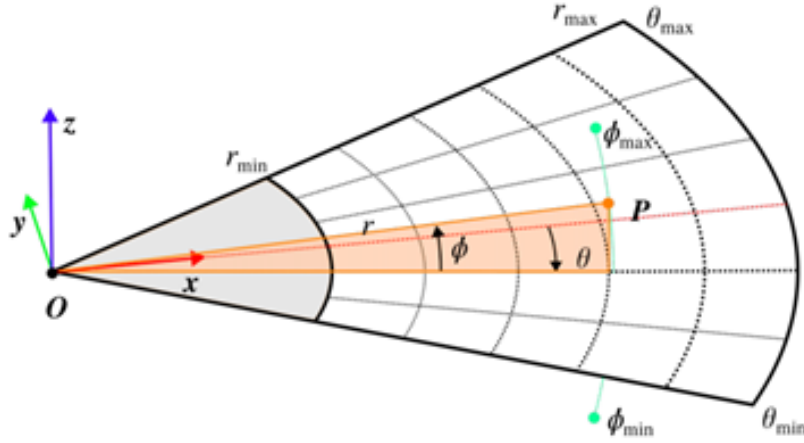


图 2-1: 单个声纳图像的几何结构，点 P 由范围 r 、仰角 Φ 和方位角 θ 表示。

$r_{\min}, r_{\max}, \Phi_{\min}, \Phi_{\max}, \theta_{\min}, \theta_{\max}$ 分别是成像声纳的最大和最小范围，最大仰角和最小仰角，最大方位角和最小方位角。

成像声纳通过发出一系列类似锥形的声波信号来生成部分球面测量值，从收发器测量观察到的反射信号的飞行时间，提供了反射表面的范围 r 和方位角 θ 。但是这样的测量无法准确判断出仰角 Φ 。因此，所有来自同一个方位角和距离的检测回波都会投影到声纳图像上的同一点 $I(\theta, r)$ ，如图2-2所示。对于对应于 I 点中的某个方位角和距离的像素，像素的强度对应于该方位角和距离下所有回波反射信号的强度。

为了直观展示前视声纳的成像特征与其在水下环境感知中的应用，如图2-3所示，展示了同一时刻下水下机器人采集到的视觉图像与对应的前视声纳图像。视觉图像反映了光学成像下的场景特征，可见由于水下光线传播的限制，视觉图像中的细节信息较为模糊。而前视声纳图像则通过声波反射捕捉了场景的结构信息，

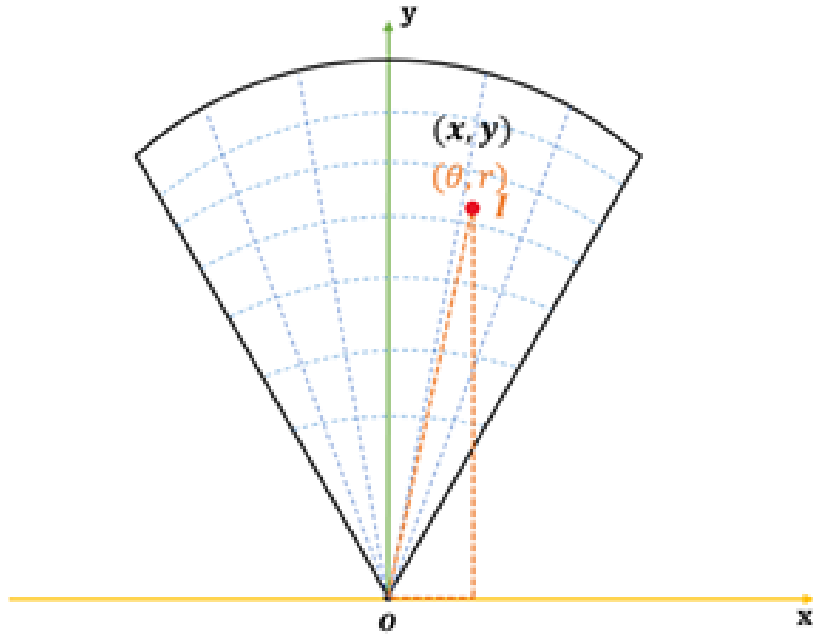


图 2-2: 真实声纳图像是整个垂直声纳数据在 2D 平面的投影。

能够清晰地显示出障碍物和地形特征。两者在信息维度和感知机制上存在显著差异。

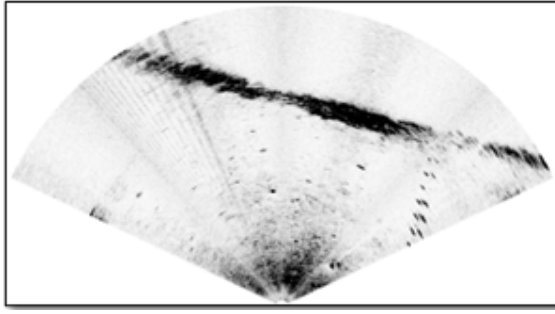


(a) 视觉图像

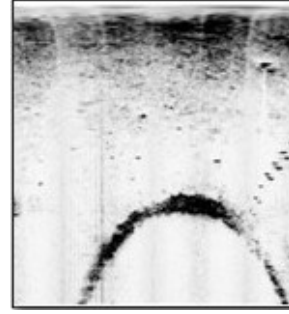
(b) 前视声纳图像

图 2-3: 水下机器人视觉图像与前视声纳图像对应示意图

需要指出的是，前视声纳图像通常可以以两种方式进行表示：一种是基于声纳测量参数，以方位角和距离为坐标轴的极坐标图像，直接对应于声纳的观测几何结构；另一种是通过坐标变换，如公式 (2-1) 和 (2-2) 所示，将极坐标数据映射到笛卡尔平面上的笛卡尔坐标图像。前者能保留声纳探测的物理意义，后者则更便于与视觉图像进行对比以及后续的图像处理任务。如图2-4所示，左侧为前视声纳的笛卡尔图像，右侧为对应的极坐标图像。可以看出，笛卡尔图像通过坐标变换展现了空间中的物体分布，而极坐标图像则更直观地展示探测的物体结构。



(a) 前视声纳笛卡尔图像



(b) 前视声纳极坐标图像

图 2-4: 笛卡尔图像与极坐标图像对应示意图

2.3 循环一致性对抗网络

循环一致性对抗网络（Cycle-Consistent Adversarial Network, CycleGAN）旨在解决图像到图像翻译的问题。该方法的核心在于能够在没有配对训练数据的情况下学习两个图像域之间的映射关系。对于前视声纳这种缺少“干净”图像作为参考的情况，适合使用 CycleGAN 进行非配对图像去噪。传统图像翻译任务（如灰度图到彩色图、边缘图到照片）通常依赖于配对的输入和输出图像样本，但 CycleGAN 通过引入循环一致性损失（Cycle Consistent loss），使其应用于许多实际场景，如艺术风格迁移，物体变形，季节转换和照片增强等。

图像到图像翻译的目标是将一个源域 X 中的图像转换为目标域 Y 中的图像。CycleGAN 针对无配对设置，仅使用源域样本和目标域样本，通过对抗训练确保生成的图像分布与目标域匹配，并使用循环一致性避免模式崩溃。CycleGAN 的模型包含两个生成器和两个判别器。生成器 $G : X \rightarrow Y$ 将源域图像转换为目标域图像，生成器 $F : Y \rightarrow X$ 则执行相反的映射。判别器 D_Y 旨在区分真实的目标域图像和生成的图像，而判别器 D_X 则区分真实的源域图像和生成的图像。这种双向对抗与映射结构如图2-6 (a)所示，通过联合训练两个生成器和判别器，实现两个域之间的循环一致性翻译。

为使生成器输出的图像在统计分布上与目标域相似，引入对抗损失函数。我们以 $G : X \rightarrow Y$ 与判别器 D_Y 为例，对抗损失 $\mathcal{L}_{GAN}(G, D_Y, X, Y)$ 定义为：

$$\begin{aligned} \mathcal{L}_{GAN}(G, D_Y, X, Y) = & \mathbb{E}_{y \sim p_Y} [\log D_Y(y)] \\ & + \mathbb{E}_{x \sim p_X} [\log(1 - D_Y(G(x)))] \end{aligned} \quad (2-3)$$

其中， G 旨在生成足够“真实”的图像以欺骗判别器 D_Y ，而 D_Y 则尽力区分生成样本与真实样本。同理，对抗损失 $\mathcal{L}_{GAN}(F, D_X, Y, X)$ 对于 $F : Y \rightarrow X$ 及其对应的判别

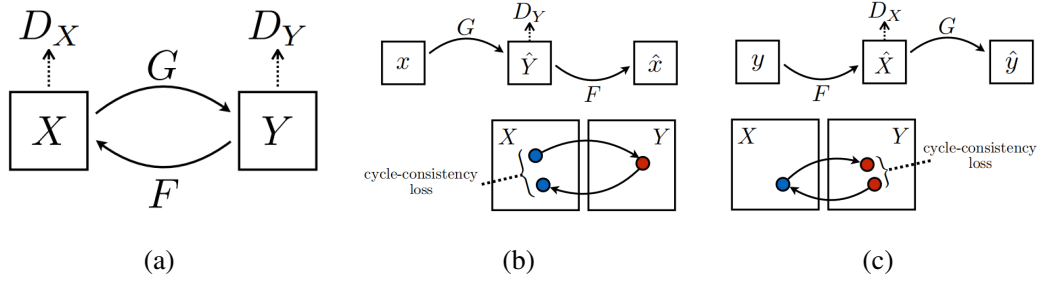


图 2-5: CycleGAN 示意图

器 D_X 也有类似的定义。总对抗损失 \mathcal{L}_{ad} 是 $\mathcal{L}_{GAN}(G, D_Y, X, Y)$ 和 $\mathcal{L}_{GAN}(F, D_X, Y, X)$ 的加权和。

由于无配对训练样本无法保证 x 与 y 的一一对应，仅使用对抗损失会导致映射关系不稳定或模式坍塌。为此，CycleGAN 引入循环一致性约束，要求经过双向映射的图像应该尽可能还原输入，即： $F(G(x)) \approx x$ 和 $G(F(y)) \approx y$ 。相应的循环一致性损失函数定义为：

$$\begin{aligned} \mathcal{L}_{cyc}(G, F) = & \mathbb{E}_{x \sim p_X} [\|F(G(x)) - x\|_1] \\ & + \mathbb{E}_{y \sim p_Y} [\|G(F(y)) - y\|_1] \end{aligned} \quad (2-4)$$

如图2-6 (b)和图2-6 (c)所示，该约束确保跨域映射保留图像的结构信息，从而提升转换的一致性与可逆性。

在某些任务（如绘画到实景照片的风格迁移）中，CycleGAN 可能在输入已经属于目标域时仍对图像进行不必要的颜色或亮度改变。为此，引入了身份映射损失，用于鼓励生成器在输入图像已经属于目标域时保持输出不变。该损失函数定义如下：

$$\begin{aligned} \mathcal{L}_{id}(G, F) = & \mathbb{E}_{y \sim p_Y} [\|G(y) - y\|_1] \\ & + \mathbb{E}_{x \sim p_X} [\|F(x) - x\|_1] \end{aligned} \quad (2-5)$$

身份映射损失在保持图像色彩一致性与亮度分布方面具有重要作用，对于保护目标物体，防止去噪过多导致信息丢失尤为关键。

CycleGAN 的总损失函数由上述三部分组成，损失函数如下：

$$\mathcal{L}(G, F, D_X, D_Y) = \mathcal{L}_{ad} + \lambda_{cyc} \mathcal{L}_{cyc}(G, F) + \lambda_{id} \mathcal{L}_{id}(G, F). \quad (2-6)$$

其中， λ_{cyc} 与 λ_{id} 分别为循环一致性与身份损失的权重系数。最终优化目标为：

$$G^*, F^* = \arg \min_{G, F} \max_{D_X, D_Y} \mathcal{L}(G, F, D_X, D_Y). \quad (2-7)$$

2.4 循环一致性对抗网络

循环一致性对抗网络 (Cycle-Consistent Adversarial Network, CycleGAN) 由 Zhu 等人于 2017 年提出^[25], 旨在解决无配对条件下的图像到图像翻译问题 (Unpaired Image-to-Image Translation)。与传统的监督式方法 (如 Pix2Pix^[26]) 不同, CycleGAN 能在没有配对训练数据的情况下学习两个图像域之间的映射关系。这一特性使其特别适用于缺乏“干净”参考图像的场景, 例如前视声纳图像去噪任务中, 真实环境下几乎不可能同时获得同一场景的有噪声与无噪声配对图像。

传统的图像翻译任务 (如灰度图到彩色图、素描到照片、昼夜转换等) 往往依赖成对的训练样本 (x_i, y_i) 。而在声纳成像等非理想观测条件下, 配对样本的获取代价极高甚至无法实现。CycleGAN 通过引入循环一致性约束 (Cycle Consistency Loss) 与身份映射约束 (Identity Loss), 在保证结构一致性的同时实现了跨域映射学习, 为非配对图像去噪提供了一种高效的可行方案。

CycleGAN 的整体结构包含两个生成器 (Generator) 和两个判别器 (Discriminator), 分别建立在两个域 X 和 Y 之间的对偶映射关系: 生成器 $G: X \rightarrow Y$ 将源域图像转换为目标域风格, 生成器 $F: Y \rightarrow X$ 执行逆向映射。判别器 D_Y 用于判别输入图像是否来自目标域 Y , 而 D_X 判别输入图像是否属于源域 X 。这种双向的对抗结构如图 2-6 (a) 所示, 通过相互约束与循环反馈, 使得生成的图像在语义与结构上保持一致。

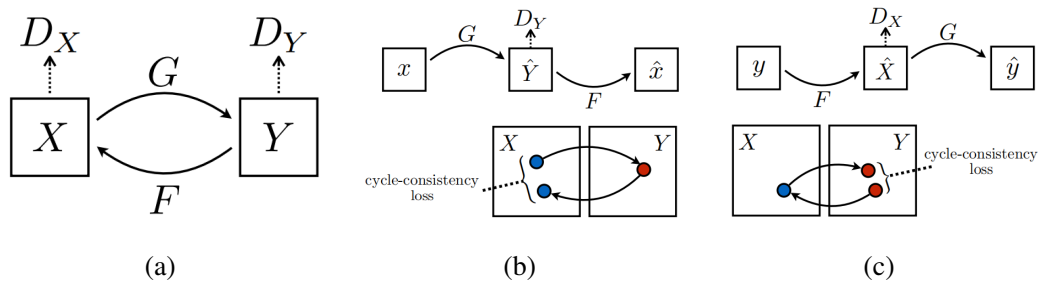


图 2-6: CycleGAN 示意图

在网络设计上, CycleGAN 的生成器采用了基于 U-Net^[28] 或 ResNet^[29] 的结构, 以便在跨域变换中保留低层特征与高层语义信息。其中, ResNet 生成器通常由 9 个残差块 (Residual Blocks) 组成, 以捕获复杂的跨域映射特征。判别器则采用 PatchGAN^[26] 结构, 通过判断每个局部图像块的真假性来提高模型的细节保真度, 从而生成更自然、更逼真的图像纹理。

为确保生成图像在分布上接近目标域, CycleGAN 引入了对抗损失函数。以 $G: X \rightarrow Y$ 与判别器 D_Y 为例, 其对抗损失定义如下:

$$\begin{aligned}\mathcal{L}_{\text{GAN}}(G, D_Y, X, Y) = & \mathbb{E}_{y \sim p_Y} [\log D_Y(y)] \\ & + \mathbb{E}_{x \sim p_X} [\log(1 - D_Y(G(x)))].\end{aligned}\quad (2-8)$$

生成器 G 旨在生成足够真实的图像以欺骗判别器 D_Y ，而判别器 D_Y 则学习区分真实样本与伪造样本。同理，对于 $F: Y \rightarrow X$ 及其对应的判别器 D_X ，定义类似的损失项 $\mathcal{L}_{\text{GAN}}(F, D_X, Y, X)$ 。两者之和构成总对抗损失 \mathcal{L}_{ad} 。

由于无配对训练样本无法保证 x 与 y 的一一对应，仅靠对抗训练容易导致模式坍塌（Mode Collapse）或结构失真。为约束生成器学习稳定的可逆映射关系，CycleGAN 引入循环一致性损失，要求经过双向映射的图像尽可能还原输入：

$$F(G(x)) \approx x, \quad G(F(y)) \approx y.$$

循环一致性损失定义为：

$$\begin{aligned}\mathcal{L}_{\text{cyc}}(G, F) = & \mathbb{E}_{x \sim p_X} [\|F(G(x)) - x\|_1] \\ & + \mathbb{E}_{y \sim p_Y} [\|G(F(y)) - y\|_1].\end{aligned}\quad (2-9)$$

该项损失确保模型在跨域转换过程中保留图像的几何与结构特征，提升了生成结果的稳定性与可逆性。在声纳图像去噪任务中，该约束有助于保持场景目标的空间一致性，避免在去噪过程中出现结构模糊或物体形态扭曲。

在部分任务（如绘画到实景照片或光照增强）中，若输入图像本身已经接近目标域，模型可能在不必要的情况下改变色彩或亮度。为缓解这一问题，引入身份映射损失（Identity Loss）以约束生成器的输出应尽量保持与输入一致：

$$\begin{aligned}\mathcal{L}_{\text{id}}(G, F) = & \mathbb{E}_{y \sim p_Y} [\|G(y) - y\|_1] \\ & + \mathbb{E}_{x \sim p_X} [\|F(x) - x\|_1].\end{aligned}\quad (2-10)$$

该项损失在保持图像色彩与亮度一致性方面具有显著作用。在声纳去噪任务中，身份约束可以防止生成器过度平滑目标结构，防止噪声抑制导致的细节丢失。

CycleGAN 的整体优化目标由三部分损失组成：

$$\mathcal{L}(G, F, D_X, D_Y) = \mathcal{L}_{\text{ad}} + \lambda_{\text{cyc}} \mathcal{L}_{\text{cyc}}(G, F) + \lambda_{\text{id}} \mathcal{L}_{\text{id}}(G, F), \quad (2-11)$$

其中， λ_{cyc} 与 λ_{id} 分别为循环一致性与身份损失的权重系数。原论文中通常取 $\lambda_{\text{cyc}} = 10$ ， $\lambda_{\text{id}} = 0.5 \lambda_{\text{cyc}}$ 。最终优化目标为：

$$G^*, F^* = \arg \min_{G, F} \max_{D_X, D_Y} \mathcal{L}(G, F, D_X, D_Y). \quad (2-12)$$

CycleGAN 的核心优势在于其无监督学习特性与结构保持能力：

- 不依赖配对样本即可实现跨域映射；

- 保留源图像的几何结构与语义一致性；
- 可扩展性强，适用于风格迁移、图像增强、语义分割前处理等多种任务。

在前视声纳图像去噪任务中，CycleGAN 能够学习噪声域与干净域之间的隐式映射关系，通过对抗学习抑制随机噪声，同时依靠循环一致性与身份约束保持目标结构完整。这为后续的结构性噪声抑制提供了可靠的先验。

2.5 声纳图像无配对去噪网络的构建与优化

本研究中，声纳图像的噪声域和干净域分别由含噪声的真实声纳图像和通过模拟仿真生成的声纳图像构成。由于实际采集的声纳图像往往受到环境和设备因素影响而包含复杂噪声，而仿真生成的图像可以在理想条件下生成，具有较高的结构与纹理清晰度，因此两者在视觉特性上存在明显域差异。为了在无配对条件下实现这两种域间的特征迁移，本阶段采用 CycleGAN 为核心框架，构建噪声域到干净域，干净域到噪声域的双向映射，通过循环一致性约束实现结构信息的保持与伪影的抑制。

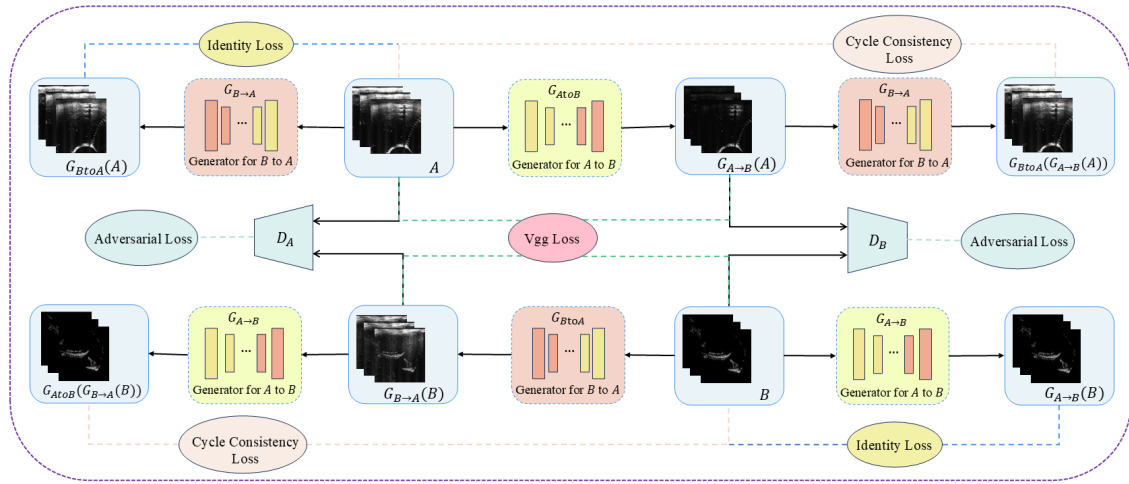


图 2-7

第 3 章 排版公式

3.1 引言

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

3.2 对物理量符号进行注释的情况

为使得对公式中物理量符号注释的转行与破折号“——”后第一个字对齐，此处最好采用表格环境。此表格无任何线条，左对齐，且在破折号处对齐，一共有“式中”二字、物理量符号和注释三列，表格的总宽度可选为文本宽度，因此应该采用`tabularx`环境。由`tabularx`环境生成的对公式中物理量符号进行注释的公式如式(3-1)所示。

$$\ddot{\rho} - \frac{\mu}{R_t^3} \left(3R_t \frac{R_t \rho}{R_t^2} - \rho \right) = a \quad (3-1)$$

式中 ρ ——追踪飞行器与目标飞行器之间的相对位置矢量；
 $\ddot{\rho}$ ——追踪飞行器与目标飞行器之间的相对加速度；
 a ——推力所产生的加速度；
 R_t ——目标飞行器在惯性坐标系中的位置矢量；
 ω_t ——目标飞行器的轨道角速度；

由此方法生成的注释内容应紧邻待注释公式并置于其下方，因此不能将代码放入`table`浮动环境中。但此方法不能实现自动转页接排，可能会在当前页剩余空间不够时，全部移动到下一页而导致当前页出现很大空白。因此在需要转页处理

时，还请您手动将需要转页的代码放入一个新的`tabularx`环境中，将原来的一个`tabularx`环境拆分为两个`tabularx`环境。

3.3 子公式

子公式编号示例：如果需要对公式的子公式进行编号，则使用`subeqnarray`环境：

$$x = a \times b \tag{3-2a}$$

$$= z + t \tag{3-2b}$$

$$= z + t \tag{3-2c}$$

式 (3-2) 中，`label`为整个公式的标签，`slabel`为子公式的标签。

3.4 本章小结

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

第 4 章 排版表格

4.1 引言

表应有自明性。表格不加左、右边线。表的编排建议采用国际通行的三线表。表中文字用宋体 5 号字。每个表格均应有表题（由表序和表名组成）。表序一般按章编排，如第 1 章第一个插表的序号为“表 1-1”等。表序与表名之间空一格，表名中不允许使用标点符号，表名后不加标点。表题置于表上，硕士学位论文只用中文，博士学位论文用中、英文两种文字居中排写，中文在上，要求中文用宋体 5 号字，英文用新罗马字体 5 号字。表头设计应简单明了，尽量不用斜线。表头中可采用化学符号或物理量符号。

4.2 普通表格的绘制方法

表格应具有三线表格式，因此需要调用 `booktabs` 宏包，其标准格式如表 4-1 所示。全表如用同一单位，则将单位符号移至表头右上角，加圆括号。表中数据应

表 4-1: 符合研究生院绘图规范的表格

Table4-1: Table in agreement of the standard from graduate school

$D(\text{in})$	$P_u(\text{lbs})$	$u_u(\text{in})$	β	$G_f(\text{psi.in})$
5	269.8	0.000674	1.79	0.04089
10	421.0	0.001035	3.59	0.04089
20	640.2	0.001565	7.18	0.04089

准确无误，书写清楚。数字空缺的格内加横线“—”（占 2 个数字宽度）。表内文字或数字上、下或左、右相同时，采用通栏处理方式，不允许用“/”、“同上”之类的写法。表内文字说明，起行空一格、转行顶格、句末不加标点。如某个表需要转页接排，在随后的各页上应重复表的编号。编号后加“（续表）”，表题可省略。续表应重复表头。

4.3 长表格的绘制方法

长表格是当表格在当前页排不下而需要转页接排的情况下所采用的一种表格环境。若长表格仍按照普通表格的绘制方法来获得，其所使用的 `table` 浮动环境无法实现表格的换页接排功能，表格下方过长部分会排在表格第 1 页的页脚以下。为

了能够实现长表格的转页接排功能，需要调用 `longtable` 宏包，由于长表格是跨页的文本内容，因此只需要单独的 `longtable` 环境，所绘制的长表格的格式如表 4-2 所示。

注意，长表格双语标题的格式。

表 4-2: 中国省级行政单位一览
Table4-2: Overview of the provincial administrative unit of China

名称	简称	省会或首府
北京市	京	北京
天津市	津	天津
河北省	冀	石家庄市
山西省	晋	太原市
内蒙古自治区	蒙	呼和浩特市
辽宁省	辽	沈阳市
吉林省	吉	长春市
黑龙江省	黑	哈尔滨市
上海市	沪/申	上海
江苏省	苏	南京市
浙江省	浙	杭州市
安徽省	皖	合肥市
福建省	闽	福州市
江西省	赣	南昌市
山东省	鲁	济南市
河南省	豫	郑州市
湖北省	鄂	武汉市
湖南省	湘	长沙市
广东省	粤	广州市
广西壮族自治区	桂	南宁市
海南省	琼	海口市
重庆市	渝	重庆
四川省	川/蜀	成都市
贵州省	黔/贵	贵阳市
云南省	云/滇	昆明市
西藏自治区	藏	拉萨市
陕西省	陕/秦	西安市
甘肃省	甘/陇	兰州市

表 4-2（续表）

名称	简称	省会或首府
青海省	青	西宁市
宁夏回族自治区	宁	银川市
新疆维吾尔自治区	新	乌鲁木齐市
香港特别行政区	港	香港
澳门特别行政区	澳	澳门
台湾省	台	台北市

此长表格 4-2 第 2 页的标题“编号（续表）”和表头是通过代码自动添加上去的，无需人工添加，若表格在页面中的竖直位置发生了变化，长表格在第 2 页及之后各页的标题和表头位置能够始终处于各页的最顶部，也无需人工调整， \LaTeX 系统的这一优点是 word 等软件所无法比拟的。

4.4 列宽可调表格的绘制方法

论文中能用列宽可调表格的情况共有两种，一种是当插入的表格某一单元格内容过长以至于一行放不下的情况，另一种是当对公式中首次出现的物理量符号进行注释的情况，这两种情况都需要调用 `tabularx` 宏包。下面将分别对这两种情况下可调表格的绘制方法进行阐述。

4.4.1 表格内某单元格内容过长的情况

首先给出这种情况下的一个例子如表 4-3 所示。`tabularx` 环境共有两个必选参

表 4-3: 最小的三个正整数的英文表示法

Table4-3: The English construction of the smallest three positive integral numbers

Value	Name	Alternate names, and names for sets of the given size
1	One	ace, single, singleton, unary, unit, unity
2	Two	binary, brace, couple, couplet, distich, deuce, double, doubleton, duad, duality, duet, duo, dyad, pair, snake eyes, span, twain, twosome, yoke
3	Three	deuce-ace, leash, set, tercet, ternary, ternion, terzetto, threesome, tierce, trey, triad, trine, trinity, trio, triplet, troika, hat-trick

数：第 1 个参数用来确定表格的总宽度，第 2 个参数用来确定每列格式，其中标为

X 的项表示该列的宽度可调，其宽度值由表格总宽度确定。标为 X 的列一般选为单元格内容过长而无法置于一行的列，这样使得该列内容能够根据表格总宽度自动分行。若列格式中存在不止一个 X 项，则这些标为 X 的列的列宽相同，因此，一般不将内容较短的列设为 X。标为 X 的列均为左对齐，因此其余列一般选为 l（左对齐），这样可使得表格美观，但也可以选为 c 或 r。

4.4.2 排版横版表格的举例

横版表格如表4-4所示。

4.5 本章小结

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

表 4-4: 不在规范中规定的横版表格
Table 4-4: A table style which is not stated in the regulation

$D(\text{in})$	$P_u(\text{lbs})$	$u_u(\text{in})$	β	$G_f(\text{psi.in})$
5	269.8	0.000674	1.79	0.04089
10	421.0	0.001035	3.59	0.04089
20	640.2	0.001565	7.18	0.04089

第 5 章 引用参考文献

5.1 引言

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

5.2 参考文献引用方法

引文标注遵照 GB/T7714-2005，采用顺序编码制。正文中引用文献的标示应置于所引内容最后一个字的右上角，所引文献编号用阿拉伯数字置于方括号“[]”中，用小 4 号字体的上角标。要求：

- (1) 引用单篇文献时，如“二次铣削^[1]”。
- (2) 同一处引用多篇文献时，各篇文献的序号在方括号内全部列出，各序号间用“，”，如遇连续序号，可标注起序号。如，…形成了多种数学模型^{[1][2]…}注意此处添加\inlinecite 中文空格 [? ?]，可以在 cfg 文件中修改空格类型。
- (3) 多次引用同一文献时，在文献序号的“[]”后标注引文页码。如，…间质细胞 CAMP 含量测定^{[1]100-197…}。…含量测定方法规定^{[1]92…}。
- (4) 当提及的参考文献为文中直接说明时，则用小 4 号字与正文排齐，如“由文献 [?] 可知”
- (5) 多^[1]引^[2]用^[3]—^[4]些^[5]参^[6]考^[7]文献以生成附录参考文献。

5.3 本章小结

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero,

nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

第 6 章 补充说明

6.1 引言

这是 HITSZTHESIS 的示例文档，基本上覆盖了模板中所有格式的设置。建议大家在使用模板之前，除了阅读《HITSZTHESIS 哈尔滨工业大学学位论文模板》^①，本示例文档也最好能看一看。此示例文档尽量使用到所有的排版格式，然而对于一些不在我工规范中规定的文档，理论上是由用户自由发挥，这里不给出样例。需要另行载入的宏包和自定义命令在文件 ‘hitszthesis.sty’ 中有示例，这里不列举。

6.2 关于数字

按《关于出版物上数字用法的试行规定》（1987 年 1 月 1 日国家语言文字工作委员会等 7 个单位公布），除习惯用中文数字表示的以外，一般数字均用阿拉伯数字。（1）公历的世纪、年代、年、月、日和时刻一律用阿拉伯数字，如 20 世纪，80 年代，4 时 3 刻等。年号要用四位数，如 1989 年，不能用 89 年。（2）记数与计算（含正负整数、分数、小数、百分比、约数等）一律用阿拉伯数字，如 3/4，4.5%，10 个月，500 多种等。（3）一个数值的书写形式要照顾到上下文。不是出现在一组表示科学计量和具有统计意义数字中的一位数可以用汉字，如一个人，六条意见。星期几一律用汉字，如星期六。邻近两个数字并列连用，表示概数，应该用汉字数字，数字间不用顿号隔开，如三五天，七八十种，四十五六岁，一千七八百元等。（4）数字作为词素构成定型的词、词组、惯用语、缩略语等应当使用汉字。如二倍体，三叶虫，第三世界，“七五”规划，相差十万八千里等。（5）5 位以上的数字，尾数零多的，可改写为以万、亿为单位的数。一般情况下不得以十、百、千、十万、百万、千万、十亿、百亿、千亿作为单位。如 345 000 000 公里可改写为 3.45 亿公里或 34 500 万公里，但不能写为 3 亿 4 500 万公里或 3 亿 4 千 5 百万公里。（6）数字的书写不必每格一个数码，一般每两数码占一格，数字间分节不用分位号“，”，凡 4 位或 4 位以上的数都从个位起每 3 位数空半数码（1/4 汉字）。“3 000 000”，不要写成“3,000,000”，小数点后的数从小数点起向右按每三位一组分节。一个用阿拉伯数字书写的多位数不能从数字中间转行。（7）数量的增加或减少要注意下列用词的概念：1）增加为（或增加到）过去的二倍，即过去为一，现在为二；2）增加（或增加了）二倍，即过去为一，现在为三；3）超额 80%，即定额为 100，现

^① 即 hitszthesis.pdf 文件

在为 180；4) 降低到 80%，即过去为 100，现在为 80；5) 降低（或降低了）80%，即原来为 100，现在为 20；6) 为原数的 $\frac{1}{4}$ ，即原数为 4，现在为 1，或原数为 1，现在为 0.25。应特别注意在表达数字减小时，不宜用倍数，而应采用分数。如减少为原来的 $\frac{1}{2}$ ， $\frac{1}{3}$ 等。

6.3 索引示例

为便于检索文中内容，可编制索引置于论文之后（根据需要决定是否设置）。索引以论文中的专业词语为检索线索，指出其相关内容的所在页码。索引用中、英两种文字书写，中文在前。中文按各词汉语拼音第一个字母排序，英文按该词第一个英文字母排序。

6.4 术语排版举例

术语的定义和使用可以结合索引，灵活使用。例如，树结构折簇过程（Tree-structured Stick-breaking process）是一种应用于狄利克雷过程抽样的算法。下次出现将是另一种格式：树结构折簇过程。还可以切换单复数例如：体细胞拷贝数变异（Somatic copy number alternation, SCNA），下次出现为：SCNAs。此处体现了 L^AT_EX 格式内容分离的优势。

6.5 定理和定义等

定理 6.1^[?]] 宇宙大爆炸是一种爆炸。

定义 6.1（霍金） 宇宙大爆炸是一种爆炸。

假设 6.1 宇宙大爆炸是一种爆炸。

引理 6.1 宇宙大爆炸是一种爆炸。

推论 6.1 宇宙大爆炸是一种爆炸。

练习 6.1 宇宙大爆炸是一种爆炸。

问题 6.1 (Albert Einstein) 宇宙大爆炸是一种爆炸。

注释 6.1 宇宙大爆炸是一种爆炸。

公理 6.1（爱因斯坦） 宇宙大爆炸是一种爆炸。

猜想 6.1 宇宙大爆炸是一种爆炸。

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

6.6 其他杂项

6.6.1 右翻页

对于双面打印的论文，强制使每章的标题页出现右手边为右翻页。规范中没有明确规定是否是右翻页打印。模板给出了右翻页选项。为了应对用户的个人喜好，在希望设置成右翻页的位置之前添加`\cleardoublepage`命令即可。

6.6.2 算法

算法不在规范中要求，在 `hitszthesis.sty` 中有相关定义，一个例子如算法6-1所示。

算法 6-1 混合重排算法

Data: A : 待重排的元素集合

α : 对多样性, 相关性作折中的权重因子

Result: A_k : a subset of A of size k

6.6.3 脚注

不在再规范^①中要求，模板默认使用清华大学的格式。

6.6.4 源码

也不在再规范中要求。如果有需要最好使用 `minted` 包，但在编译的时候需要

① 规范是指《哈尔滨工业大学（深圳）研究生学位论文撰写规范》和《哈尔滨工业大学（深圳）本科生毕业论文撰写规范》

添加“-shell-escape”选项且安装 pygmentize 软件，这些不在模板中默认载入，如果需要自行载入。

6.6.5 思源宋体

如果要使用思源字体，需要思源字体的定义文件，此文件请到模板的开发版网址 github: <https://github.com/YangLaTeX/hitszthesis> 处下载。

6.6.6 术语词汇管理

推荐使用 glossaries 包管理术语、缩略语，可以自动生成首次全写，非首次缩写。

6.7 本章小结

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

结 论

学位论文的结论作为论文正文的最后一章单独排写，但不加章标题序号。

结论应是作者在学位论文研究过程中所取得的创新性成果的概要总结，不能与摘要混为一谈。博士学位论文结论应包括论文的主要结果、创新点、展望三部分，在结论中应概括论文的核心观点，明确、客观地指出本研究内容的创新性成果（含新见解、新观点、方法创新、技术创新、理论创新），并指出今后进一步在本研究方向进行研究工作的展望与设想。对所取得的创新性成果应注意从定性和定量两方面给出科学、准确的评价，分（1）、（2）、（3）…条列出，宜用“提出了”、“建立了”等词叙述。

参考文献

- [1] RAHMAN S, LI A Q, REKLEITIS I. Svin2: An underwater slam system using sonar, visual, inertial, and depth sensor [C] // 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), [S.l.], 2019 : 1861-1868.
- [2] KAESSE M, JOHANSSON H, ENGLÖT B J, et al. Towards autonomous ship hull inspection using the Bluefin HAUV [J] , 2010.
- [3] FALLON M F, KAESSE M, JOHANSSON H, et al. Efficient AUV navigation fusing acoustic ranging and side-scan sonar [C] // 2011 IEEE International Conference on Robotics and Automation, [S.l.], 2011 : 2398-2405.
- [4] CERVENKA P, DE MOUSTIER C. Sidescan sonar image processing techniques [J] . IEEE journal of oceanic engineering, 2002, 18 (2) : 108-122.
- [5] MASCARENHAS N D. An overview of speckle noise filtering in SAR images [C] // Image Processing Techniques, First Latino-American Seminar on Radar Remote Sensing : Vol 407, [S.l.], 1997 : 71.
- [6] LOPES A, TOUZI R, NEZRY E. Adaptive speckle filters and scene heterogeneity [J] . IEEE transactions on Geoscience and Remote Sensing, 2002, 28 (6) : 992-1000.
- [7] CHANG S G, YU B, VETTERLI M. Adaptive wavelet thresholding for image denoising and compression [J] . IEEE transactions on image processing, 2000, 9 (9) : 1532-1546.
- [8] BUADES A, COLL B, MOREL J-M. A non-local algorithm for image denoising [C] // 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05) : Vol 2, [S.l.], 2005 : 60-65.
- [9] DABOV K, FOI A, KATKOVNIK V, et al. Image denoising by sparse 3-D transform-domain collaborative filtering [J] . IEEE Transactions on image processing, 2007, 16 (8) : 2080-2095.
- [10] JI Y, XIE L, XU G, et al. Sonar image denoising based on Noise2Void self-supervised learning [J] . Remote Sensing Letters, 2025, 16 (5) : 560-571.
- [11] SI C, ZHANG S, CAI Q, et al. WTCRNet: a wavelet transform and contrastive regularization network for sonar denoising by self-supervision [J] . Intelligent Marine Technology and Systems, 2024, 2 (1) : 17.

- [12] ZHANG K, LI Y, LIANG J, et al. Practical blind image denoising via Swin-Conv-UNet and data synthesis [J] . Machine Intelligence Research, 2023, 20 (6) : 822-836.
- [13] VISHWAKARMA A. Denoising and inpainting of sonar images using convolutional sparse representation [J] . IEEE Transactions on Instrumentation and Measurement, 2023, 72 : 1-9.
- [14] HONG Z, FAN X, JIANG T, et al. End-to-end unpaired image denoising with conditional adversarial networks [C] //Proceedings of the AAAI conference on artificial intelligence : Vol 34, [S.l.], 2020 : 4140-4149.
- [15] LIU M-Y, BREUEL T, KAUTZ J. Unsupervised image-to-image translation networks [J] . Advances in neural information processing systems, 2017, 30.
- [16] PARK T, EFROS A A, ZHANG R, et al. Contrastive learning for unpaired image-to-image translation [C] //European conference on computer vision, [S.l.], 2020 : 319-345.
- [17] MA Y, LIU J, LIU Y, et al. Structure and illumination constrained GAN for medical image enhancement [J] . IEEE Transactions on Medical Imaging, 2021, 40 (12) : 3955-3967.
- [18] HUANG Y, XIA W, LU Z, et al. Noise-powered disentangled representation for unsupervised speckle reduction of optical coherence tomography images [J] . IEEE Transactions on Medical Imaging, 2020, 40 (10) : 2600-2614.
- [19] ZHAO B, ZHOU Q, HUANG L, et al. Unpaired sonar image denoising with simultaneous contrastive learning [J] . Computer Vision and Image Understanding, 2023, 235 : 103783.
- [20] LIN T, HINDUJA A, QADRI M, et al. Conditional GANs for Sonar Image Filtering with Applications to Underwater Occupancy Mapping [C] //2023 IEEE International Conference on Robotics and Automation (ICRA), [S.l.], 2023 : 1048-1054.
- [21] ZHOU X, TIAN K, ZHOU Z. STGAN: Sonar Image Despeckling Method Utilizing GAN and Transformer [C] //International Symposium on Artificial Intelligence and Robotics, [S.l.], 2023 : 56-67.
- [22] HO J, JAIN A, ABBEEL P. Denoising diffusion probabilistic models [J] . Advances in neural information processing systems, 2020, 33 : 6840-6851.

- [23] BRYAN O, BERTHOMIER T, D'ALES B, et al. A diffusion-based super resolution model for enhancing sonar images [J] . The Journal of the Acoustical Society of America, 2025, 157 (1) : 509-518.
- [24] YANG Z, ZHAO J, ZHANG H, et al. A side-scan sonar image synthesis method based on a diffusion model [J] . Journal of Marine Science and Engineering, 2023, 11 (6) : 1103.
- [25] ZHU J-Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks [C] //Proceedings of the IEEE international conference on computer vision, [S.l.], 2017 : 2223-2232.
- [26] ISOLA P, ZHU J-Y, ZHOU T, et al. Image-to-image translation with conditional adversarial networks [C] //Proceedings of the IEEE conference on computer vision and pattern recognition, [S.l.], 2017 : 1125-1134.
- [27] KULLBACK S, LEIBLER R A. On information and sufficiency [J] . The annals of mathematical statistics, 1951, 22 (1) : 79-86.
- [28] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation [C] //International Conference on Medical image computing and computer-assisted intervention, [S.l.], 2015 : 234-241.
- [29] JOHNSON J, ALAHI A, FEI-FEI L. Perceptual losses for real-time style transfer and super-resolution [C] //European conference on computer vision, [S.l.], 2016 : 694-711.

攻读学士学位期间取得创新性成果

（一）发表的学术论文

- [1] XXX, XXX. Static Oxidation Model of Al-Mg/C Dissipation Thermal Protection Materials [J]. Rare Metal Materials and Engineering, 2010, 39(Suppl. 1): 520-524. (SCI 收录, IDS 号为 669JS, IF=0.16)
- [2] XXX, XXX. 精密超声振动切削单晶铜的计算机仿真研究 [J]. 系统仿真学报, 2007, 19 (4): 738-741, 753. (EI 收录号: 20071310514841)
- [3] XXX, XXX. 局部多孔质气体静压轴向轴承静态特性的数值求解 [J]. 摩擦学学报, 2007 (1): 68-72. (EI 收录号: 20071510544816)
- [4] XXX, XXX. 硬脆光学晶体材料超精密切削理论研究综述 [J]. 机械工程学报, 2003, 39 (8): 15-22. (EI 收录号: 2004088028875)
- [5] XXX, XXX. 基于遗传算法的超精密切削加工表面粗糙度预测模型的参数辨识以及切削参数优化 [J]. 机械工程学报, 2005, 41 (11): 158-162. (EI 收录号: 2006039650087)
- [6] XXX, XXX. Discrete Sliding Mode Control with Fuzzy Adaptive Reaching Law on 6-PEES Parallel Robot [C]. Intelligent System Design and Applications, Jinan, 2006: 649-652. (EI 收录号: 20073210746529)

（二）申请及已获得的专利（无专利时此项不必列出）

- [1] XXX, XXX. 一种温热外敷药制备方案: 中国, 88105607.3 [P]. 1989-07-26.

（三）参与的科研项目及获奖情况

- [1] XXX, XXX. XX 气体静压轴承技术研究, XX 省自然科学基金项目. 课题编号: XXXX.
- [2] XXX, XXX. XX 静载下预应力混凝土房屋结构设计统一理论. 黑龙江省科学技术二等奖, 2007.

本科毕业论文原创性声明

作者签名: _____ 日期: _____ 年 _____ 月 _____ 日

作者签名: _____ 日期: _____ 年 _____ 月 _____ 日

- 35 -

致 谢

衷心感谢导师 XXX 教授对本人的精心指导。他的言传身教将使我终生受益。

.....

感谢哈深 L^AT_EX 论文模板 HITSZTHESIS !