

**本科毕业论文（设计）**

**题目：** 基于深度学习的SAR图像生成算法设计与实现

学 院： 信息科学与工程学院

专 业： 人工智能

学 生 姓 名： 韩柄昱

学 号： 632007100114

指 导 教 师： 王笛

评 阅 教 师：

完 成 时 间： 2024年05月18日

重庆交通大学

CHONGQING JIAOTONG UNIVERSITY

**本科毕业论文（设计）原创性声明**

本人郑重声明：所提交的毕业论文（设计），是本人在导师指导下，独立进行研究工作所取得的成果。除文中已注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品成果。对本文研究做出过重要贡献的个人和集体，均已在文中以明确方式标明。

本人完全意识到本声明的法律后果由本人承担。

作者签名（亲笔）： 年 月 日

------------------------------------------------------------------------------------------------

**本科毕业论文（设计）版权使用授权书**

本毕业论文（设计）作者完全了解学校有关保留、使用学位论文的规定，本科生在校攻读期间毕业论文（设计）工作的知识产权单位属重庆交通大学，同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅；本人授权重庆交通大学可以将毕业论文（设计）的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编毕业设计（论文）。

作者签名（亲笔）： 年 月 日

导师签名（亲笔）： 年 月 日

摘 要

合成孔径雷达（Synthetic Aperture Radar，SAR）是一种主动微波式成像雷达，具有全天时、全天候的成像能力，在城市观察、资源勘探和灾情检测等民事、军事领域发挥着重要作用。由于SAR特殊的成像机理，基于深度学习的SAR图像解译面临着图像数据集数量匮乏和质量较低的双重挑战而对抗生成网络（Generative Adversarial Network，GAN）在生成高质量图像上有着出色的效果，因此，利用GAN生成高质量的SAR图像具有重要意义和广阔的研究前景。

本文在分析基于深度学习的SAR图像研究现状基础上，首先简要说明了SAR成像机理及图像特性并介绍了深度生成模型，然后基于非同源数据转化思想，利用GAN实现SAR图像生成任务。本文主要内容如下：

针对传统深度学习方法仅使用同源信息并且未能准确地提取目标信息的问题，使用基于图像翻译的SAR图像生成方法，实现利用光学信息辅助SAR图像生成。通过对Pix2Pix等配对翻译模型和CycleGAN等非配对翻译模型在视觉观察和评价指标的比较和分析，得出配对模型生成的SAR图像在相干斑模拟、视觉观感上更接近原始SAR图像的同时在相关指标上拥有更优异的表现。

针对图像翻译模型生成SAR图像可能会出现伪影、虚构等图像失真现象，使用基于时间偏移的SAR图像生成方法，将SAR图像生成任务转化为依据光学图像的变化而修改原始SAR图像。通过对时间偏移对抗生成网络（Temporal Shifting GAN，TSGAN）生成结果的对比和分析，TSGAN模型比图像翻译方法Pix2Pix模型在视觉效果和评价指标上有明显提高，且注意力机制的使用可以更有效地提取原始光学和SAR图像相关信息，生成更高质量的SAR图像。

关键词**：**对抗生成网络，SAR图像生成，图像翻译，时间偏移

**Design and implementation of SAR image generation algorithm based on deep learning**

Abstract

Synthetic Aperture Radar (SAR) is an active microwave imaging radar with all-weather, all-day imaging capabilities, playing a significant role in both civilian and military fields such as urban observation, resource exploration, and disaster monitoring. Due to the unique imaging mechanism of SAR, interpreting SAR images using deep learning faces dual challenges of limited and low-quality image datasets. However, Generative Adversarial Networks (GANs) have demonstrated outstanding performance in generating high-quality images. Therefore, leveraging GANs to generate high-quality SAR images is of great importance and holds broad research prospects.

This paper, based on an analysis of the current state of SAR image research using deep learning, first briefly explains the SAR imaging mechanism and image characteristics, and introduces deep generative models. It then employs the concept of non-homogeneous data transformation to utilize GANs for SAR image generation tasks. The main content of this paper includes:

Addressing the problem that traditional deep learning methods only use homogeneous information and fail to accurately extract target information, a SAR image generation method based on image translation is employed to utilize optical information to assist in SAR image generation. Through comparative analysis of paired translation models like Pix2Pix and unpaired translation models like CycleGAN in terms of visual observation and evaluation metrics, it is found that the SAR images generated by paired models are closer to the original SAR images in terms of speckle simulation and visual appearance, while also exhibiting superior performance in related metrics.

To tackle the issue of artifacts and distortions that may arise in SAR images generation by image translation models, a method based on temporal shifting is used. This approach transforms the SAR image generation task into modifying the original SAR image according to changes in optical images. Comparative analysis of the results generated by the Temporal Shifting GAN (TSGAN) shows that the TSGAN model significantly improves visual effects and evaluation metrics compared to the Pix2Pix model. Additionally, the use of attention mechanisms more effectively extracts relevant information from the original optical and SAR images, leading to the generation of higher quality SAR images.

**Key Words：**GAN, SAR image generation, image translation, temporal shifting

目 录

[第 1 章 绪论 1](#_Toc167351030)

[1.1 研究背景与意义 1](#_Toc167351031)

[1.2 国内外研究现状 1](#_Toc167351032)

[1.2.1 数据生成与图像增强研究现状 1](#_Toc167351033)

[1.2.2 SAR图像生成研究现状 3](#_Toc167351034)

[1.3 本文内容与章节安排 5](#_Toc167351035)

[第 2 章 SAR图像特性与深度生成模型概述 6](#_Toc167351036)

[2.1 SAR成像机理 6](#_Toc167351037)

[2.1.1 距离向分辨率 7](#_Toc167351038)

[2.1.2 方位向分辨率 8](#_Toc167351039)

[2.2 SAR图像特性 9](#_Toc167351040)

[2.2.1 散射特性 9](#_Toc167351041)

[2.2.2 几何特性 10](#_Toc167351042)

[2.2.3 统计特性 11](#_Toc167351043)

[2.3 深度生成模型 12](#_Toc167351044)

[2.3.1 变分自编码器 12](#_Toc167351045)

[2.3.2 对抗生成网络 13](#_Toc167351046)

[2.4 本章小结 15](#_Toc167351047)

[第 3 章 基于图像翻译的SAR图像生成方法 16](#_Toc167351048)

[3.1 图像翻译模型 16](#_Toc167351049)

[3.1.1 配对翻译 16](#_Toc167351050)

[3.1.2 非配对翻译 18](#_Toc167351051)

[3.2 实验数据集与评估指标 24](#_Toc167351052)

[3.2.1 实验数据集 24](#_Toc167351053)

[3.2.2 实验评价指标 25](#_Toc167351054)

[3.3 实验结果及分析 27](#_Toc167351055)

[3.3.1 实验参数设置 28](#_Toc167351056)

[3.3.2 实验结果分析 28](#_Toc167351057)

[3.4 本章总结 30](#_Toc167351058)

[第 4 章 基于时间偏移的SAR图像生成方法 31](#_Toc167351059)

[4.1 时间偏移对抗生成网络 31](#_Toc167351060)

[4.1.1 模型结构 31](#_Toc167351061)

[4.1.2 模型注意力机制 32](#_Toc167351062)

[4.1.3 模型损失函数 38](#_Toc167351063)

[4.2 实验数据集与评价指标 39](#_Toc167351064)

[4.2.1 实验数据集 40](#_Toc167351065)

[4.2.2 实验评估指标 40](#_Toc167351066)

[4.3 实验结果及分析 41](#_Toc167351067)

[4.3.1 实验参数设置 42](#_Toc167351068)

[4.3.2 实验结果分析 42](#_Toc167351069)

[4.4 本章小结 44](#_Toc167351070)

[第 5 章 总结与展望 45](#_Toc167351071)

[5.1 全文总结 45](#_Toc167351072)

[5.2 展望 46](#_Toc167351073)

[致 谢 47](#_Toc167351074)

[参 考 文 献 48](#_Toc167351075)

# 绪论

## 研究背景与意义

合成孔径雷达（Synthetic Aperture Radar，SAR）是一种主动式微波遥感成像系统，其利用脉冲压缩技术实现距离向的高分辨率，同时利用平台运动所产生的多普勒频移现象来合成等效大尺度孔径以实现方位向的高分辨率，最终获得目标区域的高分辨率雷达图像。同光学成像系统相比，SAR系统主要优点为不受自然环境如天气、阳光的影响，具有全天候，全天时的成像能力，同时在民用和军事等领域具有广阔的应用前景，如城市发展与分析、资源勘探、灾害监测、战场侦察等[1]。

随着人工智能特别是深度学习的发展，为SAR图像研究与处理提供了强有力的工具。不同于传统方法，深度学习方法能够更准确、高效地学习SAR图像数据内部的潜在分布，但是在训练时需要大量的SAR图像数据用于其训练和测试。由于SAR特殊的成像机理，制作一个高质量SAR图像数据集要耗费大量成本，这很大程度的限制了深度学习的遥感图像解译的发展。因此为了进一步促进深度学习在SAR图像解译方面的应用，研究SAR图像生成具有重要意义。

传统图像增强方法如旋转、剪裁和调整图像对比度、亮度等均只是对原始图像的变换、且增强后的图像和原始图像相比并未提高数据的多样性。传统SAR生成方法如电磁仿真等，其结果相比真实数据相差较大。然而，深度生成模型出现，为SAR图像生成带来全新可能，其中对抗生成网络（Generative Adversarial Network，GAN）[2]系列模型具有更出色的效果受到学者们的广泛研究。因此，本论文将重点讨论基于深度生成模型的SAR图像生成问题。

## 国内外研究现状

### 数据生成与图像增强研究现状

随着深度学习的不断发展，深度网络模型结构逐渐复杂、参数量逐渐增大，其面临的挑战也接踵而至。比如：深度学习训练需要消耗大量图形处理器（Graphics Processing Unit，GPU）资源、模型“黑箱”缺乏可解释性、和网络过拟合等问题。其中，模型训练过程中优质的训练数据缺乏导致模型过拟合、泛化能力差是目前亟待解决的问题之一。目前，数据生成与图像增强方法主要分为两种，传统方法和基于深度学习方法。传统方法主要有几何变换法，色彩空间变换法和混合图像法等[3]。

几何变换主要有如旋转、裁切、平移和添加高斯随机噪声等，其优点为操作实现较为建议，同时对于图像内容及标签不含文本信息且与相对位置无关的图像较为适用。例如Simard等[4]曾利用上述几何方法对数据集生成并利用卷积神经网络（Convolutional Neural Network，CNN）训练，证明了利用几何方法生成后的数据集可以使网络拥有更强的泛化能力。然而对于一些图像中内容及标签包含文本信息且相对位置敏感的数据集，利用几何方法会导致标签改变，则生成后的数据集将无法使用。

色彩空间变化法指在图像颜色通道中对每个通道的像素值增加或减少，以改善图像的亮度和对比度，或是对每个通道利用分布直方图均衡化整体改善。因为颜色对于图像是一个不可或缺的特征，所以色彩空间变化可能会将图像颜色信息丢失导致图像失真或造成和颜色紧密关联的标签信息失效。混合图像法是指将两幅以上图像，通过拼接、融合等方式形成新图像的方法。Ionue等人提出一种图像混合方法Sample Pairing[5]。该方法首先将两张大小为256×256的图像随机裁剪为大小224×224的图像，然后将图像进行随机水平旋转，最后将两张图像对应通道中各个像素值分别取平均融合成新图像。实验结果表明在CIFAR-10数据集上利用Sample Pairing方法后的图像相比原始图像分类错误率得到降低。传统方法优点是操作简便、有效、目前主流深度学习框架均已提供这种方法。但是对于一些特定任务和特殊数据集，传统方法有改变标签信息和数据分布、缺乏可解释性等问题，具有很大的应用局限性。

深度学习方法指利用神经网络强大的表征能力实现数据扩充与图像增强，这其中以GAN为代表的深度生成模型受到学者们的广泛关注。Relja等[6]提出CP-GAN网络，首先在无监督信息情况下，网络提取物体所对应的掩膜信息并识别，然后复制图像中的物体达到数据扩充的目的。Lee等[7]利用两个GAN网络分别实现预测生成图像物体与背景的位置和物体在背景中的形态。Frid-Adar等[8]利用DCGAN实现医学肝脏病变图像的生成。其利用包含182个CT扫描图像的数据集作为原始数据集生成新图像样本加入数据集中。该方法生成数据集相比传统方法生成数据集在敏感度和特异度有明显提高。众多研究表明，利用GAN等深度生成模型对数据进行扩充和图像增强有较好的应用前景，因此，相关学者研究GAN应用在SAR图像领域中。

### SAR图像生成研究现状

随着深度学习在SAR图像解译的研究，SAR图像生成任务受到相关学者的广泛关注。由于SAR图像特殊成像机理，相比光学图像在分辨率、图像分布和特征等方面有许多不同。本节具体介绍SAR图像生成的传统方法和深度学习方法。

1. 传统方法SAR图像生成

传统SAR图像生成方法是指利用电磁仿真技术对雷达照射目标后接受回波信号进行仿真计算生成SAR图像。Marconi等[9]指出SAR图像仿真与生成可分为基于回波信号的仿真和基于SAR图像的仿真。基于回波信号的仿真有两种实现方法，一是将初始接收的回波信号进行仿真生成模拟回波信号，并经过信号处理器进行生成图像。该方法侧重雷达散射截面积（Radar Cross Section，RCS）仿真，一般采用基尔霍夫物理光学（Kichhoff Phsical Optics，PO）[10]，几何光学（Geometrical Optics，GO）估计法[11]和积分方程法（Intergala Equation Method，IEM）[12]。而是利用雷达发射电磁波信号的局部入射角对目标SAR图像像素进行仿真，而无需对合成孔径过程进行仿真，因此在时间复杂度上具有明显优势，但对目标物体电磁特性描述结果较差[13]。目前SAR图像模拟软件基于该类方法，一般采用基于射线追踪的RaySAR与CohRaS和基于光栅法的SARViz[14]。

基于SAR图像的仿真是将非相干SAR图像同实际SAR图像相结合，仿真生成新的SAR图像。景士娴[15]等提出基于数据库SAR图像生成场景合成研究。其思想是首先利用自适应边缘检测算法将原始SAR图像进行分割使其目标区域与背景噪声分离，然后利用最大似然区域合并算法将提取到的不同目标和背景噪声合并生成新的SAR图像数据。朱柳[16]等针对分析SAR图像相干斑噪声的统计模型，提出一种基于金字塔变换法改进的EMMP（增广矩阵束）散射中心重构法，但其提取目标特征方法复杂、成本高且不同方法易造成生成图像的失真。

利用传统方法生成SAR图像通常存在生成质量不理想、耗费大量资源等问题。但随着人工智能蓬勃发展，深度学习在图像、信号处理等方面具有显著优势，近年来不断有学者研究基于深度学习的SAR图像生成。

1. 深度学习SAR图像生成

深度学习SAR图像生成主要是利用深度神经网络强大的表征能力从本质上描述数据，而深度生成模型通过学习原始数据集信息，并计算、模拟其数据分布，从而生成与原始数据分布类似的新样本。将深度生成模型应用在SAR图像数据生成，相比传统方法可避免复杂建模过程同时更准确仿真原始数据分布，达到更高质量生成效果。

Guo等[17]提出一个利用条件对抗生成网络（Conditional GAN，CGAN）生成SAR图像，该方法采用MSTAR数据集训练一个端到端模型，但生成SAR图像常受原始图像噪声干扰，使生成图像为噪声图片。因此利用杂波归一化对模型改进，提高模型的生成质量。Cui等[18]结合改进的带梯度惩罚项的WGAN（Wasserstein GAN，WGAN）实现方位角可控的SAR图像生成。提出一种基于标签导向的GAN（Label Directed GAN，LDGAN）[19]该方法不仅使用Wassertein距离代替JS散度衡量真实数据与原始数据的差异，避免训练模式崩溃问题，提高模型生成质量，同时将目标类别信息融入训练损失函数中，从而避免模型生成大量不属于原始数据集类别的图像，提高生成图像的真实性。Fan等[20]提出基于欺骗性干扰模板对抗生成网络（Deceptive Jamming Template GAN，DJTGAN）。首先该方法可以依据不同方位角、不同目标类别和图像分辨率生成高保真欺骗性干扰模板，其次采用多级特征提取机制，利用CNN有效提取原始SAR图像中目标低频和高频信息生成精细化欺骗性干扰模板，最后还利用PatchGAN作为判别器捕获原始SAR图像中局部纹理数据的分布。周驿枰学者[21]在考虑SAR图像分布基础上，提出结构相似性损失对抗生成网络SAR图像生成方法。通过利用PatchGAN改进鉴别器、利用结构相似性理论改进损失函数和利用金字塔架构网络对单张SAR图像进行内部混合生成，以提高模型生成的逼真度。冯广能学者[22]利用循环一致性生成对抗网络（Cycle-Consistent GAN，CycleGAN）通过风格迁移将图像由光学域迁移至SAR图像域，同时在训练过程中加入特征级的相似性约束和像素级的差异性约束两种方法去避免生成过程丢失SAR图像目标信息。但是该方法生成的SAR图像分辨率较低，真实感和效果有待提升。

## 本文内容与章节安排

本文以SAR图像为主要研究对象，主要研究内容基于深度学习的SAR图像生成，主要采用方法为深度生成模型。本文首先针对SAR图像生成研究现状进行分析，总结，接着讲解了SAR图像特性及深度生成模型基本原理，最后利用利用基于图像翻译方法和基于时间偏移方法，将光学目标图像转化为SAR目标图像，实现非同源数据辅助的SAR图像生成方法，同时对每种实验理论及过程进行阐述并对实验结果进行分析与总结。本文具体内容安排如下：

第一章，绪论。本章介绍基于深度学习的SAR图像生成的背景及意义，同时对数据生成及图像增强以及SAR图像生成的相关研究现状进行分析与总结，最后介绍本文主要内容与章节安排。

第二章，SAR图像特性与深度生成模型概述。本章首先介绍SAR成像机理，并简要SAR的距离分辨率和方位分辨率，接着讲解SAR图像散射特性、几何特性和统计特性，最后介绍深度生成模型和对抗生成网络变种。

第三章，基于图像翻译的SAR图像生成方法。本章首先介绍5种图像翻译基本模型，接着介绍本章实验所用数据集及评价指标，最后对生成图像结果可视化分析，并利用评价指标进行评价模型生成表现及性能。

第四章，基于时间偏移的SAR图像生成方法。本章首先介绍时间偏移对抗生成网络模型结构，并删除模型使用注意力机制原理及变化加权损失函数，接着介绍本章使用数据集和针对引入而变化图而改进的加权评价指标，最后对模型生成图像结果进行可视化分析，同时对比图像翻译方法及时间偏移方法优劣及不同注意力机制对模型生成效果影响。

第五章，总结与展望。本章首先对本文主要实验进行归纳，并再次基础上针对总结实验过程中发现问题，为下一步实验提供改进的方向和目标。

# SAR图像特性与深度生成模型概述

合成孔径雷达作为一种主动微波遥感雷达传感器，其成像相比传统光学成像在成像机理、成像特性等方面具有显示不同。而随着基于深度学习的SAR图像解译发展对高质量且多数量的数据需求越来越大，因此如何有效地对SAR图像数据集进行扩充是一个挑战。为了更好实现依据光学数据集辅助SAR图像生成，进一步挖掘光学数据集和SAR图像数据集特征，深度生成模型可以很好满足上述要求。本章首先简要介绍SAR图像成像原理何其图像特性，然后分析深度生成模型原理，最后给出一些常用SAR图像生成质量的评价方法。

## SAR成像机理

SAR作为一种相干成像雷达，其成像核心为分析回波信号的多普勒频移现象和处理回波信号中掺杂地形信息，提取目标的后向散射系数。[23]与光学成像相比，SAR成像的图像拥有更丰富的细节，更高的对比度和更复杂的图像分布。SAR主要采用有源主动的侧视雷达，其成像机理模型如图2.1所示。SAR的分辨率指雷达在目标方向可以区分目标的最小距离，为距离向分辨率和方位向分辨率两种。如图，将与天线飞行方向同向的称为距离向，与飞行方向垂直的波束发射方向称为方位向。因此，若要获得高分辨率图像则需要将提高上述两个方向分辨率。下面将阐述SAR如何提升两个方向分辨率。

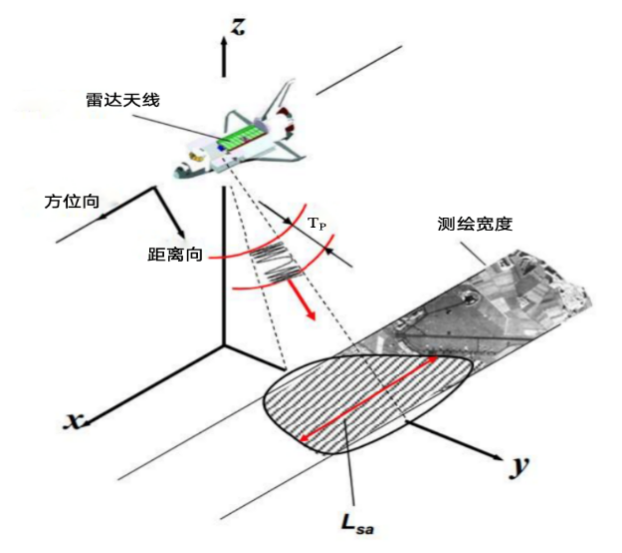


图2.1 合成孔径雷达的成像机理模型图

### 距离向分辨率

距离向分辨率指雷达在距离向可以区分目标的最小距离，其计算公式如下：



其中，为距离向分辨率为光速，为雷达发射信号的脉冲宽度。由公式可知，若想提高距离向分辨率即减小，则需减少脉冲宽度。由雷达作用距离方程可知[24]，若减小雷达发射信号的脉冲宽度则会降低使雷达峰值功率，从而导致缩短雷达最大作用距离。雷达短时间内发射高能量信号十分困难，因此，提升雷达距离分辨率面临着雷达最大作用距离与高之间分辨率的矛盾。为了克服上述矛盾，目前雷达主要发射具有大时宽带宽积的线性调频信号（Linear Frequency Modulation，LFM），并利用脉冲压缩技术解决作用距离和分辨率间的矛盾。LFM信号表达式如下：



其中，表示脉冲宽度，表示LFM信号的调频斜率，表示信号载频，表示矩形窗函数，定义如下：



由式可得其相位，对时间求导可得LFM信号瞬时频率：



则LFM信号得带宽表达式如下：



结合式，根据模糊函数理论可得LFM信号互相关函数，求解其第一过零点[25]，可推导距离分辨率为：



可以看出，发射LFM信号的雷达距离向分辨率和发射信号带宽有关，带宽越大，分辨率越高

### 方位向分辨率

方位向分辨率指在雷达方位向上可以区分目标的最小距离，其表达式如下：



其中为方位向分辨率，为雷达信号波束宽度，为雷达斜距。通常SAR被安装在机载平台或星载平台上，斜距非常大。所以若想提高雷达的方位向分辨率，只能减少波束宽度。根据雷达天线理论，波束宽度和雷达天线孔径长度呈如下关系：



其中，为天线旁瓣相关的尺度因子，为发射信号波长，为SAR的孔径。为了便于后续分析，这里设。由式和式可将方位向分辨率表示为：



由式可知若想提高方位向分辨率可以通过降低雷达最大作用距离，降低发送信号波长，或增加雷达天线孔径。但降低雷达最大作用距离和发射短波长信号会限制雷达应用场景，因此选择增加雷达天线孔径来提升方位向分辨率。目前SAR被主要安装在卫星和飞机平台上，由于平台本身限制无法安装大孔径雷达天线，因此提升方位向分辨率面临着雷达天线孔径长度与高分辨率的矛盾。为了克服上述矛盾，学者提出合成孔径概念，利用平台运动产生的多普勒频移和信号处理方式合成等效大孔径。根据合成孔径理论的方位向最大分辨率可以表示为：



因此，SAR的方位向最大分辨率仅和雷达天线孔径相关，且天线孔径越短雷达的方位向分辨率越高，与雷达作用距离和雷达发射信号波长无关。

## SAR图像特性

由于SAR特殊成像机理导致SAR图像和光学图像相比具有非常多不同的特性例如散射特性、几何特性和统计特性等。为了更好理解SAR图像特性，本节将从上述三个特性分析SAR图像，以更好实现SAR图像生成。

### 散射特性

SAR通过接收发射信号照射到多个随机分布的散射体目标所反射的回波信号进行相干叠加成像。由于发射信号的电磁波波长远小于可分辨单位单元（图像单位像素）的大小，所以一个可分辨单元内可包含多个散射体，因此其回波也会有细微差别。同时，回波信号进行相干叠加中包含着各个时间的雷达回波的相位，导致实际目标的散射回波强度在散射系数的基础上随机起伏，产生斑点噪声，在图像上表现为灰度值发生剧烈变化，这称为相干斑噪声干扰。如图2.2所示相干叠加原理。

|  |  |
| --- | --- |
| （a）单元内多个地物散射回波 | (b)相干叠加示意图 |

图2.2 SAR相干叠加原理示意图

相干斑是相干成像系统的固有现象，由于是回波信号矢量叠加形成，所以对于完全发展的相干斑，在SAR信号处理中通常建模为乘性噪声模型，可表述如下：



其中，代表图像像素坐标，代表雷达接受到回波的真实强度，代表理想状态下，不受相干噪声所影响的雷达接收到的回波的强度，也就是理论上清晰照片，代表相干斑噪声参数。

### 几何特性

SAR作为一种侧视雷达其图像本质上为距离向图像，即图像前面位置受雷达发射信号传播距离的影响。因此，SAR图像的空间几何并非和地面几何一一对应，具体可观察到透视缩短（Foreshortening）、背波延长（Elongation）、顶底倒置和阴影（Shadow）等现象[26]。图2.3给出雷达几何特性产生原理。

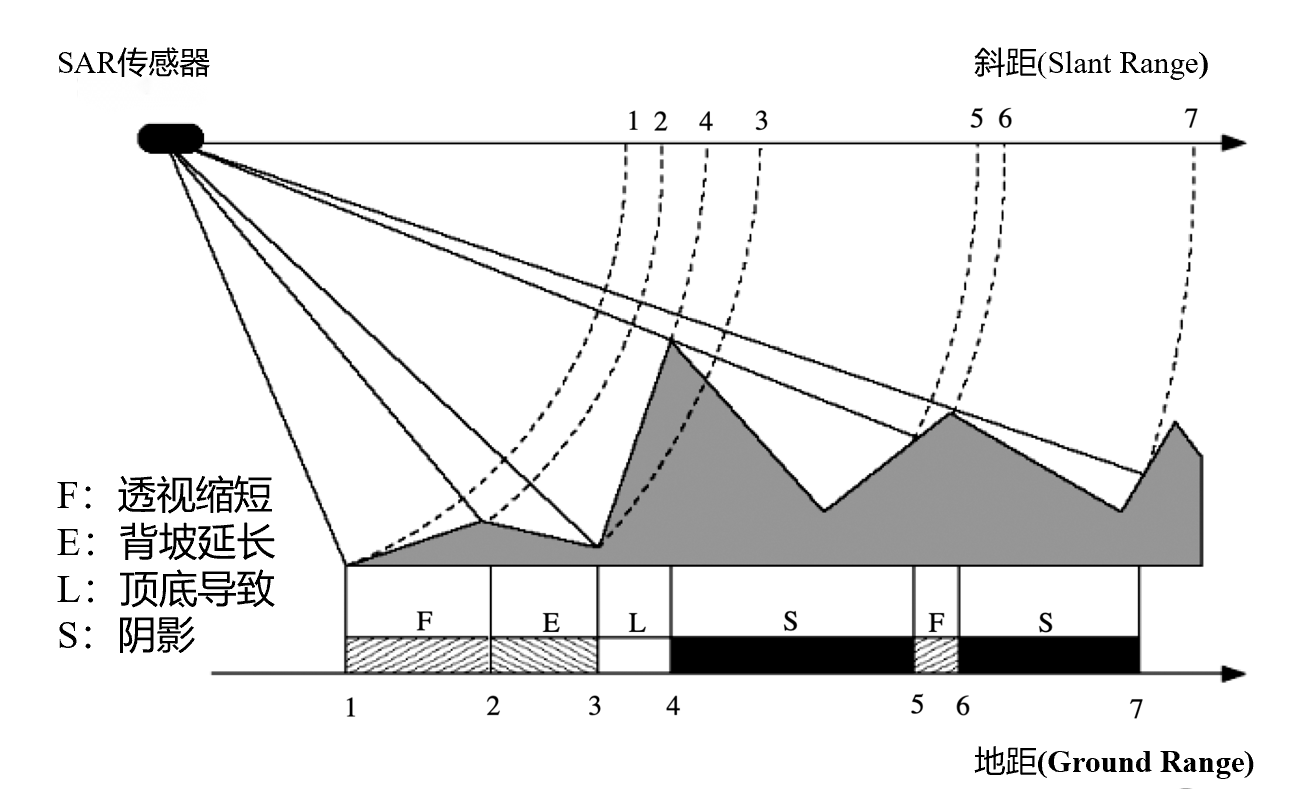


图2.3 雷达几何特性示意图

* 1. 透视收缩

透视缩短指目标在雷达图像上的尺寸相较本身的尺寸有所收缩。例如在高山区域中，朝向雷达的坡面上两点之间的斜距要比其两点之间地距要小，因此导致坡面在SAR图像上长度缩短和在SAR图像中通常具有较高的灰度值。

* 1. 背坡延长

背坡延长和透视收缩相反，指背向雷达的坡面在SAR图像长度伸长同时具有较低的灰度值。

* 1. 顶底倒置

顶底倒置指雷达图像上目标物体底部和顶部被错位记录，其产生原因为当雷达发射电磁波到达目标物体顶部的斜距小于到达底部斜距，也就是说雷达先接收到目标底部的回波信号，就会出现顶底倒置现象。

* 1. 阴影

阴影现象指在含有高大的建筑物、高山等地形，常常出现目标阻挡雷达信号导致目标区域无法被雷达信号照射，所以该区域信号无法反射并使雷达接收，导致SAR图像该区域亮度低，呈现一片黑色。

图2.4[21]所示埃及金字塔的SAR图像明显发生上述提到透视收缩、顶底倒置和阴影等几何现象。



图2.4 埃及金字塔俯视光学图像和侧视SAR图像示意图

### 统计特性

SAR作为一种相干成像系统其回波信号呈明显随机性，利用概率和统计知识分析SAR随机信号成为学者们重要研究方向。目前，描述SAR回波信号的统计分布模型可分为两种，面向全发展相干斑模式的统计模型和面向部分发展相干斑的统计模型[27]。第一种常用于低分辨率SAR图像，同时假设每个分辨单元的各个散射中心和每个散射中心回波信号的幅度和相位均相互独立，提出对于单视复数据（Single Look Complex，SLC）的SAR回波幅度服从Rayleigh分布，强度服从指数分布。对于*L*视的多视SAR图像其回波强度服从Gamma分布，幅度服从Nakagami分布。第二种应用在分辨率较高的SAR图像，学者们利用混合概率模型和高自由度的复合概率模型拟合拥有较高分辨率的SAR回波信号分布，例如Fisher概率模型，广义Gamma-Rayleigh模型等。

## 深度生成模型

在深度学习中，模型一般可分为判别模型和生成模型。若将模型表示成条件概率模式，那么判别模型则直接对条件概率建模，学习在已给输入特征*X*情况下标签*Y*的分布。常见的判别模型有：决策树、支持向量机（Support Vector Machine，SVM）、感知机和k近邻等。因为判别模型无法学习训练数据本身特性，因此将学习过程简化，而生成模型则先对联合概率分布进行建模，然后利用条件概率公式求得后验概率。条件概率公式如下：



相比判别模型，生成模型可以更深层次学习数据分布，挖掘数据深层次信息，常见的生成模型有：朴素贝叶斯算法模型、受限玻尔兹曼机和深度置信网络等。本节将重点讲解两种深度生成模型变分自编码器和对抗生成网络。

### 变分自编码器

变分自编码器（Variational AutoEncoder，VAE）[28]是在自编码器的基础上结合变分推断的思想，利用两个神经网络（生成网络和推断网络）建模两个概率分布。VAE在结构上和自编码器类似，其中编码器为推断网络，目的是将一个高维空间变量映射到低维空间隐变量，得到变分分布；解码器成为生成网络，目的是将低维隐变量重新映射回高维空间可观测变量上，得到概率分布。相比较自编码器，VAE两个网络输出的是分布而不是编码，VAE结构如图2.5所示。

推断网络的优化目标是使变分分布最大化与概率分布相似，因此可以利用评价两个分布相似性指标KL散度。当KL散度值越小代表上述两个分布越相似，故需要找到一组最优推断网络参数：





图2.5 VAE网络结构图

其中，由于KL散度在计算过程中需要后验概率分布，一般情况难以计算。所以，采用变分推断思想间接计算后验分布。根据变分推断理论，在给定样本情况下，其对数似然函数可被估计为：



其中函数为对数似然函数的证据下界：



结合式（2.14）和式（2.13），推断网络目标函数为：



生成网络优化目标同推断网络类似，也是找到最优参数使证据下界最大化。故VAE的目标函数为：



VAE虽然可以学习到数据的内在分布并生成多样的样本，但应用在图像生成上会有生成的图像较为模糊，同时VAE训练过程较为复杂，训练时间较长。

### 对抗生成网络

GAN作为生成模型中最热门的模型之一，受到学者们的广泛研究。GAN的提出受到博弈论影响，由一个生成器*G*和一个判别器*D*组成。生成器输入一个随机噪声，产生一个假的样本，而判别器输入真实样本和生成样本，输出判别结果，两者互相博弈，互相优化最终进入一个纳什均衡状态。当进入该状态后，生成器和判别器同时达到最优，同时改变生成器或判别器均会将模型性能下架，此时标志着训练完成。GAN结构图如图2.6所示。

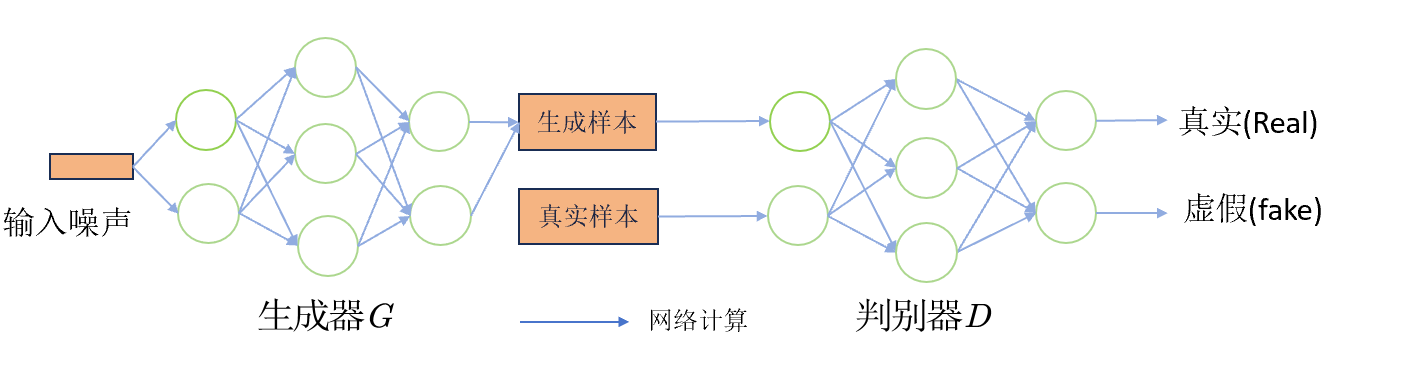


图2.6 GAN结构图

GAN的判别器*D*目标是区分输入样本是生成器生成样本还是真实样本，是一个二分类问题，若是生成器生成样本判断为0，若是真实样本则判断为1，在该类问题中常用交叉熵损失函数，其表达式如下：



其中，为输入样本，为样本所对应类别，表示样本被预测为第*i*个类别的概率。

GAN生成器*G*的目标是生成判别器无法判断生成样本是由生成器生成的还是源于真实数据，也就是说希望让判别器将自己产生的样本判断为真实样本。结合*G*和*D*目标，GAN网络的损失函数可定义为：



其中，表示样本集，表示随机噪声。而在训练过程中，样本来自于真实样本即或来自于生成器生成即。*V(D,G)*则表示生成样本和真实样本之间的差异大小。GAN的训练分为两步，首先固定生成器，最大化判别器参数；然后固定判别器，最小化生成器参数，这样GAN损失函数则可看作为最小最大化游戏（Minimax Game）[30]。

根据判别器目标函数，可以进一步得出最优判别器：



将最优判别器带入式后，GAN损失函数可变为：



其中为KL散度，为JS散度。所以在GAN中，当判别器达到最优后，生成网络的优化目标为最小化生成分布和真实分布之间的JS散度，即当JS散度为0时，两者分布相同。此时，最优生成器对应损失为。

## 本章小结

本章主要介绍了SAR图像和深度生成模型的相关理论基础知识。首先简要分析对SAR图像成像原理及两个方位的分辨率进行，然后介绍了SAR图像特性包括散射特性、几何特性和统计特性，最后阐述了深度生成模型原理。

# 基于图像翻译的SAR图像生成方法

利用传统SAR图像生成方法如基于回波信号仿真和基于SAR图像仿真，存在生成过程较繁琐，生成效果较差，生成效率较低等问题，而GAN网络有生成图像质量较高、生成过程简易等优势。本章将介绍基于图像翻译的SAR图像生成方法，利用光学域图像作为非同源信息辅助生成任务，将光学图像转化为目标SAR图像。本章首先讲解目前主流的图像翻译GAN模型，接着介绍本章实验数据集和实验评价指标，最后利用生成图像视觉分析和评价指标对比主流网络模型的生成效果。

## 图像翻译模型

图像翻译（Image-to-Image Translation，I2IT）任务是指利用一个端到端的模型实现源域（如光学域）到目标域（SAR域）转变。经过这一过程，不仅可以生成目标域照片，同时可以结合源域特征，减少异构数据之间非线性差异[33]。根据优化方法和数据集组织的不同，I2IT任务可以分为配对翻译（Paired）和非配对翻译（Unpaired）方法，本节将重点介绍这两种方法的主流模型.

### 配对翻译

1. Pix2Pix模型

Pix2Pix[34]模型是基于CGAN的图像翻译模型，该模型不仅在图像生成领域相比其他模型有更为逼真的图像生成效果，也为图像翻译领域模型发展提供一个通用解决方案和模型框架。Pix2Pix模型结构如图3.1所示。

Pix2Pix模型包含生成器*G*和判别器*D*，其生成器*G*网络为U-Net模型。该网络采用编码器-解码器（Encoder-Decoder）结构，在Encoder部分，采用多层网络模块堆叠，每个模块包含一个卷积层、一个实例归一化层和一个线性整流单元（Rectified Linear Unit，ReLU）激活函数构成，实现将输入光学图像进行编码。Decoder部分结构和Encoder类似，将卷积层换成转置卷积实现上采样，最后一层激活函数更换为Tanh函数。同时U-Net在每一层的Encoder和Decoder间加入跃级连接（Skip Connection）使对应图像的低级特征和高级特征充分融合，这样解码器在解码过程中可以将高层次信息充分利用。其判别器*D*网络为PatchGAN，该网络将生成图像和真实图像等分成Patch，分别判断每个Patch之间的真假后取平均，可以更好的对图像局部进行判断。

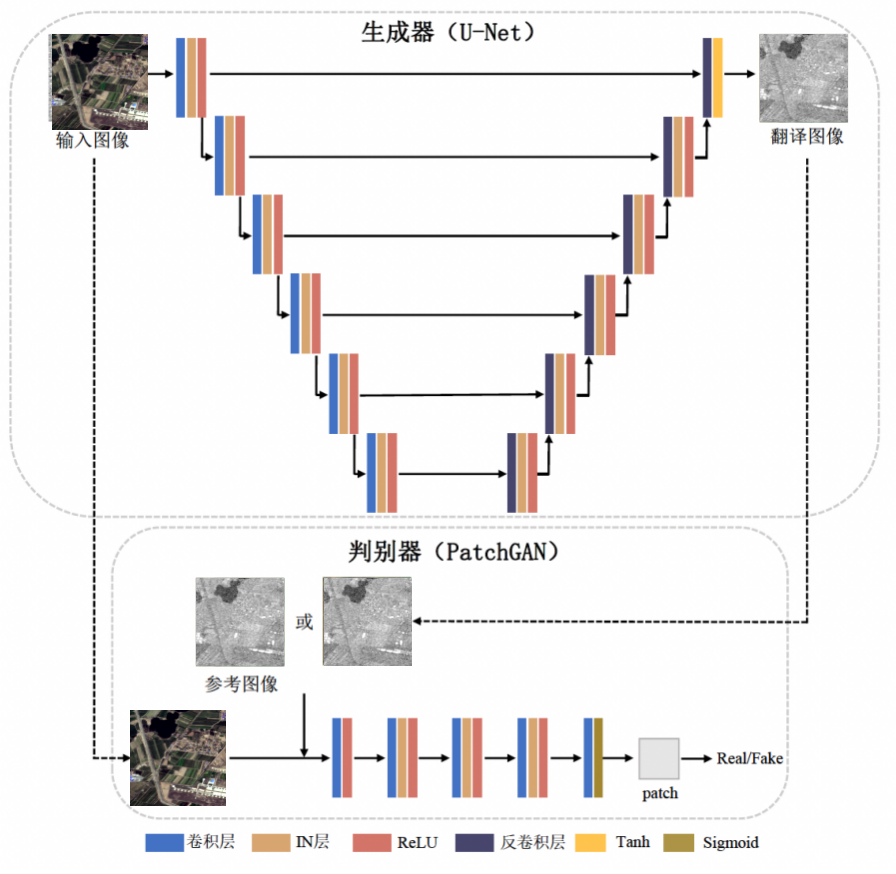


图3.1 Pix2Pix 结构图

Pix2Pix在损失函数上，相比传统CGAN Loss，Pix2Pix采用L1 Loss约束源域图像和目标域图像相似性。Pix2Pix损失函数表达式如下：



1. BicycleGAN模型，

BicycleGAN[35]提出是为了改进Pix2Pix模型输出风格单一，输入图像和生成图像仅可以实现一对一转换。BicycleGAN结合CGAN思想，输入为随机噪声和目标域图像，同时在模型结构中，组合训练两个网络：cVAE-GAN和cLR-GAN（conditional Latent Regressor GAN，cLR-GAN）实现在给定不同输入噪声情况下生成不同风格的目标域图像。Bicycle GAN的训练流程如图3.2所示：

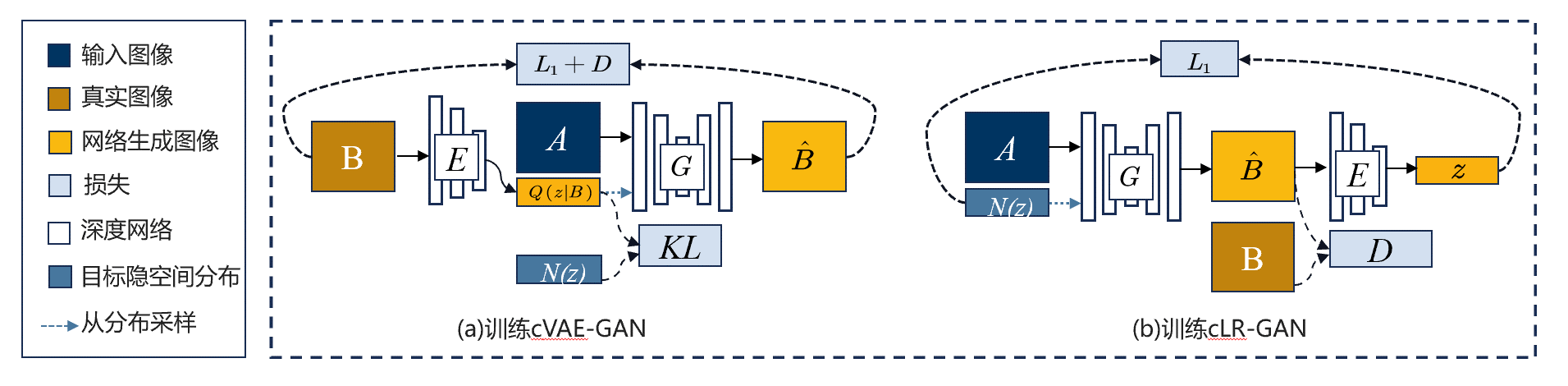


图3.2 BicycleGAN训练流程图

如图3.2，对应到本章图像生成任务上，*A*为输入光学图像，*B*为目标域SAR图像，为网络生成图像。对于训练cVAE-GAN，第一步是将目标域图*B*通过一个编码器*E*编码为一个隐空间向变量，随后利用KL散度逼近已定义好随机噪声分布，第二步将图像*A*和隐空间向量送入生成器*G*合成生成照片，第三步送入判别器*D*，并结合GAN损失和L1损失优化编码器E和生成器*G*。对于cLR-GAN则与cVAE-GAN不同，其第一步是利用图像*A*和随机噪声送入生成*G*中生成图像，第二步先将生成图像和真实目标域图像*B*送入判别器，然后将送入编码器编码为隐向量，最后利用L1损失计算噪声和隐空间编码。Bicycle GAN的结构两个GAN的生成器*G*、编码器*E*参数共享，结合上述定义最后损失函数可表达为：



### 非配对翻译

1. CycleGAN模型

由于现实中的配对数据集难以收集，为了解决例如Pix2Pix等模型需要配对数据集，提出循环对抗生成网络（Cycle-Consistent GAN，CycleGAN）[36]。CycleGAN的无配对训练，极大拓宽了图像翻译应用领域例如图像压缩、图像去雾、图像超分辨等。以本章内容为例，假设光学图像域为*X*，SAR图像域为*Y*，*X*域至*Y*域生成器为*G*，*Y*域至*X*域生成器为*F*，和分别为上述两个转化过程的判别器，CycleGAN的目标则实现，相比较传统的GAN，CycleGAN需要将生成的*Y*域图像恢复至*X*域，即，以实现“循环”。CycleGAN思想如图3.3所示。

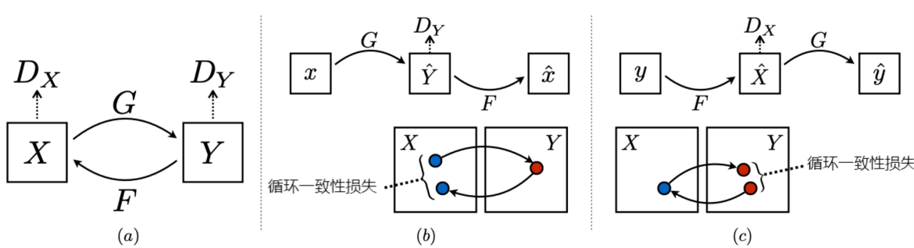


图3.3 CycleGAN映射思想

CycleGAN在模型结构上采用两个生成器和两个判别器组成，生成器结构采用残差网络（ResNet），判别器采用PatchGAN，基本结构如图3.4所示.。

为了实现G：和F：两个映射，CycleGAN提出循环一致性损失（Cycle-Consistency Loss）约束生成图像和放入生成器G和F可以尽可能和原始图像和相似。该损失表达式如下：



结合式，CycleGAN损失函数可以定义为：



其中为GAN损失。

1. Attn-CycleGAN模型

Attn-CycleGAN[37]是针对CycleGAN的改进。CycleGAN在训练过程中会出现生成器*G*比鉴别器*D*要弱得多，这会导致训练模式崩溃、梯度消失等现象导致降低生成图像。因此Attn-CycleGAN提出一个基于空间注意力机制（Spatial Attention，SA）的GAN框架，其中SA加入可以引导生成器生成更逼真的图像，具体Attn-CycleGAN框架如图3.5所示。

该框架在CycleGAN两个生成器*G*和*F*，两个判别器和基础上，加入注意力迁移模块*T*。*T*模块任务是生成注意力图*M*，*M*展示判别器重点关注的区域。最后将图*M*和输入空间结合，指导生成器生成图像，以改进I2IT任务。在Attn-CycleGAN中，采用循环一致损失，则两个生成器共享一个注意力迁移模块T，输出和，并分别和输入和输入结合。Attn-CycleGAN生成器G过程可用公式描述为：



其中，为第k次生成的第*i*张图像，为生成器参数。

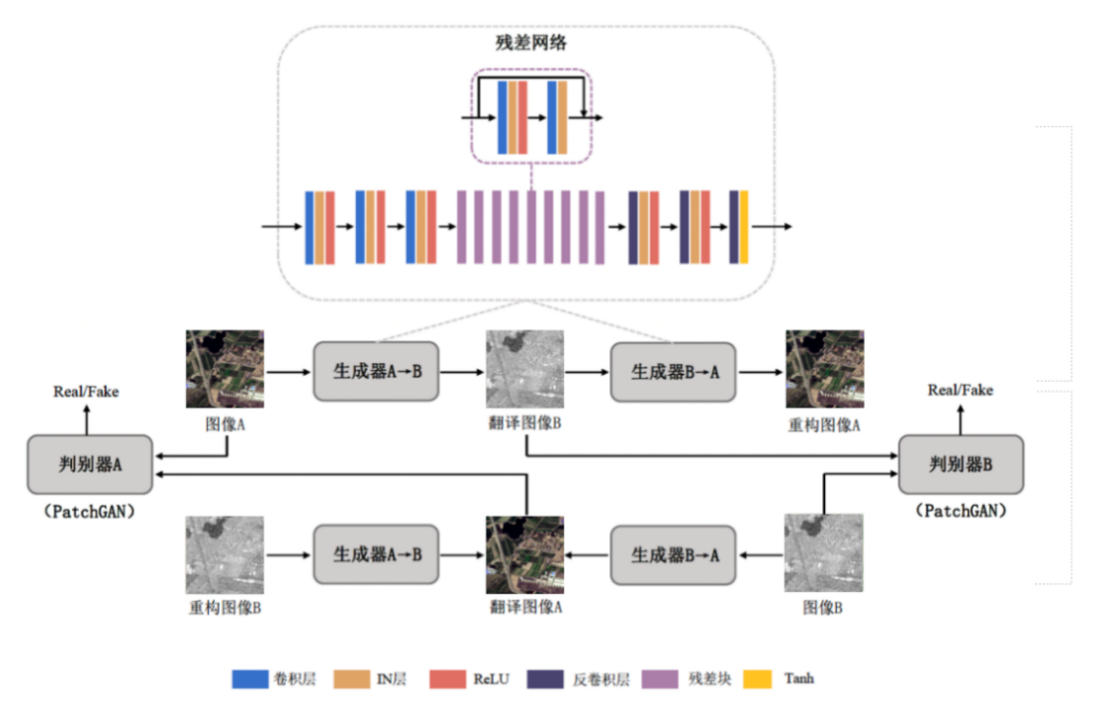


图3.4 CycleGAN结构图

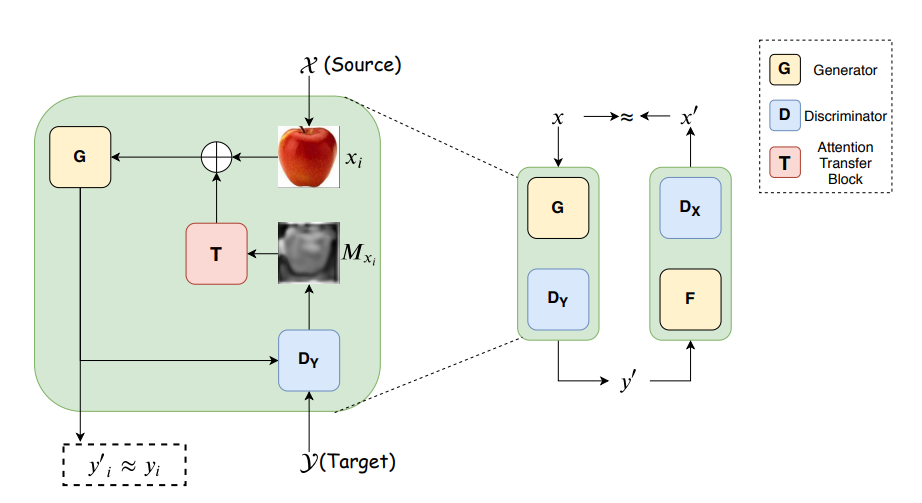


图3.5 Attn-CycleGAN框架图

Attn-CycleGAN的注意力图*M*是在判别器*D*计算过程中得到的，具体以可训练注意力模块为例（Trainable Attention Module，TAM）其判别器*D*结构如图3.6所示。其中A代表注意力计算分支，*T*代表判别器主干分支，对于*D*，将传统残差模块替换为简单的卷积模块。鉴别器最初几层卷积用于提取基本信息，随后将特征送入注意力计算分支计算注意力图和主干分支计算主干输出，最后将及注意力计算分支取平均得到注意力图*M*，将注意力图和主干输出软加权，然后通过一个卷积层计算判别器判断值，作为*D*输出。

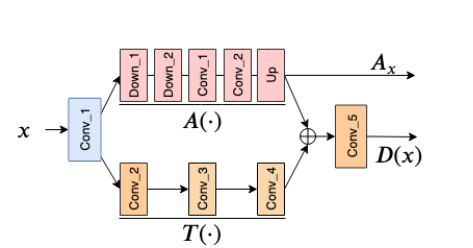


图3.6 Attn-CycleGAN判别器结构图

1. MUNIT模型

MUNIT（Multimodal UNsupervised Image-to-image Translation，MUNIT）[38]是一个非监督图像翻译，其思想是将图像翻译过程视为两个不同域联合分布迁移问题，如果知道目标域图像的边缘分布，便可利用参数学习方法推出目标图的图像联合分布，得到生成图像结果。MUNIT假设，一幅图像可被分为两种空间：内容空间和图像风格空间，其中源域和目标域内容空间是共享，但是风格空间为相互独立，内容空间相较于风格空间属于更高维空间，表现图像的结构和位置信息，而风格空间属于低维信息，表现图像的色彩与纹理信息。两种空间互相独立，而在图像翻译过程中仅需将不同域图像的内容空间信息和风格空间信息相互翻译，可以得到翻译后的图像。MUNIT迁移过程可表达为：



MUNIT网络通过图3.7展示其编码和解码的过程。在图(a)中，可以看到两个自编码器和解码器，它们各自对应不同风格图像的分布域。MUNIT网络翻译过程首先是将图像分解为内容空间和风格空间两个部分，然后再将两个空间特征结合起来进行输出，确保最终生成的图像能够还原成原始图像。此时，和CycleGAN循环一致性损失类似图像的重构损失被最小化，同时模型的编码和解码能力也得到最大化。图像风格迁移的过程涉及将图像和分解为各自的内容和风格特征，然后通过交换这些特征并融合解码，从而实现风格转换后的图像输出。最后，在编码解码过程之后，还有一个额外的编码步骤，这表示经过风格迁移的图像会被重新分解为内容和风格特征，这些特征需要与原始的内容和风格特征一起，形成重建损失，进一步增强编码器的编码能力。

|  |  |
| --- | --- |
| (a)域内重建 | (b)跨域翻译 |

图3.7 MUNIT编码解码过程

MUNIT模型网络架构如图3.8所示。左边为编码器，被分为编码风格和内容两个部分。这两部分在经历下采样处理后，分别进入不同的网络模块。风格编码器旨在捕捉表面层次的信息，因此它将下采样得到的特征图通过全局池化和全连接层处理，以提取出最终的风格特征。而内容编码器则专注于抽取更深层次的高维信息，通过多层残差网络处理下采样得到的特征图，以获取内容特征。在将风格和内容特征相互融合后，这些特征被送入解码器中。在解码前，风格特征首先通过MLP层，接着进入AdaIN网络，以获得控制风格的关键参数。随后，结合AdaIN层的残差模块，风格特征通过上采样和卷积层的处理被解码回原图像空间，生成经过风格转换的图像。此外，在鉴别器*D*中将生成的风格图像与原始图像一同输入鉴别器进行区分，并通过损失函数来指导整个网络的训练过程。在鉴别器部分，模型采用PatchGAN结构，其优势在于移除全连接层，转而使用全卷积层，从而增强对图像全局特征的控制能力，使得模型能够更精确地区分原始和风格转换后的图像，进一步提升了模型的性能。

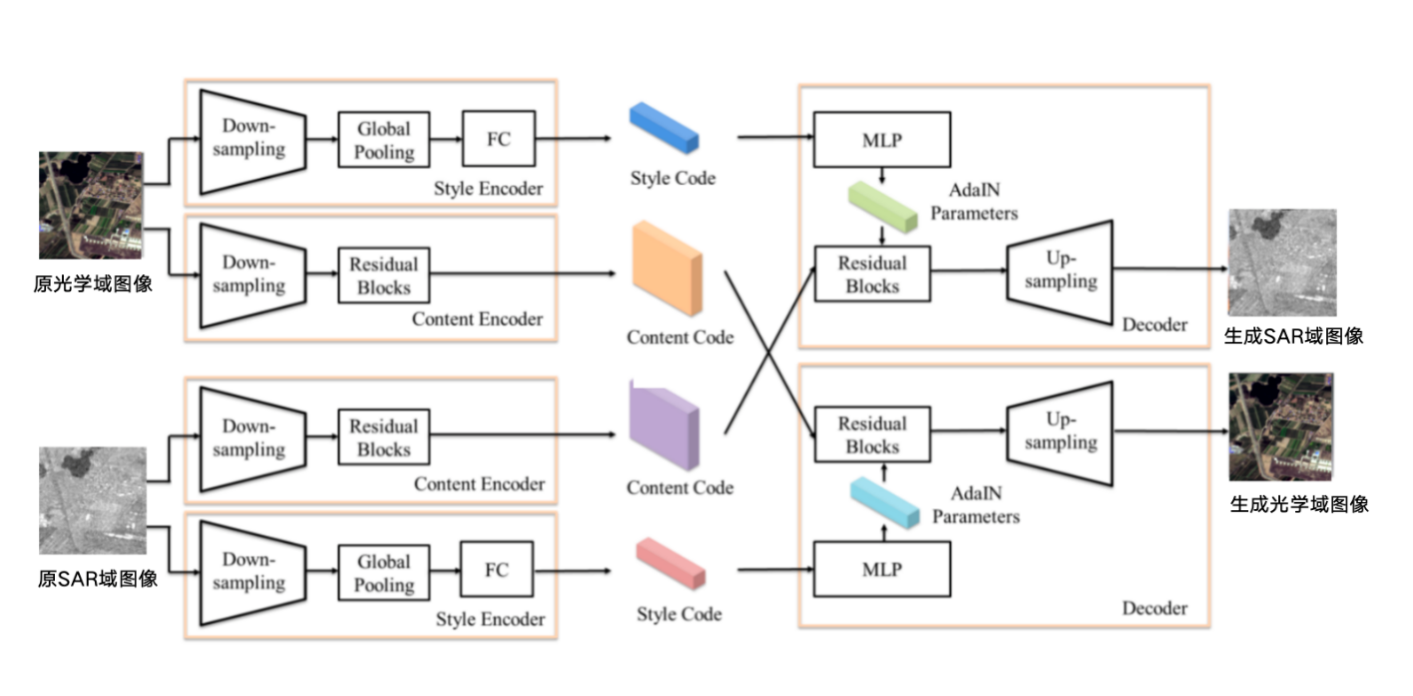


图3.8 MUNIT结构图

MUNIT模型的损失函数和CycleGAN相似，在传统生成对抗损失外，类似循环一致性损失，加入图像重构损失、风格重建损失、内容重建损失，上述三种损失表达式为：



结合，MUNIT模型的总体损失为：



## 实验数据集与评估指标

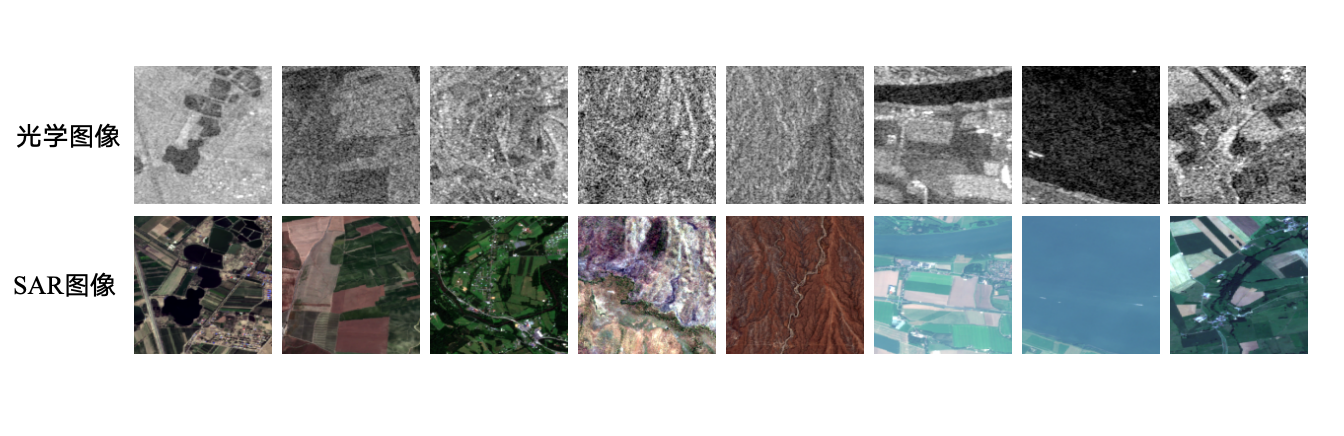
本节主要在对本章图像翻译模型所使用数据集来源及构成进行简要介绍的同时，为了进一步评价生成SAR图像质量，简要说明几种客观评价指标。这几种评价指标一方面对模型生成质量定量分析，另一方面对衡量生成图像和真实图像的差异性。

### 实验数据集

本章实验选取慕尼黑大学朱晓香教授团队发布的公开遥感数据集SEN1-2[39]。该数据集图像来源欧空局Sentinel卫星，在时间上囊括春夏秋冬每一个季节，在空间上利用随机ROI采样，在全球范围内选取感兴趣区域（Region Of Interest，ROI）。同时由于城市区域相比农村等其他区域拥有更复杂的地物信息，因此人为选取一些城市区域加入数据集中。最终SEN1-2数据集包括282384对光学和SAR配准图像，且每对图像均在像素层面完成严格对准与配齐。为了进一步提升数据集质量，该数据集还人为剔除含有云层遮挡图像和颜色扭曲图像。

SEN1-2中的SAR图像部分采用Sentinel-1卫星图像。Sentinel-1由2014年发射的欧空局哥白尼计划观测卫星，由两颗卫星组成，其携带C波段的SAR天线，可以连续提供全天时、全天候的图像。本数据集采用干涉宽幅宽带（Interferometric，Wide swath，IW）模型作为地物信息主要采集方式，该方式下所得到的地距多视影像（Ground Range Detected，GRD）产品可最高拥有距离分辨率为20m、方位分辨率为5m。同时为了简化操作，极化方式为仅为垂直极化（VV）同时对数据进行辐射校正。SEN1-2中的光学图像部分采用Sentinel-2卫星图像，其包含两颗在同一极地轨道卫星，相位角彼此相差180°。由于光学卫星仅包含红色、绿色、蓝色通道，因此只选取卫星2-4波段生成真实的RGB图像。SEN1-2所有图像，均未进行相干斑滤波处理，将两种图像分辨率统一裁剪为256×256。

对于本章实验，选取5515对SEN1-2 Summer图像用于实验，将训练集和测试集比例设置为8：2，该数据集共包含城市居民、山区、荒漠、农田等种类，具体数据集展示如图。SAR图像和光学图像大小均为256×256，SAR图像为灰度图像，光学图像为3通道RGB图像。实验数据集部分展示如图3.9.

图3.9 Sen1-2数据集展示

### 实验评价指标

图像生成评价指标是评价图像生成质量的重要环节，通常来说，图像生成评价一般从主观和客观两种方式评价。主观方式主要依据专家经验对生成图像和真实图像进行对比，依据人眼主观感受评价图像生成质量。客观方法则按照图像本身特征进行定量评估。主管方法更贴切人类视觉，和SAR图像质量较一致，但是容易受到环境影响且评价耗费时间；客观方法虽然更省时省力、结果更稳定，但结果会和人眼感官不能很好匹配[40]。

本章实验为SAR图像生成，由于SAR图像和光学图像有些许区别，因此从传统图像质量评价和基于SAR图像成像机理的质量评价。

介绍评价方法前，首先给出图像的均值、方差的计算公式：



其中，为图像的像素。对于SAR图像，图像均值可以翻译SAR所包含目标对象的平均后向散射特性而方差表示SAR图像的不均匀性，当SAR图像包含的目标越多、地形越复杂其方差值会越大。

1. 均方误差

均方误差（Mean Square Error，MSE）是图像评价指标中全参考图像评价最常见的评价指标。改评价指标计算真实图像和生成图像像素值之间的差异，以统计学视角评价生成图像质量。MSE越小，代表真实图像和生成图像间相似度越高，图像生成的质量越好。MSE计算公式如下：



其中，代表生成图像的像素点，代表真实图像对应像素点，*M*和*N*分别代表图像的长和高，*MN*乘积即为图像相似总数。

1. 峰值信噪比

峰值信噪比（Peak Signal-to-Noise Ratio, PSNR）是用于衡量信号指标，值峰值信号和其背景噪声的能力比。应用于图像领域，可以衡量真实图像和生成图像之间的失真程度。PSNR计算公式如下：



其中，L代表生成图像中每个像素点理论上最大值，比如8位像素图像中，L=255。PSNR不同于MSE，越高的PSNR表示图像生成的质量越好。

1. 结构相似性

结构相似性（Structural Similarity Index Metric, SSIM）是通过亮度、对比度和结构三个方面充分衡量真实图像和生成图方面像的相似程度。和MSE和PSNR相比，SSIM以人眼感官所能提取的图像较低层次信息为原理，其判断结果更符合人眼对图像质量直觉评估。SSIM计算公式如下：



其中，分别为图像的亮度、对比度和结构，其对SSIM自评分影响有三个影响因子、和决定。分别为真实图像和生成图像的均值，其计算公式，为真实图像和生成图像的方差，为真实图像和生成图像的协方差。通常，计算SSIM会将，，则公式可以改写为：



1. 等效视数

由于SAR图像具有特殊的散射特性，因此选用等效视数作为评价SAR图像生成和真实SAR图像的指标。等效视数反应图像中斑点噪声的强弱强度，可以很好体现SAR图像的噪声强弱的对比。等效视数越小，代表图像的乘性噪声越弱。等效视数的计算公式如下：



其中和同SSIM一样为图像的均值和标准差。

1. 辐射分辨率

辐射分辨率表示图像中像素点的灰度值的区分能力。辐射分辨率和等效视数相关，和PSNR类似，但用来描述图像的辐射质量。同一SAR传感器所照射目标的辐射分辨率为恒定值，因此当生成图像与原始图像的辐射分辨率约接近，代表生成质量越高。辐射分辨率的计算公式如下：



## 实验结果及分析

本节将对所使用基于图像翻译的SAR图像生成实验结果进行简要分析，首先将阐述实验中所用平台及平台具体参数设置，其次，针对上述所介绍的图像模型进行实验对比和结果分析，视觉评价各个模型的生成图像的效果，并利用评价指标衡量模型的生成能力。

### 实验参数设置

本章实验系统平台为Ubantu16.04，使用单块NVIDIA RTX3060显卡，实验模型基于PyTorch深度学习框架搭建，Python版本为3.9。所有模型输入SAR图像大小为256×256，输入光学图像大小为256×256×3，而模型生成的SAR图像大小为256×256。对于Pix2Pix模型，训练过程采用200个Epoch，其余模型采用100个Epoch，Batch Size设置为1。优化器选用Adam，其两个超参数指数衰减因子和设置为0.5和0.999。模型初始学习率设置为0.0002，对于Pix2Pix前100个Epoch和其余模型的前50个Epoch采用固定学习率，剩余Epoch利用线性学习率衰减策略，将学习率逐步衰减至0。

### 实验结果分析

本节将对上节介绍的5种图像翻译模型应用结果及其在SEN1-2数据集上的图像生成结果进行对比并利用图像评价指标进行分析。图像生成结果可视化示例如图3.10所示。

通过图3.10红框中选中的重点区域进行对比和分析，Pix2Pix和BicycleGAN这两种配对生成模型在视觉效果上优于CycleGAN、Attn-CycleGAN和MUNIT等非配对生成模型。这是因为配对生成模型在SAR图像生成任务中通过像素级配对可以更精确地提取信息，从而产生轮廓清晰、细节丰富、纹理精细的图像。对比重点区域在生成结果，生成图像在水体、农田和河流等相对平滑表面的图像方面表现更出色，但在更复杂的建筑细节、更清晰的农田轮廓和更精细的土壤纹理生成方面表现不佳。

尽管Pix2Pix在视觉效果上表现最佳，但在目标密集的区域，如居民区，其生成的图像可能会变得模糊，并且在均匀的区域可能产生网格状噪声。而BicycleGAN会不充分学习图像的纹理和边界信息，导致生成的图像在这些方面出现模糊，或者产生不正确的样式。这可能是由于BicycleGAN在追求生成多样性的过程中，牺牲样式的准确性。对于非配对模型，其生成图像在视觉上未出现明显的相干斑噪声。CycleGAN的生成图像更多是像光学图像转化为灰度图，并没有学习到不同光学图像的内容、纹理等信息，相比真实SAR图像相差更远，与之相比，虽然Attn-CycleGAN加入注意力机制可以生成更详细特征，但是其生成图像出现明显颜色失真现象，图像分辨率更低。MUNIT模型在生成中感官最差，图像对比度、亮度较低，对于目标轮廓及物体信息均较为模糊，同时出现严重残影和颜色失真。

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
| 光学图像 | 参考图像 | Pix2Pix | BicyleGAN | CycleGAN | ASGIT | MUNIT |

图3.10所有模型在SEN1-2生成图像示例

由于视觉观察不能全面评价生成图像质量，利用上节介绍到的客观评价指标对5种模型生成的SAR图像进行评价。实验选用1103张图像作为测试，并对结果取平均，评价指标结果如表3.1所示。

表3.1 所有模型生成图像评价指标表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 数据来源 | SSIM | MSE | PSNR | ENL |  |
| Ground Truth | —— | —— | —— | 25.66 | 7.347 |
| Pix2Pix | **0.1763** | 0.0198 | 17.17 | **25.24** | **7.491** |
| BicycleGAN | 0.1606 | **0.0192** | **17.38** | 31.77 | 7.891 |
| CycleGAN | 0.0864 | 0.0721 | 12.68 | 32.06 | 6.871 |
| Attn-cycleGAN | 0.1281 | 0.0534 | 13.78 | 46.15 | 7.795 |
| MUNIT | 0.0741 | 0.0730 | 12.11 | 16.49 | 8.732 |

对比表3.1中结果也可说明配对模型相比非配对模型具有更出色的性能。Pix2Pix模型在SSIM结果更低，ENL和上更接近真实SAR图像。BicycleGAN在MSE和PSNR两个指标表现更出色。这可能由于BicycleGAN在生成器结构上采用cVAE-GAN相比U-Net生成图像更清晰，而cLR-GAN生成的随机噪声没有约束和真实样本一致，导致生成图像和真实图像类别出现差异。

配对方法由于在像素点级别进行翻译图像，对输入光学图像和输出SAR图像之间建立像素级的一对一映射。这种配对生成使得生成的SAR图像更加准确地反映输入图像的结构和细节，可保持光学域和SAR域数据之间的一致性和相关性，在对地物特征及纹理信息的学习至关重要。

## 本章总结

本章利用图像翻译模型实现借助光学图像这类非同源域信息来辅助SAR图像生成实验。首先介绍5种主流图像翻译模型结构，接着介绍传统图像质量评价指标SAR图像成像机理的质量评价指标，同时介绍本章使用数据集，最后对5种模型图像生成结果可视化并做评估。结果显示，由于配对方法可以建立光学图像域和SAR图像域的一对一映射，导致基于配对方法生成模型Pix2Pix和BicycleGAN的生成图像质量比非配对方法的CycleGAN、Attn-cycleGAN和MUNIT的生成图像质量有显著提升。但所有模型在轮廓、具体细节信息的生成仍有待改善。

# 基于时间偏移的SAR图像生成方法

上一章基于图像翻译的SAR图像生成方法实验结果表明传统图像翻译模型虽然可以很好学习光学图像和SAR图像的特征，同时生成的SAR图像在观感上具有和真实图像由较高的相似性，但是在具体建筑物细节、颜色误差、边缘轮廓生成等方面仍与真实SAR图像有较大差异，且生成图像的评价指标表现不理想。因此，本章将采用TSGAN模型，一个基于Pix2Pix的双分支CGAN模型，考虑输入SAR图像应如何根据光学数据随时间的变化进行修改，以便生成所需时间的新SAR图像。本章首先介绍TSGAN模型框架、模型所用的注意力机制和损失函数，接着介绍本章所用实验数据集和改进的实验评价指标，最后对模型生成结果进行评估。

## 时间偏移对抗生成网络

根据上章Pix2Pix在基于图像翻译方法取得较为优异的表现，但图像会存在模糊和虚构问题，因此本章使用时间偏移对抗生成网络（Temporal Shifting GAN，TSGAN）来改善上述问题。TSGAN是基于Pix2Pix架构的图像生成方法。本节首先介绍TSGAN模型的基本架构，接着具体阐述模型所使用注意力机制，最后对解释模型损失函数。

### 模型结构

为了解决Opt2SAR任务中出现图像失真、伪影等现象，使用一个基于Pix2Pix架构的双分支CGAN模型TSGAN [41]。Reyes等人[42]曾提出，由于SAR观测目标表面地形与散射系数的不同和SAR仪器间参数不同，导致将光学图像转为SAR图像并不是简单的一对一关系，因此出现生成的SAR图像与真实SAR图像有较大差异。TSGAN模型思想是利用时刻为的Sentinel-1图像（）以及时刻的Sentinel-2的图像（）数据作为输入，生成时刻的Sentinel-1的图像（）。因此将“一幅光学图像翻译为SAR图像”的问题转化为“原始SAR图像将如何根据光学图像的改变而变化”。

TSGAN模型结构和Pix2Pix模型类似，生成器结构如图4.1所示。生成器采用两个编码器分别对输入的光学和SAR图像进行编码，随后在瓶颈层（Bottleneck）将二者编码融合。将融合后的编码通过解码器上采样，和Pix2Pix使用的U-Net编码器类似，在两个编码器均加入跃级链接（Skip-Connection）到解码器。由于在两个时间变化较小的区域生成的SAR图像会存在模糊现象，因此在第一个跃层链接后的上采用层替换为利用1×1的卷积核消除该现象。由于光学图像相比SAR图像需要更大的感受野，因此在编码器中，处理光学图像采用5×5的卷积核，处理SAR图像采用3×3的卷积核。

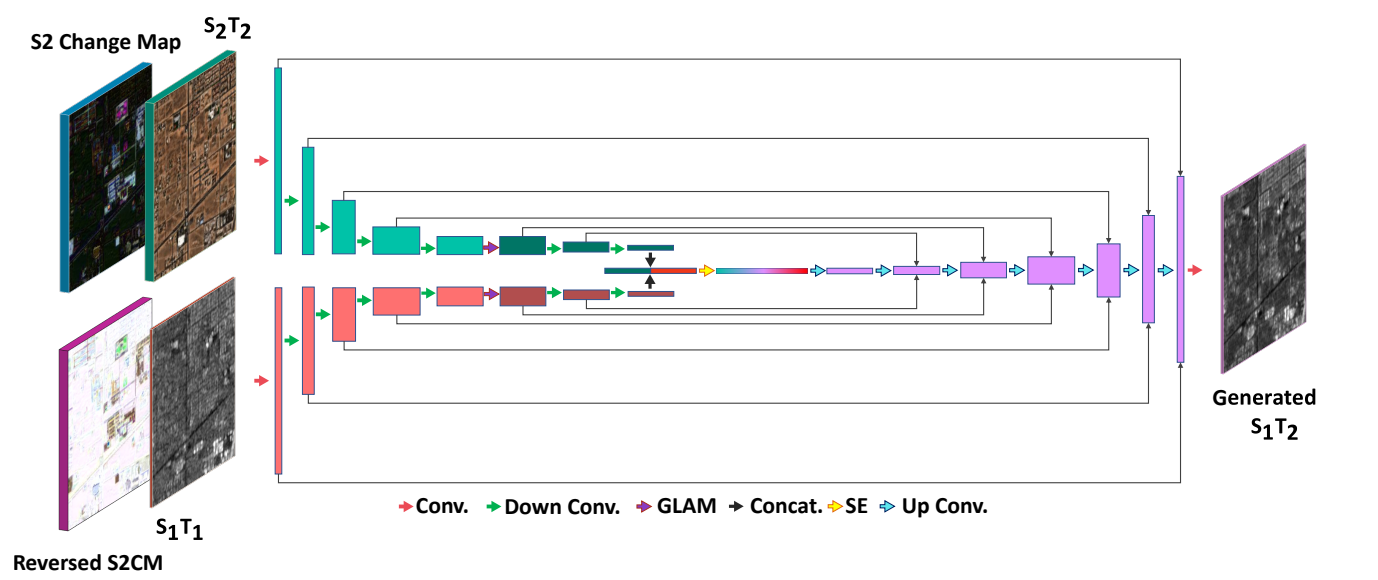


图4.1 TSGAN生成器结构图

对于TSGAN的判别器，其结构图如图4.2所示。判别器采用和Pix2Pix的PatchGAN并将结构修改为双解码器，每个解码器输入生成的，经过下采样后生成30×30的Patch判别生成图像。

### 模型注意力机制

随着Transformer[43]和ViT（Vision Transforerm）[44]在深度学习应用，注意力机制在更有效提取特征，提高模型能力方面有着显著效果。目前由许多学者研究基于注意力机制的GAN模型，因此在TSGAN，引入两种不同的注意力机制挤压-激励激活函数（Squeeze-and Excitation，SE）。和全局-局部注意力机制（Global-Local Attention Module，GLAM）。

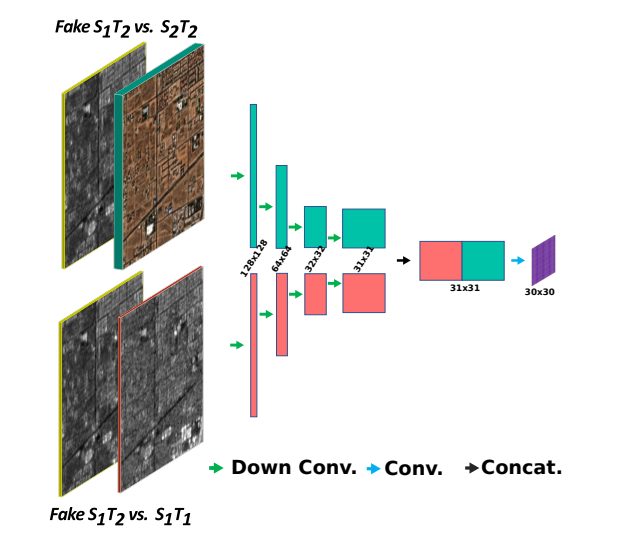


图4.2 TSGAN判别器结构图

1. SE注意力机制

挤压-激励激活函数(SE)是在SENet中[45]提出，其计算流程如图4.3所示。SE注意力机制的思想为建立图像间不同通道的关系，通过全局信息池化和特定通道的建模来学习每个通道的重要性，之后根据这些重要性对特征图进行重新加权，从而增强模型对于任务相关特征的关注程度。

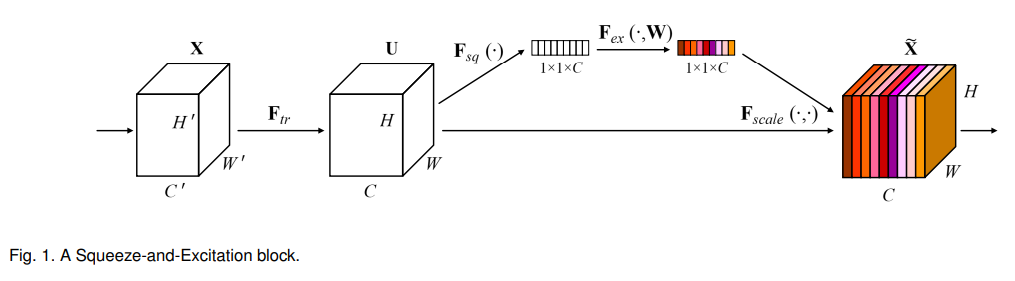


图4.3 SE注意力机制计算流程图

SE可分成三步，第一步为压缩（Squeeze），利用全局平均池化（Global，Average Pooling，GAP）将特征图（）变为一个的特征向量，也就是求特征图中每个通道特征值的平均值，该步骤可用公式表达为：



其中，为特征图第c个通道的特征，为通道*c*特征取平均。

第二步为激励（Excitation），将输入到一个小型的全连接网络中，该网络由两个全连接层组成。第一个全连接层的输出大小为，r为压缩比例，采用激活函数为ReLU，第二个全连接层的输出大小恢复为激活函数为Sigmoid。则该过程可用公式表达为：



其中为通道从的重要性得分，、和、为第一个和第二个全连接层的权重和偏差。

第三步为加权（Scale），将前面的得到的注意力重要性得分重新加权到每个通道的特征上，得到最终输出，则该步骤可用公式表达为：



SE注意力机制可以更加关注必要特征，增强特征的表示能力，同时减少非必要特征信息，增强模型的性能。

1. GLAM注意力机制

GLAM注意力机制[46]是结合全局和局部的注意力机制，以及空间和通道级别的特征，GLAM注意力机制的计算示意图如图4.4所示。下面，将讲述这四个注意力机制的计算流程。



图4.4 GLAM注意力机制计算图

* 1. 局部注意力机制

局部注意力机制（Local attention），其注意力内容是利用通道注意力机制和空间注意力机制分别提取一个一维度特征图和一个二维特征图。

局部通道注意力机制（Local channel attention）[50]重点捕捉局部通道的信息，将特征图提取成一个一维通道特征图。其计算过程为：首先针对一个的特征图，将其经过一个全局平均池化（GAP），压缩为的特征向量。然后利用一个尺寸为*k*的一维卷积沿着通维度*c*对捕捉器空间注意力，*k*的大小取决于跨通道相互作用程度。最后经过一个sigmoid激活函数，得到大小为的局部通道注意力图。局部通道注意力机制计算流程如图4.5所示。

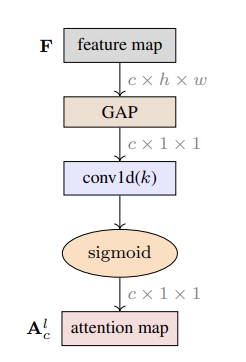


图4.5 局部通道注意力计算流程图

局部空间注意力机制（Local spatial attention）[49]重点关注特征在不同尺度下局部空间信息，将特征图提取为一个二维特征图。其计算过程为，对一个的特征图，首先利用一个1×1的卷积将初始特征图的通道数减少至，从而获得一个新特征图，大小为，然后利用三个大小分别为3×3、5×5和7×7，的卷积核提取不同尺度的空间语义信息。该过程在实现过程中，利用3×3的膨胀卷积（Dilated Convolution）而膨胀率（Dilated Rate）分别设为1，2，3。最后将卷积后的结果和拼接为一个的新特征图。最后，将经过一1×1的卷积将通道数减少为1得到局部空间注意力图。局部空间注意力机制计算流程图4.6所示。

基于上述过程，结合卷积块注意力机制（Convolutional Block Attention Module，CBAM）的残差链接，局部注意力特征图计算过程为，首先利用局部通道注意力图对初始特征图进行在通道维度进行加权，接着利用局部空间注意力图在空间维度进行加权，两个加权过程后均加入残差链接最后得到的局部注意力部特征图，具体其计算用公式可表达为：



其中，表示哈达玛积（Hadamard product），即两个同维度矩阵间对应元素相乘。

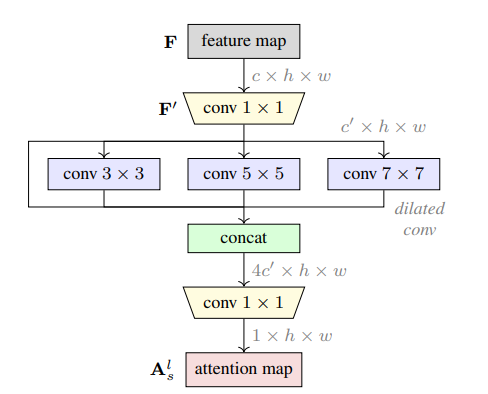


图4.6 局部空间注意力机制计算流程图

* 1. 全局注意力机制

全局注意力机制（Global attention）其注意力内容是通道注意力机制和空间注意力机制，提取两个捕获全局两两通道间和两两空间相互作用的二维矩阵。

全局通道注意力机制（Global channel attention）[50]，利用通道间的自注意力实现重点关注特征图通道的全局信息。其计算过程和局部通道注意力机制类似为，对一个的特征图，首先经过GAP模块压缩为的特征向量，之后将分别经过一个尺寸为*k*的一维卷积和sigmoid激活函数，得到尺寸为的查询向量和键向量，将整型为作为值向量，然后计算全局通道注意力得分，计算公式如下：



最后，全局通道注意力得分和值向量做矩阵乘法，并整型为得到全局通道注意力特征图，其计算流程如图4.7所示。

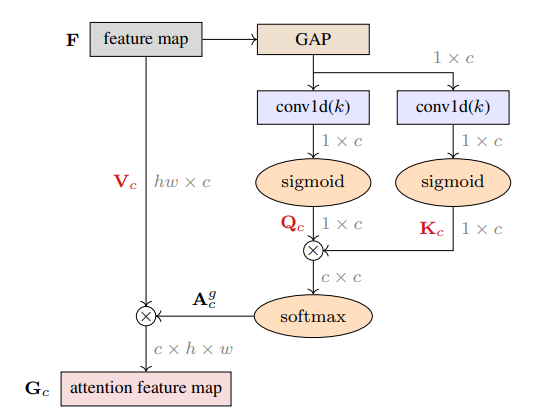


图4.7 全局通道注意力计算流程图

全局空间注意力机制（Global spatial attention），利用非局部滤波思想实现特征图间的自注意力机制。其计算过程为对一个的特征图首先利用3个1×1的卷积将通道数减少为，然后将空间维度进行展平为，得到三个大小均为的向量，分别为查询向量，键向量和值向量。和全局通道注意力机制计算方法一样，计算其注意力得分，计算公式如下：



最后，将和相乘，并在空间维展开成，并利用一个1×1卷积将通道数增加至*c*，得到尺寸为的全局空间注意力特征图，其计算流程如图4.8所示。

根据上述理论，全局注意力特征图的计算过程为，首先利用全局通道注意力特征图对初始特征图进行加权得到，接着利用全局空间注意力特征图对进行加权，并加入残差连接，得到全局注意力特征图，具体可用公式表达为：



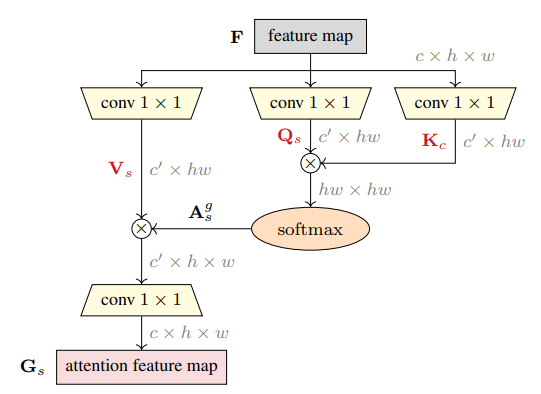


图4.8 全局空间注意力机制计算流程图

* 1. 全局-局部注意力机制

由上文可得到局部注意力部特征图和全局注意力特征图，最后利用softmax函数得到一组可学习权重分别加权到,,得到大小为全局局部注意力特征图，计算公式如下：



### 模型损失函数

模型的损失函数在模型训练过程中起着至关重要的作用，通过度量模型生成图像和真实图像之间的差异优化模型。TSGAN模型中，输入的和如果时间相差非常近，则会出现该目标区域在该时间段内没有发生重大改变，导致已更改区域和未明显更改区域之间严重不平衡。这种不平衡类似二分类之间的类不平衡会使模型过分关注未明显变化区域，导致模型在生成图像时仅将复制粘贴，而不是从的光学数据中学习目标地区潜在变化。因此，为了解决上述问题，TSGAN提出全新损失函数变化加权L1损失（Change Weighted L1，CWL1）和反转变化加权L1损失（Reverse Change Weighted L1，RCWL1），其思想是利用成本敏感学习（cost-sensitive learning ）[50]，为每组类别分配不同的惩罚权重，通过将更大的权重值赋给代表性不足类别，增加在其学习过程的重要性，提高模型关注变化小的区域。加权L1损失计算公式如下：



其中，表示生成照片的像素值，表示真实照片的像素值， 为每个像素的权重。TSGAN采用两种不同的权重图，变化权重图CWM和反转变化权重图RCWM。CWM计算方法为使用大津算法从和变化概率图获取二值变化图。RCWM计算方式如下所示：



最后，将CWM和RCWM得到的权重带入公式，得到RCWML1和CWL1，则TSGAN的总体损失可用公式表达为：



其中，为模型训练超参数，较大的则模型将重点考虑变化的区域，而较低的则模型会出现无法充分利用光学图像信息，导致“直接复制”情况发生。

## 实验数据集与评价指标

本节将对根据时间偏移任务所使用的双时态SAR/光学数据集进行介绍，接着针对任务所采用的加入变化权重后所改进的评价指标WSSIM和WPSNR进行说明。

### 实验数据集

TSGAN模型使用的数据集也是由Sentinel-1数据和Sentinel-2数据成对构成。该数据集包含全球范围46个城市区域，30个城市区域为训练集，16个城市区域为测试集。选取是聚集时间为2019年，时间为2021年。结合第二章所指出的SAR图像特性，SAR图像在海拔差较大区域会出现较大误差，因此将数据集的ROI限定在城市较为平坦区域。由于SAR图像在同一ROI和不同入射角下，会出现不同区域，因此在选择Sentinel-1图像上，对于一个ROI区域由多个轨道可以选择时，则选用平均入射角最高的轨道，同时确保和时刻的选取轨道数相同。通过上述措施，增加数据集的多样性，以提高模型的泛化能力和鲁棒性。最后，为了使图像适应模型输入，将数据集图像大小分块为256×256，并用自适应方法找到最佳水平和垂直部分，数据集部分展示如图4.9：



图4.9本章实验所用数据集部分展示

### 实验评估指标

为了更好评估模型生成效果，对第三章介绍MSE、PSNR、SSIM三个评价指标进行改进，结合变化权重图的权重，使评价指标与TSGAN相匹配。结合公式和，假设M=N，则加权后的WMSE，WPSNR计算表达式为：





其中，为变化权重图每个像素的权重。

对于SSIM评价指标，传统SSIM计算方法很难关注到全局信息，因此TSGAN将使用WSSIM评价指标[51]。WSSIM引入权重计算矩阵使评价时更多关注到全局的相似性信息。在计算中首先里利用高斯分布生成一个11×11的卷积核，接着利用该卷积核与目标卷积，实现利用权重矩阵对目标加权。引入后，均值，标准差和协方差计算公式如下：







其中，为第*i*个像素的权重矩阵，和分别为生成图像和真实图像第*i*个像素。

将权重矩阵 引入变化权重图CWM，将权重矩阵和CWM各元素相乘，便得到加权后的CWM记作，最后结合公式可得WSSIM 计算公式如下：



其中，分别为生成图像和真实图像，，为CWM中第*i*个像素。

WPSNR和WSSIM可解决传统评价指标方法对模型局部信息的关注度小，增加变化权重图对模型局部信息的影响，可以更好评估时间偏移SAR图像生成任务的结果。

## 实验结果及分析

本节主要对TSGAN模型在时间偏移任务的图像生成结果进行评估。首先通过视觉观察生成SAR图像质量和跟真实SAR图像差异，然后对模型在评价指标上和使用不同注意力机制结果进行评价。

### 实验参数设置

本章实验硬件参数上和上一章相同，在超参数设置上保持一致，BatchSize设置为4，训练10个Epoch，优化器为Adam，初始学习率为设置为0.0001同时采用线性学习率衰减策略将学习率逐步衰减至0。损失函数设置为5。

### 实验结果分析

本节将介绍TSGAN模型应用在SAR图像生成结果，并对不同注意力机制对模型效果影响进行评估，实验结果将利用居民区、沙漠土壤、农田等复杂结构的图像进行可视化比较与分析。如图4.9展示TSGAN模型生成结果、真实图像和此时刻所对应的光学图像。为了更好对比模型引入变化权重图后对变化明显区域和变化不明显区域的生成，同时展示和的变化图衡量TSGAN模型对两个区域的生成结果。

根据图4.10分析表明，TSGAN模型可以较好学习时刻SAR图像的轮廓、纹理等信息，并利用光学图像和时刻SAR图像之间变化差异生成经过时间偏移变化后时刻的SAR图像。分析居民区部分生成结果表明，TSGAN模型对较大且密度小的建筑物可以很好提取目标特性，生成的SAR图像目标区域亮度较准确、轮廓较为清晰。但当建筑物分布密集且形状较小，无法有效学习目标的散射特性，导致生成图像在目标区域的亮度较低。

对于时间偏移后有明显变化的区域，由变化图和变化图对比生成变化图，TSGAN模型对目标表面不平整区域有较好的变化生成效果，但是如果目标是较平滑区域未建造的土地和变化前后色彩相差较小区域，则变化更依赖光学图像的变化，会出现较差的生成效果。

为了进一步衡量TSGAN模型性能和两种不同注意力机制对TSGAN的作用，利用上一章所介绍的WPSNR和WSSIM评价指标，对生成SAR图像整体进行评估。同时为了关注时间平移在未改变区域和改变区域的效果，对上述两块区域分别使用评价指标评价，命名为UC-WPSNR、UC-WSSIM和C-WPSNR、C-WSSIM。评价指标结果如表4.1所示。

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  | 生成 | 变化图 | 变化图 | 生成变化图 |

图4.10 TSGAN生成结果可视化

由表4.1数据分析得出，相比Pix2Pix模型，TSGAN模型在各指标上有均有显著提升，证明通过时间偏移生成方法要比传统翻译方法有更出色生成效果。同时UC和C两类指标提升可以解释在不明显变化区域，模型可以通过输入前一时刻SAR图像重新生成该区域，而不是仅依靠光学图像；在变化区域上相比较通过光学图像和变化图差异直接生成该区域，利用前一时刻的SAR图像针对变化图权重直接修改可以获得更出色生成效果。

表4.1 Pix2Pix及不同注意力模型的评价指标结果表

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 模型 | SE | GLAM | WSSIM | WPSNR | UC-WPSNR | UC-WSSIM | C-WPSNR | C-WSSIM |
| V1.1 | × | × | 0.588 | 22.21 | 22.76 | **0.580** | 16.07 | 0.358 |
| V1.2 | 🗸 | × | 0.572 | 22.48 | 23.02 | 0.549 | 16.18 | 0.328 |
| V1.3 | 🞨 | 🗸 | **0.594** | 21.40 | 22.95 | 0.571 | 16.11 | 0.353 |
| V1.4 | 🗸 | 🗸 | 0.581 | **22.65** | **23.15** | 0.561 | **16.66** | **0.360** |
| Pix2Pix | —— | —— | 0.308 | 17.49 | 17.67 | 0.312 | 14.84 | 0.266 |

对比本文所介绍的两种注意力机制，当模型不含两种注意力机制，在未明显变化区域的SSIM值更高，这可能是因为模型没有充分学习到输入SAR图像的信息，导致“直接复制”现象的出现。对比SE和GLAM两种注意力机制，两种注意力机制同时使用在变化区域整体指标且图像的PSNR较优秀，证明注意力机制可以有效地关注并提取输入光学图像和变化图的信息，对有新增建筑物等明显变化区域生成效果较好，生成的SAR图像质量较高。对于使用单个注意力，GLAM注意力机制在提取特征能力上要优于SE注意力机制，这可能因为GLAM兼顾全局和局部注意力，提高模型对变化图和输入SAR图像提取信息的能力。但GLAM的全局注意力模块在训练中偶尔会关注输入映射中的随机的单个点，这会导致模型性能出现下降，模型可靠性降低[41]。

## 本章小结

本章利用基于时间偏移的SAR图像生成模型TSGAN进行SAR图像生成实验。本章先是介绍TSGAN模型结构、使用的注意力机制原理及针对变化图的损失函数CWL1，接着对实验所用数据集及针对时间偏移任务改进的评价指标讲解，最后对不同注意力机制TSGAN可视化及评价指标分析。

# 总结与展望

## 全文总结

合成孔径雷达（SAR）能够实现全天时，全天候的成像能力，在民用和军事领域上具有显著的应用价值。然而，随着深度学习的不断发展，基于深度学习的SAR图像解译面临着图像数据集数量匮乏和质量较低的双重挑战，因此，SAR图像生成对于推动SAR图像解译具有重大意义。本文参考相关学者工作，复现深度学习算法并对比，实现基于深度学习的SAR图像生成。首先，本文全面分析了国内外关于图像生成、基于传统方法和深度学习方法的SAR图像生成的研究现状，并简要说明了SAR图像生成的意义；接着，本文介绍了SAR成像机理并数学推导了SAR的距离分辨率和方位分辨率，同时分析了SAR图像的散射特性、几何特性及统计特性；随后简要阐述了深度生成模型原理并介绍几种对抗生成网络变种；最后针对上述研究内容，采用非同源数据转换生成SAR图像思想，将光学目标图像转化为SAR图像，利用基于图像翻译方法和基于时间偏移方法实现。

针对传统深度学习方法仅使用同源信息和提取能力不足的问题，使用基于图像翻译的SAR图像生成方法，实现光学信息辅助生成SAR图像。经过对配对模型和非配对模型两种方法比较和分析，得出如Pix2Pix模型等配对生成方法能够在像素级对图像进行匹配和翻译，生成的SAR图像在视觉上更接近真实SAR图像，并且相关评价指标上表现更出色，相较于CycleGAN模型等非配对生成方法，拥有更优异的表现。

针对图像翻译模型生成SAR图像方法会出现复杂目标的生成图像失真和虚构的现象，使用基于时间偏移的SAR图像生成方法，将SAR图像生成任务转化为依据光学数据变化进而生成新的SAR图像数据。该方法利用基于Pix2Pix模型所改进的时间偏移对抗生成网络且结合不同注意力机制。通过对比分析，TSGAN模型比图像翻译方法Pix2Pix模型在视觉效果和评价指标上有明显提高，且可以有效解决模型“直接复制”现象发烧。同时，SE和GLAM等注意力机制的使用，可以提高模型对输入图像和变化图的信息提取能力，使模型生成的图像在视觉上更接近真实图像。

## 展望

本文重点研究了基于深度学习的SAR图像生成，但研究过程仍存在一些不足和值得探讨的问题，有待进一步研究，主要包括以下几个方面：

（1）图像生成结果上，尽管基于图像翻译的SAR图像生成可以视觉上与真实SAR图像相近的生成SAR图像，但模型生成在评价指标数值上仍偏低且模型在生成上针对一些图像纹理、结构、信息不复杂的简单图像，当面对复杂场景下如居民区存在大量建筑物、目标地区散射体信息复杂等难以实现精确生成，因而导致该区域存在模糊、边界不清晰等现象。这一现象原理可能由于数据集分辨率不足，模型对复杂信息的提取能力不足。所以后续利用更高分辨率数据集，改进模型对复杂区域信息的学习能力值得进一步研究。

（2）数据集选择上，仅选择农田、城市等类别数据，未能涵盖更多数据，导致模型泛化能力和鲁棒性较低且生成结果仅做定量的指标评价，未能根据具体SAR图像解译任务进行实验。因此，后续选择拥有类别更多样的数据集以提高模型的泛化能力且针对生成数据集在图像解译方面进行应用以进一步评价生成图像质量值得进一步深入探讨。

（3）在模型选择上，主要采用经典的生成模型，然而，对于一些当前流行的图像生成模型，例如扩散模型，可以应用于SAR图像的生成上。此外，由于硬件的限制，未能使用更复杂的模型或在更大规模的数据集上进行实验。因此探索如何将不同目标的特有属性和散射特性结合起来，同时更大规模的数据集上应用更复杂的模型以期望能够生成高质量的SAR图像可以开展更多相应的研究。

致 谢

时光飞逝，转眼间我的本科学习生活即将结束。回顾这段旅程，一路上有老师的指导、同学的帮助、家人的鼓励与支持，借此机会，向所有帮助过我的人表示最衷心的感谢。

感谢本论文指导老师王笛副教授。王老师以认真的工作态度指导我的毕设工作，更以乐观开朗的态度、高尚的师德和无私的关怀激励我不断前行，每次同王老师交谈都能给予我很多启发。其次，感谢大创指导老师杨祥立副教授。杨老师严谨的治学态度、诲人不倦的师德对于本人在课题研究、学业以及未来发展等方面给予详细的指导和无微不至的照顾。

感谢信息科学与工程学院大数据与人工智能系梁宗保教授、刘新龙老师、杨庆老师、田朋老师等在学业方面给予的指导和帮助；感谢辅导员孔言老师、李世辉老师在本科学习生活中给予的照顾和关怀；感谢马克思主义学院傅红教授对本人四年以来的引导、鼓励和一路支持。

感谢“求人不如求己”队友和1236室友王玮桦、曾垣淞、李佳霖、王康松、溥吉满等，同学郭芷静、彭鹏铭、谭超、杨宇航等，四年本科时光同你们度过留下许多美好的回忆，大家在相互帮助、鼓励和学习中一起度过了这段最美好的时光，你们是我求学探索道路上的最重要的力量，祝你们未来一切顺利。

感谢我的母亲，母亲身上追求完美、坚韧的品格始终是我学习的榜样，是求学道路上最坚强的后盾，无论我遇到多大的挑战，都给予了我无条件的爱和支持，让我可以自信地面对困难，无惧挑战。

最后，感谢百忙之中抽出时间参加论文评审和答辩的老师，感谢你们对我的论文提出宝贵的意见。

参 考 文 献

1. 皮亦鸣，杨建宇，付毓生. 合成孔径雷达成像原理[M]. 成都: 电子科技大学出版社, 2007
2. Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[J]. Advances in neural information processing systems, 2014, 27.
3. 范黎.基于生成对抗网络的图像数据增强技术研究及应用[D].浙江大学,2022.
4. Simard P, Steinkraus D, Platt C. Best practices for convolutional neural networks applied to visual document analysis[J]. In ICDAR, 2003, 3:958–962.
5. Inoue H. Data augmentation by pairing samples for images classification[J]. arXiv preprint arXiv:1801.02929, 2018.
6. Arandjelović R, Zisserman A. Object discovery with a copy-pasting gan[J]. arXiv preprint arXiv:1905.11369, 2019.
7. Lee D, Liu S, Gu J, et al. Context-aware synthesis and placement of object instances[J]. Advances in neural information processing systems, 2018, 31.
8. Frid-Adar M, Klang E, Amitai M, et al. Synthetic data augmentation using GAN for improved liver lesion classification[C]. 2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018). IEEE, 2018: 289-293.
9. Gelautz M, Frick H, Raggam J, et al. SAR image simulation and analysis of alpine terrain[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 1998, 53(1): 17-38.
10. Ulaby F T, Moore R K, Fung A K. Microwave remote sensing: Active and passive. Vol. II. Radar remote sensing and surface scattering and emission theory[M]. Reading, MA: Addison-Wesley, 1982.
11. Franceschetti G, Iodice A, Riccio D, et al. SAR raw signal simulation for urban structures[J]. IEEE Transactions on Geoscience and Remote Sensing, 2003, 41(9): 1986-1995
12. Xu F, Jin Y Q. Imaging simulation of polarimetric SAR for a comprehensive terrain scene using the mapping and projection algorithm[J]. IEEE Transactions on Geoscience and Remote Sensing, 2006, 44(11): 3219-3234.
13. Qiu L, Ding B, Hu H. Bistatic SAR Data Processing Algorithms[M]. USA: Wiley, 2013.
14. 王哲.区域卷积神经网络和生成对抗模型在遥感图像解译中的应用[D].西安电子科技大学,2019.
15. 景士娴, 许小剑. 基于数据库的 SAR 图像场景合成研究[C]. 全国 DSP 应用技术联合学术会 议，2008: 47-51.
16. 朱柳. 高分辨 SAR 场景分布的重建方法研究[D]. 成都: 电子科技大学, 2013, 21-58.
17. Guo J, Lei B, Ding C, et al．Synthetic aperture radar image synthesis by using generative adversarial nets[J]．IEEE Geoscience and Remote Sensing Letters, 2017, 14(7): 1111－1115．
18. Cui Z, Zhang M, Cao Z, et al. Image data augmentation for SAR sensor via generative adversarial nets[J]. IEEE Access, 2019, 7: 42255-42268
19. Cao C, Cao Z, Cui Z. LDGAN: A synthetic aperture radar image generation method for automatic target recognition[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 58(5): 3495-3508.
20. Fan W, Zhou F, Tian T. A deceptive jamming template synthesis method for SAR using generative adversarial nets[C]. IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium. IEEE, 2020: 6926-6929.
21. 周驿枰.基于生成对抗网络的SAR图像生成与数据扩充研究[D].西安电子科技大学,2020.
22. 冯广能.基于深度生成学习的SAR图像仿真研究[D].电子科技大学,2021.
23. 保铮, 刑孟道, 王彤. 雷达成像技术 [M]. 中国北京 : 电子工业出版社, 2005.
24. 邢孟道, 王彤, 李真芳. 雷达信号处理基础 [M]. 中国北京 : 电子工业出版社, 2008
25. 师君.双基地SAR与线阵SAR原理及成像技术研究[D].电子科技大学,2009.
26. 桑成伟, 苏鑫, 孙洪. 合成孔径雷达图像信息解译与应用技术合成孔径雷达图像信息解译与应用技术 [M]. 中国北京 : 科学出版社, 2020.
27. 刘新龙. 高分辨率 SAR 图像的统计多层网络与稳健性分类方法研究[D]. 武汉大学, 2020.
28. Kingma D P, Welling M. Auto-encoding variational bayes[J]. arXiv preprint arXiv:1312.6114, 2013
29. Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks[J]. arXiv preprint arXiv:1511.06434, 2015.
30. 李航. 机器学习方法[M]. 清华大学出版社, 2022.
31. Chen X, Duan Y, Houthooft R, et al. Infogan: Interpretable representation learning by information maximizing generative adversarial nets[J]. Advances in neural information processing systems, 2016, 29.
32. Mirza M, Osindero S. Conditional generative adversarial nets[J]. arXiv preprint arXiv:1411.1784, 2014.
33. Zhao Y, Celik T, Liu N, et al. A Comparative Analysis of GAN-based Methods for SAR-to-optical Image Translation[J]. IEEE Geoscience and Remote Sensing Letters, 2022, 19: 1-5.
34. 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. 2017: 1125-1134.
35. Zhu J Y, Zhang R, Pathak D, et al. Toward multimodal image-to-image translation[J]. Advances in neural information processing systems, 2017, 30.
36. 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. 2017: 2223-2232.
37. Lin Y, Wang Y, Li Y, et al. Attention-based spatial guidance for image-to-image translation[C] Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 2021: 816-825.
38. Huang X, Liu M Y, Belongie S, et al. Multimodal unsupervised image-to-image translation[C] Proceedings of the European conference on computer vision (ECCV). 2018: 172-189.
39. Schmitt M, Hughes L H, Zhu X X. The SEN1-2 dataset for deep learning in SAR-optical data fusion[J]. arXiv preprint arXiv:1807.01569, 2018.
40. 马一鲲.基于cGAN的SAR到光学图像翻译[D].兰州大学,2023.
41. Rangzan M, Attarchi S, Gloaguen R, et al. TSGAN: An Optical-to-SAR Dual Conditional GAN for Optical based SAR Temporal Shifting[J]. arXiv preprint arXiv:2401.00440, 2023.
42. Fuentes Reyes M, Auer S, Merkle N, et al. Sar-to-optical image translation based on conditional generative adversarial networks—Optimization, opportunities and limits[J]. Remote Sensing, 2019, 11(17): 2067.
43. Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[J]. Advances in neural information processing systems, 2017, 30.
44. Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16x16 words: Transformers for image recognition at scale[J]. arXiv preprint arXiv:2010.11929, 2020.
45. Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C] Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 7132-7141.
46. Song C H, Han H J, Avrithis Y. All the attention you need: Global-local, spatial-channel attention for image retrieval[C] Proceedings of the IEEE/CVF winter conference on applications of computer vision. 2022: 2754-2763.
47. Wang Q, Wu B, Zhu P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks[C]. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020: 11534-11542.
48. Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[C]. Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 1-9.
49. Wang X, Girshick R, Gupta A, et al. Non-local neural networks[C]. Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 7794-7803.
50. Zhou Z H, Liu X Y. On multi‐class cost‐sensitive learning[J]. Computational Intelligence, 2010, 26(3): 232-257.
51. Gu K, Zhai G, Yang X, et al. Structural similarity weighting for image quality assessment[C] 2013 IEEE International Conference on Multimedia and Expo Workshops (ICMEW). IEEE, 2013: 1-6.