

**本 科 毕 业 设 计（论文）**



**题目:**  **基于深度学习的数码迷彩伪装生成技术研究**

**姓 名 刘嘉玄**

**学 院 计算机学院（国家示范性软件学院）**

**专 业 网络工程**

**班 级 2019211314**

**学 号 2019211514**

**指导教师 赵志诚**

**2023 年 5 月**

基于深度学习的数码迷彩伪装生成技术研究

摘 要

随着军事侦察技术的不断发展演进，迷彩技术作为躲避军事侦察的有效技术手段也在不断进行着升级演变。伴随数字图像技术飞速发展，数码迷彩技术逐渐成为主流迷彩技术。常规的数码迷彩图案可以让遮蔽物更好地融入到背景中，使得侦察方只能感知破碎模糊的信息，从而保护了隐蔽方的位置信息。

在深度学习未被广泛应用之前，数码迷彩图像的生成主要依赖手工设计和一些传统的机器学习算法。大部分生成算法主要分为三个流程：提取主要作战环境图片的基础色、将搜集的原始图像进行斑块化、生成最终的数码迷彩图像。鉴于军事保密性等原因，目前所公布的传统的数码迷彩生成算法仍然存在主色提取不理想、由自然背景图像到数码迷彩图像“风格迁移”能力差等问题。本文采用深度学习技术，针对上述问题进行了研究，主要工作有：1）提出了一种基于循环生成对抗网络的数码迷彩图像生成算法；2）构建了一个数码迷彩数据集，缓解了现有公开数据集缺乏的问题；3）在自制数据集上对模型进行了训练和测试，实验结果验证了本文算法的有效性；4）搭建了一个数码迷彩图像实时生成系统。

**关键词** 数码迷彩 深度学习 生成对抗网络 循环一致性 结构相似性

**Research on generation technology of digital camouflage**

**based on deep learning**

**ABSTRACT**

With the continuous development and evolution of military reconnaissance technology, camouflage technology as an effective technical means to evade military reconnaissance is also constantly upgrading and evolving. With the rapid development of digital image technology, digital camouflage technology has gradually become the mainstream camouflage technology. The conventional digital camouflage pattern can better integrate the concealment into the background, so that the reconnaissance party can only perceive the broken and blurred information, thereby protecting the location information of the hidden party.

Before deep learning was widely used, the generation of digital camouflage images mainly relied on manual design and some traditional machine learning algorithms. Most generation algorithms are mainly divided into three processes: extracting the base color of the main combat environment picture, plaqueing the collected raw image, and generating the final digital camouflage image. In view of military secrecy and other reasons, the traditional digital camouflage generation algorithm currently published still has problems such as unsatisfactory main color extraction and poor "style transfer" ability from natural background images to digital camouflage images. In this paper, this paper uses a cyclic generative adversarial network in the field of deep learning to generate digital camouflage images, and solves the problem of lack of existing public datasets by self-made datasets, and verifies the effectiveness of the model by training and testing the model on self-made datasets.

【参考中文】

**KEY WORDS** digital camouflage deep learning Generative Adversarial Nets(GAN)

Cycle Consistency Structure Similarity

目 录

[第一章 引言 1](#_Toc135164468)

[1.1 研究背景和意义 1](#_Toc135164469)

[1.2 研究现状 2](#_Toc135164470)

[1.3 研究内容 3](#_Toc135164471)

[1.4 章节安排 3](#_Toc135164472)

[第二章 基础理论与相关工作 4](#_Toc135164473)

[2.1 数码迷彩伪装技术 4](#_Toc135164474)

[2.1.1 空间混色原理 4](#_Toc135164475)

[2.1.2 视错觉效应 4](#_Toc135164476)

[2.2 卷积神经网络 5](#_Toc135164477)

[2.2.1 卷积层 5](#_Toc135164478)

[2.2.2 池化层 6](#_Toc135164479)

[2.2.3 全连接层及输出层 7](#_Toc135164480)

[2.2.4 激活函数 7](#_Toc135164481)

[2.3 循环神经网络 8](#_Toc135164482)

[2.4 生成式对抗网络 9](#_Toc135164483)

[2.4.1 生成器模块 9](#_Toc135164484)

[2.4.2 鉴别器模块 10](#_Toc135164485)

[2.4.3 损失函数 10](#_Toc135164486)

[2.5 本章小结 10](#_Toc135164487)

[第三章 基于深度学习的数码迷彩生成 11](#_Toc135164488)

[3.1 引言 11](#_Toc135164489)

[3.2 经典数码迷彩生成算法 11](#_Toc135164490)

[3.2.1 基于数码迷彩斑点概率分布模型的数码迷彩生成 11](#_Toc135164491)

[3.2.2 基于K-means的数码迷彩生成 12](#_Toc135164492)

[3.2.3 基于马尔科夫随机场和金字塔模型的的数码迷彩生成 12](#_Toc135164493)

[3.3 循环生成对抗网络 13](#_Toc135164494)

[3.3.1 生成网络结构 14](#_Toc135164495)

[3.3.2 判别网络结构 15](#_Toc135164496)

[3.3.3 损失函数 16](#_Toc135164497)

[3.4 本章小结 17](#_Toc135164498)

[第四章 实验与性能分析 18](#_Toc135164499)

[4.1 数据集构建 18](#_Toc135164500)

[4.2 实验设置 19](#_Toc135164501)

[4.3 评价指标设置 19](#_Toc135164502)

[4.4 消融实验 20](#_Toc135164503)

[4.5 实验结果分析 21](#_Toc135164504)

[4.6 本章小结 22](#_Toc135164505)

[第五章 总结与展望 23](#_Toc135164506)

[5.1 本文工作总结 23](#_Toc135164507)

[5.2 未来工作展望 23](#_Toc135164508)

[参考文献](#_Toc135164509)

[致 谢](#_Toc135164510)

[附 录](#_Toc135164511)

# 引言

## 研究背景和意义

迷彩（Camouflage）是一种让物体或人员在特定的环境中难以被发现的掩饰技术，通过融入周围环境以干扰敌对方的视觉识别来实现隐藏和混淆。迷彩技术源于自然界，生物以本身的皮肤颜色来迷惑捕食者的视觉，并提高了它们在环境中的生存率。在古代的狩猎过程中，原始人民往往会涂上颜料，以保证在林地中不被猎物发现。随着科技的发展，迷彩技术也得到了不断地改进和拓展。迷彩技术的发展既是社会进步的体现，也向我们展示了人类从自然中学习的不竭创造力和探索精神。

迷彩伪装技术作为对抗敌方军事侦察最常见和最有效的手段，通过在服装、设备、车辆等物品上涂上特殊颜料或材料来达到模仿环境或欺骗敌人的伪装效果。常见的迷彩伪装技术包括单色迷彩、变形迷彩和数码迷彩等。其中，单色迷彩是使用同一种颜色的迷彩涂料或服装，使人或设备与周围环境颜色相似，掩盖其存在感，使敌人难以发现。但由于单色迷彩的颜色局限性，其伪装效果比较有限，目前仅在大型固定军事设备上有所应用。变形迷彩则是利用不同的颜色和形状，将地形、植被等元素融入到伪装中。这种伪装方式被广泛应用于丛林、草原等地形的作战。数码迷彩是利用计算机技术制作的一种伪装方式，其原理是将目标物的图像通过计算机模拟程序与目标环境合成，使人或设备在外形、色彩和纹理等方面与环境相一致，避免被发现。相比传统的迷彩技术，数码迷彩技术具有更高的适用范围，能够应对更加复杂的作战环境，帮助士兵或军事设施躲避各类侦查，以提高军队反侦察能力。在近现代战争中，数码迷彩愈发成为了不可或缺的关键技术，可以大大减少士兵的无效伤亡，保障各类军备的安全性。因此，发展好数码迷彩技术，对于国家层面安防的高科技化建设具有重大的意义。

鉴于军事相关研究工作具有极高的保密性，基本上没有任何国家对于关键性的算法研究进行公开分享。目前公开的生成算法主要有三个流程。第一步是提取背景图片主色，根据颜色的主成分，选取最具代表性的颜色，作为数字迷彩的主要颜色。第二步是将背景图片斑块化，主要步骤是将背景图像分为若干个块，并为每个块选择相应的颜色或数字图案，用来生成数字迷彩。在这个过程需要考虑块的大小和形状，根据实际情况进行动态调整，以达到最优效果。第三步是生成数码迷彩，主要步骤是将选定的主色和块信息组合在一起，生成数字迷彩。由于现有技术的局限性，传统的数码迷彩生成算法只能处理单一的背景图像，无法适应不同的自然环境。因此如何实现更高效是数码迷彩生成方法，成为了现在的研究热点。

## 研究现状

数码迷彩技术研究一直是国内外军事领域非常重要的一项课题，大量学者深入该领域进行了多角度、全方位的研究。数码迷彩技术诞生于1976年，美国陆军军官蒂莫西·R·奥尼尔 (Timothy R. O'Neill) 提出[9]，由方块色块组成的图案可以提供比传统的棕色和绿色斑块更有效的伪装效果，即创建一个复杂的小正方形（2 英寸）图案，在近距离观察者会看到林地图案，而在远距离时小像素会合并成更大的块，从而给出更大范围内的林地图案的外观。

20世纪末期，伴随数字图像技术飞速发展，数码迷彩技术获得了井喷式发展。1996年，加拿大温带丛林数字迷彩作战服研制成功，在测试中取得了非常不错的隐蔽效果。随后美国研究出了MARPAT（Marine Pattern Camouflage）数码迷彩，这种迷彩伪装可以同时用于多种作战环境。我国对于数码迷彩技术的研究可以追溯到20世纪80年代，技术专家率先在反雷达伪装网络使用了“双纹理伪装”技术。2009年，在纪念建中华人民共和国60周年的国庆阅兵式上，数码迷彩图案出现在了13支陆军主战部队演示方阵中。

在传统的数码迷彩生成技术的三个步骤中，主色提取的常见方法主要是各种聚类算法(K-means、FCM等)及峰值点算法等。这些方法采用聚类学习的方式，获得最常出现的颜色，将其作为后续要生成的数码迷彩的主要颜色，减小被伪装物与周围环境之间的颜色差异。针对斑块化图片处理环节，目前主要采用根据阈值、区域或边缘进行分割处理等策略。该方法的主要目的是计算不同区域数码迷彩斑块的大小并识别需要模糊的边界，从而实现伪装目标与背景之间的边界隐藏。喻钧等人利用K-means算法，根据常见作战背景图像提取主色，设计生成数码迷彩图案[2]。徐英等人提出了一种基于HIS模型和量化颜色直方图的迷彩颜色选择方法，对量化的迷彩颜色使用给定阈值的策略进行非均匀化处理[3]。蔡云骧等采用基于金字塔结构的分层模糊C均值聚类算法和基于分形布朗运动的分形维数估计方法提取背景图像的纹理和主要颜色特征[4]，生成数码迷彩图案。贾琦等人结合实际背景特征，基于金字塔模型和马尔可夫随机场设计了数码迷彩图案[5]。

随着深度学习技术的不断发展，计算机视觉领域得到了飞速的发展，有更多的研究人员开始关注如何使用神经网络优化数码迷彩生成方法，基于深度学习的数码迷彩制作也诞生了许多重要的成果。Alfimtsev等利用深度卷积生成对抗网络（DCGAN）来生成具有不同风格的迷彩样本，并使用Fast R-CNN等图像识别模型评估其效果[6]。Yang等人则提出了基于对抗性自编码器（AAE）模型的数码迷彩制作模型[7]，解决了传统基于特征空间变形生成斑块时描述不精确、复制困难等问题。同时，Yang等还使用了形态学处理技术来进一步提高伪装效果。除此之外，还有许多研究正在探索使用深度学习技术来提高数码迷彩制作的质量与效率。然而，基于深度学习的数码迷彩制作也面临着一些挑战。首先，由于训练样本的数量和质量对结果具有重要的影响，因此训练数据的匮乏在一定程度上阻碍了数码迷彩生成技术的发展。其次，深度学习模型的实时性和运行效率是需要突破的瓶颈。第三，数码迷彩的伪装效果需要考虑各种环境变化，如光照、天气等影响，仍具有较大难度。

## 研究内容

本文将设计一个循环生成对抗网络。通过自制数据集训练循环对抗生成网络模型，使该模型可以根据背景图像实时自适应生成数码迷彩图案，用于更好地实现军事伪装，躲避军事侦察。

综上，本文的主要工作可总结如下：

（1）针对深度学习模型需要大量数据进行训练以保证模型适用性，及缺少公开数码迷彩数据集问题。设计并创建了一个数据集，数据集包含数草地、城市、丛林、沙漠、山地、水系、雪地七种常见的自然环境类型800张自然环境图片，及800张对应数码迷彩图像。

（2）设计循环生成对抗网络模型，在构建的数据集上进行训练。通过选取不同损失函数组合及单层网络类型，确定表现最优模型结构，实现从背景图像域到数码迷彩图像域的良好映射。

（3）在上述模型基础上，本文将建立一个数码迷彩实时生成系统：用户可上传任意类型自然图像，并可对于所上传自然图像进行选取裁剪，系统将根据用户需求实时生成对应的数码迷彩图像。

## 章节安排

本论文共分为五章，各章节主要内容分别如下：

第一章为引言。在该章节中，首先阐述了迷彩技术的发展历程以及其在军事、安全和文化领域中的研究意义。随后，详细叙述了数码迷彩技术在现代迷彩领域中的研究进展和应用现状。最后，简要概述了本论文的研究内容，以及各章节的内容安排。

第二章为基础理论与相关工作。该章节着重介绍了本文使用的循环对抗生成网络模型，及其相关的基础知识。

第三章为基于深度学习的数码迷彩生成技术。该章节首先介绍了数码迷彩生成现有优秀模型，接着解释了循环对抗神经网络的整体架构。

第四章为实验与效果分析。该章节先描述了数据集的构建思路及特点，紧接着阐述了基础模型参数以及对比实验的设置方式，并详细阐述了相关的评级指标。随后介绍了消融实验及实验结果。最终，经过对比实验结果的分析得出了最优解决方案。

第五章为总结与展望。该章节首先总结了本文的研究成果，随后指出了当前数码迷彩技术仍需改进的研究难点，同时提出了对未来数码迷彩技术发展进行展望的想法。

# 基础理论与相关工作

## 数码迷彩伪装技术

数码迷彩技术是通过不同色块的排列组合，使得不同颜色间的边缘变得模糊、破碎，并利用空间混色原理和视错觉效应，将斑块“马赛克”化，使得伪装目标的边界变得模糊、破碎，从而更好地融入背景之中，从而大大增强了伪装效果。本节将分别介绍空间混色原理和视错觉效应。

### 空间混色原理

空间混色原理是通过调整混合图像不同通道的像素值，使得混合图像呈现出与原始图像不同的颜色和亮度。

一般情况下，空间混色通过将两个或多个图像的红、绿、蓝三种颜色通道进行混合得到混合图像。最常见的方法是对每个颜色通道进行加权平均，计算方法如式2-1、2-2及2-3所示：

其中，、、分别代表原始图像A的R、G、B通道像素值，、、分别代表原始图像B的R、G、B通道像素值，、、、、、是对应通道的权重。

### 视错觉效应

视错觉是指人类感知世界时发生的一种现象，即在感知对象时出现错觉或误解。常见的视错觉包括不相等、尺度、色彩、运动、方向等。通过巧妙地运用视错觉技术，数码迷彩可以使观察者产生一种错觉，让观察者看不清楚隐藏的物体提高数码迷彩的隐蔽性。这被称为“视域消除”，即在环境中难以被视觉定位。在诸多视错觉设计思路中，尺度错觉是被广泛应用的一种技术，它是指当两个相同大小的物体在不同的尺度下，显示不同大小的现象。在数码迷彩中，通过使用滤镜和图像处理技术，可以放大或缩小某些区域，以产生尺度错觉，从而隐藏或模糊一些目标。除此之外，色彩错觉也被广泛用于数码迷彩中。色彩错觉是指物体的颜色会受到周围颜色的影响，使其看起来不同于实际的颜色。在数码迷彩中，通过运用颜色的对比和协调，可以产生一种天然的色彩错觉，使目标物与环境融合在一起。

在迷彩设计中，从视觉效应的应用角度还可将其分为静态视错觉、动态视错觉和全息视错觉。静态视错觉是利用单独的颜色、形状和图案等视觉元素来掩盖人体轮廓。这种技术通常用于山林、草地等环境中，其中的森林迷彩就是很好的例子。动态视错觉是利用颜色和图案的变化来掩盖人体轮廓，例如草原、沙漠、城市和田野等环境中常见的迷彩。全息视错觉利用多个颜色和图案的变化来制造一种错觉，从而对观察者的视觉产生影响。这种技术通常用于水中或野外的环境中。

## 卷积神经网络

人工神经网络最早诞生于上世纪40-50年代，它的出现是受到了人脑中的神经元的启发，通过模仿神经元的活动方式来模拟人脑工作。

形式神经元模型早在1943年就由沃伦·麦卡洛克和皮茨创造，该模型通过一些物理元器件（如电阻等）构造，即M-P模型。1958年，Roseblatt又在此基础上研究出了感知机模型[13]，至此最早的前馈神经网络架构便已初具。随后人工神经网络经历了不断发展迭代的过程。1980年，Fukushima提出了神经认知机（Neocogniron）模型[25]，这被学界认为是卷积神经网络（CNN）的雏形。1990年，Lecun.等在此工作基础上，创建了一种基于反向传播算法的手写数字识别系统，并在美国邮政部门的手写邮政编码数据集上取得了非常好的分类结果[26]，CNN的基础框架就此被正式确立，并在图像处理领域取得了突破性进展。尤其论文中提到在训练模型时使用的反向传播算法（Back Propagation, BP）更是使得CNN的性能得到了极大的提升。2006年，深度学习奠基人Hinton提出了基于卷积神经网络的多层感知机模型（MLP）[27]，用于图像分类任务中的问题。然而，这个模型仍然面临着训练和计算效率低、模型复杂等问题，在当时并没有获得广泛认可。直到2012年，Alex等提出了深度卷积神经网络模型（AlexNet）用于在ImageNet数据集上的识别任务[28]，该模型采用了卷积、池化、Dropout等方法，处理对计算机来说相当庞大的图像数据集，使得准确率得到了大幅提升，引领了CNN的研究和应用进一步发展的浪潮。

相比很多复杂的深度学习模型，CNN的结构并不复杂，其大致组成包括卷积层、池化层、全连接层、输出层以及在其中起着很关键作用的激活函数，下面将对它们进行详细介绍。

### 卷积层

卷积层的主要任务是对输入图像的特征进行提取。卷积层的层数又被称为卷积神经网络的深度，深度越大通常能带来更好的特征提取以及模型的表现，即可以解决更为复杂图像处理任务。但是，深度过大也可能出现梯度消失或爆炸等问题，所以在实践中需要控制合适的深度。

卷积层主要是通过卷积操作对输入数据进行处理，提取出数据中的特征信息。具体来说，卷积层会定义一组卷积核，每个卷积核都是一个小矩阵，它的作用如同过滤器。卷积核会在输入数据的局部区域上进行滑动操作，将每个区域的数值与卷积核中对应位置上的数值相乘并求和，得到一个单一的输出值。随着卷积核在输入数据上的滑动，就可以逐渐遍历整个输入数据，得到一组输出值，这个过程就称为卷积操作。在卷积层中，卷积核尺寸、图像填充以及步长是最重要的三个参数，它们决定了当前层最终输出的特征图大小。卷积核尺寸为，步长为1时的卷积操作如图2-1所示。

图表, 图示

描述已自动生成

图 2‑1 步长为1时卷积核工作过程

### 池化层

如果卷积层直接进行叠加，经过多层卷积后，特征图的维度会变得越来越大，导致需要更多的计算开销，同时也会对模型的训练效率产生极大的影响。因此，每个卷积层卷积操作结束后，再进入下一个卷积层之前都会进入一个池化层。池化层的主要作用是对上一个卷积层的输出进行下采样操作，即在主要特征信息被保留的情况下减少特征图的维度，提高模型的稳定性。

池化层主要进行的池化操作有最大池化或平均池化两种方式，通过对输入特征图中的每个局部区域进行处理，将其进行压缩成一个单独的值。具体来说，最大池化操作可以直接选择局部区域中的最大值作为输出，而平均池化操作则将局部区域的值进行平均，并将平均值作为输出。例如，一个2x2的最大池化层可以将相邻的4个卷积层输出按照2x2的大小分成4个小区域，然后在每个小区域中选择最大的值（最大池化）或者区域平均值（平均池化）作为该区域的输出，该过程如图2-2所示。这样可以将输入特征图的空间维度缩小一半，同时保留最显著的特征信息。

图示

描述已自动生成

图 2‑2 池化操作示意图

### 全连接层及输出层

卷积神经网络中，在所有的卷积层和池化层结束后，会通过全连接层将前面各层中提取出来的局部特征信息整合成全局特征，提高模型性能。全连接层的全部神经元和前一层都有连接，以此进行特征信息的整合，从而更好地表示图像或其他类型的输入信号。

具体地说，全连接层中的每一个节点都会对输入进行加权处理，并产生一个输出。这个权重是通过训练来确定的，目标是最小化网络的训练误差，从而提高预测性能。通过反向传播算法，网络可以更新全连接层中的所有权重和偏置，从而最小化损失函数，提高模型的预测准确率。

在一个或多个全连接层后连接的是卷积神经网络最终的输出层，这个输出层通常是一个分类器，用于将卷积神经网络处理后的特征映射到目标类别或回归目标。其形式和具体使用情况，根据不同的任务而有所不同。

### 激活函数

卷积层和全连接层的操作本质上还属于线性操作，然而现行操作的堆叠仍属于线性操作，这将大大限制深度卷积神经网络的表达能力，因此需要依赖激活函数提供非线性变化。此外，负值对于某些任务是不必要甚至有害的，例如图像分类任务中，一张图像中的背景对分类是没有贡献的，去除背景的负值就可以提升模型性能。而激活函数就可以将负值输入映射成非负值，优化网络。下面介绍三种CNN中常用的激活函数。

1. Sigmoid激活函数

Sigmoid激活函数的计算如式2-4所示，由其可知Sigmoid函数的输出范围在0到1之间，可以被解释为概率，因此特别适用于二分类任务中的输出层。

1. ReLU激活函数

ReLU函数的输出在时，输出值等于，在时，输出值为0。其优点是收敛速度比Sigmoid函数快，而且ReLU函数在输入大于0时，梯度为1，可以加速反向传播算法的计算速度。

1. Leaky ReLU激活函数

Leaky ReLU是ReLU函数的一种变种。在ReLU中，当输入小于0时，激活函数的输出将直接变为0。而在Leaky ReLU中，当输入小于0时，激活函数的输出将不再是0，而是输入的一个小比例，这样设置可以避免出现“死亡神经元”的问题。当ReLU处于负半区时，梯度为0，导致权重不能更新，从而影响网络的学习能力。而Leaky ReLU能够在负半区产生非零输出，从而使梯度不为0，避免了这个问题出现。

## 循环神经网络

随着科学家们开始用神经网络处理和解决更为多元化的数据和任务，前馈神经网络架构的缺陷也被暴露出来，即该网络只考虑输入信息，网络当前的输出只与当前的输入有关，而不考虑与前面输入处理的数据是否存在关联性。尤其在自然语言处理（NLP）领域，输入数据即文本数据往往具有较强的序列性，需要结合上下文的信息去理解文本含义。在计算机视觉的视频处理领域，同样存在相似的问题，视频数据的结构特点决定了数据时间存在较高的时序性，而理解这种时序性也就成为了理解视频数据的关键。1982年John Hopfield创造性研究出了霍菲尔德网络[14]，该网络可以实现将输入的残缺数据重建成完整数据。在1986年，Elman等人收到该网络的启发，提出了用于处理序列数据的循环神经网络（recurrent neural network，RNN）[15]。RNN网络的出现，成功改进了传统全连接神经网络无法较好发掘输入数据时序性的问题。

RNN网络在结构上与传统的全连接神经网络一样，分为输入层、隐藏层和输出层。与之不同的地方在于，RNN拥有一个记忆单元，即隐藏状态。RNN的隐藏层状态会随着时间更新，当前时刻的隐藏层不仅受当前时刻输入层影响，也会受到前一时刻的隐藏层影响，从而在处理序列数据时更加灵活和高效。通过这样的设定，神经网络就有了“记忆”，可以像人类处理文本信息一样，联系上下文进行分析。RNN的结构示意图如图2-3所示。

图示, 示意图

描述已自动生成

图 2‑3 RNN网络结构图

在RNN网络结构中，输入层负责接收序列数据，隐藏层则将序列数据中的信息编码成隐藏状态，输出层则根据隐藏状态输出相应的结果。在隐藏层中，每个时刻RNN会从序列数据中提取特征，用于更新隐藏状态。在后续的时间步中，RNN会把前一个时间步的隐藏状态作为输入改写后通过激活函数输出新的隐藏状态。当前时刻t的隐藏状态公式如式2-7所示：

时刻t的输出计算公式如式2-8所示：

其中表示当前时间步的输入，是当前时间步的隐藏状态，是前一时间步的隐藏状态，代表输入神经元的权重矩阵，是循环神经元的权重矩阵，代表隐藏层偏置项，代表输出层的偏置项，代表输出神经元的权重矩阵，是激活函数。

## 生成式对抗网络

海量数据的出现方便了深度学习算法模型的发展，使得深度学习领域出现了质的飞跃。深度学习模型分为两个主要发展方向：判别式模型和生成式模型。判别式模型的发展由于BP、Dropout等算法的发明获得了极大的成功，出现了诸如线性回归、决策树、支撑向量机SVM、神经网络等非常成熟高效的模型。生成式模型由于模型建构更加复杂，所以一直没有得到比较好的发展。直到2014年，GoodFellow提出了生成对抗网络模型（GAN）[17]，使得生成模型领域重获新生，在计算机视觉和自然语言处理领域大放异彩。GAN模型实现的过程概括来说就是“无中生有”。整个GAN模型共分为两个主要部分，生成器模块和鉴别器模块，其整体结构如图2-4所示。下面对这两个模块分别进行详细介绍。

图示

描述已自动生成

图 2‑4 GAN模型原理示意图

### 生成器模块

GAN生成器的内部结构通常采用基于深度学习模型的卷积神经网络（CNN）或递归神经网络（RNN）架构。对于CNN结构，一个典型的GAN生成器通常由多个卷积层、反卷积层和归一化层等组成。而RNN结构则通常由多个LSTM（长短时记忆）层、GRU（门控循环单元）层和Dropout层等组成。此外，GAN生成器还通常使用一些常见的激活函数，如ReLU和Leaky ReLU。

在GAN生成器的构建过程中，通常采用反向卷积（Deconvolution）或上采样（Upsampling）的方式，将随机噪声向量（通常为100维向量）逐层扩张，直到生成器最后输出一个与原始图像大小一致的矩阵。

### 鉴别器模块

鉴别器是一个二分类器，其内部结构通常为一个卷积神经网络（CNN）、全连接层（FC），以及一个sigmoid激活函数。整个鉴别器CNN部分的作用是从原始输入数据中提取特征，比如图像中的纹理、轮廓、颜色等信息，以便进行分类任务。其中，卷积层会不断提取图像的特征，池化层会对特征进行下采样，而全连接层会将最后一层卷积的输出结果映射到二分类最终输出。输入一般是一张图片，输出是图片所属真假类别的二分类概率。

举个例子，如果我们将GAN用于生成手写数字，鉴别器将学习如何识别真实的手写数字，当它看到生成器生成的带有数字的图像时，鉴别器需要检测图像的像素、灰度级别、曲线等特征，以判断它是否真实的手写数字。如果鉴别器判断为假，它将返回一个0输出，并向生成器返回反馈。如果鉴别器判断为真，它将返回一个1输出。

### 损失函数

GAN模型训练时的损失函数如式2-6所示：

其中，G代表生成器，D代表判别器，代表真实数据，代表真实数据概率密度分布，代表了输入的随机数据，服从高斯噪声分布。因此所量化的是生成数据和真实数据的差异性。代表生成器要在判别器最大化真假数据交叉熵损失的情况下最小化该损失。表示将真实数据输入到鉴别器中，代表根据输入的随机数据生成虚假数据。该损失函数代表了生成器和鉴别器进行零和博弈的过程，最终通过训练模型降低损失函数值，当训练中的GAN模型达到纳什均衡时，我们就得到了该情景下训练完成的GAN模型。

## 本章小结

本章介绍了空间混色原理及视错觉效应，分析了两种技术在数码迷彩设计过程中起到的作用。随后介绍了使用深度学习框架解决数码迷彩生成问题中涉及到的相关技术原理，包括卷积神经网络、循环神经网络和生成式对抗网络的原理。

# 基于深度学习的数码迷彩生成

## 引言

在深度学习未得到快速发展的阶段，传统的机器学习算法在数码迷彩生成问题上也起到了比较重要的作用。3.2中分别介绍了几种经典的数码迷彩生成算法，虽然相比如深度学习的方法，这些传统方法有一定的缺陷，但在原理剖析的层面上，这些方法的可据性更强，依然能为整个课题的研究提供诸多指导意见。3.3节提出了本文解决数码迷彩生成文题的基础模型框架，并对该模型框架进行了详细的剖析介绍。

## 经典数码迷彩生成算法

### 基于数码迷彩斑点概率分布模型的数码迷彩生成

秦雷等人结合了空间混色原理和概率分布模型，建立了一套基于数码迷彩斑点概率分布的概率模型[18]。设某位置出现数码单元的概率为P，斑点外部点与斑点边界的距离为S，根据秦雷等人的研究两者之间关系如式3-1所示：

其中，代表了变化率，通过调整变化率的值，可以设计出不同破碎效果的数码迷彩图案。该模型可以实现从主色斑点到数码迷彩的转换，同时保证原有主色斑点之间的相对位置不发生较大变化。该方法的算法流程图如图3-1所示。

图示, 示意图

描述已自动生成

图 3‑1 基于数码迷彩斑点概率分布模型的数码迷彩生成算法流程图

首先根据原有的大斑块迷彩初步数码化，生成1类代表图像，原始图像的像素点对应该图像的数码单元。随后，对1类代表图像进行颜色分层处理，从而将其转换为索引图像，即不同颜色的图像分别处在不同的图层。再根据所建立的线性概率模型计算出每个像素点是处于斑点内抑或斑点外，斑点内外分别记作1和0处理。第三步将前面得到的二值图像进行腐蚀膨胀处理以消除图像中的噪声。第四步通过图像欧式距离变换对多张二值图像进行处理合成得到距离矩阵。第五步根据距离矩阵和概率模型计算得到斑点位置设计图，然后将原有迷彩中的斑点颜色填充进去得到彩色的数码迷彩图像。

该方法很好地解决了将传统大斑块迷彩图案转换成数码迷彩的过程，得到了不错的实验效果，但对于如何将原始自然背景图案直接转换成数码迷彩图案的问题并不能较好解决。

### 基于K-means的数码迷彩生成

主色提取是传统数码迷彩生成过程中非常重要的一个环节，比较传统的处理方式有颜色直方图方法，但是这种方法在解决较为复杂的背景图像时表现较差，无法快速准确提取出正确的主色。喻钧等人就将聚类算法运用到了主色提取这一环境上，采用了聚类算法中应用最多最广泛的K-means算法来提取背景图片的主色[2]。其基本思路如下：首先将输入图像转换为RGB颜色空间，然后将RGB颜色空间值转换成CIE 1976 Lab的颜色空间值，该颜色空间相比原始的RGB颜色空间有着色域更宽阔不受设备限制等优点。第二步，使用K-means算法对颜色空间值构成的颜色向量进行聚类，将图像分为K个类别。这里K代表生成的数码迷彩颜色数量。第三步，对原始图像中的每个像素点的颜色进行分类，各自归并到K类中最接近的颜色种类中，生成颜色序号分布图。第四步，根据迷彩单元大小及图像大小等尺寸信息对图片进行重新处理，分出不同区域板块，对各自板块内的颜色种类序号进行平均化处理，用平均后的颜色序号代替同板块所有其余颜色序号，形成新的颜色序号分布图。第五步，根据颜色序号分布图，向对应板块填充对应主色。最后一步，将图片的CIE 1976 Lab的颜色空间值在转换回RGB颜色空间，形成最终的数码迷彩图。

该方法巧妙解决了传统数码迷彩生成过程主色提取部分不稳定不准确的缺点，对于特定自然风景图片生成数码迷彩图案解决效果较好，但缺点在于生成出的数码迷彩图案专一性太强，不能更好地适应当前地貌下其余场景，即不能实现“风格”模仿，同时K-means算法运算过程略慢，对于实时性变换需求不能很好解决。

### 基于马尔科夫随机场和金字塔模型的的数码迷彩生成

在数字迷彩中，纹理模型也是十分重要的一个组成部分，利用纹理模型可以使数字迷彩产品具有更加准确和真实的外观。常见的纹理模型通常分为描述模型和再生模型，再生模型在数码迷彩的设计生成中起到了至关重要的作用，它可以通过描述和一组随机数按照要求重新生成一幅图像。马尔科夫随机场（Markov Random Field，简称MRF）是一种常见的概率图模型，常用于建模具有空间依赖关系的随机变量。鉴于马尔科夫随机场模型的特点，贾其等用此模型结合金字塔模型构建了一套数码迷彩生成模型[5]。金字塔模型是一种计算机视觉中的图像处理方法，它利用了图像的多尺度特性，可以实现对图像不同尺度下的特征提取和匹配。其基本思想是将原始图像不断缩小，得到一系列尺度不同的图像金字塔层，从而可以在不同尺度上检测和匹配图像中的特征，该模型的流程框架如图3-2所示。

图示

描述已自动生成

图 3‑2 基于马尔科夫随机场和金字塔模型的数码迷彩生成模型结构图

具体的算法流程是：第一步，通过聚类算法对背景图像的主色进行提取，同时得到主色面积的比例信息。第二步，创建一个随机的噪声矩阵。第三步，对获得的噪声矩阵使用傅里叶变换，得到随机噪声矩阵的离散傅里叶形式。第四步，通过确定相邻域大小及频域系数，再根据前面变换后的噪声矩阵得到马尔科夫随机场。第五步，在对马尔科夫随机场进行傅里叶逆变换得到二维纹理矩阵。第六步，将主色比例信息和背景主色信息加入到二维纹理矩阵中，获得基础迷彩图案。最后一步，通过金字塔分解的方式得到最终的数码迷彩图案。

该模型初步实现了背景图案到数码迷彩的生成，但根据实验效果不难发现，由于主色提取部分提取主色信息过于简略化，损失了很多色彩，导致该方法生成出的数码迷彩图案效果欠佳，但相比传统手工设计依然起到了提升效率的优势。

## 循环生成对抗网络

在生成对抗网络（GAN）模型被研究出来后，深度学习领域产生了一阵有关生成网络的热潮。GAN模型也在热潮中不断发展改进，围绕GAN模型的研究层出不穷。Alec Radford等首先将深度卷积神经网络与GAN模型相结合，研究出了DCGAN模型[19]。该模型通过加入卷积神经网络（CNN）使得模型在处理图像数据时能够更好的提取图像特征，同时采用了无池化层和批规范化的方法，减小了信息损失，加速了模型的收敛和泛化性能。2017年，Zhu.等提出了CycleGAN 模型[20]，该模型相比DCGAN模型有了比较大的提升，可以支持不同领域之间图像的转换。在训练时，CycleGAN结合了循环一致性损失和对抗损失，考虑了生成图像和原始图像之间的相互关系，从而获得了更好地效果。更为重要的是，CycleGAN模型不再需要配对的训练数据，这极大便利了数码迷彩设计时缺少规范数据集的劣势。因此，本课题采用了CycleGAN模型来实现数码迷彩生成任务。接下来将对本文使用的CycleGAN模型网络进行详细的介绍。CycleGAN模型如同两个拼接起来的GAN模型，整体结构如图3-3所示。

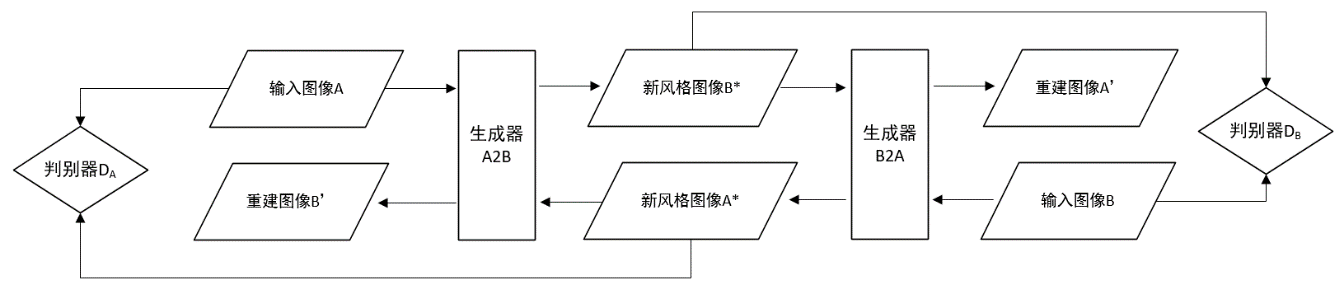


图 3‑3 CycleGAN模型结构图

生成器G1 （A2B）按照CycleGAN模型的设计通过一个编码器、一个转换器和一个解码器构成，其中编码器将输入A中的图像纹理特征提取出来，转换器通过新风格的图像B的生成使得纹理特征得到重构和转换，最后解码器将转换后的特征重构成目标图像B'。同样地，生成器G2 （B2A）也由编码器、转换器和解码器构成，可以从输入B中提取出纹理特征，生成新风格的图像A，并通过传输模块将图像A传输到生成器G1 （A2B）进行重建操作。判别器DA用于判断生成图像A的真实性，通过学习真实图像和生成图像之间的差异来提高自身的判别效果。同理，判别器DB用于判断生成图像B\*的真实性。通过这样的结构，我们可以实现从自然背景图像到数码迷彩图像的转换，并最终得到一个训练好的生成器A2B。

### 生成网络结构

生成器是CycleGAN模型中的关键组成部分之一，负责将输入域的图片转换为输出域的图片。它的内部结构包含三个部分：编码器、转换器和解码器。编码器由三层卷积神经网络组成，主要负责提取输入图片的特征信息，并通过降采样将图片尺寸变小。每层卷积包括一个2D卷积层、一个正则化（IN）层和一个ReLU激活函数。转换器采用了6个残差块来将编码器提取到的特征进行转换，实现了从输入域到输出域的映射。每层残差块内部包含两层卷积层，每个卷积层后跟一个IN层，第一个IN层后面接一个ReLU激活函数。解码器用于将转换器输出的特征图还原为最终的输出图片。它也由三层卷积神经网络组成，包括前两个卷积层各包括一个2D 转置卷积层、一个IN层和一个ReLU激活函数，这些卷积层使用反卷积（转置卷积）操作实现上采样的效果。最后一层是一个2D 卷积层，使用tanh作为激活函数，将最终的特征图还原成输出域中的图片。整个生成器内部结构的设计旨在实现从输入域到输出域的无缝映射，同时尽可能保留输入图片中的尽可能多的信息，具体结构如表3-1所示。

表 3‑1 CycleGAN模型生成网络结构

|  |  |
| --- | --- |
| 模块 | 结构 |
| encoder | 7×7 Conv-Norm-ReLU |
| 7×7 Conv-Norm-ReLU |
| 7×7 Conv-Norm-ReLU |
| ResNet层 | Residual Block(Conv-Norm-ReLU-Conv-Norm) |
| Residual Block(Conv-Norm-ReLU-Conv-Norm) |
| Residual Block(Conv-Norm-ReLU-Conv-Norm) |
| Residual Block(Conv-Norm-ReLU-Conv-Norm) |
| Residual Block(Conv-Norm-ReLU-Conv-Norm) |
| Residual Block(Conv-Norm-ReLU-Conv-Norm) |
| decoder | 3×3 Deconv-Norm-ReLU |
| 3×3 Deconv-Norm-ReLU |
| 7×7 Conv-tanh |

### 判别网络结构

鉴别器的内部判别网络通常由几个核心的模块构成，其中包括卷积层以及激活函数、实例归一化和全连接层（FCN）等。在具体实现时，首先通过一个4×4大小、步长为2、填充大小为1的卷积层，对输入数据进行特征提取操作，生成包含64个通道的特征图。接着，经过实例归一化处理和激活函数操作，对特征图进一步加工处理。然后，将上一层的特征图传入下一层卷积层，进行卷积、实例归一化和激活函数操作，将得到的结果传递到后续层级。最终，输出一个特征图，通过一个大小为4×4的全连接层，生成一个1维实数，用于表示所生成的图像是否为真实图像，具体内部结构如表3-2所示。

|  |  |
| --- | --- |
| 模块 | 结构 |
| 卷积层 | 4×4 Conv-LReLU, stride = 2, filter = 64 |
| 4×4 Conv-Norm-LReLU, stride = 2, filter = 128 |
| 4×4 Conv-Norm-LReLU, stride = 2, filter = 256 |
| 4×4 Conv-Norm-LReLU, stride = 1, filter = 512 |
| 4×4 Conv, stride = 1, filter = 1 |

表 3‑2 CycleGAN模型判别网络结构

### 

### 损失函数

CycleGAN模型的损失函数是CycleGAN模型的设计中非常重要的一环，该模型能实现非配对图片数据的无监督训练就是通过损失函数来进行规避和约束的。CycleGAN模型的损失函数在该模型的设计中扮演着至关重要的角色，正是通过损失函数的约束，该模型才可以实现非配对图片数据的无监督训练。该损失函数的设计旨在优化生成器和判别器的性能，并在循环一致性损失和对抗损失之间平衡，从而实现最佳的图像转换效果。基础版本的损失函数包括两个部分，一个是对抗性损失函数，另一个损失函数是循环一致性损失。对抗性损失与常规的GAN模型的损失函数一致，这部分损失函数可以使生成器生成的数据分布接近真实数据分布，其计算公式如式3-2所示：

其中，y 为真实数码迷彩图像域，x 为背景图像域中的数据， 和分别代表真实真实数码迷彩图像域的分布和背景图像域的分布。公式中的 代表生成器 G 生成的假数据 y 被判别器 判别为真实数据的概率， 则代表生成器 G 生成的假数据被判别器 判别为假数据的概率。目标是最大化 ，从而让生成器 G 生成的假数据更接近于真实数据。

循环一致性损失函数可以防止生成器G与F相互矛盾，即两个生成器生成数据之后还能变换回来近似看成。G通常以x为输入，用于生成虚假的y图像， F通常以y为输入，用于生成虚假的x图像。在将x输入到G后，会生成虚假的y图像，将该虚假的y图像再输入到F中，会生成更虚假的x图像。在理想情况下，更虚假的x图像应该与原始的x图像近乎相同。由于这构成了一个循环，所以被称为循环一致性损失。在训练过程中，经常会遇到的一种情况是，通过x生成的虚假y越来越倾向于成为一个可以欺骗判别器的值，生成器渐渐发现，无论输入的x如何，只要生成的图像与y越像，就越能欺骗判别器，因此只需生成与y相似的图像。但这种y并不是我们所期望的，我们希望保留x中的内容，仅改变其中的样式。因此，设计了这种循环一致性损失来实现此目的，其计算公式如式3-3所示：

其中，和 分别表示转换前后的图像之间的差异，采用 L1 距离进行度量。在优化模型的过程中，使得函数最小化，就可以约束生成器 G 和 F 使得转换后的图像可以通过两个方向的转换实现相互映射，从而提高模型的生成效果。最终整体的损失函数计算公式如式3-5所示：

## 本章小结

本章首先介绍了数码迷彩技术生成问题的经典算法，包括基于数码迷彩概率分布模型的数码迷彩生成算法、基于K-means的数码迷彩生成算法以及基于马尔科夫随机场和金字塔模型的的数码迷彩生成算法，并指出了这些算法所存在的问题。随后详细介绍了本研究课题基于深度学习的数码迷彩生成技术的核心框架模型循环生成对抗网络模型及其内部具体结构，包括模型整体架构、生成网络结构、判别网络结构以及损失函数。

# 实验与性能分析

## 数据集构建

鉴于公开的迷彩数据集非常少，对于深度学习这种需要大量数据进行训练的技术不是十分友好，为完成模型的训练构建了一个自制数据集。数据集中包含草地、城市、丛林、沙漠、山地、水系、雪地七种常见的自然环境类型。训练集有820张自然背景图片及820张数码迷彩图片，测试集中有200张自然背景图片，每种类别的数据如表4-1所示，图片示意图如图4-1所示。

表 4‑1 数据集数据

|  |  |  |
| --- | --- | --- |
| 数据集属性 | 自然环境类型 | 数量（自然背景+迷彩） |
| 训练集 | 草地 | 100+100 |
| 城市 | 130+130 |
| 丛林 | 150+150 |
| 沙漠 | 100+100 |
| 山地 | 160+160 |
| 水系 | 100+100 |
| 雪地 | 80+80 |
| 测试集 | 草地 | 20+20 |
| 城市 | 30+30 |
| 丛林 | 40+40 |
| 沙漠 | 20+20 |
| 山地 | 50+50 |
| 水系 | 20+20 |
| 雪地 | 20+20 |

绿色的草地上

描述已自动生成地图

描述已自动生成卡通人物

中度可信度描述已自动生成山上的岩石

描述已自动生成

图 4‑1 草地自然背景图（左1）草地迷彩图（左2）

山地自然背景图（右1）山地迷彩图（右2）

## 实验设置

在CycleGAN模型的生成网络部分，转换器采用了6个残差块来将编码器提取到的特征进行转换，实现了从输入域到输出域的映射。2020年，Zhang.等人在ResNet网络的基础上进行改进，研究出了ResNeSt网络[24]。ResNeSt网络相比ResNet网络有四个方面的改进：第一点，引入了Split-Attention机制，将输入特征按照通道分组，并在组内进行channel attention，之后再将分组后的特征连接起来。这样可以使得模型更加关注不同的特征，并且学习到更加有效的特征表达。第二点，增加了更多的直接通道（residual connection），这些直接通道会跨过不同层级之间的split attention模块并且连接到后面的不同分组中，加速了训练和推理的过程。第三点，引入了复合缩放（Compound Scaling）技术，即同时调整网络的深度、宽度和分组参数，从而权衡准确性、速度和资源占用。第四点，采用了更高效的记忆布局模式，即分组卷积的方式，有效地减少模型存储需求，提高了性能并降低了内存占用。通过以上四点改进，使得ResNeSt网络可以在使用更少的参数和更少的计算量下实现比ResNet更好的性能。

鉴于同等情况下，ResNeSt网络比ResNet网络拥有更好的性能，因此尝试将CycleGAN模型的生成网络部分的RestNet网络更换为ResNeSt，并与原有RestNet网络性能相比较。实验部分将优化后的CycleGAN模型、优化前的CycleGAN模型以及传统使用K-means算法生成的数码迷彩的方法进行对比实验，评价三种生成方法的有效性。

## 评价指标设置

数码迷彩图像的有效性是一个比较难以衡量的问题，传统的人工评价法具有较强的主观性。考虑到数码迷彩的设计作用是为了在所提供的背景图片环境中实现隐蔽，因此在理论上，评估生成出的数码迷彩图与原背景图在结构上的相似性可以较好地反映出数码迷彩图的有效性，两者相似度越高，证明算法生成出的数码迷彩图越有效。

在图片相似性评估领域中，通常采用三种算法进行比较。第一种是基于直方图的算法，通过计算RGB 三个通道的直方图，然后利用这些直方图比较两张图片的相似性。这种方法具有简单易懂、计算量较低等优点，但是精度相较于其他算法较低。第二种算法是基于感知哈希的算法，该算法通过将图片处理成一个二进制值，来代表图片的特征。比较两张图片的二进制值，就能计算它们之间的相似程度。然而该算法对颜色和光照变化较为敏感，适用范围较小。在本课题的自制测试集中，很多场景中涉及到了复杂的光影和颜色组合，因此该算法亦不能胜任本课题的效果评价算法。第三种算法是基于结构相似性的算法（SSIM），通过比较两张图片在亮度、对比度及结构方面的相似性，来评估它们的相似程度。由于SSIM算法通常能够比较准确地模拟人眼对图像的感知，因此在图像质量评估的任务中表现较好。此外，它的计算简单、速度快，是本课题评价算法的不二之选。

具体原理方面，SSIM从亮度、对比度及结构三个维度来评估两张图片的相似性，因此在具体计算方面也分成了三个函数部分。描述两张图片的亮度对比关系的亮度函数：

描述两张图片对比度关系的对比度比较函数如4-2所示：

描述两张图片的结构对比关系的结构对比函数如式4-3所示：

、分别代表图像和的平均灰度，、分别代表图像x和y的标准差，代表图像x和y的协方差，其中常数部分，其中是视觉常数，是像素值的动态范围，的目的是为了保证分母为0是整体式子结构的稳定性。最终将这三个式子根据重要性比例调整，再精简化可得到最终的公式如式4-4所示：

## 消融实验

图示

描述已自动生成在CycleGAN模型的论文中提到，加入恒等映射损失，有助于鼓励在原始图像域和迁移图像域的映射上保留输入和输出之间的颜色组成。因此，可以在输入图像和生成的数码迷彩图像、输入图像和重建图像之间增加恒等映射损失函数来保留输入图片的颜色组成与输出图片的颜色组成一致，恒等映射损失加入位置如图4-2所示，其计算公式如式4-5所示：

图 4‑2 恒等映射损失添加位置示意图

为验证次损失函数的有效性，设置消融实验进行测试，具体实验设置及结果如表4-2所示。

表 4‑2 消融实验结果

|  |  |
| --- | --- |
|  | SSIM得分（200张平均） |
| 无 | 0.580 |
| 有 | 0.643 |

消融实验中，在训练集上训练了基础的CycleGAN模型以及引入损失后的CycleGAN模型，并在测试集上进行了测试。根据实验结果可以发现，引入损失后，测试集的SSIM平均得分相比原有模型提升了10.9%，这证明了引入损失的有效性。这可能是由于SSIM评分中色彩相似度占据比较重要的部分，因此生成图片的效果有了一定的提升，可视化实验结果如图4-3所示。

森林中间有树

描述已自动生成

图片包含 电子, 电路, 游戏机

描述已自动生成

图 4‑2 原始背景图像（左） 未加入数码迷彩（中） 加入数码迷彩（右）

文字图案

中度可信度描述已自动生成

## 实验结果分析

在4.5节中验证了损失函数的有效性，因此本小节的实验默认已经添加了该损失函数，模型总的损失函数如式4-6所示：

本小节的实验设置如4.2节中所提，验证CycleGAN模型相比传统K-means算法提取主色后再随机生成数码迷彩迷彩的方法是否有效果提升，同时验证将RestNet网络更换成ResNeSt网络是否有效果提升，具体实验设置及结果如表4-3所示。

表 4‑3 实验结果

|  |  |  |
| --- | --- | --- |
| 模型 | ResNet/ResNeSt | SSIM得分（200张平均） |
| CycleGAN | ResNet | 0.643 |
| CycleGAN | ResNeSt | 0.702 |
| K-means | - | 0.601 |

由表4-3结果，可以发现通过ResNeSt改进的CycleGAN模型在自制数据集的测

试集上效果表现最好，相比传统K-means算法SSIM得分提升了16.8%，相比基础的CycleGAN模型（加入色彩保持损失）SSIM得分提升了9.2%，证明了ResNeSt改进的有效性及CycleGAN模型对于处理数码迷彩生成问题的有效性。

卡通人物

中度可信度描述已自动生成文字图案

中度可信度描述已自动生成文字图案

中度可信度描述已自动生成森林中间有树

描述已自动生成

表 4‑3 实验结果

## 本章小结

本章对于CycleGAN模型进行了改进，优化了生成网络中的RestNet网络，替换成了性能更好的ResNeSt网络。随后介绍了SSIM算法判断两张图片相似性的原理，并以此为评价算法。同时，根据CycleGAN论文的指导，设置了消融实验，验证了恒等映射损失函数的有效性。最后，在测试集上分别计算优化后的CycleGAN模型、优化前的CycleGAN模型以及传统使用K-means算法生成的数码迷彩的方法生成的数码迷彩与原有背景图像之间的SSIM得分均值，验证了经ResNeSt网络改进后的CycleGAN模型生成效果最好。

# 总结与展望

## 本文工作总结

军事技术的发展对于一个国家而言有着至关重要的作用，随着侦查技术的不断发展，只有迷彩技术能同步发展才能更好的保护军事设施及参战人员。本文针对数码迷彩生成问题进行了深入透彻的研究，取得了一些阶段性的成果。

本文提出了使用优化后的CycleGAN模型来解决数码迷彩生成问题，通过向传统基础的CycleGAN模型中加入色彩补偿函数解决了生成的数码迷彩图像存在颜色缺失的问题，又将CycleGAN的生成网络中负责提取图像特征的ResNet残差网络模块替换成了ResNeSt网络，在自制数据集上相比传统算法及基础的CycleGAN模型有了较好的效果提升，验证了整体模型在数码迷彩生成问题上的可行性。

## 未来工作展望

目前，尽管通过ResNeSt网络优化后的CycleGAN模型已经在自制数据集的测试集部分取得了较好的效果，但在在生成数码迷彩图像过程中仍然存在效果不太理想的情况，特别是在原始背景图像过于复杂时，这可能与当前使用的数据集的数量还不足有关。因此，在下一步的研究中，我们应该致力于进一步优化数据集，在训练集中添加更多复杂场景的数据，以此优化模型的训练效果。同时，我们也应该考虑天气、季节等因素对于自然背景图像的影响，这些因素在数据集的设计和数码迷彩图像的生成过程中也发挥着至关重要的作用。

此外，当前模型只能处理单张自然背景图像进行数码迷彩图像生成，即一对一模式。然而，实际作战情况是多样的，不能完全依靠一张图片来进行全面描述。因此，未来我们需要对模型进行优化，并实现多对一生成模式，以提高迷彩效果。总的来说，未来的迷彩技术将会变得更加自适应，这意味着我们需要在数据集收集和设计方面更加注重环境的多样性，以使我们的模型表现更加稳健，生成更好的数码迷彩图像。

参考文献

1. 郑万里,杨萍,闫少强,吴丰轩,燕松.军事伪装技术研究现状及发展趋势分析[J].现代防御术,2022,50(01):81-86.
2. 喻钧,双晓.仿造数码迷彩的设计方法[J].应用科学学报,2012,30(04):331-334.
3. 徐英.基于背景代表色提取的迷彩伪装颜色选取算法[J].光电工程,2007(01):100-103+144.
4. 蔡云骧,薛士强,周志勇,黄云.基于分形布朗模型的数码迷彩图案生成方法研究[J].兵工学报,2016,37(01):186-192.
5. 贾其,吕绪良,吴超,荣先辉.马尔科夫随机场和金字塔模型用于数字迷彩图案设计[J].应用科学学报,2012,30(06):624-628.
6. Alfimtsev A N,Sakulin S A, Loktev D A, et al. Hostis Humani ET Mashinae: Adversarial CamouflageGenerationJ . Journal of Advanced Research in Dynamical and Control Systems, 2019, 11(2-SpecialIssue): 382-392
7. Yang X,Xu WD,Jia Q, et al.Research on extraction and reproduction of deformation camouflage spot based on generative adversarial network model - SaanceDirect[J]. Defence Technology,2020,16(3);555-563.
8. 张立新.基于自然背景的数码迷彩设计及伪装效果评价[J].西安工业大学学报,2019,39(03):358.
9. Wikipedia.Multi-scale camouflage[EB/OL].https://en.wikipedia.org/wiki/Multi-scale\_camouflage. [2021-04-22].
10. Friskovec M , Gabrijelcic H , Simoncic B . Design and Evaluation of a Camouflage Pattern for the Slovenian Urban Environment[J]. Journal of Imaging ence and Technology, 2010, 54(2):20507-1-20507-11(11).
11. Qi, JIA, Zai-yuan, et al. Research on Optimization Method and Experiment of Digital Camouflage Pattern[C]// 2018.
12. Smerity. Dithering: An Overview[EB/OL]. https://smerity.com/articles/2013/dithering.html.
13. Rosenblatt, F. The perceptron: a probabilistic model for information storage and organization in the brain.[J]. Psychological Review, 1958, 65:386-408.
14. Hopfield, J. J . Neural networks and physical systems with emergent collective computational abilities[J]. Proceedings of the National Academy of Sciences of the United States of America, 1982, 79(8):2554-2558.
15. Elman J L . Finding Structure in Time[J]. Cognitive Science, 1990, 14(2):179-211.
16. Hochreiter S , Schmidhuber J . Long Short-Term Memory[J]. Neural Computation, 1997, 9(8):1735-1780.
17. Goodfellow I , Pouget-Abadie J , Mirza M , et al. Generative Adversarial Nets[C]// Neural Information Processing Systems. MIT Press, 2014.
18. 秦雷胡江华付天齐. 一种新的数码迷彩生成方法[J]. 光电技术应用, 2013, 000(005):33-35,45.
19. Radford A , Metz L , Chintala S . Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks[J]. Computer ence, 2015.
20. Zhu J Y , Park T , Isola P , et al. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks[J]. IEEE, 2017.
21. Friskovec M , Gabrijelcic H . Development of a Procedure for Camouflage Pattern Design[J]. Fibres & Textiles in Eastern Europe, 2010, 18(4):p.68-76,5.
22. Xiao H , Qu Z , Lv M , et al. Fast Self-Adaptive Digital Camouflage Design Method Based on Deep Learning[J]. Applied Sciences, 2020, 10(15):5284.
23. Zhou W , Bovik A C , Sheikh H R , et al. Image quality assessment: from error visibility to structural similarity[J]. IEEE Trans Image Process, 2004, 13(4).
24. Zhang H , Wu C , Zhang Z , et al. ResNeSt: Split-Attention Networks[J]. 2020.
25. Fukushima K . Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position[J]. Biological Cybernetics, 1980, 36(4):193-202.
26. Lecun Y , Boser B , Denker J , et al. Backpropagation Applied to Handwritten Zip Code Recognition[J]. Neural Computation, 1989, 1(4):541-551.
27. Hinton G E , Salakhutdinov R R . Reducing the Dimensionality of Data with Neural Networks[J]. Science, 313.
28. Krizhevsky A , Sutskever I , Hinton G . ImageNet Classification with Deep Convolutional Neural Networks[J]. Advances in neural information processing systems, 2012, 25(2).

# 致 谢

行笔至此，甚是激动！这篇论文的完成初步标志着我的本科生涯即将告一段落，回顾本科四年的生活，首先第一个想到的词就是感谢，在这个过程中收获了太多的帮助，寥寥数语难表满腔感激之情。

首先要感谢我的父母，在无数个日子里，给予了我太多太多的关怀和鼓励，这些关怀和鼓励让我能一直充满勇气的面对各种各样的苦难和挫折。其次我要感谢我的导师赵志诚老师，赵老师是我学术上的领路人，为我打开了学术视野。在整个论文的撰写过程中，从选题、开题报告、中期报告以及毕业论文的撰写，每一个环节都倾注着赵老师的心血。同时，我还要感谢在毕业论文完成过程中曾帮助过我的苏菲老师、赵衍运老师和崔哲老师，正是因为有了诸位老师的指导，我才能更好的完成我的毕业论文。

除此之外，我想感谢赵睿宁师姐。睿宁师姐既是我来到MCPRL实验室的引路人，又是我的小导师。我永远无法忘记无数个日子里，师姐对我的耐心教导，面对我无数的提问，师姐永远都是耐心细致的为我解答。默默付出的背后，师姐仅用了一句“这是实验室的传承精神”回答了我，希望我也可以将这份精神传承下去。

最后，我想感谢大学四年的两位辅导员和身边的朋友们，大家在学习上、生活上对我曾有过太多的帮助。对此，如今的我无以为报，唯有怀着这份感激之情，更加努力地学习，争取有朝一日能成为一个对社会真正有用的人，不枉诸位于我身上的付出。

# 附 录

社会、健康、安全、法律、文化以及环境等非技术因素影响评估

本科毕业设计课题除了要达到毕业要求之外，还需要结合非技术因素如社会、健康、安全、法律、文化以及环境等因素进行考虑。下面将从制造成本、系统的供电问题、废弃物处理与循环再利用、绿色节能四个方面对本课题的内容进行评估。

在制造成本方面，本课题的研究内容主要是行人再识别算法的研究与创新，因此所有研究内容仅在软件上进行实现与验证。只需要一台个人计算机进行代码编写，一台具有高性能GPU的服务器提供深度学习计算能力即可。

在系统的供电问题上，本课题只在系统运转时需要对服务器进行供电，以提供深度学习算力。在应用过程中需要消耗一定的电量。

在废弃物处理与循环再利用方面，本课题实现的内容均在软件上完成，实现过程无废弃物的产生，因而不会对环境产生危害。

在绿色节能方面，在课题完成过程中，仅在模型训练时需要使用一部分计算机资源，此时电量消耗较多，在模型训练完成后，电量消耗将会大大减少。当系统不运行时，仅需要完成对计算机和服务器的正常供电即可，消耗电能较少，比较节能环保。

综上所述，本课题的实现满足低成本、低功耗、无污染等要求，符合法律规范，可通过社会、健康、安全、法律、文化以及环境的评估。