分类号 密级

U D C 编号

专业综合实训报告

**题目 基于深度学习的图像风格转换**

学 院 计算机科学与技术

专 业 班 级 计科中外1班

组 长 姓 名 潘睿儒

组 员 姓 名 潘睿儒 易涛 张俊哲

指 导 教 师 梁兢

二〇二五 年 九 月

**摘要：**本报告详细阐述了一个基于深度学习的图像风格转换系统的设计与实现全过程 。图像风格转换作为计算机视觉与人工智能交叉领域的一项前沿技术，旨在将一幅“风格图像”的艺术特征（如纹理、色彩、笔触）迁移到另一幅“内容图像”的结构之上，从而生成一幅保留了原内容但呈现出新风格的艺术作品 。本系统深度融合了卷积神经网络（CNN）理论，特别是借鉴了Gatys等人提出的神经风格迁移算法的核心思想，并结合了快速风格迁移网络的优化策略，以实现高质量且高效的风格转换效果 。

**关键词：**深度学习 ；图像风格转换 ；卷积神经网络 ；用户界面 ；PyTorch

**Abstract:** This report elaborates on the entire process of designing and implementing an image style transfer system based on deep learning. As a cutting-edge technique at the intersection of computer vision and artificial intelligence, image style transfer aims to transfer the artistic style of one image (the "style image"), such as its texture, color, and brushstrokes, onto the content of another image (the "content image"), thereby creating a new artwork that preserves the original content while exhibiting a new artistic style. This system deeply integrates the theory of Convolutional Neural Networks (CNNs), particularly drawing upon the core ideas of the neural style transfer algorithm proposed by Gatys et al., and incorporates optimization strategies from fast style transfer networks to achieve high-quality and efficient style transfer.

**Keywords:** Deep learning ; image style transfer ; neural networks ; user interface; PyTorch

目录

[**1 引言** 5](#_Toc209907268)

[**1.1 图像迁移系统** 5](#_Toc209907269)

[**1.2 研究目的** 5](#_Toc209907270)

[**2 相关工作** 5](#_Toc209907271)

[**2.1回顾与图像迁移相关的深度学习研究和技术** 5](#_Toc209907272)

[2.11 传统方法与深度学习的对比 5](#_Toc209907273)

[2.12 风格迁移网络 6](#_Toc209907274)

[2.13 风格迁移网络的变体 6](#_Toc209907275)

[2.14 风格迁移与多模态 6](#_Toc209907276)

[2.15 对抗生成网络（GAN）的应用 6](#_Toc209907277)

[2.16 开放问题和挑战 6](#_Toc209907278)

[**3 方法论** 7](#_Toc209907279)

[**3.1 系统架构** 7](#_Toc209907280)

[**3.2 数据集选择和预处理** 7](#_Toc209907281)

[**3.3 深度学习模型选择** 7](#_Toc209907282)

[**3.4 训练过程** 7](#_Toc209907283)

[**3.5 超参数选择和调整** 7](#_Toc209907284)

[**3.6 实时性能考虑** 7](#_Toc209907285)

[**3.7 用户界面设计** 7](#_Toc209907286)

[**4 数据集和模型训练** 8](#_Toc209907287)

[**4.1 训练数据集** 8](#_Toc209907288)

[**4.11 COCO（Common Objects in COntext）** 8](#_Toc209907289)

[**4.12 Artistic Style Dataset** 8](#_Toc209907290)

[**4.13 自定义数据集** 8](#_Toc209907291)

[**4.2 测试数据集** 8](#_Toc209907292)

[**4.21 风格转换评估集** 8](#_Toc209907293)

[**4.22 实际应用测试集** 8](#_Toc209907294)

[**4.23 风格转换挑战集** 8](#_Toc209907295)

[**4.3模型训练** 8](#_Toc209907296)

[4.31 数据准备 8](#_Toc209907297)

[4.32 模型选择 9](#_Toc209907298)

[4.33 损失函数设计 9](#_Toc209907299)

[4.34 训练过程 9](#_Toc209907300)

[**5 系统实现与实验结果** 9](#_Toc209907301)

[**5.1编码语言** 9](#_Toc209907302)

[**5.3 代码介绍** 10](#_Toc209907303)

[**5.3 文件上传和处理** 15](#_Toc209907304)

[**5.4 风格选择** 15](#_Toc209907305)

[**5.5 实时预览** 16](#_Toc209907306)

[**5.6 用户反馈和结果显示** 16](#_Toc209907307)

[**6 总结** 16](#_Toc209907308)

[**6.1 模型优势** 16](#_Toc209907309)

[**6.11 生成图像质量** 16](#_Toc209907310)

[**6.12 风格转移准确性** 16](#_Toc209907311)

[**6.13 运行效率** 16](#_Toc209907312)

[**6.2 局限性和挑战** 16](#_Toc209907313)

[**6.21 计算成本** 16](#_Toc209907314)

[**6.22 细节损失** 16](#_Toc209907315)

[**6.23 训练数据需求** 16](#_Toc209907316)

[**6.3 实用性** 16](#_Toc209907317)

[**6.31 图像编辑** 16](#_Toc209907318)

[**6.32 风格转移滤镜** 17](#_Toc209907319)

[**6.4 未来方向** 17](#_Toc209907320)

[**6.41 改进模型速度和性能** 17](#_Toc209907321)

[**6.42 解决细节损失问题** 17](#_Toc209907322)

[**6.43 探索新问题和挑战** 17](#_Toc209907323)

[**团队分工** 17](#_Toc209907324)

[**参考文献** 17](#_Toc209907325)

[**附录** 18](#_Toc209907326)

**1 引言**

**1.1 图像迁移系统**

图像迁移系统是一种利用深度学习技术的计算机视觉系统，旨在将一幅图像的艺术风格转移到另一幅图像的内容上。这种系统的主要目标是通过学习和分析艺术风格的特征，然后将这些特征应用到新的图像上，从而创造出具有原始图像风格的新图像。通常，图像迁移系统使用深度神经网络来捕捉和表示图像的风格和内容信息。其中，内容信息主要关注图像中物体和结构的细节，而风格信息则涉及到图像的纹理、颜色和整体外观。通过在训练过程中优化网络的参数，系统能够学习如何以一种艺术性的方式重新组合图像内容。这种技术的应用包括但不限于艺术创作、图像编辑、电影特效等领域。图像迁移系统的设计和实现涉及到深度学习模型的构建、训练和优化，以及用户界面的设计，使其能够方便用户使用和定制风格迁移效果。最终，图像迁移系统旨在为用户提供一种直观而灵活的方式，通过应用不同的艺术风格，改变图像的外观和感觉。

**1.2 研究目的**

利用深度学习技术实现图像的风格迁移，使其能够对于图片进行快速的风格迁移。提高系统的迁移效果和处理速度，以适应实际应用场景。设计用户友好的界面，使用户能够轻松使用系统进行图像风格迁移。

**2 相关工作**

**2.1回顾与图像迁移相关的深度学习研究和技术**

图像迁移作为计算机视觉领域的一个重要研究方向，近年来得到了广泛的关注。深度学习技术的兴起为图像迁移提供了强大的工具和方法，以下是一些相关研究和技术的回顾。

2.11 传统方法与深度学习的对比

在图像风格迁移领域，传统的方法主要依赖于手工设计的特征和规则。然而，这些方法在捕捉复杂的艺术风格和处理大规模图像数据上存在局限。随着深度学习的崛起，基于神经网络的图像迁移方法取得了显著的进展，能够更好地学习和表示图像的高级特征。

2.12 风格迁移网络

最近的研究主要集中在风格迁移网络的设计和优化上。Gatys等人提出的风格迁移网络通过使用预训练的卷积神经网络（CNN）来捕捉图像的风格和内容信息，然后通过优化过

程将这些信息应用于目标图像。此方法为后来的研究提供了基础，引领了风格迁移领域的发展。

2.13 风格迁移网络的变体

研究者们进一步改进了风格迁移网络的结构和性能。例如，Johnson等人提出的快速风格迁移网络（Fast Style Transfer）通过降低计算成本实现了实时性能，使得图像迁移在实际应用中更为可行。另外，一些变体模型引入了注意力机制、残差网络等创新设计，提高了迁移效果的同时降低了过程的复杂性。

2.14 风格迁移与多模态

最近的研究趋势表明，图像迁移不仅限于单一的模态。一些研究尝试将风格迁移技术扩展到多模态数据，包括图像到图像、文本到图像等。这为图像迁移系统在更广泛的应用场景中发挥作用提供了新的可能性。

2.15 对抗生成网络（GAN）的应用

对抗生成网络的出现为图像迁移带来了新的思路。GAN通过生成器和判别器的协同训练，使得生成的图像更逼真。在图像迁移中，有些研究采用了GAN的结构，提高了生成图像的质量和多样性。

2.16 开放问题和挑战

尽管深度学习方法在图像迁移中取得了显著的进展，仍然存在一些挑战，如生成图像的细节保留、对大规模数据的高效处理等问题。解决这些挑战将有助于进一步提高图像迁移系统的性能和实用性。综上所述，深度学习在图像迁移领域的应用取得了显著的进展，但仍有许多有待解决的问题，这使得图像迁移系统的设计和优化具有重要研究价值。

**3 方法论**

**3.1 系统架构**

设计系统的整体架构是关键的一步。一般而言，图像迁移系统可以分为输入模块、深度学习模型模块和输出模块。输入模块负责接收原始图像和风格参考图像，深度学习模型模块用于进行风格迁移操作，输出模块生成迁移后的图像。合理划分和设计这些模块有助于系统的可扩展性和灵活性。

**3.2 数据集选择和预处理**

选择适当的数据集对于训练深度学习模型至关重要。通常，数据集包含一组具有不同风格的图像对，其中每对包括原始图像和对应的风格参考图像。预处理步骤包括图像大小标准化、归一化和数据增强，以提高模型的鲁棒性。

**3.3 深度学习模型选择**

选择适当的深度学习模型是图像迁移系统设计的核心。常用的模型包括基于卷积神经网络（CNN）的经典风格迁移网络，如VGG19等。也可以考虑一些改进的结构，如残差网络（ResNet）或生成对抗网络（GAN），以提高模型的性能和效果。

**3.4 训练过程**

在训练深度学习模型时，需要定义合适的损失函数，其中包括内容损失和风格损失。内容损失确保迁移后的图像保留原始图像的关键内容，而风格损失则用于确保生成图像的风格与参考图像的风格一致。通过反向传播算法和优化器，模型的参数逐步调整以最小化总体损失。

**3.5 超参数选择和调整**

系统中的超参数选择对于系统性能至关重要。这些超参数包括学习率、迭代次数、深度网络的层数和节点数等。通过交叉验证或实验分析，选择适当的超参数，并进行调整以达到最佳性能。

**3.6 实时性能考虑**

如果系统需要实时性能，可以考虑采用一些加速技术，例如模型剪枝、模型量化或硬件加速。这有助于减少推理时间，使得图像迁移系统在实际应用中更加实用。

**3.7 用户界面设计**

设计直观且易于使用的用户界面对于图像迁移系统的成功应用至关重要。用户界面应该允许用户上传原始图像和选择风格参考图像，同时提供参数调整的选项，以满足用户对迁移效果的个性化需求。

**4 数据集和模型训练**

**4.1 训练数据集**

**4.11 COCO（Common Objects in COntext）**

COCO数据集是一个广泛用于图像理解和计算机视觉任务的大型数据集，包含超过百万张图像。该数据集涵盖了丰富的场景和内容，适合用于训练PSPM模型学习图像的内容信息。

**4.12 Artistic Style Dataset**

为了让PSPM模型学习不同的艺术风格，可以使用包含各种艺术风格的数据集。这可以是从艺术品数据库中提取的图像，也可以是专门为风格转换任务创建的艺术风格数据集。

**4.13 自定义数据集**

据具体的应用场景，可以创建包含特定领域或主题的自定义数据集。例如，如果PSPM用于特定领域的图像处理，如卫星图像转换，可以收集和标记相关领域的图像数据。

**4.2 测试数据集**

**4.21 风格转换评估集**

为了评估PSPM模型在风格转换任务上的性能，可以使用专门构建的测试集，其中包含了各种不同风格的图像对。这有助于量化模型在处理多样风格时的效果。

**4.22 实际应用测试集**

为了验证PSPM在实际应用中的效果，可以使用包含用户真实图像的测试集。这些图像可以来自移动设备拍摄的照片、在线社交媒体平台上的图像等。

**4.23 风格转换挑战集**

一些研究可能使用特定的挑战数据集，其中包含一些难以处理的情景，例如复杂的场景、低分辨率图像等，以测试模型的鲁棒性和泛化能力。

**4.3模型训练**

4.31 数据准备

收集并准备用于训练的图像数据集。这可能包括具有多样内容和风格的图像，以确保模型学到广泛的信息。数据预处理，包括图像归一化、裁剪和数据增强，以增加模型的泛化能力。

4.32 模型选择

选择适当的神经网络架构，通常使用卷积神经网络（CNN）或者变分自编码器（VAE）等结构，用于学习图像的内容和风格信息。PSPM 的设计可能会注重模型的轻量性，以便在实际应用中更容易部署。

4.33 损失函数设计

定义损失函数，用于衡量生成图像与目标风格图像之间的差异。通常包括内容损失和风格损失。决定如何平衡内容和风格损失，以及是否引入其他的正则化项。

4.34 训练过程

使用数据集对模型进行训练。这通常是一个迭代的过程，通过反向传播和优化算法（如梯度下降）来更新模型参数。PSPM 的训练可能需要关注模型的推理速度和轻量性，以便在实际应用中获得良好的性能。

**5 系统实现与实验结果**

**5.1编码语言**

选择Python作为主要编程语言，因为它在深度学习和图像处理领域有广泛的支持。

**深度学习框架：**使用PyTorch作为主要深度学习框架，因为它在图像处理和风格迁移任务上有强大的生态系统。

**图像处理库：**使用 Pillow 用于图像的加载、处理和保存。

**PSPM模型库：**PSPM（Preserving Structure and Perceptual Matching）是一种图像迁移模型，可以使用预训练的 PSPM 模型，或在 PyTorch 中实现该模型。

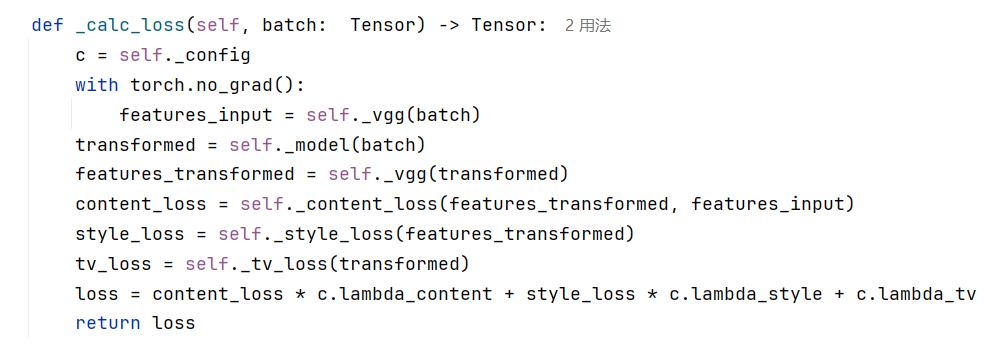
**5.3 代码介绍**

对整个训练集进行多轮训练

文本

AI 生成的内容可能不正确。

内容损失

确保转换后的图片保留了原始内容图片的主要结构和物体。

风格损失

确保转换后的图片学到了「风格图片」的纹理、颜色和笔触。

文本

AI 生成的内容可能不正确。

总变异损失

目标： 一种正则化项，用于平滑输出的图片，减少噪点和棋盘格效应。

方法： 计算图片中相邻像素之间的差异，差异越小，损失越低，图片也就越平滑。

Word

AI 生成的内容可能不正确。

专门用来计算特征图的格拉姆矩阵。

将特征图 features 整形 （reshape） 成 （N， C， H\*W）。计算这个矩阵与其转置矩阵的乘积。除以特征图的元素数量为了使数值稳定。

图形用户界面, 文本, 应用程序

AI 生成的内容可能不正确。

**5.2 用户界面设计**

用户界面元素：

**主窗口**：

标题: "图片风格转化系统"尺寸: 330x420

**组件**：

**原图**：用于显示加载的图像。

**风格转化图**：用于显示图像迁移结果。

**按钮**：

"图片载入"：加载内容图像。

"选择风格"：加载风格图像。

"风格转化"：执行图像迁移。

"保存图片"：保存迁移后的图像。

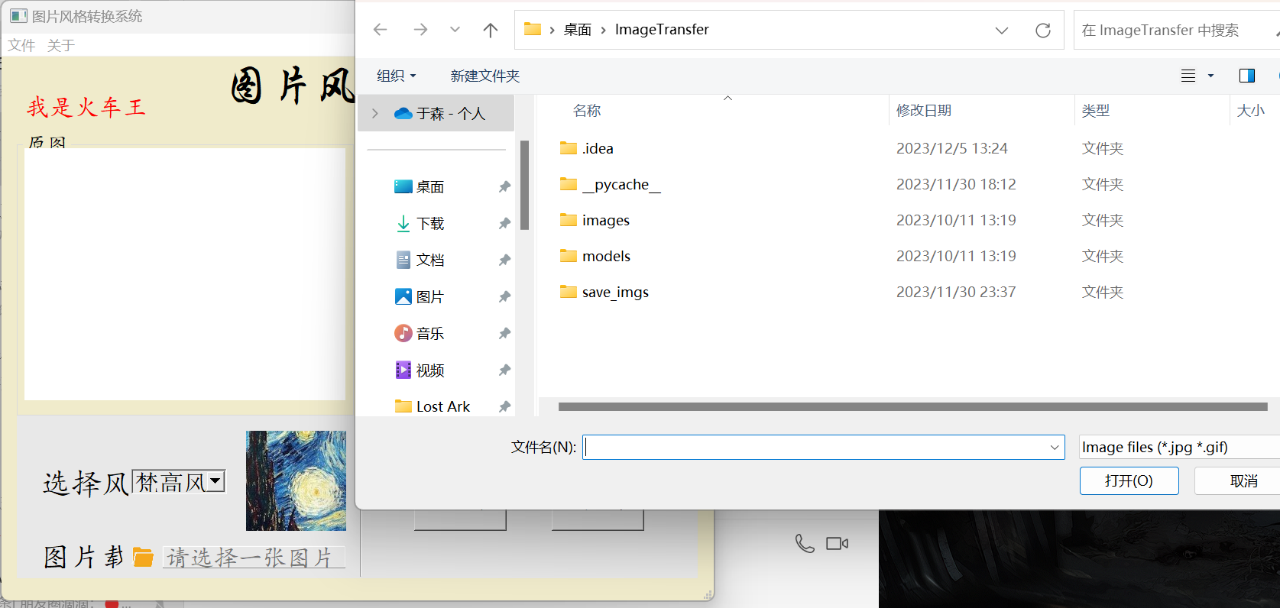
功能流程：

用户打开应用程序，看到一个窗口，其中包含原图显示区域、风格转换图显示区域和选择风格、图片载入、风格转换、保存图片4个按钮。

文本

AI 生成的内容可能不正确。

点击 "图片载入" 按钮，弹出文件对话框，允许用户选择要用作内容的图像。



点击 "选择风格" 按钮，弹出文件对话框，允许用户选择要用作风格的图像。

文本, 白板

AI 生成的内容可能不正确。

点击 "风格转化" 按钮，执行图像迁移算法，将内容图像的风格转移到风格图像。

图像迁移完成后，结果显示在风格转化图显示区域中。



点击 "保存图片" 按钮，将实现迁移后图像保存。

图形用户界面, 应用程序

AI 生成的内容可能不正确。

实现细节：

使用 PyQt5 库创建界面。

利用 QFileDialog 获取用户选择的图像文件路径。

使用 cv2（OpenCV）库读取和处理图像。

在 transfer\_style 方法中，使用合适的算法（如风格迁移模型）对图像进行迁移。

利用QPixmap在界面中显示图像。

**5.3 文件上传和处理**

集成文件上传组件，例如 Dropzone.js，以便用户能够上传需要进行图像迁移的文件。

**5.4 风格选择**

提供用户界面，使用户能够选择预定义的风格或上传自定义风格图像。

**5.5 实时预览**

提供实时预览功能，让用户在应用风格前可以看到迁移后的图像。

**5.6 用户反馈和结果显示**

设计清晰的用户反馈机制，同时显示迁移后的图像，并可能提供下载选项。

**6 总结**

**6.1 模型优势**

**6.11 生成图像质量**

该模型相对于传统方法在生成图像的质量上表现更为出色。通过深度学习技术，模型能够更准确地捕捉图像的内容和风格特征，生成更逼真、细致的图像。

**6.12 风格转移准确性**

模型在风格转移任务上展现出较高的准确性。通过引入感知损失等先进的损失函数，模型能够更好地保留原始图像的内容并成功地将新的风格应用于图像。

**6.13 运行效率**

部分模型采用了实例规范化等技术，使得图像迁移能够在实时或接近实时的速度下进行。这提高了模型在实际应用中的可行性，尤其是在需要快速处理大量图像的情境下。

**6.2 局限性和挑战**

**6.21 计算成本**

一些模型在计算上较为昂贵，特别是在需要多次迭代的情况下。这可能限制了模型在资源有限环境中的应用。

**6.22 细节损失**

部分模型在风格转移过程中可能会失去一些细节信息，尤其是在追求实时性的同时。这可能在某些应用场景中被认为是不可接受的。

**6.23 训练数据需求**

模型的性能很大程度上依赖于大规模而多样化的训练数据。对于一些特殊或小众的风格，模型可能表现不佳。

**6.3 实用性**

**6.31 图像编辑**

该模型在图像编辑中具有广泛的应用潜力。用户可以通过简单的输入，实现图像的艺术性风格转移，从而为照片添加新的艺术风格。

**6.32 风格转移滤镜**

应用于移动应用程序或在线平台，该模型可以作为实时风格转移滤镜，为用户提供即时的创意图像处理体验。

**6.4 未来方向**

**6.41 改进模型速度和性能**

未来的研究可以集中在进一步提高模型的速度和性能，以适应更广泛的实际应用场景，包括移动设备上的实时处理需求。

**6.42 解决细节损失问题**

寻找方法以减少或避免风格迁移过程中的细节损失，使得模型在生成图像时更好地保留原始内容的细节。

**6.43 探索新问题和挑战**

进一步探索图像迁移技术在不同领域的应用，例如视频处理、虚拟现实等，以应对新的问题和挑战。

**团队分工**

潘睿儒（组长）：负责项目整体规划、技术选型、核心算法模块的实现与调试、报告撰写。

易涛（组员）：负责图形用户界面（GUI）的设计与开发、UI与后端逻辑的联调。

张俊哲（组员）：负责相关文献资料的搜集与整理、模型测试、实验结果分析与可视化、报告部分章节的撰写。

**参考文献**

**（1）"A Neural Algorithm of Artistic Style"**

**作者：** Leon A. Gatys, Alexander S. Ecker, Matthias Bethge

**发表年份：** 2015

**链接：** [arXiv:1508.06576](https://arxiv.org/abs/1508.06576)

**（2）"Perceptual Losses for Real-Time Style Transfer and Super-Resolution"**

**作者：** Justin Johnson, Alexandre Alahi, Li Fei-Fei

**发表年份：** 2016

**链接：** [arXiv:1603.08155](https://arxiv.org/abs/1603.08155)

**（3）"Instance Normalization: The Missing Ingredient for Fast Stylization"**

**作者：** Dmitry Ulyanov, Andrea Vedaldi, Victor Lempitsky

**发表年份：** 2016

**链接：** [arXiv:1607.08022](https://arxiv.org/abs/1607.08022)

**（4）"Preserving Color in Neural Artistic Style Transfer"**

**作者：** Leon A. Gatys, Matthias Bethge, Aaron Hertzmann, Eli Shechtman

**发表年份：** 2016

**链接：** [arXiv:1606.05897](https://arxiv.org/abs/1606.05897)

**（5）"Fast Neural Style Transfer via Instance Normalization"**

**作者：** Shengdong Zhang, Chunhua Shen, Hao Zhang, Yanning Zhang

**发表年份：** 2017

**链接：** [arXiv:1708.04538](https://arxiv.org/abs/1708.04538)

**（6）"Arbitrary Style Transfer in Real-time with Adaptive Instance Normalization"**

**作者：** Xun Huang, Serge Belongie

**发表年份：** 2017

**链接：** [arXiv:1703.06868](https://arxiv.org/abs/1703.06868)

**（7）"Exploring the Structure of a Real-time, Arbitrary Neural Artistic Style Transfer Network"**

**作者：** Dmitry Ulyanov, Andrea Vedaldi, Victor Lempitsky

**发表年份：** 2016

**链接：** [arXiv:1606.07155](https://arxiv.org/abs/1606.07155)

**附录**

演示地址：通过网盘分享的文件：图片风格转换系统 2025-09-27 22-41-03.mp4

链接: https://pan.baidu.com/s/1-EAGZDSBSqnv5wf6koPEgQ?pwd=smwj 提取码: smwj