氏　　　　　　名　彭 錦程　〈※枠内は 12 ポイント〉
授　与　学　位　博士（工学）
学位授与年月日　令和 6 年 3 月 14 日
学位授与の根拠法規　学位規則第４条第１項
研　究　科　専　攻　秋田県立大学大学院システム科学技術研究科
博士後期課程総合システム科学専攻

学位論文題目　Image Processing Based on Improved Denoising Diffusion Probability Model and its Application in Medical Imaging
(改良ノイズ除去拡散確率モデルに基づく画像処理と医用画像への応用)

主　指　導　教　員　陳 国躍

論　文　審　査　委　員　主査　陳 国躍

副査　西口 正之　　陳 延偉（立命館大学）

猿田 和樹　　寺田 裕樹

※本研究科以外に所属する場合はその所属を括弧書きすること

論 文 内 容 要 旨　※(20 ポイント)

※用紙体裁：A4 縦、左右余白 20㎜、上下余白 30㎜、1 行 45 文字、40 行
※文字：MS 明朝 10.5 ポイント
※文字数： 4,500 文字程度

Super-resolution reconstruction technology, as a fundamental task in image processing, plays a crucial role in the field of computer vision, super-resolution reconstruction refers to the process of transforming a sequence of low-resolution and blurry pixel values into a high-resolution image through certain image processing algorithms. This process enriches the detailed information while preserving the original image structure, enlarging the image size, and supplementing high-frequency components to make the overall image clearer, super-resolution reconstruction technology has been widely used in various fields such as security surveillance, facial recognition, and medicine.

Previously, some super-resolution reconstruction techniques obtained low-resolution datasets by degrading images using simple bicubic interpolation, these methods were effective only for specific types of image degradation, whereas in the real world, various processes lead to image quality degradation. Hence, there is a need to enhance low-resolution datasets. Moreover, the latest denoising diffusion probability model (DDPM) defines a Markov chain for both forward diffusion and reverse

inference processes, it gradually adds random noise to the data and then reconstructs the desired data from the noise during the reverse diffusion process. This network exhibits excellent properties in generating realistic image texture features, a large body of literature suggests that DDPM has surpassed traditional generative adversarial networks in the synthetic domain, demonstrating outstanding performance in unconditional and conditional image generation. It excels in various computer vision research domains such as image synthesis, image translation, image restoration, image colorization, and image composition. However, the original DDPM model suffers from drawbacks such as large model parameter count, high GPU memory consumption, and being limited to processing images of only up to 256x256 in size due to device constraints. Additionally, the reverse inference time of the model is often excessively long.

Based on the motivations and research background outlined above, this paper aims to investigate an image processing method based on an improved denoising diffusion probability model. The focus is on enhancing the original denoising diffusion probability model for better application in image generation and super-resolution reconstruction techniques, aiming to achieve high-definition, efficient, and high-quality natural image processing through this image processing model approach. Additionally, this technology will be transferred to the field of medical image processing and other tasks such as colorizing image processing, the main results and contributions involve four aspects:

- Researching the degradation and deterioration processes of super-resolution inverse images, and augmenting the dataset with high-quality low-resolution images based on optimized degradation strategies.
- Studying the principles of denoising diffusion probability model and conditional denoising diffusion probability model super-resolution models.
- Investigating VAE and VQVAE latent generative models, integrating VAE and VQVAE latent models with denoising diffusion probability model networks, proposing an improved denoising diffusion probability model, and refining the model's image generation and super-resolution reconstruction tasks with a reasonable noise addition scheduling strategy.
- Transferring the improved model technology to medical super-resolution image processing and other computer vision image processing fields such as colorizing image diversity.

In Chapter 2, we conducted research on the degradation and deterioration processes of super-resolution inverse images, we focused on studying the factors of blur, downsampling, and noise, which lead to image degradation, based on the actual image sampling process. We aimed to consider as many real image degradation factors as possible and used a more realistic random mixed degradation strategy to augment the low-resolution dataset needed for training. Additionally, we investigated the generation processes of different latent generative models.

In Chapter 3, our focus was on studying the principles of the denoising diffusion probability model and its image generation process, we also researched the DDPM super-resolution reconstruction technique, unlike typical denoising diffusion probability models, super-resolution reconstruction requires constraining the solution space of the HR image. During the forward process, we stacked noise images of low-resolution LR and conditional high-resolution HR images together for conditional sampling. Then, we utilized the powerful parameter fitting capability of the U-net in the denoising diffusion probability model to fit the conditional feature update model. In the reverse inference process, we reconstructed the high-resolution image through iterations using the model with the low-resolution image as the guiding condition.

In Chapter 4, we first studied the noise addition scheduling strategy of the denoising diffusion probability model. By improving the noise addition scheduling strategy and introducing cosine time steps control during iterations, the noise distribution becomes more uniform, and diffusion becomes more stable. We also improved the original DDPM model by introducing VAE/VQVAE latent variable models as transitional bridges in the forward diffusion process and reverse inference, we compressed the original image into latent feature variables using the decoding module of VAE/VQVAE, synthesized it with the original DDPM model, and finally increased the resolution of the latent feature variables to match the resolution of the original generated image using the decoding module of VAE/VQVAE. This improvement method reduced the parameter count, saved GPU memory compared to the original DDPM model, and reduced inference time, the images generated by this method are more natural compared to those generated by the original model, and the high-resolution images reconstructed by VQVAE-DDPM are clearer.

Finally, in Chapter 5, experimental results demonstrated that the improved model has fewer parameters and shorter inference time. The VQVAE latent feature module we proposed enhanced feature extraction, making the semantic texture details of the images richer compared to the original model. In image generation and super-resolution tasks, the improved model generated high-quality, high-definition images. We also verified that the high-quality low-resolution images obtained through optimized degradation strategies resulted in better image reconstruction and effectively augmented the original dataset. Ultimately, our approach achieved advanced applications in medical super-resolution image processing and other computer vision image processing tasks.

論文提出者氏名	彭 錦程
論　文　題　目	Image Processing Based on Improved Denoising Diffusion Probability Model and its Application in Medical Imaging (改良ノイズ除去拡散確率モデルに基づく画像処理と医用画像への応用)
主 指 導 教 員	陳　国躍
副 指 導 教 員	猿田 和樹
論 文 審 査 委 員	主査　陳　国躍 ㊞ 副査　西口 正之 ㊞　陳 延偉（立命館大学）㊞ 猿田 和樹 ㊞　寺田 裕樹 ㊞

論文審査結果要旨

本論文は、改良ノイズ除去拡散確率モデル（DDPM: Denoising Diffusion Probability Model）に基づき、高速・高解像度に画像(医用画像)を処理すること、また画像のカラー化など多様な画像処理を実現することを目的とする。しかし、従来の DDPM モデルでは、モデルのパラメータ量が多い、GPU メモリ使用量が大きい、入力画像のサイズが小さい等、高速・高解像度に画像を処理できないという制限がある。そこで、本論文では、従来の DDPM モデルにいくつかの改良を加え、入力画像のサイズが大きい場合でも高速・高解像度の画像処理が可能な改良手法を提案した。本論文は全 6 章から構成されている。

第 1 章は序論であり、高解像度画像処理の背景、研究の目的について述べている。第 2 章では、画像の劣化と低品質化の原因、そして既存の超解像度処理技術、その応用について述べている。第 3 章では、最新のノイズ除去拡散確率モデル（DDPM）原理と超解像度処理技術について詳細に解析している。DDPM はディープラーニングの新たな展開を担う技術であり、微分可能な確率密度関数と密度比推定を組み合わせたアプローチを取り、多様な画像処理技術である。第 4 章では、オリジナルの DDPM モデルに、変分オートエンコーダモデル(VAE: Variational AutoEncoder）およびベクトル量子化変分オートエンコーダモデル（VQ-VAE: Vector Quantised-Variational AutoEncoder）のエンコーダーモジュールを使用して画像を圧縮し、より低い解像度の特徴変数を取得することで、圧縮された特徴変数を DDPM で合成し、入力画像のサイズが大きい場合でも高速・高解像度の画像処理が可能な改良手法を提案した。第 5 章では、改良型 DDPM モデルで高速・高解像度に画像(医用画像）処理できることを検証実験により確認した。最後に、第 6 章は本研究の結論であり、本研究をまとめ、関連研究課題について述べている。

以上、本論文では、高速・高解像度画像(医用画像)処理に対する改良型 DDPM モデルによる知見と成果が得られており、高速・高解像度画像処理技術の発展に寄与するところが大きいと考えられる。

よって、本論文は博士（工学）の学位論文として合格と認める。