性能と軽量さのトレードオフを考慮した画像変換深層学習ネットワークに関する研究

2024 年度

柴崎　圭

別表５
(3)

主　論　文　要　旨

No.

報告番号	(甲)乙　第　　号	氏　名	柴﨑 圭

主 論 文 題 名：
性能と軽量さのトレードオフを考慮した画像変換深層学習ネットワークに関する研究

（内容の要旨）

コンピューティング技術やビッグデータ処理技術の向上により，デジタル画像及び動画像を変換，復元する技術の需要が高まっている．深層学習は画像の変換といった複雑なタスクに対する有効な手法であるが，計算量が多くなるという問題がある．さらに，深層学習ベースの手法は多くの場合において性能と手法の軽量さの間にトレードオフが存在する．本研究では，多様な 4 つのタスクにおいて，新規モジュールを採用したネットワークを用いることと，効果的な処理スキームや損失関数を導入するという 2 つのアプローチから，タスクごとに性能と軽量さを両立した画像変換深層学習ネットワークを提案している．

第 1 章では，研究背景及び研究目的について述べている．

第 2 章では，本研究で取り扱うデジタル動画像や深層学習の基礎理論について説明し，第 3 章では，本研究で取り組むタスクである，フォトレタッチ，動画のノイズ除去，姿勢変換，非ペア画像変換に関する従来研究について概説している．

第 4 章では，フォトレタッチというタスクに対して，ラプラシアンピラミッドで分解された低周波画像に関して Axial Transformer を含む特徴抽出能力の高いブロックを採用した深層学習ネットワークを提案し，従来法に比べ計算量を大きく削減しつつ高性能な画像変換を達成した．

第 5 章では，動画のノイズ除去というタスクに対して，軽量で高効率であるモダン ConvBlock を採用し，3 次元畳み込みのような直接時間軸の相互作用を行うことなく，疑似的に時間的な関係性を捉える Pseudo Temporal Fusion を新たに提案することで，性能と軽量さを両立した動画のノイズ除去ネットワークを提案した．

第 6 章では，姿勢変換というタスクを「大まかな姿勢の変換」と「詳細なテクスチャの生成」という 2 つのサブタスクに分解し，それぞれを異なる適切なモジュールで処理する推論スキームを提案することで効率的な変換を実現した．

第 7 章では，非ペア画像変換というタスクに対して，訓練済みの salient object detection ネットワークで抽出された saliency domain を同時に変換する．変換後の画像だけではなく，変換した saliency domain や，それをもとにした画像の前景に対しても損失を計算することでネットワークの変換リソースを変換するべき前景に集中することが可能となり、ネットワークを軽量にしつつ高水準の変換性能を達成した．

第 8 章では，本研究によって得られた成果を結論としてまとめ，今後の展望を述べている．

別表６
（３）

Keio University

Thesis Abstract

No.

Registration Number	□ "KOU" □ "OTSU" No. *Office use only	Name	Kei Shibasaki

Thesis Title

A Study on Image Transformation Deep Learning Networks Considering the Trade-off between Performance and Lightweight

Thesis Summary

Improvements in computing and big data processing technologies have increased the demand for technologies to transform and restore digital and video images. Deep learning is an effective method for complex tasks such as image transformation, but it is computationally expensive. Furthermore, deep learning-based methods often involve a trade-off between performance and method lightness. In this study, we propose image transformation deep learning methods that achieves both performance and lightweight for each task by using a network with novel modules and introducing effective processing schemes and loss functions for four diverse tasks.

Chapter 1 describes the research background and objectives for the study.

Chapter 2 describes the basic theory of digital video and deep learning that we deal with in this research.

Chapter 3 outlines previous research on photo retouching, video denoising, pose transformation, and unpaired image to image transformation, which are the tasks we address in this research.

In Chapter 4, for photo retouching, we proposed a deep learning network employing blocks with high feature extraction capability, including Axial Transformer, with respect to low-frequency images decomposed by Laplacian pyramids. Compared to conventional methods, the proposed network achieves high performance image transformation while significantly reducing computational complexity.

In Chapter 5, for the task of video denoising, we employed the lightweight and highly efficient Modern ConvBlock and proposed a novel Pseudo Temporal Fusion that captures temporal relationships in a pseudo manner without direct temporal interaction like 3D convolution. We proposed a video denoising network that achieves both performance and light weight.

In Chapter 6, we decompose the task of pose transformation into two subtasks, "rough pose transformation" and "detailed texture generation," and propose a processing scheme in which each is processed by a different appropriate module to achieve efficient transformation.

In Chapter 7, we simultaneously transform saliency domains extracted by a pretrained salient object detection network for unpaired image to image transformation. By computing the loss not only for the transformed image but also for the transformed saliency domain and the foreground of the image from which it was derived, the network's transformation resources can be focused on the foreground to be transformed, achieving a high level of transformation performance while reducing the network weight.

Chapter 8 describes the conclusions and results of each task, their contribution to solving the issues raised by this study, and future prospects.