

論文 / 著書情報  
Article / Book Information

題目(和文)	多様な劣化に対応可能な画像認識ネットワークに関する研究
Title(English)	
著者(和文)	遠藤和紀
Author(English)	Kazuki Endo
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第11744号, 授与年月日:2022年3月26日, 学位の種別:課程博士, 審査員:田中 正行,奥富 正敏,蜂屋 弘之,倉林 大輔,原 精一郎
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第11744号, Conferred date:2022/3/26, Degree Type:Course doctor, Examiner:,,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	論文要旨
Type(English)	Summary

## 論文要旨

THESIS SUMMARY

系・コース： Department of, Graduate major in	システム制御 システム制御	系 コース	申請学位 (専攻分野)： Academic Degree Requested	博士 Doctor of	(工学)
学生氏名： Student's Name	遠藤 和紀		指導教員 (主)： Academic Supervisor(main)	准教授 田中 正行	
			指導教員 (副)： Academic Supervisor(sub)	教授 奥富 正敏	

### 要旨 (和文 2000 字程度)

Thesis Summary (approx.2000 Japanese Characters )

デジタルカメラ等によって撮影された被写体を、コンピュータを用いて認識する画像認識タスクは、畳み込みニューラルネットワーク (CNN) を用いた深層学習の登場により、近年飛躍的な精度向上を遂げ、多くの研究成果が報告されている。ここで入力画像に着目すると、そのような画像認識タスクの多くは、圧縮やノイズ、ぼけ等の劣化のない画像を入力画像としており、劣化のある画像 (以下、劣化画像という) を入力とした研究はさほど多くはない。しかしながら、Web 上で取得できるデジタル画像には、通常 JPEG 圧縮等の画像処理が施されていること、また自動運転車の車載カメラの映像には、半導体の温度上昇に伴い発生するノイズが含まれること等、入力画像には劣化が含まれる。したがって、画像認識タスクの社会実装では、入力画像として劣化画像を扱うことが必要であり現実的である。

さて、画像の劣化には JPEG 圧縮やノイズ等の様々な種類があり、更に各劣化に対して圧縮率やノイズレベル等の劣化の程度 (以下、劣化水準という) が存在する。そのため、CNN を用いて劣化画像を認識する最も単純な方法は、様々な劣化水準の劣化画像を学習データに加えることである。しかしながら、この方法には、全ての劣化水準に対して平均的に CNN の学習が進み、原画像や高品質画像について十分な認識精度が得られるとは限らないという問題がある。本論文の目的は、この問題の克服を企図し、劣化の種類が単一かつ既知で、劣化水準が未知である前提の下、CNN を用いて多様な劣化に対応可能な画像認識ネットワークを構築することである。

本論文は全6章で構成され、各章の概要は以下のとおりである。

第1章「緒言」では、本研究の背景や目的、本論文の構成を述べた。

第2章「劣化水準を利用したクラス分類性能の改善」では、劣化画像と劣化水準を入力として用いるクラス分類 CNN を提案した。ここで、クラス分類 CNN の前処理として、劣化水準推定 CNN を導入することで、ネットワーク全体では劣化画像のみを入力とした。第2章の提案手法は、劣化画像のみで学習したクラス分類 CNN に比べ、ほとんど全ての劣化水準に対して精度が高く、劣化水準の入力が分類精度向上に寄与することを確認した。しかしながら、高品質画像に着目すると、第2章の提案手法は、原画像のみで学習したクラス分類 CNN よりも精度が低い。そのため、多様な劣化水準に対応するためには、高品質画像とそれ以外の画像で、クラス分類 CNN を使い分ける必要がある。

第3章「アンサンブル学習を用いた多様な劣化水準に対する性能の改善」では、第2章の課題であった高品質画像に対するクラス分類性能を改善するため、原画像のみで学習したクラス分類 CNN と復元画像で学習したクラス分類 CNN からなるアンサンブルネットワークを提案した。ここで、クラス分類性能を強化するため、劣化画像を画像復元器により復元し、復元画像をクラス分類 CNN の入力とした。また、アンサンブル時の重みは、劣化画像の劣化水準推定値や復元画像の特徴量に応じてネットワークにより自動的に決定される。第3章の提案手法は、第2章の提案手法や既存手法に比べ、高品質画像についても高いクラス分類精度を示した。しかしながら、原画像に近い高品質画像の分類精度については、原画像のみで学習したクラス分類 CNN に劣ることがある。

第4章「原画像のクラス分類精度を維持した consistency regularization に基づくマルチタスク学習」では、第3章の課題であった原画像に近い高品質画像に対するクラス分類性能を改善するため、クラス分類と劣化水準推定のマルチタスク学習を用いたクラス分類ネットワークを提案した。Consistency regularization に基づき、「原画像のみを学習したクラス分類 CNN (ソースネットワーク) が出力する原画像の特徴量」と「劣化画像を学習するクラス分類 CNN (ターゲットネットワーク) が出力する劣化画像の特徴量」が一致するよう、ターゲットネットワークを学習した。ターゲットネットワークは、ソースネットワークと同一の特徴量抽出器を有し、入力された劣化画像に応じて特徴量を補正する構造を持つ。更に、ターゲットネットワークのクラス分類器をソースネットワークと同一とすることで、原画像のクラス分類性能を再現できる構造とした。第4章の提案手法により、原画像のクラス分類性能を損なうことなく、多様な劣化水準の劣化画像を分類できることを示した。また、劣化水準のマルチタスク学習は、クラス分類性能の向上に寄与することを確認した。

第5章「セマンティックセグメンテーションへの応用」では、第4章の提案手法を劣化画像のセマンティックセグメンテーションに拡張した。実験に用いたセマンティックセグメンテーションネットワークは、クラス分類ネットワークと異なり、画像特徴量の間出力値や空間情報をデコーダーで再利用する。そのため、デコーダーを固定せず再学習することで、既存手法よりも性能が改善することを確認し、第4章の提案手法がセマンティックセグメンテーションへも応用可能であることを示した。

最後に、第6章「結言」では、本論文のまとめと今後の課題を述べた。

備考：論文要旨は、和文 2000 字と英文 300 語を 1 部ずつ提出するか、もしくは英文 800 語を 1 部提出してください。

Note : Thesis Summary should be submitted in either a copy of 2000 Japanese Characters and 300 Words (English) or 1copy of 800 Words (English).

注意：論文要旨は、東工大リサーチリポジトリ(T2R2)にてインターネット公表されますので、公表可能な範囲の内容で作成してください。

Attention: Thesis Summary will be published on Tokyo Tech Research Repository Website (T2R2).

(博士課程)  
Doctoral Program

## 論文要旨

THESIS SUMMARY

系・コース： システム制御 系  
Department of, Graduate major in システム制御 コース  
学生氏名： 遠藤 和紀  
Student's Name

申請学位 (専攻分野)： 博士 (工学)  
Academic Degree Requested Doctor of

指導教員 (主)： 准教授 田中 正行  
Academic Supervisor(main)

指導教員 (副)： 教授 奥富 正敏  
Academic Supervisor(sub)

要旨 (英文 300 語程度)

Thesis Summary (approx.300 English Words)

Image recognition has been progressed remarkably by using a convolutional neural network. Many studies of image recognition assume that input images are clean without any degradation. However, there are usually some degradations in practical applications, such as autonomous driving, security camera, etc. Moreover, degradation levels, which represent degrees of degradation, are generally unknown. Therefore, it is necessary for an image recognition network to cope with several levels of degradation. This thesis aims to construct an image recognition network of degraded images under several levels of degradation.

This thesis has proposed three approaches as follows. The first approach was to improve the classification performance of degraded images by using a degradation level as an additional input. This approach underperformed a classification network trained with only clean images for high-quality images. Therefore, in order to deal with various degradation levels, it was necessary to use two classification networks properly depending on image qualities though it was inconvenient in practice.

Secondly, to overcome this inconvenience, an ensemble network composed of two classification networks was proposed: trained with clean images only and trained with restored images. The restored images were acquired from degraded images by a restoration network, and ensemble weights were inferred depending on an estimated degradation level and features of a restored image. The proposed ensemble network could classify degraded images well for almost all degradation levels. However, it sometimes showed lower classification accuracy for high-quality images, which were like clean images, than a classification network trained with only clean images.

Finally, not to low the classification performance of such high-quality images, this thesis proposed a convolutional neural network that had an adjustment structure to deal with the features of degraded images. The proposed network was trained by multi-task learning based on consistency regularization to classify degraded images and estimate degradation levels. The proposed network could classify degraded images without sacrificing the classification performance of clean images. Moreover, the proposed network was extended to the semantic segmentation of degraded images.

備考：論文要旨は、和文 2000 字と英文 300 語を 1 部ずつ提出するか、もしくは英文 800 語を 1 部提出してください。

Note: Thesis Summary should be submitted in either a copy of 2000 Japanese Characters and 300 Words (English) or 1copy of 800 Words (English).

注意：論文要旨は、東工大リサーチリポジトリ(T2R2)にてインターネット公表されますので、公表可能な範囲の内容で作成してください。

Attention: Thesis Summary will be published on Tokyo Tech Research Repository Website (T2R2).