

令和 4 年 6 月 23 日現在

機関番号：33924

研究種目：基盤研究(C)（一般）

研究期間：2019～2021

課題番号：19K12129

研究課題名（和文）タスク指向画像処理のための深層学習と、遠方歩行者認識への応用

研究課題名（英文）Deep Learning for Task-Driven Image Processing and its Application to Distant Pedestrian Detection

研究代表者

浮田 宗伯 (Ukita, Norimichi)

豊田工業大学・工学（系）研究科（研究院）・教授

研究者番号：20343270

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：タスク指向画像処理の実例として、「遠方歩行者検出（微小物体検出）」をタスクとして、このタスク性能を向上させるための前処理として「画像超解像」に焦点を当てて研究を進めた。これら画像超解像と微小物体検出を独立に解くのではなく、1つの深層学習ネットワークで統合的に学習し、微小物体検出というタスクの性能が向上するという目的関数を新たに提案することで、人が画像を見た際の奇麗さを追求した画像超解像ではなく、機械（人工知能）が画像の意味を理解するための画像超解像（画像処理）を実現した。

研究成果の学術的意義や社会的意義

タスク指向超解像のための基礎技術として、汎用的な画像超解像を研究し、世界的な協議会でも上位入賞する性能を実現した。画像超解像は、人が鑑賞する画像を生成するという従来型のタスクにおいても、記録媒体に保存する画像ファイル容量の圧縮や、Youtubeや遠隔会議における映像配信など、多様な応用において実用的な技術である。

また、本研究で実現した微小物体検出は、車載画像における遠方物体検出などセキュリティや安全を目的にした多様な分野に波及する技術である。

研究成果の概要（英文）：As an examples of task-driven image processing, we put our focus on tiny object detection and image super-resolution as a task and an image processing algorithm, respectively. Different from previous approaches in which these image processing and task are achieved independently, our proposed framework (i) integrates these two sub-problems in a single neural network and (ii) trains this network so that the reconstructed super-resolution image is optimized for improving the tiny object detection task. With this framework, we realize image super-resolution for generating images that are understandable by artificial intelligence instead of image super-resolution for images that are observed by human.

研究分野：画像認識

キーワード：画像超解像 物体検出 画像処理

1. 研究開始当初の背景

ボケ除去や超解像などの画像処理の目的は「人が画像を観察する際の印象向上」と「計算機（人工知能）による画像認識性能の向上」に大別される。画像の研究初期において、画像処理の結果は画像認識に大きな影響を与えていた。しかし、処理結果が認識に使われるだけで、この2つの技術は独立に研究されていた。更に、近年の深層学習による画像特徴の最適化により、画像認識の前処理としての画像処理は存在感を失ってきた。しかし、画像認識同様、深層学習により躍進した画像処理が、画像認識に与える影響は小さくないと考える。以上の議論を踏まえ、本研究では以下の2点に焦点を当てる。

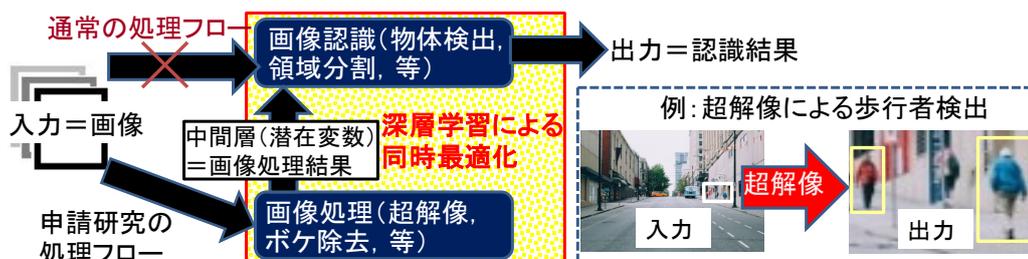


図1：タスク指向画像処理の深層学習の最適化

【人の為の画像処理と人工知能の為の画像処理の違い】画像処理では、昔から近年の深層学習に至るまで、PSNRなどの画質基準が評価されてきた。しかし、人の視覚的印象とPSNRは必ずしも一致しないことが分かっている。また物体認識能力についても、人間の知覚と深層学習に代表される人工知能の間には大きな違いがある。

そこで申請者は、画像認識のような人工知能の性能向上とした場合、それに適した画像処理の評価はPSNRのような画質基準が最適ではないと考える。すなわち、「情報欠損・劣化した画像の復元」という不良設定問題である画像処理において、「各領域の正確な復元がどの程度必要または不要か」という復元内容の取舍選択は、認識問題（例：物体検出、行動認識）や認識対象（例：人の検出、車の検出）によって異なると予測する。よって、画像処理最適化の段階で、画像認識というタスクの性能向上を考慮するタスク指向画像処理が重要と考える（図1）。

【画像処理と画像認識の統合最適化の有用性】深層学習は、事象間の複雑な関係もモデル化できる。そこで、画像処理と画像認識を独立に最適化するのではなく、同時最適化することで両者の性能を向上できると予測する。この相互最適化のため、図1の「入力＝画像」→「中間層（潜在変数）＝画像処理結果」→「出力＝認識結果」という構造によって、画像処理結果に相当する中間層を教師付き学習で最適化することにより、深層学習の可読性向上や多層の同時最適化による局所解の回避という課題に取り組む。

2. 研究の目的

【人のための画像処理と人工知能のための画像処理の違い】画像認識（人工知能）のための画像処理の評価基準と、人の視覚のための画像処理の評価基準とが異なるという予測を実験的に検証する。検証実験には、画像処理と画像認識の統合最適化モデルを利用する。

【深層学習による画像処理と画像認識の統合最適化】入力画像から復元画像を獲得し、その復元画像に基づいて画像認識する一連の処理を、一つの深層学習で統合的に最適化する方法論が、本研究の技術目標である。復元画像を深層学習で潜在変数として扱うことで、深層学習の可視化や局所解回避を実現しつつ、「画像復元という画像処理」と「タスクである画像認識」の有機的な統合は、本研究の学術的独自性である。

上記研究の実例として、運転支援や自動運転に応用できる「車載カメラの超解像による遠方歩行者の検出」を研究する。

【遠方歩行者の超解像】安全な衝突回避には、歩行者が遠く、車載映像中で小さく映っている時点での車両制御が望ましい。また、車載映像はブレなどの画像劣化を含む。そこで、従来倍率を超えた超解像と劣化復元を目指す。

【遠方歩行者の検出】飛び出し等の危険行動を認識できると、安全に車両制御できる。本研究では、従来技術の歩行者検出と比べ大幅に難易度のあがる「低解像画像な遠方歩行者像での歩行者検出」を実現する。

3. 研究の方法

【概要】タスク指向画像処理のための深層学習では、画像処理の種類に対する汎用性を検証するため、超解像と車載カメラの動きに起因するブレの除去を例に研究を進める。更に、超解像は、静止画から動画へと対象を拡張する。画像認識（タスク）に対する汎用性の検証のためには、人領域の検出に提案手法を適用する。

【画像処理】タスク指向画像処理の目的は、最終的な画像認識の性能向上であるが、中間層の画像処理結果も定量的に評価する。

【画像認識】歩行者の位置の認識（すなわち、歩行者検出）を検証する。

【人の為の画像処理と人工知能の為の画像処理の違い】タスク指向画像処理の結果を、画質基準の PSNR と認識成功率という 2 つの基準で評価し、その関係性を検証する。

【深層学習による画像処理と画像認識の統合最適化】タスク指向画像処理の有無が画像認識結果に与える影響を定量的に評価する。

【実用に向けた速度検証】高い計算リソースによる巨大な深層学習の性能限界を示すと共に、車載応用を視野に入れ、リーズナブルな計算リソースでの実時間処理の可能性も検討する。具体的には、認識性能と計算コスト（例：深層学習器のサイズ）のトレードオフを検証する。

4. 研究成果

【画像処理】

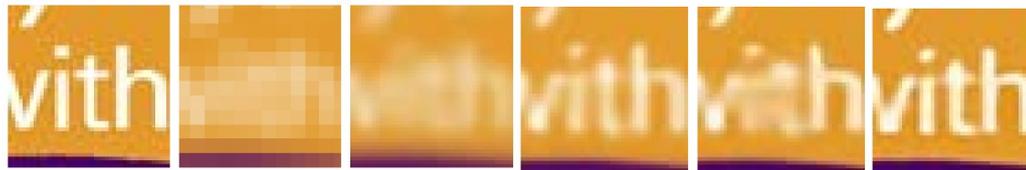
本科研費研究開始前に提案していた静止画超解像のための手法（Deep Back-Projection Network; DBPN）を、動画超解像手法（Recurrent Back-Projection Network; RBPN）に拡張した。RBPN では、超解像のターゲットとなるフレームを DBPN と同じネットワークで復元すると同時に、そのターゲットフレームに時間的に隣接するフレームの超解像特徴量を計算し、複数連続フレームの超解像特徴量を再帰型ニューラルネットワークによって逐次統合していくことにより、ターゲットフレームの超解像精度を向上させた。その精度は、従来の類似他手法と比べて、定性的にも定量的にも上回った。その超解像結果を図 2 に示す。



(a) 画像補間 (b) 従来法 TOFlow (c) 従来法 VSR_DUF (d) RBPN : 3フレーム入力 (e) RBPN : 6フレーム入力 (f) 正解画像

図 2 : 動画超解像の出力画像例

任意かつ未知の動きボケを含んだ低解像度画像の超解像のためには、そのボケカーネルを低解像度画像から推定し、そのボケカーネルを条件付けして画像超解像を実現する手法（Kernelized Back-Projection Network）を提案した。この手法も、従来の類似他手法と比べて、定性的にも定量的にも上回った。その超解像結果を図 3 に示す。



(a)正解画像 (b)低解像度画像 (c) DBPN (d) 従来法IKC (e) 従来法KOALA (f) KBPN

図3：ボケを含んだ低解像度画像からの超解像画像例

【画像認識】と【深層学習による画像処理と画像認識の統合最適化】

提案するタスク指向型超解像は、超解像ネットワークとタスクネットワーク（本研究では物体検出ネットワーク）が直列に接続されている。超解像ネットワークと物体検出ネットワークは、それぞれ独立に事前学習しておく。事前学習された2つのネットワークを接続し、最後に物体検出ネットワークの学習のための損失関数を、物体検出ネットワークを介して、その前段の超解像ネットワークまでエンドトゥエンドで統合学習する。この統合学習によって、物体検出精度を向上できることを確認した。精度向上した物体検出例を、図4に示す。

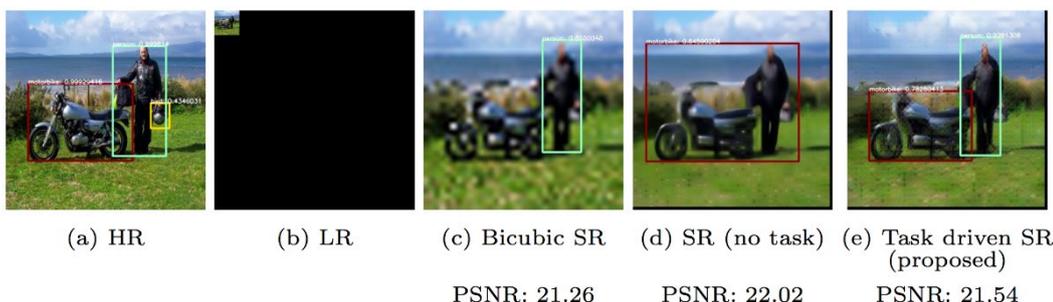


図4：タスク指向型物体検出の成功例。(a)に示す元の高解像度画像であれば、正しく物体検出可能。しかし、(b)の低解像度画像からは物体検出不可能。一般的な画像補間だけで画像拡大した(c)では、人だけが検出できた。物体検出とは独立に学習された超解像 DBPN によって得られた画像(d)では、「人とバイク」を含めて「人」として誤検出してしまった。これに対して提案するタスク指向型超解像による物体検出(e)では正しい物体検出が可能になった。

【人の為の画像処理と人工知能の為の画像処理の違い】

人の為の画像処理（超解像）と人工知能の為の画像処理（超解像）の結果を、一般的な画像復元結果の定量的指標である PSNR によって評価してみた。PSNR は、大きいほど精度が良いことを示す。その結果例を図4に示す。単純な画像補間による画像拡大(c)では、PSNR = 21.26 になったが、提案する画像超解像 DBPN によってこれを上回る PSNR = 22.02 が得られた。これに対して、タスク指向型超解像(e)の結果は、PSNR = 21.54 のように一般的な画像復元評価指標は下回ってしまった。これは人の為の超解像精度は落ちてしまったと解釈される。一方、前述のように(e)の結果によって物体検出性能は向上した。すなわち、人工知能の為の画像処理結果は向上したといえる。このことから、人の為の見た目向上の画像処理と、人工知能が画像の内容を理解するための画像処理とでは、精度向上の基準が異なる可能性が示唆されたといえる。

【実用に向けた速度検証】

超解像や物体検出に限らず、ニューラルネットワークの実行速度はその構成パラメータ数に大きく依存する。パラメータ数が少ないほど実行速度は高速である。一方、パラメータ数が大きいほど、性能は向上（超解像であれば復元画像精度は向上）する。そこで、提案する静止画超解像手法 DBPN のパラメータ数と復元精度の関係を、類似他手法と比較した。その結果を図5に示す。DBPN はそのパラメータ数を変化させた DBPN-SS, DBPN-S, DBPN-M, DBPN-L, D-DBPN (パラメータ数が小から大順にならべた)の結果を示す。図5からわかるように、DBPN はパラメータ数と復元精度のトレードオフにおいて、いずれの類似他手法よりも上回っていることが確認できる。

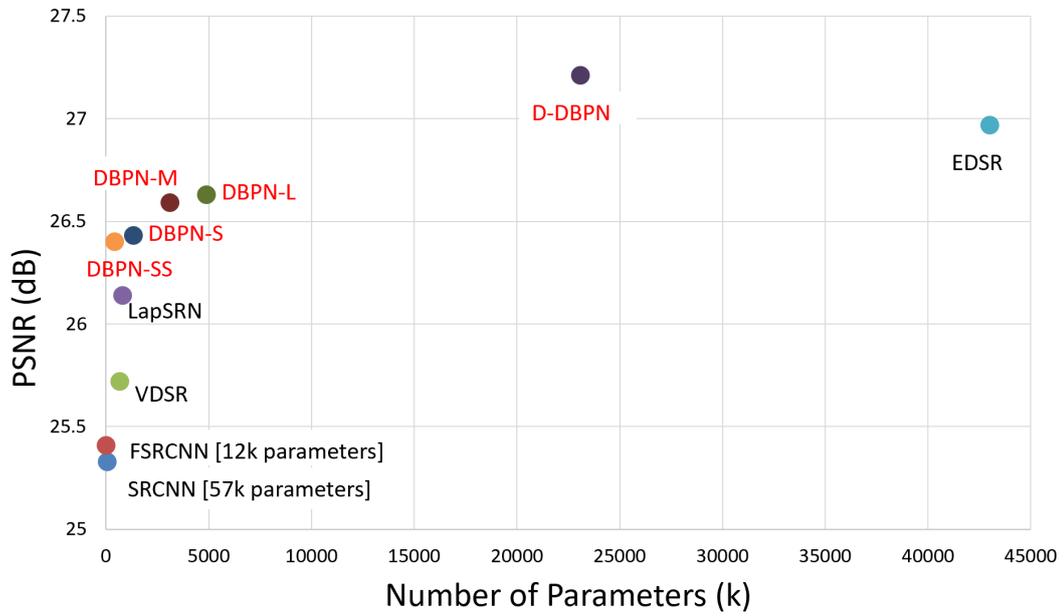


図 5 : 提案する静止画超解像DBPNのパラメータ数（実行速度に影響）と復元性能の関係性。パラメータ数は少ないほど高速である。

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 2件/うち国際共著 1件/うちオープンアクセス 2件）

1. 著者名 Kodai Shimosato and Norimichi Ukita	4. 巻 9
2. 論文標題 Multi-modal Data Fusion for Land-subsidence Image Improvement in PSInSAR Analysis	5. 発行年 2021年
3. 雑誌名 IEEE Access	6. 最初と最後の頁 141970-141980
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/ACCESS.2021.3120133	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 Muhammad Haris, Greg Shakhnarovich, and Norimichi Ukita	4. 巻 43
2. 論文標題 Deep Back-Projection Networks For Single Image Super-Resolution	5. 発行年 2021年
3. 雑誌名 IEEE Transactions on Pattern Analysis and Machine Intelligence	6. 最初と最後の頁 4323-4337
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/TPAMI.2020.3002836	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する

〔学会発表〕 計12件（うち招待講演 6件/うち国際学会 8件）

1. 発表者名 Muhammad Haris, Greg Shakhnarovich, and Norimichi Ukita
2. 発表標題 Space-Time-Aware Multi-Resolution Video Enhancement
3. 学会等名 IEEE Conference on Computer Vision and Pattern Recognition（国際学会）
4. 発表年 2020年

1. 発表者名 Kazutoshi Akita, Muhammad Haris, and Norimichi Ukita
2. 発表標題 Region-dependent Scale Proposals for Super-Resolution in Object Detection
3. 学会等名 The Fourth IEEE International Conference on Image Processing, Applications and Systems（国際学会）
4. 発表年 2020年

1. 発表者名 Sanghyun Son, ..., N. Ukita
2. 発表標題 AIM 2020 Challenge on Video Temporal Super-Resolution
3. 学会等名 Advances in Image Manipulation workshop and challenges on image and video manipulation Workshop (国際学会)
4. 発表年 2020年

1. 発表者名 Dario Fuoli, ..., N. Ukita, et al.
2. 発表標題 AIM 2020 Challenge on Video Extreme Super-Resolution: Methods and Results
3. 学会等名 Advances in Image Manipulation workshop and challenges on image and video manipulation Workshop (国際学会)
4. 発表年 2020年

1. 発表者名 Kai Zhang, ..., N. Ukita, et al.
2. 発表標題 NTIRE 2020 Challenge on Perceptual Extreme Super-Resolution: Methods and Results
3. 学会等名 New Trends in Image Restoration and Enhancement Workshop (国際学会)
4. 発表年 2020年

1. 発表者名 浮田宗伯
2. 発表標題 機械学習を用いた低解像度画像データの高解像化と電子顕微鏡画像などへの応用
3. 学会等名 応用物理学会・東海支部 基礎セミナー「AIを用いた画像処理・認識技術の進展」(招待講演)
4. 発表年 2020年

1. 発表者名 浮田宗伯
2. 発表標題 AIを用いた超解像
3. 学会等名 日本学術振興会 167委員会 ナノプロ-ブテクノロジー研究会 (招待講演)
4. 発表年 2021年

1. 発表者名 Muhammad Haris, Greg Shakhnarovich, and Norimichi Ukita
2. 発表標題 Recurrent Back-Projection Network for Video Super-resolution
3. 学会等名 IEEE Conference on Computer Vision and Pattern Recognition (国際学会)
4. 発表年 2019年

1. 発表者名 Norimichi Ukita
2. 発表標題 Image and Video Super-resolution and their Application to Tiny Object Detection
3. 学会等名 3D Computer Vision and Graphics Workshop (招待講演) (国際学会)
4. 発表年 2019年

1. 発表者名 浮田宗伯
2. 発表標題 画像の超解像による微小構造解析の可能性
3. 学会等名 格子欠陥フォーラム「格子欠陥とマテリアルインフォマティクス」(招待講演)
4. 発表年 2019年

1. 発表者名 Muhammad Haris, Greg Shakhnarovich, and Norimichi Ukita
2. 発表標題 Recurrent Back-Projection Network for Video Super-resolution
3. 学会等名 画像の認識・理解シンポジウム (招待講演)
4. 発表年 2019年

1. 発表者名 Norimichi Ukita
2. 発表標題 Image and Video Super-resolution for Learning Visual Representations: Application to Tiny Object Detection
3. 学会等名 International Workshop on Symbolic-Neural Learning (招待講演) (国際学会)
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

静止画超解像 https://www.toyota-ti.ac.jp/Lab/Denshi/iim/ukita/selection/PAMI2020.sisr.html 動画時空間超解像 https://www.toyota-ti.ac.jp/Lab/Denshi/iim/ukita/selection/CVPR2020.stsr.html 動画超解像 https://www.toyota-ti.ac.jp/Lab/Denshi/iim/ukita/selection/CVPR2019.vsr.html 静止画超解像 https://www.toyota-ti.ac.jp/Lab/Denshi/iim/ukita/selection/CVPR2018.sr.html

6. 研究組織			
	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	M u h a m m a d H a r i s (Muhammad Haris) (60816643)	豊田工業大学・工学(系)研究科(研究院)・ポスドクトラル研究員 (33924)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8 . 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------