

欠損画素復元ニューラルネットワークにおける 回帰モデルと分類モデルの性能評価

小椋 清孝[†] 橋口 仁哉[‡] 伊藤 信之[†]

岡山県立大学情報工学部[†] 岡山県立大学大学院情報系工学研究科[‡]

1. 背景

近年、4K/8K などの超高精細映像が一般的に使用されるようになり、これまで非圧縮で行われてきたデスクトップ PC などの映像送信側機器と、ディスプレイなどの映像受信側機器との間の伝送に、軽圧縮コーデックと呼ばれる低処理量・低遅延の圧縮技術が使用され始めている。

本研究室では、より低遅延な伝送圧縮手法として、欠損画素復元手法[1]の研究を行っている。この手法では、映像の一部の画素を欠損させ、非欠損画素のデータのみを伝送することで伝送量の削減を行い、送られてきたデータから深層学習モデルを用いて欠損画素値を推論し補間を行うことで映像を復元する(図1)。送信側での圧縮処理がほぼ不要であることから、本手法では送信側での遅延を大幅に削減できる。また、このモデルは、回路化しリアルタイム伝送に使用することを前提としているため、復元用ニューラルネットワークは複雑なものを使用せず、中間層が1~2層程度の浅いネットワークを用いる。これまでは回帰モデルを用いる手法について研究を進めてきた。

今回、これと異なるアプローチとして、この「画素値を求める」という問題を分類問題(8bit色深さの場合 256 クラス)と考えた分類モデルについて検討した。分類モデルはモデル圧縮手法として量子化を適用しやすく、回帰モデルと同程度の復元性能であれば、より小型の回路として実装できる可能性がある。本研究の目的は、欠損画素復元ニューラルネットワークを分類モデルで構築し、従来の回帰モデルとの性能比較を行うことである。2画素×2画素のブロックを単位とする欠損パターンにおいて各色画面で1画素を欠損した1/4欠損画像の復元を対象とし、扱う画像の色空間はRGB、色深さは8bitとする。

2. 方法

本研究で検討するモデルは、図2に示すように、

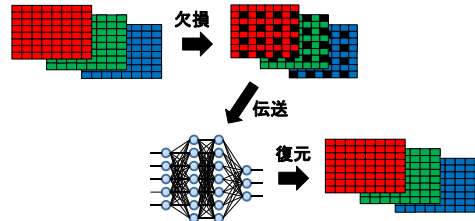


図1 欠損画素復元手法による伝送量削減

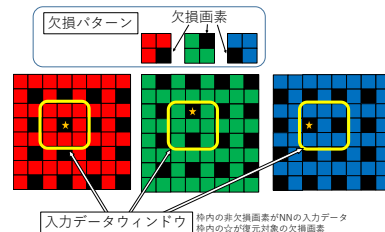


図2 欠損パターンと入力データウィンドウ

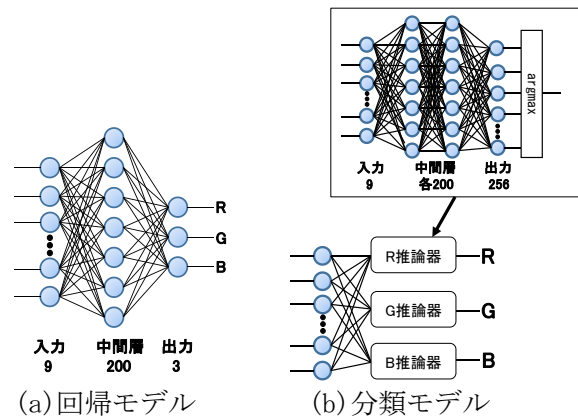


図3 欠損画素復元ニューラルネットワーク

0番目の色画面(R)の復元対象画素を中心にした周囲3画素×3画素の範囲内(他の色画面ではこれと同じ座標範囲内)の非欠損画素の画素値を入力とし、欠損画素(R,G,B)を予測する全結合型ニューラルネットワークである。回帰モデルは0~255の欠損画素値を出力する3出力モデルとなり、分類モデルでは画素値をクラスとして扱うため、256クラス×3が出力となる。ただし、回帰モデルでは1個のネットワークでR,G,Bの3画素を同時に推論する構成であるのに対し、回帰モデルでは、R,G,Bごとに独立した推論ネットワークを持つ構成である。ネットワークの中間

Performance Evaluation of Regression and Classification Model on Missing Pixel Restoration Neural Network

Kiyotaka Komoku[†], Jinya Hashiguchi[‡] and Nobuyuki Itoh[†]

[†]Faculty of Computer Science and Systems Engineering, Okayama prefectural University

[‡]Graduate School of Computer Science and Systems Engineering, Okayama prefectural University

層は、回帰モデルが1層200ニューロン、分類モデルが2層各200ニューロンである。非線形関数は両モデルともReLUを用い、学習時の評価関数は回帰モデルが平均二乗誤差、分類モデルがソフトマックスクロスエントロピーである。学習回数は、回帰モデルが100epoch、分類モデルが1000epochである。

学習で用いる画像はPASCAL VOCデータセット[2]を使用し、訓練用に200枚、テスト用に50枚を用いた。これらの画像はおおよそ300画素×500画素程度のサイズである。

学習済みモデルの評価は、50枚のテスト画像を欠損・復元させた際のPSNR(Peak Signal to Noise Ratio)の平均値で行った。

3. 結果と考察

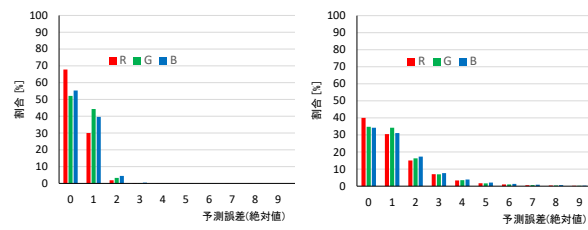
学習済みの各モデルを用いて、50枚のテスト画像を欠損・復元した際のPSNRの統計量を表1に示す。平均PSNRは分類モデルが53.12dB、回帰モデルが53.48dBとなり、回帰モデルのほうが0.36dB上回る結果となった。一方、標準偏差は、それぞれ2.02dB、2.91dBとなり、分類モデルの方のばらつきが小さくなった。50枚のPSNRを個別に比較したところ、分類モデルの方が高いPSNRの画像もあり、一律に回帰モデルの性能が高いとはいえない結果となった。

ここで、分類モデルで復元性能が最も高かった画像(2007_003565.jpg)と最も低かった画像(2007_003101.jpg)について、予測誤差の度数分布を調べた(図4)。横軸は予測誤差の絶対値、縦軸は予測対象の全欠損画素数に対する割合(%)であり、RGB各面の予測誤差分布を示している。予測誤差0の部分が正解値を予測した割合となる。この図より、正解値を外した場合でもその近傍値が予測されていることがわかる。分類モデルでは出力の0~255のクラスはそれぞれ独立であり、学習時に予測が外れた場合のペナルティは予測値から近くても遠くても同等である。これより、予測が外れた場合に全く正解値とは無関係のクラスばかりが予測される可能性も考えられたが、この結果を見ると、近接クラス間は類似であるというクラス間の関係性が学習により構築されていると考えられる。

図5は、同じ画像について回帰モデルでの予測誤差分布を調べた結果である。図4(a)と図5(a)を比較すると、図5(a)の回帰モデルの方が正解値が多く、復元性能が高いことがわかる。一方、図4(b)と図5(b)では同じような分布となっている。PSNRの大小の違いによる誤差分布の傾向が図4と図5で異なるが、今回の分類モデルは学習終了時も学習曲線に若干の傾きがあり、十分に

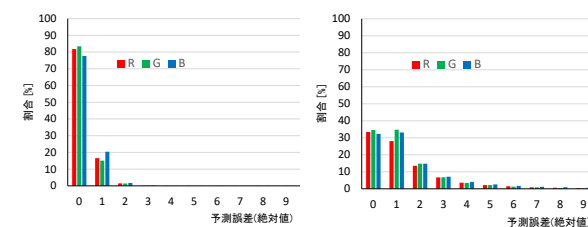
表1 各モデルで復元したテスト画像50枚のPSNRの統計量[dB]

	平均値	標準偏差	最大値	最小値
分類モデル	53.12	2.02	56.76	47.33
回帰モデル	53.48	2.91	63.22	46.37



(a) 2007_003565.jpg (56.67dB) (b) 2007_003101.jpg (47.33dB)

図4 分類モデルの予測誤差分布



(a) 2007_003565.jpg (59.92dB) (b) 2007_003101.jpg (46.37dB)

図5 回帰モデルの予測誤差分布

学習が収束したと言い切れない状態のため、これが分類モデルと回帰モデルの本質的な違いかどうかについては現時点では不明であるが、もしそうであれば、両モデルでは異なるアルゴリズムで予測値を推論しているということも考えられる。

4. まとめ

欠損画素復元ニューラルネットワークとして分類モデルを構築し、回帰モデルとの比較を行った。50枚のテスト画像を用いた評価では、回帰モデルが平均PSNRで0.36dB高いという結果となったが、標準偏差は分類モデルの方が0.89dB小さくばらつきが少ないという結果となった。今後は、分類モデルの学習の高速化及びさらに詳細な回帰モデルとの比較・評価を行う予定である。

参考文献

[1] 橋口, "ニューラルネットワークを用いた欠損画素復元回路の設計", 2022年度電気・情報関連学会中国支部連合大会, R22-24-08, 2022
 [2] M. Everingham et.al., "The PASCAL Visual Object Classes (VOC) Challenge", Int. J. Computer Vision, 88(2), 303-338, 2010