

encoder-decoder モデルによる回転変換の色の影響に関する検討

Consideration of Color Effects of Rotational Transformation by the Encoder-Decoder Model

岡本 紗季¹⁾ 神野 健哉¹⁾
Saki Okamoto Kenya Jin'no

概要

画像処理や音声認識など様々な場面で用いられている学習方法に表現学習がある。我々はこれまで、encoder と decoder 間にスキップ接続を有する encoder-decoder モデルを用いて表現学習を行い、モデル内で抽出されている特徴を明らかにするために、各層の出力である潜在変数の役割について着目してきた。立体画像を用いた回転変換の場合において、各潜在変数は役割を持っているが、背景の色を固定した回転を色で識別している場合のモデルでしか検討できていない。そこで本稿では、背景の色を固定せず色の影響をなくした場合の潜在変数の役割を検討する。

1 まえがき

画像処理や音声認識など様々な場面で用いられている学習方法に表現学習 [1] がある。表現学習はデータの特徴を自動で抽出する学習方法である。我々はこれまで、encoder と decoder 間にスキップ接続を有する encoder-decoder モデルを用いて表現学習を行い、モデル内で抽出される特徴を明らかにするため、各層の潜在変数の役割について着目してきた [2]。利用したモデルを図 1 に示す。潜在変数は図 1 に示すように浅い層を

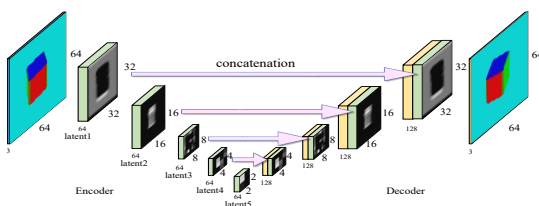
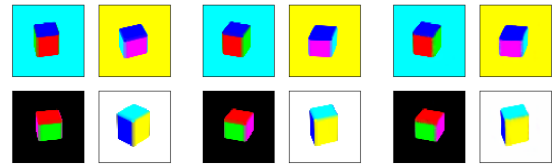


図 1 モデル

latent1 として順に latent5 とする。背景の色を固定し物体の回転変換では latent1 が非回転部分、latent2 が回転部分の低周波成分、latent3 以降が回転部分の高周波成分など生成画像の質を向上させるための情報を抽出していると推測した。しかし、背景の色を固定した画像で学習すると、回転を色のみで識別し、各潜在変数の役割も色の影響を受けている可能性がある。背景の色を固定しない画像で学習をすると、回転を色と形で識別していること

1) 東京都市大学大学院総合理工学研究科情報専攻
Informatics, Graduate School of Integrative Science and Engineering, Tokyo City University



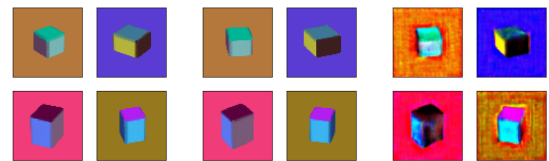
(a) 入力画像 (b) 正解画像 (c) 出力画像

図 2 学習データ (基本 8 色)

を確認している [3]。そこで本稿では、背景の色を固定せず色の影響を無くすようにした場合の潜在変数の役割を検討する。

2 学習データの色の追加

先行研究 [3] ではホワイト、シアン、マゼンダ、イエローにランダムな値のノイズを付与した 4 色を使って立体画像を作成して学習をした。回転はできているが、学習に含まれていない未知の色では正しく色が出力されないものが存在する。そのため、基本 8 色の RGB の画像で学習を行い、様々な組み合わせの RGB からランダムに 4 色選択した立体画像で予測をした。物体は直方体とした。学習データの例を図 2、テストデータの例を図 3 に示す。



(a) 入力画像 (b) 正解画像 (c) 出力画像

図 3 テストデータ (様々な組み合わせ)

図 2 より、正解画像と比較すると出力画像は正しく回転可能であるが、学習データに存在する色の影響を受けていることを確認した。このことから、未知の色を予測させるためには学習データにある程度の数の色を学習させる必要がある。しかし、背景の色を固定しない画像で学習をすると物体を認識することができると考えられる。

3 潜在変数の役割の確認

前節より、少ない色から正しく予測をするのは難しいため、色のバリエーションの多くして学習をさせなければ

ば、色の影響を受ける可能性がある。このため、推測した潜在変数の役割が周波数成分に基づくか色情報に基づくかを判断することができない。そこで、画像をグレースケールで読み込み学習したデータに対して潜在変数の役割を考える。

3.1 様々な組み合わせの RGB を用いた画像

図 3 のテストデータで用いた画像をグレースケール化して学習を行う。入力画像と正解画像、予測した出力画像を図 4 に示す。

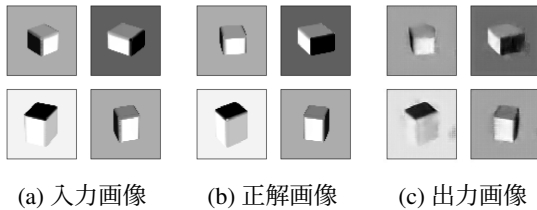


図 4 カラー画像をグレースケール化

正解画像に近い出力画像が得られているため、潜在変数を 1 つずつ結合させない場合と 1 つずつ結合させた場合の実験を行う。前者を潜在変数を削減した場合、後者を潜在変数を残した場合とする。削減したい潜在変数をゼロにすることで潜在変数の削減を行い、実験結果を図 5 に示す。

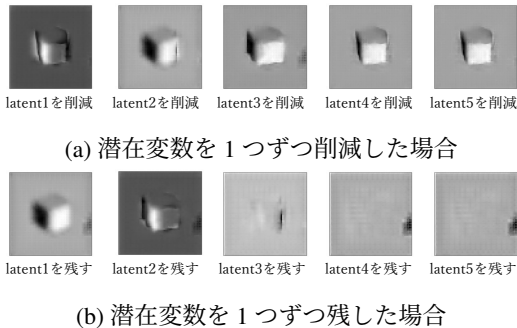


図 5 潜在変数の変更による出力画像の比較

latent1 を削減すると背景の色が変化する。latent2 と latent3 をそれぞれ削減すると、回転前や回転途中の画像が出力される。latent4 以降をそれぞれ削減すると深い層になるほど図 4(c) に近づいている。このことから、latent1 は変化しない部分、latent2 では変化する部分を抽出していると言える。

3.2 物体の輪郭のみの画像

画像の情報をさらに減らしても同じ結果が得られるかを確かめるため、輪郭のみ色を変えた画像を作成する。色は白と黒のみ使用し、背景が白、輪郭が黒の画像と色を反転させた画像を作成する。入力と正解画像、予測して得られた出力画像を図 6 に示す。

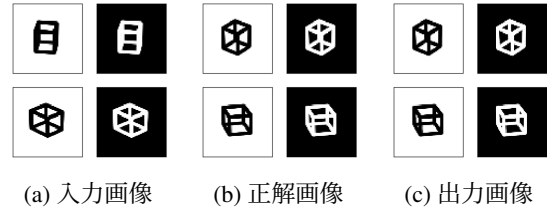


図 6 物体の輪郭のみの画像

背景の色に関係なく、正解画像に近い画像が出力できている。カラー画像と同様に潜在変数の削減した場合と残した場合の実験を行う。その結果を図 7 に示す。

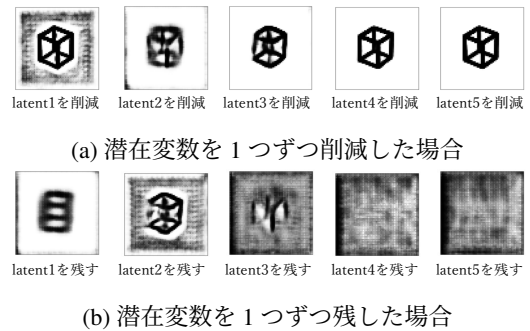


図 7 潜在変数の変更による出力画像の比較

latent1 を削減すると、背景の情報が失われている。latent2 と latent3 をそれぞれ残すと回転前の物体との差分が出力されているように見える。これらの結果から latent1 が非回転部分、latent2 と latent3 では回転部分の情報が抽出されているが、latent2 と latent3 でどのような順番で抽出されるのかは確認できなかった。

4 まとめ

色の影響をなくしても latent1 では非回転部分、latent2 では回転部分、latent3 以降は抽出しきれなかった情報を抽出していることが示唆された。回転部分において低周波成分と高周波成分がどの潜在変数で抽出されているかは明らかにできていない。今後は周波数特性を測定できる画像等を用いて潜在変数の役割の検討を進める。

謝辞

本研究の一部は JSPS 科研費 20K11978, 23K11266, 23H03387, 東北大学電気通信研究所共同プロジェクト研究, 東京都市大学重点推進研究未来知能ユニットの助成によるものです。

参考文献

- [1] Bengio, Yoshua, Aaron Courville, and Pascal Vincent. "Representation learning: A review and new perspectives." IEEE transactions on pattern analysis and machine intelligence 35.8 (2013): 1798-1828
- [2] 岡本紗季, 神野健哉, "画像データセットを用いた潜在変数の役割の検討," JSAI2023, 4Xin1-57, 2023.
- [3] 岡本紗季, 神野健哉, "encoder-decoder モデルに与える色の影響に関する検討," 電子情報通信学会 2023 NOLTA ソサイエティ大会, NLS-38, 2023.