

深層学習モデル TransUNet と組み合わせ最適化を用いた歪んだ 2 次元コード復号方式に関する研究

濱田 悠樹 * 若松 健斗 * 新川 翔貴 * 鞍津輪 一希 * 神園 誠 * 比良 祥子 * 向田 真志保 * 小野 智司 *
* 鹿児島大学

1 はじめに

2 次元コードは、多くの情報を格納できることから生産、物流、販売等の幅広い分野において活用されている。また、平面や一様な曲面上に印刷されることを想定しており、2 次元コードを復号するソフトウェアやアプリケーションの多くは射影変換などの機能を備えている。このため 2 次元コードを斜めから撮影した場合や、2 次元コードが円柱形状に印刷された場合においても復号が可能である。しかし、2 次元コードが紙や布に印刷され、図 1 に示すように、激しく変形することで 2 次元コードが不均一に歪んでしまう場合は、復号が困難になる。これは、2 次元コードが一般的な復号処理において、2 次元コードの四隅にある位置検出パターンをもとに計算されるサンプリンググリッド（図 2）の交点をサンプリング点とし、サンプリング点の画素の明暗を取得することで、2 次元コードの復号を行うためである。

このため著者らは、紙や布に印刷された 2 次元コードに不均一な歪みが含まれる場合でも復号可能な復号方式を提案した [7]。本手法は、2 次元コードを構成するモジュールのエッジ検出に深層学習モデルを用いて、各エッジが歪んでいない 2 次元コード内においてどの位置にあるか相対的な位置（以下 ID）を認識することで歪みのない 2 次元コードを復元する。しかし、歪みが激しいために自己遮蔽が生じてしまい、視認できない領域が存在する場合、遮蔽領域付近に存在するエッジに誤った ID を割り当ててしまい、復号成功率が低下することがある。

本研究では、不均一な歪みを含んだ 2 次元コードの復号成功率向上を目的として、先行研究 [7] に対して深層学習モデルの TransUNet を組み込む。これにより、遮蔽領域が存在するほど歪みが激しい場合でも、2 次元コードを構成するモジュールエッジの推定が可能である。評価実験により、先行研究と比較して復号成功率が向上することを確認した。

2 関連研究

これまで、緩やかに歪曲した 2 次元コードを、射影変換等の単純な画像処理による歪み補正や包絡線を用いて復号する方式が提案されている [2, 5, 6]。また、不均一で複雑な歪みを含む 2 次元コードの復号を可能とするため、組み合わせ最適化を用いた 2 次元コード復号方式も提案されている [7]。この方式は、図 3 に示すように、モジュールを構成する各エッジが 2 次元コード内で上または左から何本目の位置に存在するか推定し、推定されたエッジの位置をもとに歪んだ 2 次元コードを構成する全モジュールの位置を推定し、歪みのない 2 次元コードを再構成する。

3 提案手法

3.1 基本アイデア

本研究では、先行研究 [7] の改良を行い、先行研究と比較し復号成功率の高い 2 次元コード復号手法を提案する。

アイデア 1: TransUNet を利用した 2 次元コードのモジュールエッジ推定。

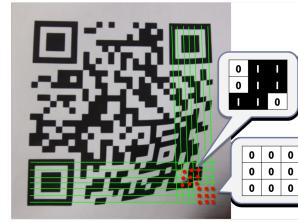


図 1: 復号失敗の例

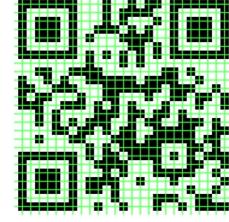


図 2: サンプリンググリッド

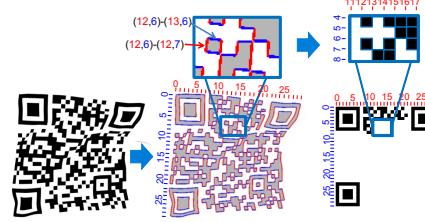


図 3: エッジの ID 推定による 2 次元コード復元

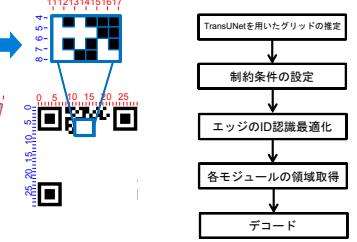


図 4: 復号の処理手順

提案手法は 2 次元コードの歪みをより正確に認識するため、2 次元コードを構成するモジュールエッジ推定に TransUNet を用いる。

アイデア 2: モジュールエッジ ID の推定を行う最適化。

提案手法では不可視のモジュール境界を含めたモジュールエッジに対して、各縦横エッジがそれ左または上から何番目に存在するかを示す ID の推定を行う。モジュールエッジの ID を推定する組み合わせ最適化において、局所最適解が発生する。このため複数の修復処理を用いて局所解への収束を抑制する。

3.2 処理手順

提案する方式の処理手順を図 4 に示すとともに、以下で各処理について述べる。本手法は TransUNet を用いて 2 次元コードにおける縦のモジュールエッジ、横のモジュールエッジ、モジュールエッジの交点、遮蔽領域および背景の 5 クラス分類を行う。その後、2 次元コード内のモジュールエッジの位置を推定することにより、2 次元コードの歪みを補正する。

(1) TransUNet を用いたモジュールエッジの推定

セマンティックセグメンテーションタスクで用いられる深層学習モデル TransUNet を用いて、2 次元コードからモジュールのエッジ検出を行う。検出されたエッジを水平方向と垂直方向の 2 種類に分別し、それについて以下の処理を行う。詳細は 3.3 節に後述する。

(2) 制約条件の設定

2 次元コードに遮蔽領域が無いと仮定し、エッジの左右または上下の位置関係から、エッジの ID の制約

提案手法	従来手法 1	従来手法 2
249/250	242/250	99/250

条件を設定する。

(3) エッジの ID 認識最適化

山登り法を適用し、エッジ ID の推定を行う。詳細は 3.4 節で後述する。

(4) 各モジュールの領域取得

最適化されたエッジと ID から補助線を作成し、各モジュールの領域を推定する。次に、算出された各モジュールの領域に着目し、平均輝度値から明暗情報を取得する。

(5) デコード

取得した各モジュールの明暗情報をから 2 次元コードを復元し、デコード処理を行うことでデコード結果を取得する。

3.3 モジュールのエッジ検出

提案手法は 2 次元コードの歪みを正確に認識するため、2 次元コードを構成するモジュールエッジの推定に若松ら [3] が提案した領域分割手法を用いる。この手法はネットワークアーキテクチャに TransUNet [1]、損失関数に UFL [4] を採用している。これにより、2 次元コードの大域的特徴とクラス間の不均衡性を考慮したモジュールエッジの検出が可能となる。先行研究でモジュールエッジ検出に Unet を用いていたのに比べ、提案手法では TransUNet を用いることで推定可能な領域が先行研究と比較し多くなる。学習には、仮想的に作成した 30,240 枚のデータセットを用いた。また、学習データの画像サイズは 512×512 に統一した。

3.4 ID 認識最適化

組み合わせ最適化における個体の評価は、同じ ID 同士のエッジを結ぶ補助線に着目し、エッジと補助線の角度、エッジと補助線の交叉回数による評価、ならびにエッジ間の間隔による評価に基づいて行う。また、隣接するエッジに同じ ID が割り当てられるのを防ぐため、接続線の角度にもとづく評価を、接続線の長さによって重み付けしている。組み合わせ最適化中に局所解へ陥ってしまった場合に、5 つの修復操作によって局所解を脱する。

4 評価実験

提案手法の有効性を検証するため、合成データを用いて実験を行った [7]。合成データは、3 次元 CG レンダリングによって生成した歪んだ 2 次元コード 250 枚である。比較対象として、先行研究の手法 [7]（従来手法 1）と株式会社デンソーウェーブの”クルクル - QR コードリーダー”¹（従来手法 2）を用いた。提案手法と従来手法 1においては、1 枚の画像ごとに 5 回復号を行い復号成功率が 50% 以上となった場合を復号成功とした。また、従来手法 2においては距離と角度を自由に動かし約 5 秒間撮影を行い、復号に成功した場合を復号成功とした。2 次元コードの歪み量と遮蔽量については、先行研究と同様に定義した。

合成データ 250 枚を復号した場合の復号成功枚数を表 1、提案手法と従来手法 1 を用いた成否結果を図 5, 6 に示す。これらの結果から、従来手法 1, 2 と比較し提案手法の復号成功率が向上していることが分かる。これはエッジ検出

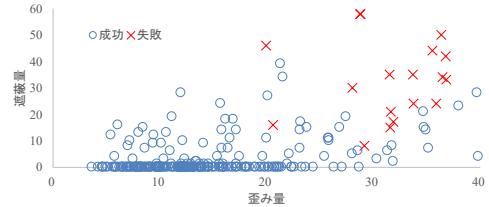


図 5: 提案手法による復号結果

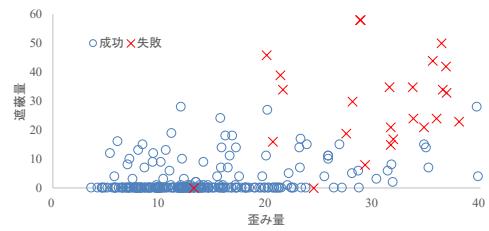


図 6: 従来手法 1 による復号結果

手法に TransUNet を用いたことで、検出されるモジュールエッジの領域が増加し、認識可能なモジュール数が増加したためと考える。

5 結論

本研究では、遮蔽領域が存在する 2 次元コードに対して頑健な復号方式を提案した。提案手法では、2 次元コードを構成するモジュールエッジの検出可能な領域を増加させるため、エッジ検出に TransUNet を使用した。実験から、従来手法と比較し歪みに対して頑健であることを確認した。今後、本方式の実用性向上に向けて処理時間短縮可能な手法を検討する。

参考文献

- [1] Chen Jieneng, Lu Yongyi, Yu Qihang, Luo Xiangde, Adeli Ehsan, Wang Yan, Lu Le, Yuille Alan L., and Zhou Yuyin. Transunet: Transformers make strong encoders for medical image segmentation. *arXiv preprint arXiv:2102.04306*, 2021.
- [2] Eisaku Ohbuchi, Hiroshi Hanaizumi, and Lim Ah Hock. Barcode readers using the camera device in mobile phones. In *2004 international conference on cyberworlds*, pages 260–265. IEEE, 2004.
- [3] Kento Wakamatsu and Satoshi Ono. Transunet with unified focal loss for class-imbalanced semantic segmentation. In *AROB-ISBC-SWARM 2023*, 2023.
- [4] Michael Yeung, Evis Sala, Carola-Bibiane Schönlieb, and Leonardo Rundo. Unified focal loss: Generalising dice and cross entropy-based losses to handle class imbalanced medical image segmentation. *Computerized Medical Imaging and Graphics*, 95:102026, 2022.
- [5] 中野裕哉 and 大矢誠. 1017 強い画像歪みを伴う qr コードの補正 (gs4-1 計測・制御). In *北陸信越支部総会・講演会 講演論文集 2010.47*, pages 379–380. 一般社団法人 日本機械学会, 2010.
- [6] 水本裕之 and 濱口陽太郎. 二次元コード読取装置, 2004. 特許出願公開番号 特開 2004-362053.
- [7] 濱田悠樹, 若松健斗, 新川翔貴, 鞍津輪一希, and 比良祥子 and 小野智司. 歪んだ 2 次元コードの復号における組合せ最適化手法の改善. In *情報処理学会 第 85 回全国大会 (2023)*, 2023.

¹<https://www.denso-wave.com/ja/system/qr/product/reader.html>