Vol.2022-CDS-34 No.6
2022/6/10

Semantic Segmentationによる電子帳票画像の改ざん検知

黒澤 匠雅[1,a)]　高橋 柊[1]　鈴木 康央[1]

概要：近年，企業は業務の IT 化によって競争力維持・強化を図る Digital Transformation(DX) を推進しており，その重要な要素として紙文書のペーパーレス化（電子化）を進めている．紙文書や帳票をスキャンデータにより電子化することで業務の簡素化・効率化が期待出来る一方で，電子化された画像データは安易に改ざん可能であり，改ざんされた電子帳票画像は不正請求に利用される懸念がある．本研究では，PSPNet を用いた Semantic Segmentation モデルを応用した電子帳票画像における改ざん検知手法を提案する．提案手法が多様な改ざん手法や極小領域における改ざんを高精度に検知可能であることを示す．また改ざん手法ごとの訓練データセット生成方法について議論する．

Detecting Tampering with Electronic Form Images Using Semantic Segmentation Model

Abstract: In recent years, companies have been promoting Digital Transformation (DX), which aims to maintain and strengthen competitiveness through promoting IT in business operations. One of the critical elements of DX is the paperless (digitization) of paper documents. Digitization of documents through scanned data is expected to simplify and streamline operations, but digitized image data is easily tampered with, and there are problems such as falsifying the amounts on the image and making unauthorized claims. In this study, we propose a tamper detection method for electronic form images based on the Semantic Segmentation Model using PSPNet. We show that the proposed method can detect various tampering techniques in tiny areas with high accuracy. We also discuss how to generate training data sets for each tampering method.

Keywords: Semantic Segmentation，Tampering Detection, Electronic Form Image

1. はじめに

近年，紙文書を中心とした事務作業をペーパーレス化（電子化）することによる，業務の簡素化・効率化および意思決定の迅速化が推進されている．企業においては，業務における IT 利用を推進することにより競争力維持・強化を図る Digital Transformation(以下，DX) における重要な要素として紙文書・書類のデジタル化を進めている．さらに，COVID-19 のパンデミック危機により在宅勤務者が増加したことで，2020 年の紙消費量が急激に低下していることが報告されており，ペーパーレス化の動きは加速している [1]．日本においては，2022 年より改正電子帳簿保存法[*1]が施行され国税関係の帳簿・書類のデータ保存について、抜本的な見直しが実施されている．特に請求書、領収書、注文書など国税関係の書類についてはスキャナ保存による保存方法を選択することが可能となり，経費精算業務のペーパーレス化が可能となっている．

紙文書をスキャンデータにより電子化することで企業は業務の DX を推進可能であるが，電子化された画像データは安易に改ざんされ得る．改ざん者は数字や文字、言葉を簡易な Copy-Paste 操作で書き換えることが可能であり、改ざん画像は不正請求などに利用される懸念がある．本研究では請求書、領収書などをスキャンして電子化した電子帳票画像における改ざんを検知することを目的とする．電子帳票画像の改ざんを検知することが出来れば，電子画像のデメリットである耐改ざん性を補完することができる．電子帳票画像からの改ざん検知には以下のような問題が存在する．

[1] SAS Institute Japan Ltd.
Roppongi Hills Mori Tower 11th floor
6-10-1 Roppongi, Minato-ku, Tokyo 106–6011, Japan
[a)] takuma.kurosawa@sas.com

[*1] 国税庁 電子帳簿保存法関係
https://www.nta.go.jp/law/joho-zeikaishaku/sonota/jirei

改ざん手法の多様性
改ざん者は画像編集ソフトを用いることで単純な Copy-Paste 操作だけでなく，多様な改ざん手法を用いることが出来る．

極小領域における改ざん
風景写真や人物写真などの一般画像における画像改ざんと比較し，帳票画像の改ざんでは極めて小さい領域における改ざんが意味を持つ．例えば，領収書上における金額に 0 を添えることで改ざん画像は改ざん前と比較して 10 倍の価値を持つ．

データセット
一般画像における改ざん検知モデルを構築し既存研究と比較するためのデータセットは多く存在するが，帳票画像の改ざんを検知するためのデータセットは少ない．また，実際に発生した帳票画像に対する改ざんデータは公開されていない．

多様な改ざん手法や極小領域における改ざんを高精度に検知可能であれば，フォーマットを問わず電子帳票画像における改ざん検知が可能となる．本研究では，Pyramid Scene Parsing Network(以下，PSPNet) [2] を用いた Semantic Segmentation モデルを応用した電子帳票画像における改ざん検知手法を提案する．PSPNet を用いることで，極小領域における改ざん箇所を高精度に検知可能であることを示す．また，モデルの訓練データに用いる改ざん画像を，非改ざん画像から疑似的に生成する手法を提案する．

本研究の貢献は以下の通りである．

(1) モデルの訓練データに用いる改ざん画像を，非改ざん画像から疑似的に生成する事で実際の改ざん画像に対して高精度に改ざん検知が可能となる事を示す．

(2) PSPNet を用いた Semantic Segmentation モデルが，複数の改ざんパターンや極小領域の改ざんに対して高精度で改ざん検知可能であることを示す．

本稿では，次章にて関連研究について述べ本研究との差分について説明する．次に，3 章にて提案手法の詳細について説明する．4 章では実データを用いた実験を行い既存手法と提案手法の精度について比較する．最後に 5 章にて本研究のまとめおよび今後の課題について述べる．

2. 関連研究

画像データの改ざん検知については，多くの手法が提案されている．改ざん検知の対象としては，風景画像や人物画像などの一般画像に対する改ざん検知と，より対象を限定した文章画像に対する改ざん検知が存在する．

(1) 一般画像における改ざん検知
Zhou ら [3] は，物体検出手法を応用した画像改ざん検知手法を提案している．改ざん領域と非改ざん領域間の境界における不自然さおよびノイズ変化に着目した Faster R-CNN を用いることで高精度に改ざん領域の Bounding Box を予測可能であることを示している．Wu ら [4] はピクセルレベルで改ざんの有無を判定する Semantic Segmentation のアプローチを採用し，Copy-Move によって改ざんされた画像の中から類似箇所を検出する BusterNet を提案している．Semantic Segmentation による画像改ざん検知は特徴抽出層と異常検知層を組み合わせることで多様な改ざん手法に対する検知を可能にした ManTra-Net [5] や，敵対的生成ネットワークによる Splice 検知手法 [6]，LSTM にリサンプリング特徴を入力して Encoder-Decoder 構造を組み合わせる検知手法 [7] など深層学習を用いた多様な発展的手法が提案されている．また改ざん画像のデータ量が少ないという課題に対し，Huh ら [8] は非改ざん画像のみを学習データとして用いた改ざん検知手法を提案している．一般画像の中でも特に人物画像の改ざんについては，画像や動画中のある人物の顔を他者の顔と入れ換えたり，表情を変えたりすることができる DeepFake が社会的な問題となっており [9]，人物画像の改ざん検知に特化した手法が数多く提案され [10–12]，またベンチマークデータセット作成も盛んに行われている [13, 14].

(2) 電子帳票画像における改ざん検知
一般画像における改ざん検知と比較して，電子帳票画像における改ざん検知の取り組みは非常に少ない．ICPR2018 において電子帳票画像の改ざん検知コンペティション [15] が実施されたが，優勝した手法は画像データセット全体から Steganographic な特徴を活用することで検知率 100%を達成しており，主催者はその原因を評価データに用いた電子帳票画像の撮影状況が統一化されていたためであると考察している．今回我々が作成したデータセットは画像ごとに撮影状況が異なり，ICPR2018 のコンペティションでは削除されていた背景画像が含まれていることから，問題はより難しく現実的になっている．James ら [16] は Optical Character Recognition/Reader (OCR) を用いて画像中の文字を検出した後，検出領域から構成されるグラフ構造の特徴から改ざんの有無を判定する手法を提案している．しかし電子帳票は印字が不明瞭であることが少なくなく，OCR による文字検出精度は十分でない．そのため，本研究では画像データとしての特徴のみを用いる Semantic Segmentation のアプローチを採用し，印字が不明瞭な画像に対しても適用可能な手法を提案する．

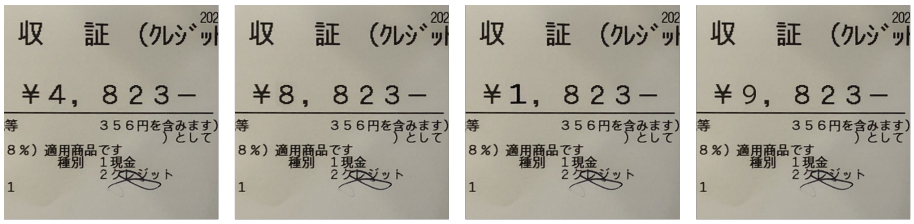

図 1 異なる改ざん手法によって改ざんを施された画像の例．左から，改ざん前画像，Copy-Move，Splice，Pixel-Manipulation による改ざん画像．

Fig. 1 Examples of images tampered with by three different tampering methods. From left to right, the image before tampering, and the tampered images by Copy-Move, Splice, and Pixel-Manipulation.

3. 提案手法

本研究では，PSPNet を用いた Semantic Segmentation モデルを応用した電子帳票画像における改ざん検知について検討する．提案手法は以下の 2 つのステップから構成される．

1. 訓練データの生成
我々が調査した範囲においては，モデルの学習に耐えうる十分な量の文章や帳票画像における改ざんデータは一般公開されていない．そのため，モデルの訓練データに用いる改ざん画像を，非改ざん画像から疑似的に生成する．

2. Semantic Segmentation モデルの構築
ピクセルレベルで電子帳票画像中の改ざん箇所を判定する Semantic Segmentation モデルを構築する．Semantic Segmentation モデルを用いることで，多様な改ざん手法や極小領域における改ざんを高精度に判定することが期待できる．

3.1 訓練データの生成

画像改ざん検知のための Semantic Segmentation モデルのベンチマークとして利用可能な改ざん画像のデータセットに COVERAGE [17]，Columbia dataset [18]，CASIA [19] がある．いずれのデータセットも，人物や風景などの一般画像に対する改ざんを対象としており，文章や帳票の画像は含まれていない．また，データセット中の改ざん画像における改ざん手法が限定されており，多様な改ざん手法に対するモデルの評価が出来ない．本研究は，文章や帳票画像を撮影・スキャンしたデータを対象とし，多様な改ざん手法に対してロバストな改ざん検知モデルを構築することを目的とする．James ら [16] は文章画像に対する改ざん手法を Copy-Move(CM), Splice(SP), Pixel-Manipulation(PM) の 3 つに分類している．本研究ではこれら 3 手法のすべてを対象とし，非改ざん画像から改ざん画像を疑似的に生成することでモデルの訓練データを生成する．

図 1 に各改ざん手法ごとの改ざん例を示す．Copy-Move は，画像内の他の部位から画像をコピーし，改ざん対象部分に移動して上書きする手法である．コピー元となる画像が同一画像内から取得されることから，フォント，色および解像度が改ざん前と近しくなり，人の目で改ざん箇所を判定することは難しい．Splice は他の画像における部位を，改ざん対象部分に移動して上書きする手法である．Copy-Move と比較し，他の画像から改ざん箇所をコピーするため，注意深く観察することでフォント，色および解像度の違いから改ざん箇所を人の目で判定することが容易である．Pixel-Manipulation は画像のピクセル値を任意のピクセル値に書き換えることで画像を改ざんする手法である．具体的には，新たなテキストを配置して元画像を上書きすることで，改ざん者の意図にあった改ざんの実施などがある．本研究では，Web から収集したレシート・領収書の撮影画像から訓練用の改ざん画像を生成する．マスキング処理などの編集がなされている画像を除いた 432 枚の画像を対象とする．収集データに対して，改ざん手法ごとに以下のステップで疑似的に改ざんデータを生成する．

共通のステップ

Step A. 改ざんを施す画像の大きさ $size_i = (x_i, y_i)$ に占める改ざん領域の割合を対数正規分布 $\Lambda(\mu, \sigma^2)$ から無作為に生成し，切り抜き領域の大きさ $size_c = (x_c, y_c)$ を式 1 によって定める．

$$\begin{aligned} W, H &\overset{i.i.d.}{\sim} \Lambda(\mu, \sigma^2) \\ x_c &= p(x_i \times w) \\ y_c &= p(y_i \times h) \end{aligned} \tag{1}$$

ただし，ここで w および h は確率変数 W および H の実現値であり，$p : \mathbb{R}^+ \to \mathbb{N}$ である．

Step B. 切り抜き領域全体が改ざんを施す画像内に収まるように，切り抜き領域の位置を一様分布を用いて無作為に定める．

Step C. 改ざんを施す前後の画像におけるピクセル値の差分からマスク画像を作成する．

Copy-Move/Splice

Step I. 改ざん手法が Splice の場合は，改ざんを施す画像以外の画像から，改ざん元とする画像を無作為に 1 枚選択する．

Step II. 共通のステップ：Step A. を実施する．

Step III. 共通のステップ：Step B. を実施する．

Step IV. 切り抜き領域全体が改ざんを施す画像内に収まるように，貼り付け先の候補となる領域を一様分布を用いて無作為に N 個選択する．

Step V. 切り抜き領域内のピクセルの中央値 m_c と貼り付け先の候補となる領域内のピクセルの中央値 $m_1, \ldots, m_N$ を求め，式 2 より貼り付け先を決定する．

$$\arg\min_i ||m_c - m_i||_2 \tag{2}$$

Step VI. 貼り付け先の領域に切り抜き領域のピクセル値を上書きする．

Step VII. 共通のステップ：Step C. を実施する．

Pixel-Manipulation

Step I. 文字数の集合 $C_{PM} \subseteq \mathbb{N}$ およびフォントの種類の集合 F_{PM} から改ざんに用いる文字数およびフォントの種類を無作為に選択し，ランダムな文字列を生成する．

Step II. 共通のステップ：Step A. を実施する．

Step III. 共通のステップ：Step B. を実施する．

Step IV. 生成した文字列を切り抜き領域内に収まるフォントサイズで書き込む．

Step V. 共通のステップ：Step C. を実施する．

図 2 に疑似生成した訓練データの例を示す．改ざん箇所の大きさは確率的に決定されるため，様々な大きさをとることが分かる．また，Copy-Move および Splice では切り取り領域のピクセル中央値と貼り付け先のピクセル値の差が最小となる箇所で改ざんを実施するため，改ざん箇所とその周辺ピクセルで色合いの差が小さくなっている．

3.2 Semantic Segmentation モデルの構築

本研究では改ざん箇所の検知に Semantic Segmentation モデルを用いる．Semantic Segmentation は入力画像のピクセルごとにラベルやカテゴリを付与する深層学習の一種である．モデルのアーキテクチャには PSPNet を用いる．PSPNet では Encoder-Decoder 間で Spatial Pyramid Pooling により周辺コンテクストを得ることで，高精度な Semantic Segmentation を可能にしている．

図 3 に本研究で利用したモデルアーキテクチャの概要を示す．Unel ら [20] は高解像度画像において極小領域を判定する際に，モデル入力画像を分割することで判定精度が向上することを示している．本研究では，入力画像を 624×624 の window をスライドさせ Cropping/Padding することで分割された部分画像をモデルの入力データとする．また，分割されて出力されるマスク画像は再合成し，重複部分における予測値に平均値を用いることで，入力画像サイズと同じサイズのモデル出力を生成する．

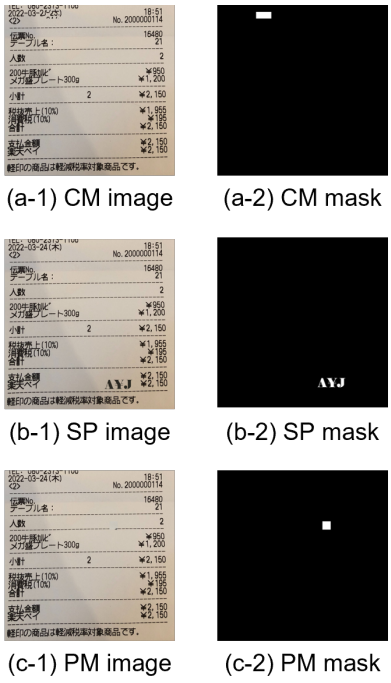

図 2 改ざん画像と対応するマスク画像の例．(a-1)，(b-1) および (c-1) は，それぞれ Copy-Move, Splice, Pixel-Manipulation による改ざん画像であり，(a-2)，(b-2) および (c-2) は対応するマスク画像である．

Fig. 2 Examples of tampered images and corresponding mask images. (a-1), (b-1) and (c-1) are tampered images by Copy-Move, Splice and Pixel-Manipulation, respectively, and (a-2), (b-2) and (c-2) are their corresponding mask images.

4. 実験

提案手法を用いた際の電子帳票画像の改ざん検知の性能を確認するために，非改ざん電子帳票画像 60 枚の各画像に対して 3 種類の改ざん（Copy-Move, Splice, Pixel-Manipulation）を施した計 180 枚の改ざん画像に対して予測を行い，IoU および F1-Score を用いて評価した．ここで，IoU は改ざん領域の座標の集合 G およびモデルの予測した改ざん領域の座標の集合 P を用いて $G \cap P / G \cup P$ として得られる値であり，F1-Score は適合率と再現率との調和平均 $2 \times$ 適合率 $\times$ 再現率/(適合率 + 再現率) である．さらに，Semantic Segmentation モデルをベースとした既存の画像改ざん検知手法との比較として，多様な改ざん手法に対応した改ざん検知手法である ManTra-Net との精度比較を行った．

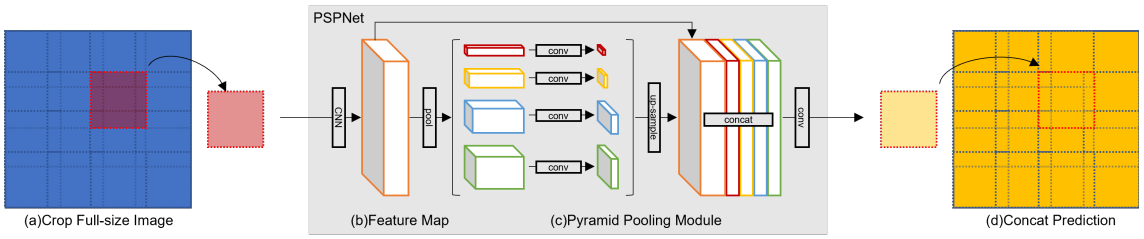

図 3 提案手法のモデル概要．入力画像は適当なサイズのウィンドウ幅で Cropping/Padding をすることで分割される．ネットワークは PSPNet を採用した．出力は分割画像の平均値を用いて元の大きさの画像に復元している．

Fig. 3 Model overview of our method. Input images are splitted by cropping/padding with an appropriate window width. PSPNet is used as the network. The output image is restored to its original size using the average value of the splitted images.

表 1 本実験において作成した検証用画像に対する評価結果．

Table 1 Evaluation results for the images created in this experiment.

method	IoU	F1-Score
Our Method	0.331	0.425
ManTra-Net	0.122	0.263

(1) 実験条件

改ざん領域の大きさを制御する対数正規分布 $\Lambda(\mu, \sigma^2)$ の確率密度関数 f_X は式 3 で与えられる．

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}x} \exp\left(-\frac{(\log x - \mu)^2}{2\sigma^2}\right) \qquad (3)$$
$$0 < x < \infty$$

ただし，$-\infty < \mu < \infty$ および $\sigma > 0$ である．本実験では，精度検証用に作成した改ざん画像における改ざん領域の大きさの要約統計量を元に対数正規分布のパラメータ (μ, σ) を決定しており，Copy-Move では $(-7.8, 1)$，Splice では $(-5.8, 0.7)$，Pixel-Manipulation では $(-6, 1.5)$ としている．また，学習に用いたデータは精度検証用に用いた電子帳票画像とは別の未改ざんの電子帳票画像 432 枚に対して前述の 3 種類の改ざんを用いて作成した改ざん画像 25,860 枚である．

(2) 実験結果

精度検証用に作成した改ざん画像 180 枚に対する各手法の評価結果を表 1 に示す．IoU および F1-Score については提案手法が既存手法より高い結果となっており，本実験で用いた検証用の画像については，提案手法が優れていることがわかる．図 4 に無作為に選択した 3 種類の改ざん手法に対するモデルの予測結果を示す．提案手法では，いずれの改ざん手法に対しても改ざん部分を予測出来ている．一方，ManTra-Net による予測は Splice, Pixel-Manipulation に対しては適切に予測出来ているが，Copy-Move に対しては予測が出来ていない．また，本実験で検証に用いた画像に対する ManTra-Net による予測結果では，いずれの改ざん手法に対しても画像全体にわたり斑に改ざんを誤検知する特徴がある．改ざん種類別の IoU を図 5 に示す．Copy-Move は同一画像内から改ざん元となる画像を

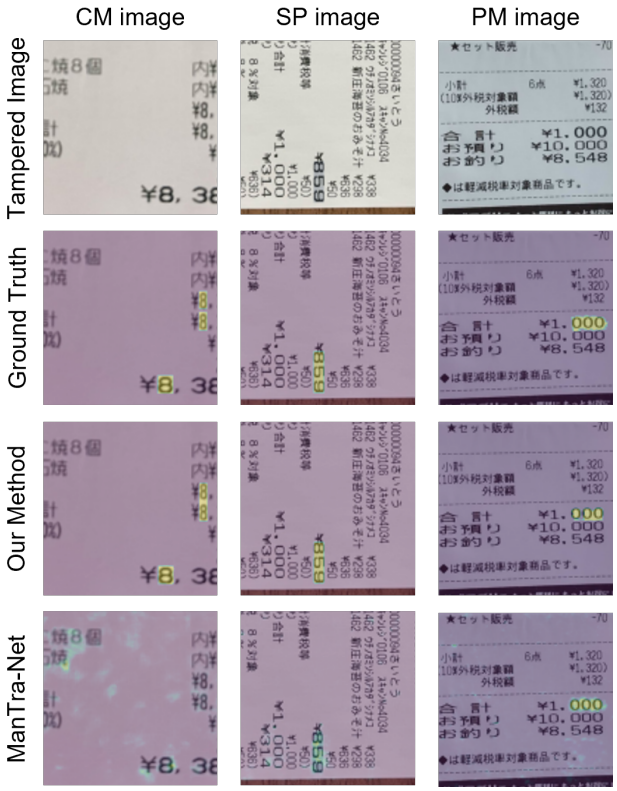

図 4 本実験において作成した検証用画像に対する予測結果．左から，Copy-Move，Splice，Pixel-Manipulation による改ざん画像の例．また，上から改ざん画像，正解画像，提案手法による予測画像，ManTra-Net による予測画像である．

Fig. 4 Prediction results for the images created in this experiment. From left to right, examples of tampered images by Copy-Move, Splice, and Pixel-Manipulation are depicted. From the top are the tampered image, the ground truth image, the image predicted by the our method, and the image predicted by ManTra-Net.

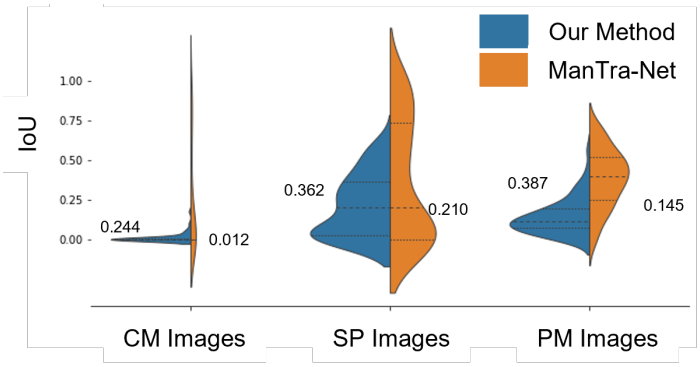

図 5 各画像ごとの IoU の改ざん手法別の分布．左から Splice, Copy-Move, Pixel-Manipulation による改ざん画像に対する IoU の分布が示されている．ヴァイオリンプロットの左側は提案手法を表し，右側は ManTra-Net を表している．

Fig. 5 Distribution of IoU for each image by tampering method. From left to right, the distribution of IoU for images tampered with by Splice, Copy-Move, and Pixel-Manipulation are shown. The left and right sides of the violin plot represent the proposed method and ManTra-Net, respectively.

取ってくるため，他の改ざん手法と比べ，改ざん検知の難易度が高く，いずれの手法においても IoU 平均は低い．Splice のように改ざん検知が人間でも比較的容易な改ざん手法では，提案手法では IoU が 1 に近く高精度で改ざん検知を実施出来ていることがわかる．Pixel-Manipulation については多くの画像について提案手法の IoU は ManTra-Net のものより高くなっている．いずれの改ざん手法に対しても，提案手法の IoU は ManTra-Net よりも平均的に高く，提案手法が有効であるといえる．

（3） 実験結果の考察

提案手法と ManTra-Net は同一の画像に対して必ずしも類似した予測結果を返すわけではない．図 4 の予測結果の例に示すとおり，提案手法による予測が良い場合であっても ManTra-Net による予測が良いわけではない．本実験の結果では，IoU および F1-Score の観点からは提案手法の方が優れていたが，この要因として，ManTra-Net が画像全体にわたり斑に改ざんピクセルを誤検知するようなモデルあることが考えられる．このようなモデルでは IoU の分母部分である実際の改ざん領域とモデルが予測した改ざん領域との和集合が大きくなるため，IoU は小さくなる．また適合率が小さくなるため F1-Score も同様に低く評価される．

電子帳票画像の改ざんは画像内の一部の数値または文字に対して行われることが多く，結果として改ざん領域が画像内に占める割合は極めて小さいものになる．そのため誤ってクラスを予測されたピクセルが僅かであっても IoU は大きく減少する傾向にある．モデル

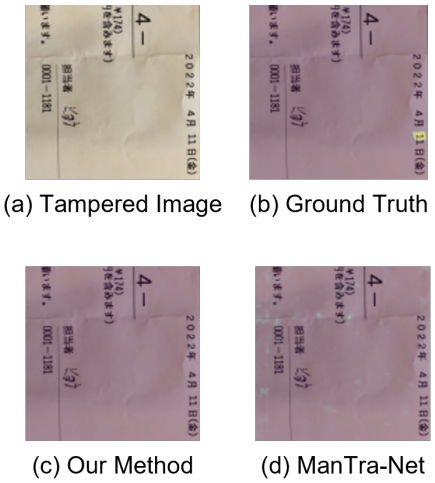

図 6 IoU が 0 である Copy-Move による改ざんが施された画像の例．それぞれ，(a) 改ざん画像，(b) 正解画像，(c) 提案手法による予測結果，(d)ManTra-Net による予測結果を示している．

Fig. 6 Examples of images tampered by Copy-Move with IoU of 0. Each of them shows (a) a tampered image, (b) a ground truth image, (c) a prediction result by the our method, and (d) a prediction result by ManTra-Net, respectively.

の実務適用を考えたとき、運用方法によっては，予測が改ざん領域内のピクセル全てを網羅している必要はなく，対象としている画像内の実務上着目すべき箇所に改ざんが存在するか否かが分かれば良い状況が想定される．そのような状況下では，IoU がある閾値以上となる画像を正解とする評価指標も有用である．例えば，IoU が 0.20 以上の画像を正しく改ざんが検知できた正解としたとき，本実験における正解率は，提案手法では 0.567 であり，ManTra-Net では 0.244 である．

（4） 提案手法の限界

提案手法および ManTra-Net ともに改ざんを検知できない画像が存在する．このような画像は図 5 に示すとおり Copy-Move を用いて作成された画像に多く存在する．Copy-Move による改ざんの検知はコピー元と貼り付け先が隣接している場合，フォント，色および解像度がほとんどなどしくなり，また改ざん領域の境界部分におけるピクセル値の変動も自然なものとなるため，人の目にはもちろん現状のモデルにおいても改ざんを検知することは出来ていない．

5. まとめ

本研究では，入力画像および出力画像を加工し Semantic Segmentation を用いて電子帳票画像の画像改ざんを検知する方法を提案した．また，典型的な 3 種類の改ざん手法（Copy-Move, Splice, Pixel-Manipulation）によるモデル訓練用改ざん画像を非改ざんから疑似的に生成する手法を提

案した．生成した疑似改ざん画像を用いて訓練された提案モデルが，既存手法と比較して IoU および F1-Score が高く，改ざん箇所以外の誤検知が少ないことを示した．一方，いずれの手法においても検知できない画像が Copy-Move を用いて改ざんされた画像に多く存在するため，モデルの更なる改良の必要性があることが示唆された．

参考文献

[1] Goel, A., Grünewald, F., Lingqvist, O. and Vainberg, G.: Graphic-paper producers: Boosting resilience amid the COVID-19 crisis, *McKinsey & Company* (2020).
[2] Zhao, H., Shi, J., Qi, X., Wang, X. and Jia, J.: Pyramid scene parsing network, *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2881–2890 (2017).
[3] Zhou, P., Han, X., Morariu, V. I. and Davis, L. S.: Learning rich features for image manipulation detection, *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1053–1061 (2018).
[4] Wu, Y., AbdAlmageed, W. and Natarajan, P.: BusterNet: Detecting Image Copy-Move Forgery With Source/Target Localization, *European Conference on Computer Vision (ECCV)*, Springer (2018).
[5] Yue Wu, W. A. and Natarajan, P.: ManTra-Net: Manipulation Tracing Network For Detection And Localization of Image ForgeriesWith Anomalous Features (2019).
[6] Kniaz, V. V., Knyaz, V. and Remondino, F.: The point where reality meets fantasy: Mixed adversarial generators for image splice detection, *Advances in Neural Information Processing Systems*, Vol. 32 (2019).
[7] Bappy, J. H., Simons, C., Nataraj, L., Manjunath, B. and Roy-Chowdhury, A. K.: Hybrid lstm and encoder–decoder architecture for detection of image forgeries, *IEEE Transactions on Image Processing*, Vol. 28, No. 7, pp. 3286–3300 (2019).
[8] Huh, M., Liu, A., Owens, A. and Efros, A. A.: Fighting fake news: Image splice detection via learned self-consistency, *Proceedings of the European conference on computer vision (ECCV)*, pp. 101–117 (2018).
[9] Tolosana, R., Vera-Rodriguez, R., Fierrez, J., Morales, A. and Ortega-Garcia, J.: Deepfakes and beyond: A survey of face manipulation and fake detection, *Information Fusion*, Vol. 64, pp. 131–148 (2020).
[10] Nguyen, H. H., Fang, F., Yamagishi, J. and Echizen, I.: Multi-task learning for detecting and segmenting manipulated facial images and videos, *arXiv preprint arXiv:1906.06876* (2019).
[11] Sabir, E., Cheng, J., Jaiswal, A., AbdAlmageed, W., Masi, I. and Natarajan, P.: Recurrent convolutional strategies for face manipulation detection in videos, *Interfaces (GUI)*, Vol. 3, No. 1, pp. 80–87 (2019).
[12] Li, Y. and Lyu, S.: Exposing deepfake videos by detecting face warping artifacts, *arXiv preprint arXiv:1811.00656* (2018).
[13] Korshunov, P. and Marcel, S.: Deepfakes: a new threat to face recognition? assessment and detection, *arXiv preprint arXiv:1812.08685* (2018).
[14] Rossler, A., Cozzolino, D., Verdoliva, L., Riess, C., Thies, J. and Nießner, M.: Faceforensics++: Learning to detect manipulated facial images, *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 1–11 (2019).
[15] Artaud, C., Sidère, N., Doucet, A., Ogier, J.-M. and Yooz, V. P. D.: Find it! fraud detection contest report, *2018 24th International Conference on Pattern Recognition (ICPR)*, IEEE, pp. 13–18 (2018).
[16] James, H., Gupta, O. and Raviv, D.: OCR Graph Features for Manipulation Detection in Documents, *arXiv preprint arXiv:2009.05158* (2020).
[17] Wen, B., Zhu, Y., Subramanian, R., Ng, T.-T., Shen, X. and Winkler, S.: COVERAGE – A NOVEL DATABASE FOR COPY-MOVE FORGERY DETECTION, *IEEE International Conference on Image processing (ICIP)*, pp. 161–165 (2016).
[18] Ng, T.-T., Hsu, J. and Chang, S.-F.: Columbia image splicing detection evaluation dataset, *DVMM lab. Columbia Univ CalPhotos Digit Libr* (2009).
[19] Dong, J., Wang, W. and Tan, T.: Casia image tampering detection evaluation database, *2013 IEEE China Summit and International Conference on Signal and Information Processing*, IEEE, pp. 422–426 (2013).
[20] Unel, F., Ozkalayci, B. O. and Cigla, C.: The Power of Tiling for Small Object Detection, *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 582–591 (2019).