

書類記入領域自動抽出手法の提案

A method of extracting entry areas from application form images

片岡 えり[†] 松本 光弘[†] 白木 宏明[†]

Eri Kataoka Mitsuhiro Matsumoto Hiroaki Shiraki

1 概要

金融機関や役所等では、顧客の申請等の手続きにおいて未だに紙書類（以下、帳票）を用いた事務処理が行われており、不備を判定するために必要なルールの決定、帳票の不備判定、記入内容のPCへのデータ入力を人手で行っている。これらを効率化するためには、不備判定ルールの定義、記入内容の不備判定、記入内容の登録の全ての処理について自動化を目指す必要がある。そこで本稿では、不備判定ルールの定義を効率化するための一機能として、帳票の中から顧客が記入した領域（以下、記入領域）のみを自動で抽出する手法を提案する。本手法では畳み込みニューラルネットワーク（CNN）を用いて帳票の画像を学習し、学習で獲得した特徴を帳票のテンプレートとみなす。これにより、帳票からテンプレートを除くことで記入領域のみを抽出できる。

2 背景と課題

金融機関や役所等では、帳票を用いた事務処理が人手で実施されている。例えば金融機関で口座を開設する場合、顧客は口座開設申込書に氏名や住所等の情報を記入し、申請する。金融機関では記入内容の不備を確認し、不備があれば顧客に返却、不備がなければ内容を手入力し、登録する。

帳票を用いた事務処理を効率化するためには、帳票をスキャンして得た画像（以下、帳票画像）を入力として(1)不備判定ルールの定義、(2)記入内容の不備判定、(3)記入内容の登録の3処理を効率化する必要がある。処理の流れは図1に示す通りである。

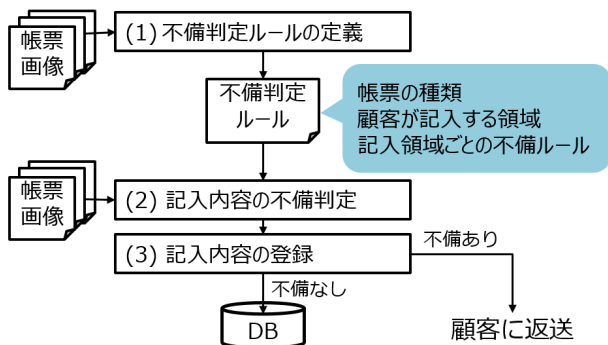


図1. 帳票を用いた事務処理の効率化

事務処理効率化を妨げる一因として、不備判定ルールの定義が煩雑であることが考えられる。不備判定ルールの定義では、記入内容の不備判定のため、顧客が記入する領

域を手動で設定しなければならないが、これらの作業は帳票の種類及び記入領域が多いほど時間がかかる。帳票のテンプレートを保持している場合は半自動で抽出可能だが、他社から送られてくる帳票を対象とする場合、事前に保持していないため、送付された記入済みの帳票のみから記入領域を定義しなければならない。

3 提案手法

上記課題を解決するため、書類記入領域自動抽出手法を提案する。本手法は、記入済みの帳票画像を用いてCNNによる分類学習を行う。学習によって得られる特徴が書類のテンプレートを表現することから、帳票のテンプレートを事前に用意することなく、テンプレート領域を獲得できる。そして獲得したテンプレート領域を帳票画像から取り除くことで、人手で記入された領域を自動で抽出する。

提案手法の処理の流れを図2に示す。まず前処理として、帳票画像の補正と学習を行う。次に、補正画像に対して記入領域の候補となる領域を矩形として求めるとともに、学習結果を用いてテンプレート領域を抽出する。そして、記入領域の候補からテンプレート領域と重複した領域を取り除く。最後に、複数枚の帳票に対して抽出した領域を結合し、記入領域を求める。本稿では提案手法の中心となる前処理とテンプレート抽出処理について、3.1節および3.2節で詳細を述べる。

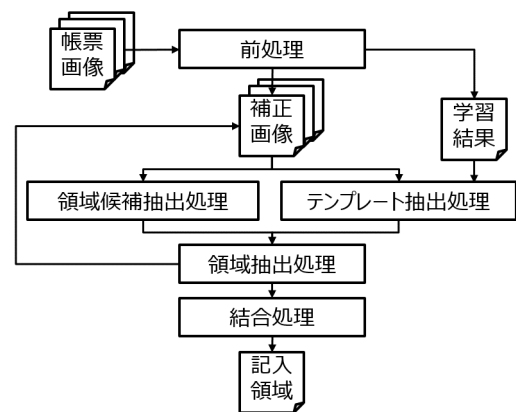


図2. 提案手法概要

3.1 前処理

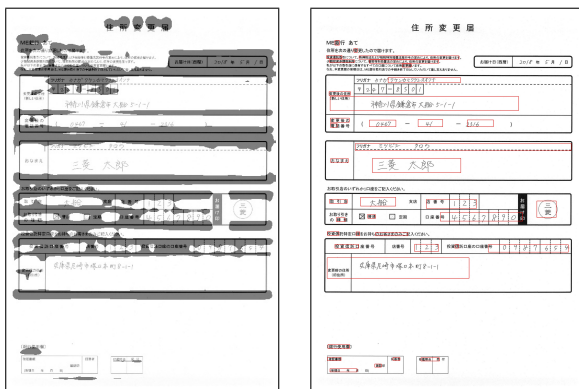
記入済みの帳票画像に対して、回転不変位相限定相関法（RIPOC）を用いて回転および上下左右の位置ずれを補正する。また、補正後の帳票画像を圧縮した画像を学習データとして、CNNを用いた分類学習を行う。

[†]三菱電機（株）情報技術総合研究所

3.2 テンプレート抽出処理

前処理の分類学習結果を用いて、分類に寄与した画像領域を抽出し、これを帳票のテンプレートとみなす。本稿では、Grad-CAM [3] を用いて領域を推定する。

例えば、ある帳票画像を学習し、テンプレート抽出処理を行った領域を重畳表示したところ、図 3 (a) のようになった。これより、帳票の枠線やタイトルがテンプレート領域として抽出できていることがわかる。また、テンプレート領域を用いて記入領域を抽出した結果を図 3 (b) に示す。赤枠で示した矩形が抽出した記入領域であり、記入されている領域上に矩形が出力できていることがわかる。



(a) テンプレート領域 (b) 記入領域抽出結果
図 3. 提案手法適用例

4 評価

提案手法を実装し、記入領域となる矩形の抽出精度を評価した。本評価では、1 つの正解領域に対して抽出領域が重ならない場合を抽出失敗、それ以外を抽出成功とみなし、ある帳票に含まれる正解領域のうち、抽出に成功した領域の割合を抽出精度として求めた。なお、抽出結果の領域は、結合処理の結合条件を調節し、出力する正解領域の個数を 20・30・40 に設定して求めた。例えば図 3 (b) の場合、正解領域の個数を 40 個に設定した。結果、14 個の正解領域に対して 12 個が結果領域と紐づいており、精度は 86% となる。

本評価では、某金融機関の実際の A4 サイズの記入済み帳票を 10 種類、各種 50 枚を用いた。また、それぞれの種類の帳票について、50 枚のうち 25 枚を学習、残り 25 枚を抽出精度の評価に用いた。帳票画像の大きさは 2048×2896 ピクセルとし、512×724 ピクセルにリサイズし、グレースケールに変換したものを学習に用いた。学習モデルの構築、学習、および Grad-CAM は Keras [2] を、バックエンドには Tensorflow [1] を用いて実装した。本評価における学習モデルは図 4 に示す通りである。今回は学習データの数が少ないため、畳み込み層とプーリング層の組み合わせを 3 層と全結合層を組み合わせで構築した。テンプレート抽出処理では、全結合層の直前のプーリング層を対象に Grad-CAM を適用し、テンプレート領域を求めた。

評価の結果、平均 65% の精度で記入領域を抽出できることが確認できた。詳細を表 1 に示す。また、帳票の種

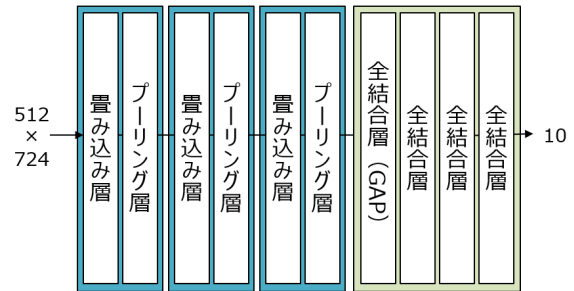


図 4. 評価で使用した学習モデル

類ごとに精度のばらつきが大きくなることがわかった。

本評価で用いた帳票は、記入が任意である項目が含まれる帳票がある。任意の項目は空欄である場合が多く、記入領域を抽出できずに精度が著しく低下した。これは、本手法における領域候補抽出処理で、記入領域の候補として抽出できないためである。そこで、必須項目のみを抽出精度評価の対象とした場合、任意項目が含まれている帳票 (2)(3)(5)(10) の精度は表 1 内の括弧内に示す通りとなり、平均 80% の精度が得られることが確認できた。

表 1. 評価結果

帳票	記入領域の個数		
	20	30	40
(1)	67%	67%	67%
(2)	34%(70%)	34%(70%)	34%(70%)
(3)	45%(60%)	50%(80%)	50%(71%)
(4)	88%	88%	82%
(5)	27%(84%)	27%(84%)	27%(84%)
(6)	100%	100%	100%
(7)	59%	59%	59%
(8)	86%	86%	86%
(9)	100%	100%	100%
(10)	50%(75%)	46%(69%)	50%(59%)

5 まとめと今後の課題

本稿では、CNN を用いた学習と Grad-CAM によって帳票のテンプレートとなる領域を求め、手書きの帳票から記入領域のみを抽出する手法を提案し、精度を評価した。結果、65% の精度で記入領域を抽出できることが分かった。今後は、テンプレート抽出処理の精度を向上するとともに領域候補抽出処理の手法を再考し、記入領域抽出精度の向上を目指す。

参考文献

- [1] Martín Abadi, et al. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org.
- [2] François Chollet, et al. Keras. <https://github.com/keras-team/keras>, 2015.
- [3] Ramprasaath R. Selvaraju, et al. Grad-CAM: Why did you say that? visual explanations from deep networks via gradient-based localization. *CoRR*, Vol. abs/1610.02391, 2016.