

# 画像解析と文字認識に基づく実用化に向けた 行政文書の DX 化に関する研究

白川 優羽<sup>1, a)</sup> KIM CHANNA<sup>1, b)</sup>  
鹿嶋 雅之<sup>1, c)</sup> 福元 伸也<sup>1, d)</sup> 渡邊 睦<sup>1, e)</sup> 西 正満<sup>2, f)</sup>  
坂元 哲史<sup>3, g)</sup> 西 保幸<sup>3, h)</sup> 福永 育郎<sup>3, i)</sup>

**概要:** 現在、鹿児島市役所では行政文書が紙媒体で数万件以上保管されており、時間経過による文書の劣化や必要文書の検索が困難といった問題が存在する。そこで、本研究は鹿児島市役所との共同で行い、行政文書をデータ化し、管理システムの構築・実用化を目的とする。撮影した行政文書に頂点検出を用いた背景の除去、文字座標及び直線検出を用いた必要部分の切り出し等の前処理を行う。その後、切り出した画像に生成 AI 等の外部 API を用いた文字認識を実行し、認識精度は最大 82%程度となっている。最後に認識結果をデータとして取得・管理を行う。今後の課題として、手書き文字の認識精度の向上、個人情報の保護を踏まえた実用化に向けた開発が挙げられる。

**キーワード:** 画像補正・幾何補正・画像推定、画像認識・理解、テキスト処理

## Research on the Digital Transformation of Administrative Documents for Practical Application Based on Image Analysis and Character Recognition

**Abstract:** Kagoshima City Hall stores many administrative documents in paper form, causing issues such as deterioration and difficulty in retrieval. This study aims to digitize these documents in collaboration with the city and develop a management system. Preprocessing techniques, including background removal using edge detection and cropping based on character coordinates and line detection, were applied. Character recognition was then performed using external APIs such as generative AI, achieving a maximum accuracy of approximately 82%. The recognized text was managed as structured data. Future challenges include improving handwritten character recognition and ensuring personal information protection.

**Keywords:** Image correction/geometric correction/image estimation, Image recognition/understanding, Text processing

### 1. 研究の背景と目的

現在、多くの自治体で DX 化が推し進められているが、人材不足等の原因で DX 化を進めることができず、データベースの活用等を進めることができていない自治体も多く存在する[1]。鹿児島市役所においても、境界確定調書と呼ばれる土地の境界を確定する文書が数万件以上紙媒体で保管されている。しかし、紙媒体での保管は、時間経過による劣化や外的要因による破損などが発生する懸念点がある。

また、数万件以上の資料を紙媒体で保管する場合、必要資料を探し出すことに非常に時間がかかってしまい、業務が長時間化してしまうという問題点も存在する。

本研究は、鹿児島市役所との共同で行い、境界確定調書と呼ばれる行政文書の情報をデータで管理し、行政文書の DX 化を目指すことを目的とする。行政文書を撮影した画像に画像解析による前処理と文字認識を実行し、必要情報の読み取りを行う。行政文書から読みとった情報でデータベースを作成し、管理するシステムの構築を目指す。また、行政文書のデータ化・管理システムをアプリケーションとして市役所で運用を行ってもらい、行政文書管理業務の DX 化による効率・利便性の向上を目指す。

### 2. 関連研究

本研究では、境界確定調書に前処理と文字認識を実行し、必要情報の読み取りを行う。本章では、文書への前処理と文字認識に関連する研究についてまとめていく。

1 鹿児島大学大学院理工学研究科情報科学専攻協創情報コース  
Department of Information Science, Graduate School of Science and Technology, Kagoshima University  
2 鹿児島大学大学院理工学研究科技術部  
Department of Technology, Graduate School of Science and Engineering, Kagoshima University  
3 鹿児島市建設局道路部道路管理課管理係  
Kagoshima City Construction Bureau, Road Department, Road Management Section, Management Section  
a) k0399564@kadai.jp  
b) channakimdis@gmail.com  
c) kashima@ibe.kagoshima-u.ac.jp  
d) fukumoto@ibe.kagoshima-u.ac.jp  
e) mutuyas256@gmail.com  
f) nishi@eng.kagoshima-u.ac.jp  
g) sakamoto-t49@city.kagoshima.lg.jp  
h) nishi-y21@city.kagoshima.lg.jp  
i) fukunaga-i26@city.kagoshim.lg.jp

## 2.1 文化遺産のデジタル文書に対する前処理による認識精度の変化に関する研究

Mariana Dias らは、デジタル形式で保存された文化遺産の文書に対し、画像処理や最適化手法等、様々な前処理を実行し、それらによって認識精度がどのように変化するかの検証を行っている[2]。結果としては、二値化、バイラテラルフィルタ、オープニング処理が認識精度に貢献するという結果になり、特にバイラテラルフィルタとオープニング処理が、文書タイプを問わずに精度向上に貢献する結果となっている。また、最適化手法である NSGA-II によるパラメータ最適化によって、各画像処理アルゴリズムのパラメータを最適化することによって、文字認識精度の向上に貢献するという結果になっている。

本研究では、余白部分の切り取り段階で、グレースケール変換、二値化、画像の回転補正、モルフォロジー変換の前処理を実行する。その後、一度文字認識を実行し、取得した文字の座標から必要部分の抽出を行うことで認識精度の向上と不必要な情報の除去を目指す。

## 2.2 U-NET によるセマンティックセグメンテーションを用いることによる手書きテキストの文字認識の精度向上に関する研究

Shruti Patil らは、手書きテキストに U-NET を用いたセマンティックセグメンテーションによって、テキストの手書き文字と印刷文字の領域を分類し、切り出すことによって、文字認識の認識精度の向上を図っている[3]。文字認識は Google cloud vision で実行しており、結果としては、セグメンテーションによる切り出しありの場合の方が、切り出しなしの場合より認識精度が約 2.7 倍向上するという結果になっている。

本研究では、文書全体に Google Cloud Vision による文字認識を実行することでテキストの座標を取得し、取得した座標を用いてテキストの切り出しを行うことで、不必要情報の除去を行う。

## 2.3 n-gram 出現回数の特徴量とした機械学習による文字認識後の誤り検出システムに関する研究

Shafqat Mumtaz Virk らは n-gram 出現回数という単語を n 文字の連続部分列に分割したものの出現回数特徴量とした機械学習モデルによって文字認識後の誤り検出システムの作成を行った。分割した文字の出現回数を特徴量とすることで一般化性能の向上とモデルの計算効率を向上させ、ICDAR2019 における誤り検出スコアを7つの言語で上回る結果を示している[4]。

本研究では、複数のシステムでの文字認識結果を組み合わせや、市役所の行政文書であることを活用した後処理による精度向上を目指す。

## 2.4 大規模言語モデルと従来の光学文字認識における手書き文字認識の認識精度の比較に関する研究

Seorin Kim らは、Chat gpt4v と Claude 3.5 Sonnet の大規模言語モデルと EasyOCR, Keras, Pytesseract, TrOCR の従来の文字認識モデルでの手書き文字認識モデルの認識精度の比較を行っている[5]。ベルギーの相続申告書のスキャン画像に文字認識を実行し、大規模言語モデルの方が従来の文字認識モデルよりも手書き文字の認識精度が高いという結果となっている。また、Chatgpt4v は行ごとに入力された際の認識精度が高く、Claude 3.5 Sonnet はスキャン画像全体を入力された際の認識精度が特に高いという傾向が見られている。

また、本研究では以下の研究を引き継ぐ形で研究を行っている。

## 2.5 画像解析による行政文書の DX 化に関する研究

鹿儿島市役所の行政文書の DX 化を図るために、アクリル板による歪みの除去と黒背景による前処理の補助を目的とした撮影台の作成を行った[6]。

また、撮影した行政文書に二値化やモルフォロジー変換による輪郭強調とノイズの除去、回転補正による傾きの軽減等を行い、背景部分の切り取りを行っている。その後、切り取りを行った画像の登録番号部分を切り出し、切り出した画像に Google cloud vision で文字認識を実行し、登録番号部分の読み取りを行っている。認識エラーの際に番号が連番であることを利用し、前後の番号と比較して後処理による修正を実装することで、認識精度を 95.54%まで向上させている。

しかし、行政文書の切り取りの際に文書に付箋が貼ってある場合には、切り取りが失敗するといった問題点も存在した。

## 3. 研究手法

本研究では図 1 にある境界確定調書と呼ばれる土地の境界を決定する資料を扱う。

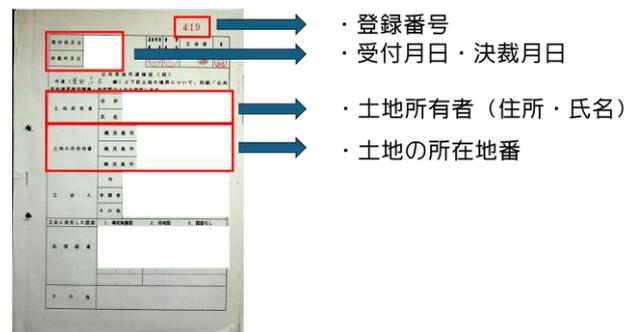


図 1 行政文書と必要情報

Figure 1 Administrative Documents and Required Information.

境界確定調書における、登録番号, 受付月日, 決裁月日, 土地所有者の住所, 氏名, 土地の所在地番の情報を読み取り, これらの情報を管理・検索できるシステムの作成を目指す。

また, 本研究で取り扱う境界確定調書においては, 個人の氏名や住所等の個人情報が含まれている。そのため, 本論文中に示す行政文書のデータは, 個人情報漏洩を防ぐために, 一部加工したものとなっている。

### 3.1 処理手順

初めに, 行政文書の撮影をドキュメントカメラ「CZUR.ET18Pro」で行う。撮影時には, 外部からの光が入らないように暗幕で覆い, 図2で示す撮影装置を用いる。文書と背景部分をより明確にするために, 黒い板の上に行政文書を置き, 上からアクリル板で抑えることによって文書の歪みの軽減を図る。撮影装置は現在鹿児島市役所に設置しており, 文書の撮影を行ってもらっている。



図 2 行政文書撮影装置

Figure 2 Administrative Document Imaging Device.

その後, 撮影した文書に前処理を実行することで行政文書のみを切り取り, 背景部分の除去を行う。背景の除去を行った画像に文字認識を実行し, 必要情報の読み取りを行い, 読み取った情報を SQL に保存することで, データベースとして管理を行う。

### 3.2 前処理段階

Python の OpenCV を用いて, 撮影した行政文書に画像処理を行い, 背景部分の除去を行う。

#### 3.2.1 二値化とモルフォロジー変換

初めに, 撮影した画像にグレースケール変換と二値化を実行し, 行政文書の輪郭の強調を行う。二値化の際の閾値

は 100 として二値化を行った。その後, モルフォロジー変換として, オープニング処理とクロージング処理を実行し, ノイズの除去を行う。

#### 3.2.2 頂点検出

次に, 画像内の行政文書の輪郭の検出を行う。画像内で四角形の輪郭の検出を行い, 検出した四角形の輪郭の内, 最も面積が大きいものの座標を取得する。この時取得した座標が, 行政文書の頂点となる。この時, 行政文書に付箋が存在する場合には, 付箋が頂点の一つとして反応してしまい, 図3のように背景切り取りでエラーが発生してしまう。



図 3 付箋による背景切り出しの失敗

Figure 3 Failure to extract background due to sticky notes.

このエラーを防ぐために, 検出する頂点は行政文書の右上以外の3点のみとする。右上の頂点は, 図4のように右上以外の頂点座標から計算して求める。

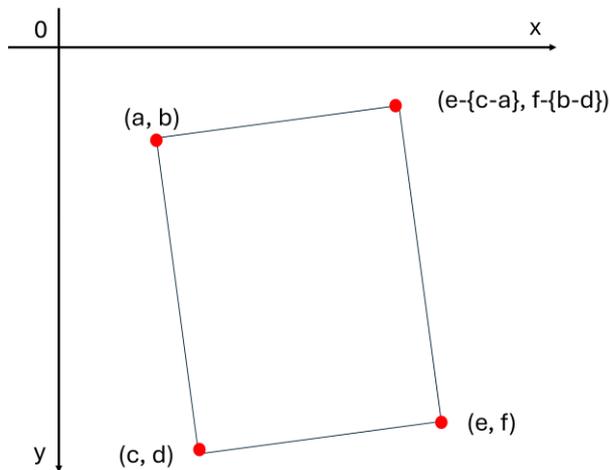


図 4 右上の頂点の計算

Figure 4 Calculation of the upper-right vertex.

#### 3.2.3 回転補正と行政文書の切り取り

検出した頂点の座標から行政文書の傾きを求め, 回転補正を行うことで行政文書の傾きの軽減を行う。回転補正により, 行政文書の頂点の座標がずれてしまうため, 先程と同様にもう一度四角形の輪郭の検出を行い, 回転補正後の行政文書の頂点座標を取得する。最後に, 検出した行政文書の頂点の座標を基準に, `cv2.boundingRect` で行政文書の

切り取りを行い、図5のように背景部分の除去を行う。

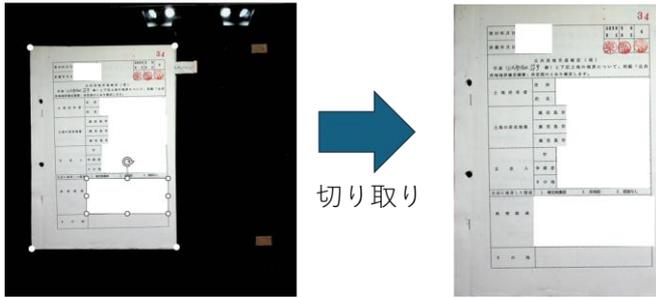


図5 背景部分の除去  
Figure 5 Background removal.

### 3.3 文字認識段階

背景の除去を行った行政文書から文字認識に必要な部分のみを切り出し、文字認識を実行する。

#### 3.3.1 必要部分の切り出し

背景の除去を行った行政文書に、Google cloud vision で文字認識を実行する。Google cloud vision には認識した単語の座標を取得する機能が存在する。この機能を利用して、「受付月日」、「決裁月日」、「土地所有者」、「土地の所在地番」という文字の座標を取得する。認識した座標の x 座標の最大値と y 座標を基準点とし、各項目で基準点から切り出す範囲を調整することによって、必要部分の切り出しを行う。

必要な単語の座標を取得するためには、Google cloud vision による行政文書全体の文字認識の結果から、必要な単語を探し出す必要がある。しかし、Google cloud vision では意図しない部分で単語を区切ってしまう場合がある。(例として、「受付月日」を「受付」と「月日」の二つの単語として扱ってしまう場合がある)。従って、必要単語の検索は、「受付月日」、「付月日」、「月日」、「日」の順番で単語が見つかるまで頭文字を1つずつ減らしながら検索を行い、必要情報の座標を取得する。また、「土地の所在地番」においては、「番」という単語が文書内で頻繁に出てくることが原因で誤検出が発生する為、頭文字ではなく末尾から文字を減らして検索を行う。

必要単語が正しく文字認識ができなかった場合は、行政文書のフォーマットが固定であることを利用し事前に設定した座標で切り出しを行う。これにより必要部分の切り出しを行うが、いくつかの行政文書ではずれが発生してしまう。

また、「受付月日」と「決裁月日」、「土地所有者」と「土地の所在地番」のように、同じ文字が含まれている場合に、誤った単語の座標を取得してしまう場合が存在する。必要な単語の取得は、行政文書の上から行うため、「受付月日」と「土地所有者」の情報を取得した後に、図6のようにこれらの単語から上の部分を切り取り、切りとった画像にも

う一度 Google cloud vision で文字認識を実行する。これにより単語の共通部分による誤った単語の取得を防いでいる。

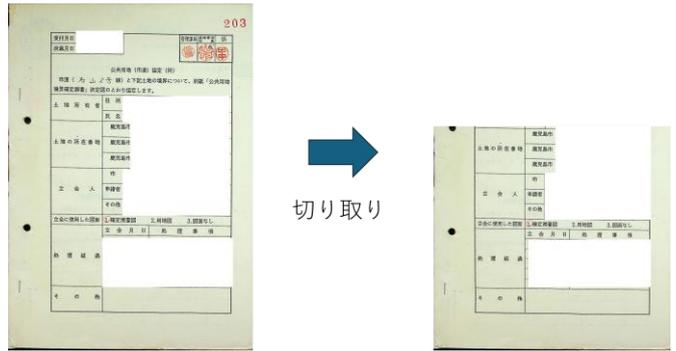


図6 単語取得エラー改善のための切り取り  
Figure 6 Cut for improving word acquisition errors.

各単語の取得した座標の x 座標の最大値と y 座標の値を基準に、各項目で切り出す範囲を決定することで、受付月日、決裁月日、土地所有者の氏名と住所、土地の所在地番の情報の部分の切り出しを行う。この時、取得した座標情報のみでの切り出しでは、図7のように文書のずれや傾きによって、切り出しによる文字の欠損が発生してしまう。

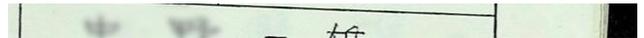


図7 切り出しによる文字の欠損  
Figure 7 Character loss due to cut.

認識の必要がある文字は、行政文書であることから矩形の枠に囲まれている。このことを利用し、文字の座標情報を取得後に文字の範囲より少し広い範囲で切り出しを行う。その後、切り出した画像に図8のように OpenCV の Canny 法で直線の検出を行う。画像内の緑の線が検出された横線、赤の線が検出された縦の線となっている。

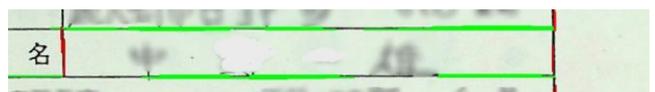


図8 直線検出  
Figure 8 Line detection.

検出された直線のうち、縦と横の直線それぞれの最も外側の直線の座標を取得し、取得した座標を用いてもう一度切り出しを行う。この手法により、図9のように文字の欠損がない状態での切り出しに成功した。



図9 直線検出による切り出し  
Figure 9 Cut by straight line detection.

以上の直線検出を用いた切り出しを用いて、図10のように受付月日、決裁月日、土地所有者の氏名と住所、土地の所在地番の切り出しを行う。また、登録番号については、

必ず行政文書の右上に記載されているため、右上の座標を指定して切り出しを行う。

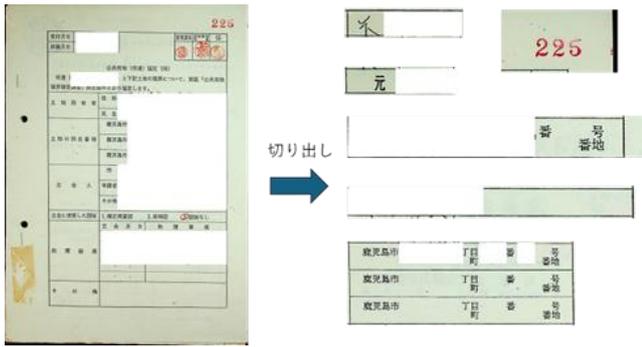


図 10 必要部分の切り出し

Figure 10 Cutting the necessary parts.

### 3.3.2 文字認識

必要情報を切り出した各画像に文字認識を実行することで、必要情報の読み取りを行う。文字認識は、Google cloud vision, Chat gpt 4v, Claude 3.5 Sonnet の 3 種類の API で実行する。また、読み取った情報は SQL にデータベースとして保存を行う。

### 3.4 データ管理

認識した必要情報は SQL にデータベースとして保存をし、図 11 のように管理システム上で、必要情報の管理、検索、閲覧ができるシステムの開発を行っている。



図 11 データ管理システム

Figure 11 Data Management System

## 4. 評価実験

必要部分の切り出しによる認識精度の変化を検証するために、行政文書における切り出しを行った場合と切り出しを行わなかった場合での認識精度の比較を行った。

### 4.1 実験方法

行政文書 30 枚における、受付月日、決裁月日、土地所有者の住所と氏名、土地の所在地番、登録番号の情報を行政

文書の切り出しありと切り出しなしで文字認識を実行し、認識精度の比較を行う。文字認識は、Google cloud vision, Chat gpt 4v, Claude 3.5 Sonnet の 3 種類で行う。認識精度は、文字認識によって読み取ることができた文字数から、行政文書に実際に記載されている文字数の割合を計算する。また、行政文書には印刷されている文字と手書きの文字が記載されているため、認識精度は印刷文字のみ、手書き文字のみ、印刷文字と手書き文字の合計の 3 種類で求めた。

## 4.2 実験結果

実験結果を以下の表 1 に示す。

表 1 文字認識の認識精度

table 1 Character recognition accuracy.

cloud_vision	印刷	手書き	印刷+手書き
切り出しあり	0.8558	0.6824	0.7763
切り出しなし	0.8802	0.7078	0.7940
gpt4v	印刷	手書き	印刷+手書き
切り出しあり	0.9199	0.7304	0.8256
切り出しなし	0.4110	0.3706	0.3908
claude	印刷	手書き	印刷+手書き
切り出しあり	0.9085	0.7322	0.8201
切り出しなし	0.8456	0.8178	0.8287

### 4.2.1 印刷文字

印刷文字は、google cloud vision は切り出しによって約 3% 認識精度が低下する結果となった。対して、chat gpt 4v は約 50%、Claude 3.5 Sonnet では約 6% 切り出しにより認識精度が向上した。印刷文字においては、おおむね認識精度が切り出しによって向上するという結果になった。

### 4.2.2 手書き文字

手書き文字では、chat gpt 4v では切り出しにより認識精度が約 36% 向上した。対して、Google cloud vision では約 3%、Cloud 3.5 Sonnet では約 6% 認識精度が低下した。手書き文字においては、切り出しにより認識精度が一部低下してしまう結果となった。

### 4.2.3 印刷文字+手書き文字

印刷文字と手書き文字の合計における認識結果は、Google cloud vision では約 2% 認識精度が低下、Chat gpt 4v においては約 43% 認識精度が向上、Claude 3.5 Sonnet では約 0.8% 認識精度が低下する結果となった。全体として、切り出しによって認識精度が向上するモデルと低下するモデルに分かれる結果となった。

## 5. 考察

### 5.1 Chat gpt 4v の認識精度

Chat gpt 4v では、切り出しによって認識精度が大きく向上するという結果となった。この認識精度の向上には、切

り出しによって切り出しなしで認識を行った際に発生したエラーが改善されたことが理由として考えられる。Chat gpt 4v を用いて、必要部分の切り出しを行っていない境界確定調書の文字認識を行うと、一部資料で「画像からの文字の認識ができません」というエラーメッセージが返され、文字認識を実行しない場合があった。これは、境界画定調書が文書として構造の複雑さや、文書に記載されている文字数が多いことから認識に失敗してしまったことによるエラーであると考えられる。対して、必要部分の切り出しを行うことによって、文書の構造をシンプルにすると同時に、読み取る必要のある文字のみに文字認識を実行することができる。このことにより、必要部分を切り出した画像での Chat gpt 4v の文字認識では、文字の認識ができないというエラーがほとんど発生せず、結果として認識精度が大きく向上するという結果につながったと考えられる。

## 5.2 必要部分切り出しと周辺情報

Google Cloud Vision と Claude 3.5 Sonnet の手書き文字において、認識精度は必要部分の切り出しによって低下する結果となった。これは、切り出しを行ったことにより周辺情報が失われたことが原因ではないかと考えられる。必要部分の切り出しを行った認識結果と切り出しを行わない場合の認識結果を比較したところ、日付と地名において切り出しなしの方が正しく認識できているにもかかわらず、切り出しを行うと認識を失敗している場合が多くみられた。AI が用いられた OCR モデルでは、文脈や構造の理解を行ったうえで文字の推定を行う。そのため、周辺情報が残っている切り出しの場合の方がより正確な認識結果となる場合があるのではないかと考えられる。例として、受付月日の切り出しありと切り出しなしの場合の認識結果を図 12 に示す。

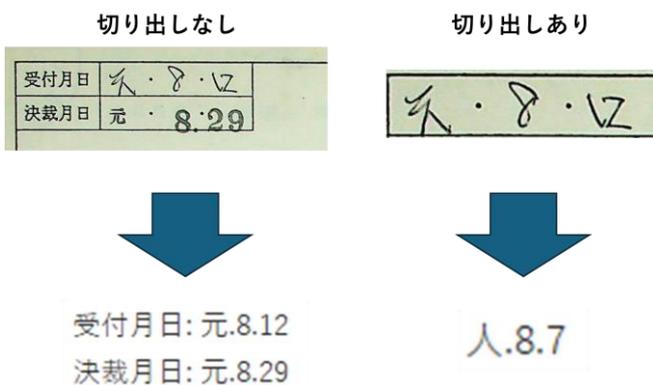


図 12 切り出しの有無における受付月日の認識結果  
Figure 12 Recognition Results for Reception Dates with and without Cutting

図 11 において、切り出しを行った場合の認識結果では、元年を表している「元」という字が「人」という字で誤認識されてしまっている。対して、切り出しを行っていない場合の認識結果では、正しく「元」という字の認識に成功

している。このことは、周辺に記載されている情報から認識しようとしている文字が年月日を表しているということから推定したうえで認識を行った結果なのではないかと考えられる。ただし、図 13 のように切り出しを行った場合の方が正確に認識できている場合もある。

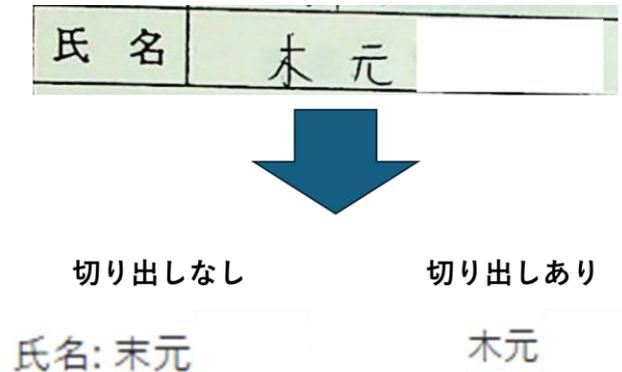


図 13 切り出しによる認識精度の向上例  
Figure 13 Example of Improved Recognition Accuracy Through Cutting

また、境界確定調書の右上にある登録番号においては、切り出しによる強調を行わなかった場合、認識されない場合が多数見られた。このように、切り出しによる文字の強調や不必要情報の除去により、認識精度が改善されている場合も多くみられた。したがって、必要部分の切り出しにおいては、不必要な情報の除去や、認識した文字の強調ができるという利点があるが、周辺情報が失われてしまうことによる推定精度の低下につながることも考慮する必要がある。

## 6. 結論

本研究は、鹿児島市役所における境界確定調書から画像解析による前処理と文字認識を用いて必要情報の読み取りを行った。前処理の段階では、輪郭検出や回転補正を行い、付箋によるエラーの改善のために、検出する頂点を3点のみに限定し、背景部分の除去を行った。文字認識段階では、必要部分の文字の座標と文字の周りの直線の座標を取得し、それらの座標を基準に必要部分の切り出しを行った。その後、Google cloud vision, Chat gpt 4v, Claude 3.5 Sonnet の API を利用し、必要情報の文字認識を行った。

切り出しによる認識精度の検証を行ったところ、Chat gpt 4v では認識精度が向上したが、Google cloud vision と Claude 3.5 Sonnet では一部認識精度が低下する結果となった。必要部分の切り出しによる不必要な情報の除去や目的の文字の強調には成功し、Chat gpt 4v での認識精度は上昇したが、同時に周辺情報の損失による推定精度低下につながり、このことにより一部認識精度が低下してしまったりと考えられる。

## 7. 今後の課題

### 7.1 認識精度の向上

鹿児島市役所の取り扱う境界確定調書には、個人の氏名など個人情報が含まれている。そのため、文字認識によって必要な情報をデータとして取得した後、市役所の職員による誤りの修正作業が行われる。そのため、認識精度を向上させることにより、修正の必要部分を限りなく少なくし、職員の作業負担の低下を目指す必要がある。

認識精度の改善手法として、後処理による認識精度の向上を図ることが可能ではないかと考える。Youssef Bassil らの研究では、Google のもしかして機能を利用した後処理を実装することにより、認識精度の向上に成功している[7]。本研究で扱う境界確定調書において、「土地の所在地番」の部分には基本的に鹿児島市内の住所が記載される。そのため、事前に鹿児島市の住所のデータを準備しておき、文書の認識結果と住所のデータの比較を行い、誤り部分の修正を行うことで、認識精度の向上を図ることができるのではないかと考える。

### 7.2 鹿児島市役所での実用化

本研究では、境界確定調書の撮影から、文字認識による必要情報の取得、データでの管理までを市役所で実用化してもらうことを目的としている。そのため、作成した行政文書への前処理、文字認識を実行するシステムを市役所の方々に扱ってもらえるアプリケーションの形にする必要がある。現在、特定のファイルに読み取りを行いたい行政文書の画像を保存してもらい、実行ボタンを押すだけで、画像の前処理、文字認識を行い、自動でデータベースに保存を行うアプリケーションの作成を並行して行っている。



図 14 文字認識アプリケーション

Figure 14 Character Recognition Application

作成したアプリケーションは、市役所の方々に実際に試験運用を行ってもらい、フィードバックを受けたうえで UI や利便性、機能などの改善を行っていくことを検討している。

また、市役所で運用を行ってもらう上で、個人情報の取り扱いが大きな課題となっている。今後の実用化の方針において、個人情報が含まれている境界確定調書を外部の API を用いて文字認識を行うことが個人情報保護の観点から許可されない可能性が存在する。そのため、今後文字認識の手法を一部変更することが必要とされる可能性がある。この問題に対する対策としては、Google の Gemma3 などの、ローカル環境で扱うことができる LLM を代替システムとして導入することを検討している。今後も、システムの有用性と市役所での運用の実現可能性を検討しながら、開発を進めていきたい。

## 参考文献

- [1] IPA 独立行政法人情報処理推進機構. DX 動向 2024 進む取り組み, 求められる成果と変革. <https://www.ipa.go.jp/digital/chousa/dx-trend/dx-trend-2024.html>, 参照 February.1,2026.
- [2] Mariana Dias, Carla Teixeira Lopes. Optimization of Image Processing Algorithms for Character Recognition in Cultural Typewritten Documents. ACM journal on Computing and Cultural Heritage, Volume 16, issue4, 2023.
- [3] Shruti Patil, Vijayakumar Varadarajan, Supriya Mahadevkar, Rohan Athawade, Lakhan Maheshwari, Shrushti Kumbhare, Yash Garg, Deepak Dharrao, Pooja Ka mat and Ketan Kotecha. Enhancing Optical Character Recognition on Images with Mixed Text Using Semantic Segmentation J. Sens. Actuator Netw. 2022.
- [4] Shafqat Mumtaz Virk, Dana Dannells and Azma Sheikh Muhammad. A Novel Machine Learning Based Approach for Post-OCR Error Detection. Proceedings of the International Conference on Recent Advances in Natural Language Processing(RANLP 2021), pp 1463~1470, September 2021.
- [5] Seorin Kim, Julien Baudru, Wouter Ryckbosch, Hugues Bersini, Vincent Ginis. Early evidence of how LLMs outperform traditional systems on OCR/HTR tasks for historical records. arXiv preprint arXiv:2501.11623. 2025.
- [6] 二石美奈. 画像解析による行政文書の DX 化に関する研究. 鹿児島大学工学部学士論文, 2024.
- [7] Youssef Bassil, Mohammad Alwani. OCR Post-Processing Error Corection Algo rithm using Google Online Spelling Suggestion. Journal of Emerging Trends in Computing and Information Sciences, Vol. 3, No. 1, 2012.