

# 帝国議会会議録におけるレイアウト解析手法の比較

飯田紗也香<sup>†</sup> 竹本有紀<sup>†</sup> 石川由羽<sup>†</sup> 高田雅美<sup>†</sup> 城和貴<sup>†</sup>

**概要:** 本稿では、OCRの精度向上のためにレイアウト解析の手法の提案を行い、その有用性を確認する。OCRの精度は文字切り出しの精度に依存し、文字切り出しの精度はレイアウト解析の精度に依存する。つまり、レイアウトが複雑になればなるほどOCRの難易度が上がる。一般的な書籍のレイアウトは文書以外に見出し、図表、ルビ、枠線などを含むものも存在し、必ずしも単純であるとは限らない。そこで、一般的なレイアウトの書籍にOCRを行うには、書籍に対してレイアウト解析を行う必要がある。レイアウト解析の方法として、帝国議会会議録に対してヒストグラムを用いた手法とSemantic Segmentationを利用した手法の提案を行う。提案手法を適用した結果に文字切り出しを行い、それぞれの結果とレイアウト解析を行わなかった場合の結果を比較し、提案手法の有用性を確認する。

**キーワード:** レイアウト解析, 帝国議会会議録, FCN(Fully Convolutional Network)

## Comparison of Layout Analysis Method in Imperial Congress Meeting

SAYAKA IIDA<sup>†</sup> YUKI TAKEMOTO<sup>†</sup>  
YU ISHIKAWA<sup>†</sup> MASAMI TAKATA<sup>†</sup>  
KAZUKI JOE<sup>†</sup>

### 1. はじめに

文書作成ソフトウェアなど、デジタルな執筆の手段が存在しなかった時代の書籍には、テキストデータが存在しない。例えば、明治から昭和初期にかけて刊行された近代の書籍がそれに該当する。国立国会図書館[1]では、国立国会図書館デジタルコレクション[2]というWebサービスで近代書籍を、Web上で画像データとして公開している。公開されている書籍画像はテキスト化されていないため、本文検索や音声読み上げが不可能などの問題がある。

書籍画像をテキスト化する方法としては、OCR(Optical Character Recognition)ソフトウェア[3]を用いることが一般的である。OCRの精度は文字切り出しの精度に依存し、文字切り出しの精度はレイアウト解析の精度に依存する。つまり、正確なOCRには正確なレイアウト解析が必要であり、レイアウトが複雑になればなるほどOCRの難易度が上がる。見開きの文書のみならずレイアウトならばレイアウト解析は容易であるが、一般的な書籍には、見出し、図表、ルビ、枠線などを含む書籍も存在し、そのレイアウトは必ずしも単純であるとは限らない。一般的なレイアウトの書籍にOCRを行うには、書籍に対してレイアウト解析を行う必要がある。

そこで、本稿ではヒストグラムによるレイアウト解析の手法とFCN(Fully Convolutional Networks)[7]によるSemantic Segmentation[8]を用いたレイアウト解析の手法を提案する。それらを帝国議会会議録に適用し、レイアウト解析を行わ

なかった場合と文字切り出しの精度を比較し、提案した手法の有用性を示す。帝国議会会議録は業者が撮影し、適切にデジタル化された書籍画像である。

以下、本稿の構成を示す。2章でレイアウト解析の既存研究を紹介し、3章でレイアウト解析の対象とする帝国議会会議録の概要を述べる。4章では、4.1節でヒストグラムによるレイアウト解析の手法の提案、4.2節でSemantic Segmentationの手法であるFCNを用いたレイアウト解析の手法の提案を行い、5章では、実験と実験結果、考察を述べる。6章でまとめについて述べる。

### 2. レイアウト解析に関する既存研究

書籍のレイアウト解析を行う既存研究を2つ紹介する。1つ目はRay Smithによる”Hybrid page layout analysis via tab-stop detection”[4]、2つ目はGranaらによる”Layout analysis and content enrichment of digitized books”[5]の中で使用されている文書領域と図領域の判別についてである。

1つ目の”Hybrid page layout analysis via tab-stop detection”について説明する。この研究ではページの列・段落構造を見つけるため、各行頭の空白部分を用いて外接短径の作成を行う。まず、各行頭の外接短径を作成し、文書画像中の文字・画像等を分類して文字の矩形のみを統合した領域、列・段落の境界であると予想される空白に沿った短径を作成する。この3つを組み合わせることで行頭の並びが直線でないレイアウトの解析を行う。例えば、文書中に挿入された図形の外周で文字列の折り返しが行われている場合、図形をよけて文書領域のみを切り出すことができる。

2つ目の”Layout analysis and content enrichment of

<sup>†</sup> 奈良女子大学  
Nara women's University



図 1 会議録の例

digitized books”について説明する。この研究は古い書籍を自動的に分析するシステムの開発を目的に、書籍のレイアウト解析を行っている。レイアウト解析は SVM を用いて文書と図、それぞれの特徴を学習する。対象とする書籍は古い書籍であるが、時代に関わらず文書領域には図領域と比較して強い反復パターンが含まれている。自己相関行列を用いて局所相関に基づく特徴を抽出し、切り出した領域が文書か図か判別する。また、抽出した画像をカテゴリーごとに分類し、関連する現代のコンテンツに結び付けユーザがコンテンツを見つける補助を行う。

紹介した2つの手法では空白、文書、図領域の抽出に画素射影ヒストグラムを用いている。ヒストグラムによる領域抽出はレイアウト解析において一般的な手法である。本稿では領域の抽出に用いる手法として画素射影ヒストグラムを用いる場合と、Semantic Segmentation を用いる場合を比較する。Semantic Segmentation は自動運転技術や医療分野に使用される、画像を画素ごとに意味づけする技術であり、一般的にレイアウト解析の領域抽出の手段として使用されていないが、画素ごとに領域抽出が可能である。よって、レイアウト解析に有用であると推測される。

### 3. 帝国議会議録

帝国議会議録[6]は、国立国会図書館議会官庁資料室が所蔵する明治23年から昭和22年の帝国議会議会期の本会議・委員会の速記録である。会議録は国立国会図書館により、帝国議会議録検索システムという Web サービスで画

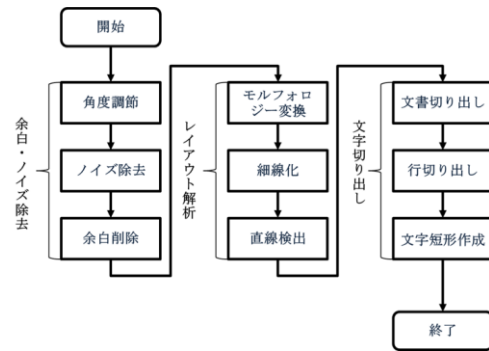


図 2 ヒストグラムによる手法の流れ

像データとして公開されている。戦後の会議録である第88回から第92回会議録については、既にテキスト化がされており、テキストの全文、法案、附則による検索ができる。しかし、それ以前のもはテキスト化されておらず、目次・索引検索、発言者検索のみ可能で、早急なテキスト化が求められている。公開されている画像は活版印刷の単ページの書籍画像で、保存状態が良く、業者が撮影を行い、適切にデジタル化された、単純なレイアウトの書籍画像である。そのため、レイアウト解析が容易である。会議録の主なレイアウトは2段から5段の枠線に区切られ、画像の右部分に題があるものと無いもの計8種類であり、図表やルビは含まない。具体的には、図1に示す左上が2段、右上が3段、左下が4段の題あり、右下が5段の題ありのレイアウトである。例外として白紙、正誤表などを含む。Webで公開されている画像の形式はマルチページの Tiff であり、色数は白黒2値、幅高は約3300\*4700pxである。

## 4. 提案手法

### 4.1 ヒストグラムによるレイアウト解析

会議録に対するレイアウト解析の手法として、ヒストグラムによるレイアウト解析を提案する。提案する処理の流れは図2のとおりである。処理は大まかに余白・ノイズ除去、レイアウト解析、文字切り出し処理に分けられる。入力する画像は白黒2値の BMP 形式の会議録画像である。

余白・ノイズ除去処理では、角度調節後、メディアンフィルタによるノイズ除去を行う。ガウシアンフィルタによる平滑化を行い、2値化し画像のかすれを軽減し、余白部分を削除する。

次に、レイアウト解析を行う。レイアウト解析には確率的ハフ変換による直線検出と画素射影ヒストグラムを用いる。ハフ変換の前処理として細線化を行う際、ノイズが細線化の妨げになるため、モルフォロジー変換によるノイズの除去を行う。ハフ変換により検出された直線の位置、長さを用いて画素射影ヒストグラムで枠線の区切りごとに外接矩形を作成し、文書部分ごとに会議録画像を切り出す。

最後に、文字短形作成処理では、まず文書部分を行ごと



図 3 文字短形作成の失敗

に分割する。まず、レイアウト解析後、切り出された文書部分の余白の削除を行う。水平画素射影ヒストグラムを用いて、ヒストグラムが 0 になる地点を行の区切りとみなし、行ごとに切り出しを行う。切り出された行に対して文字の外接短形を作成する。文字短形作成処理は、対象とする書籍が、フォントが規格化される以前の活版印刷であるため、形状や文字幅がそろっていない文字の外接短形を作成する必要がある。文字のパーツごとに求めた外接短形を上下に連結することで文字分離の対策を行っている。上下のパーツ間の距離が、文字間隔の平均の 0.3 倍以下、かつ統合後の文字の縦幅が行中の文字高の平均の 1.2 倍以下となるときに外接短形を連結する。

ヒストグラムによる会議録画像のレイアウト解析にはいくつか問題がある。まず、レイアウト解析に使用した確率的ハフ変換は、直線など数式で表現される軌跡を、その軌跡上の画素の個数により検出する手法で、形状の一部が欠損していても検出が可能であるという長所がある。しかし、書籍画像に適用したとき、印刷の濃い書籍画像や、文字数が多い行が存在すると、行を直線と誤認識するという問題がある。次に、行切り出し処理では、水平画素射影ヒストグラムが 0 になる地点で行を区切っているが、極端に行幅が狭い場合や、インクの染みが存在する場合、正確に行の区切りを検出できない。最後の文字短形作成処理では、活版印刷による文字形状のばらつきに対応するため、上下パーツの統合による外接短形の作成を行っている。図 3 に文字短形作成の失敗例を示す。図 3 左はパーツ間の距離が原因で分離した場合、図 3 中央は文字のかすれが原因で分離した場合、図 3 右は文字サイズが小さいため連結された場合である。文字短形作成処理では文字の上下のパーツ間の距離、文字間隔の平均、統合後の文字の縦幅、文字高の平均を用いて外接短形を連結しているが、この手法においては、図 3 左で示したようなパーツ間の距離が文字間隔の平均の 0.3 倍より大きい「云」や「二」などの文字が分離、図 3 中央に示したように文字がかすれている場合、パーツ間の距離が離れていると判別され分離する。また、上下の文字間隔が狭い場合や文字サイズが平均より大きい場合、連結がほとんど行われず、文字間隔が極端に広い場合や文字サイズが平均より小さい場合、図 3 右に示したように連結する必要のない外接短形を連結する。

#### 4.2 FCN を用いたレイアウト解析

会議録画像に対するレイアウト解析の手法として、FCN による Semantic Segmentation を用いたレイアウト解析を提

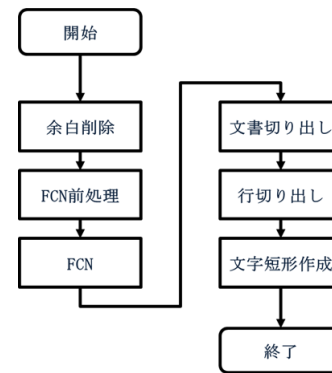


図 4 FCN を用いた手法の流れ

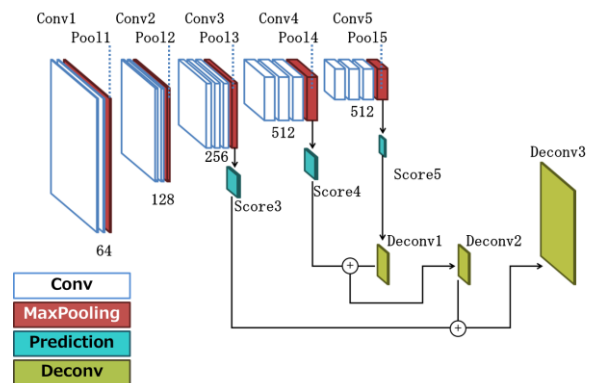


図 5 FCN8s の概要図(Conv6, Conv7 省略)

案する。提案する処理の流れは図 4 のとおりである。処理はだまかに余白削除、FCN、文字切り出し処理に分けられる。余白削除、行切り出し、文字短形作成はヒストグラムによる手法と同じ処理を適用する。

Semantic Segmentation とは画像をピクセル単位で意味づけする技術で、FCN はニューラルネットワークを用いた Semantic Segmentation の一手法である。FCN は画像を入力するとクラスごとに色分けされた画像を出力するネットワークで、CNN における完全結合層を畳み込み層に置き換えることで分類クラスではなく 2 次元マップとしての出力を可能としている。FCN の概要を図 5 に示す。本稿では図 5 中の Pool3, Pool4, Pool5 の出力する Score を合計した FCN8s を行う。また、簡略化のため Conv6, Conv7 層を省略した。FCN には Pool5 の出力のみの FCN32s, Pool4 のみの FCN16s が存在するが FCN8s が FCN を提案する文献中[7]で最も細部までセグメンテーションが可能のため、このネットワークにデータセットの画像を入力することで学習を行う。学習データセットの例を図 6 に示す。本稿では会議録画像のレイアウト解析を目的に FCN を行うため、データセットには会議録画像を使用する。データセットは図 6 左に示すような JPEG 形式のオリジナルの画像と、図 6 右に示すような PNG 形式の領域を色でクラス分けしたセグメンテーションクラス画像の 2 枚を 1 セットとする。オリジナル画像

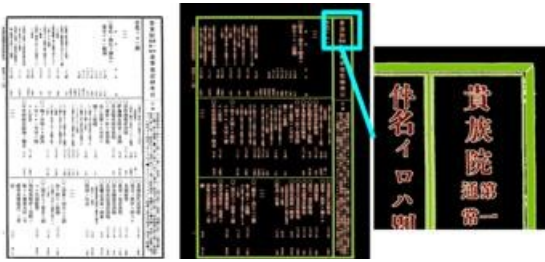


図 6 FCN の学習データセット例

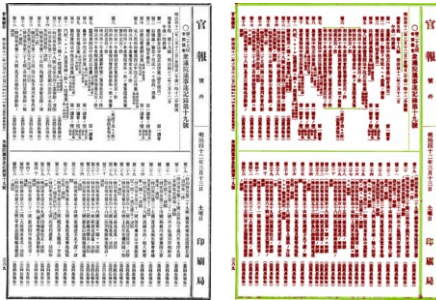


図 7 FCN 入力(左), 出力(右)例

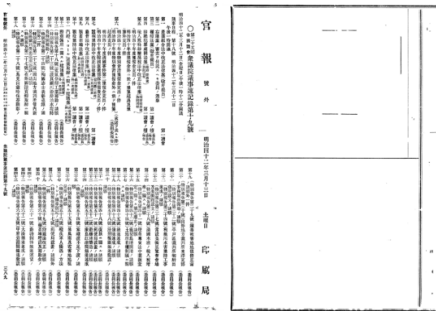


図 8 文字領域(左), 文書領域(右)画素抽出後

と、セグメンテーションクラス画像は同じ幅高の画像である。クラス分けされた画像の赤色の領域が文字領域、緑色の領域が枠線領域、その他が黒色の領域である。ネットワークにオリジナル画像とセグメンテーションクラス画像を入力することにより学習が行われる。

まず、Semantic Segmentation を行うため、FCN の学習を行う。学習データセットは幅高が 1024\*1024 の会議録のオリジナル画像と会議録のセグメンテーションクラス画像、それぞれ 186 枚である。図 7 に FCN の入出力を示す。FCN に対して図 7 左のような画像が入力された場合、図 7 右のような画像が出力される。

FCN を用いた手法の流れを説明する。FCN によるレイアウト解析の対象とする会議録画像は 4.1 節と同様の白黒 2 値の BMP 画像である。まず、会議録画像の余白削除を行った後、FCN 入力の前処理としてメディアンフィルタによるノイズ除去を行い、画像の形式を白黒 2 値の BMP から RGB カラーの JPEG へ変換する。

次に、会議録画像に対して FCN による Semantic

Segmentation を行う。ネットワークに図 7 左のような JPEG 形式の会議録画像を入力すると、図 7 右のような PNG 形式のセグメンテーションされた会議録画像が出力される。セグメンテーションされた会議録画像は図 7 に示すように領域ごとに色分けされる。図 7 右中の赤色の領域が文字領域、緑色の領域が枠線領域である。その他の領域は透過される。余白・ノイズ除去直後の白黒 2 値の BMP 形式の会議録から、図 8 に示すように文字領域、枠線領域それぞれの画素を抽出し、文字領域の画像、枠線領域の画像に分割する。枠線領域の画像の画素射影ヒストグラムを用いてレイアウト解析を行う。文字領域の画素が取り除かれているため、元の会議録画像よりも容易に枠線の位置、段数、見出しの有無が判別可能と推測される。レイアウト解析で検出された枠線の位置をもとに、文書部分ごとの外接矩形を作成し、その外接矩形に従って文字領域の画像を分割する。分割された文書部分画像に対して、4.1 節と同じ行切り出し、文字矩形作成処理を適用し、文字切り出しを行う。

## 5. 実験

### 5.1 実験方法

本稿では、帝国議会全会期からランダムに選んだ画像 11 枚を対象に文字切り出しの精度の比較を行う。4.1 節で提案したヒストグラムによるレイアウト解析の手法をヒストグラムによる手法、4.2 節で提案した Semantic Segmentation を用いたレイアウト解析の手法を FCN による手法とする。会議録画像に対して、ヒストグラムによる手法、FCN による手法を適用し文字切り出しを行った結果と、レイアウト解析を行わない場合のページ毎の正しく切り出された文字の割合、文字切り出しの失敗原因の処理毎の割合を比較する。比較した結果から、レイアウト解析の有用性を確認する。

### 5.2 結果・考察

会議録画像 11 枚に対してレイアウト解析を行わない場合、ヒストグラムによる手法、FCN による手法を用いたレイアウト解析を行う場合、それぞれに対して文字切り出しを行う。会議録画像 1 ページに対して正確に切り出された文字の割合、文字矩形作成処理の失敗、行切り出し処理の失敗、パーツ切り出し処理の失敗、FCN による文字領域のセグメンテーションの失敗の割合を表 1 に示す。レイアウト解析無しの場合、1 ページにつき平均 1.24%、ヒストグラムによる手法の場合、1 ページにつき平均 65.8%、FCN による手法の場合、1 ページにつき平均 81.7%の文字が正確に切り出された。

ヒストグラムによる手法、FCN による手法それぞれのレイアウト解析の結果を図 9 に示す。ヒストグラムによる手法の場合、4.1 節で説明した通り文字の画素が妨げとなり、枠線外など細部の文書を切り出すことができなかった。しかし、FCN による手法の場合、枠線領域のみの抽出を行う



表 1 切り出しに成功した割合と失敗した原因ごとの割合(%)

	レイアウト解析無し	ヒストグラムによる手法	FCN による手法
文字切り出し成功	1.24	65.8	81.7
文字短形作成処理失敗	0.06	10.7	11.4
行切り出し処理失敗	98.7	7.86	6.24
文書切り出し処理失敗	-	14.3	0
FCN 失敗	-	-	0.28

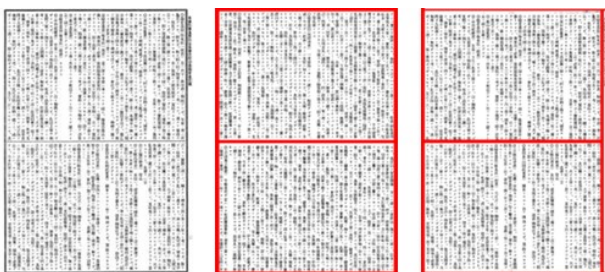


図 9 会議録画像(左)に対するヒストグラムによる手法(中央), FCN による手法(右), それぞれのレイアウト解析結果

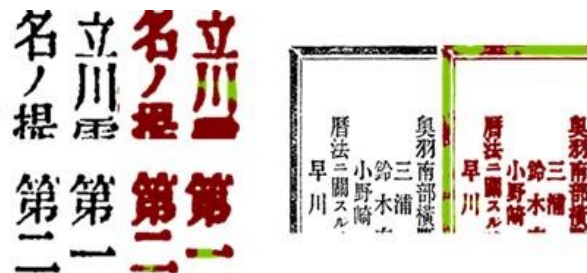


図 10 FCN の失敗

ため、細部の切り出しが容易になる。レイアウト解析において FCN による Semantic Segmentation は有用といえる。

レイアウト解析無しの場合、ほとんどの場合、行切り出し処理の時点で切り出しに失敗した。これは、会議録画像が複数の段におよぶレイアウトであることが原因であると推測される。本稿で使用した行切り出し処理では水平射影ヒストグラムが 0 になる地点を行の区切りとみなす。会議録画像のように、文書部分が複数の段に分かれているレイアウトの場合、文書部分の行間の空白部分の位置が完全に一致したときのみ水平射影ヒストグラムが 0 になる。よって全ての段で行間の空白部分の位置が一致したときのみ、

行切り出しが行われる。特に段数が多い場合、切り出しはほとんど行われない。

ヒストグラムによる手法の場合、文字短形作成処理の失敗、行切り出し処理の失敗、レイアウト解析の失敗による文書範囲の誤った切り出しが主な原因である。ヒストグラムによる手法における文字短形作成処理の失敗の原因は、4.1 節で説明したものに加え、会議録画像はインクの染みなどで上下の文字間が近くなる場合がある。また、文字間の空白の距離が必ずしも一定ではなく、レイアウトによっては行中に大きな空白を含み、文字間隔の平均値が上がり、外接短形の統合が生じやすくなる。以上の理由により、本稿で使用した文字短形作成処理では、1 つの文字として切り出すことが困難な文字が存在することが分かる。

ヒストグラムによる手法における行切り出し処理の失敗は、角度調整、行間隔が極端に狭いことが原因である。角度調整処理では、主に枠線の角度をもとに、縦方向の線がちょうど垂直になるよう画像を回転させている。そのため、枠線の角度と行の角度にずれがある場合、切り出された文書部分の文字が垂直に並んでおらず、水平射影ヒストグラムが 0 になる地点を検出できない。会議録画像は活版印刷のため、枠線、行のずれがある場合も少なくない。

ヒストグラムによる手法を用いた場合、レイアウト解析の失敗による文書範囲切り出しの失敗は、4.1 節で説明した通り、ハフ変換による直線検出で行を直線として検出し、本来枠線ではない部分を直線と認識し、外接短形作成時に文書部分を含めず文書部分を区切ったためである。

FCN を用いた手法の失敗は Semantic Segmentation の失敗、行切り出し処理の失敗、文字短形作成処理の失敗が主な原因である。Semantic Segmentation の失敗について、図 10 に示すように FCN から出力された会議録画像の領域が正しくセグメンテーションされていない場合がある。文字領域の誤認識について、例えば、「川」、「一」、「二」など、文字のパーツに直線を含む文字は、文字の一部が枠線と誤認識される場合がある。特に、対象とする文字の大きさが平均以上であるとき、この誤認識が生じやすい。FCN による枠線領域の誤認識について、画像によっては枠線中にかすれや汚れ、枠線の継ぎ目に空白部分が存在し、それらを文字領域として誤認識する場合がある。その他、インクの染み、汚れを文字や枠線と認識する場合がある。誤認識された領域はその後の処理のノイズとなる。

FCN による手法における行切り出し処理の失敗について、行切り出し処理では、事前に切り出された文書部分から余白の削除を行っている。よって、余白削除の際 FCN の文字領域の誤認識により文字の一部に欠損、かすれが生じたとき、余白ではない部分を誤ってノイズとして除去する。

FCN による手法における文字短形作成処理の失敗について、FCN を用いた手法は文字短形作成処理が原因の失敗の割合がヒストグラムによる手法に比べて高い。行切り出

し処理が原因の失敗と同様に FCN による文字領域のセグメンテーションの失敗によるものである。文字短形作成処理は文字のパーツが上下に分離しているとき、正しい外接短形による切り出しに失敗することがある。そのため、図 10 左上の「川」のような文字領域が上下に分離した文字は文字短形作成処理により分離しやすくなる。つまり、文字の一部に欠損、かすれが生じたため、本来連結されているはずの外接短形が連結されず、別の文字として分離したと推測される。図 10 左下の「一」、「二」のような左右に文字領域が分離した文字は文字短形作成処理では分離せずにかすれた文字として切り出される。

以上より、提案手法によるレイアウト解析を行い文書部分の切り出しを行うことで、文字切り出しの精度向上が認められる。また、レイアウト解析の精度はヒストグラムによる手法に比べ、FCN を用いた手法が高いといえる。現状では FCN の学習データが不足しており、正確に文字部分、枠線部分を抽出することができない。誤認識を防ぐ方法としては、学習データの増強、入力画像のノイズとかすれの除去の適用、ネットワークのフィルタサイズの変更などで Semantic Segmentation の精度向上を目指す。また、文字短形作成処理の文字パーツ間の距離が大きい文字の外接短形の統合の失敗への対策が必要である。

## 6. まとめ

本稿では、OCR の精度向上を目的にレイアウト解析の手法を 2 つ提案した。ヒストグラムによるレイアウト解析の手法と Semantic Segmentation を用いたレイアウト解析の手法を適用した場合、それぞれの文字切り出しの精度の比較を行った。

書籍の OCR の精度は文字切り出しの精度に依存し、文字切り出しの精度はレイアウト解析の精度に依存する。よって、正確な OCR には正確なレイアウト解析が必要である。レイアウトが複雑になればなるほど OCR の難易度が上がる。一般的な書籍は文書以外に、見出し、図表、ルビ、枠線などを含み、そのレイアウトは必ずしも単純であるとは限らない。一般的なレイアウトの書籍に OCR を行うには、書籍に対してレイアウト解析を行う必要がある。そこで、本稿ではヒストグラムによるレイアウト解析の手法と FCN による Semantic Segmentation を用いたレイアウト解析の手法を提案し、それぞれの有用性を示した。

提案手法の有用性の検証のため、レイアウト解析を行わずに会議録画像の文字切り出しを行う場合と、提案する 2 つの手法を適用し文字切り出しを行う場合の結果を比較した。1 つ目の提案手法はヒストグラムによる手法で、ハフ変換による直線検出と画素射影ヒストグラムを用いてレイアウト解析を行った。2 つ目の提案手法は FCN という Semantic Segmentation に特化したニューラルネットワークを使用する手法である。会議録画像を学習した FCN により

会議録画像をセグメンテーションして文字領域、枠線領域ごとに画素を抽出し、レイアウト解析を行った。

会議録画像 11 枚に対してレイアウト解析無し、ヒストグラムによる手法、FCN による手法を用いて文字切り出しを行い、それぞれの正確に切り出された文字の割合を比較した。その結果、レイアウト解析無しの場合、1 ページにつき平均 1.24%、ヒストグラムによる手法の場合、1 ページにつき平均 65.8%、FCN による手法の場合、1 ページにつき平均 81.7%の文字が正確に切り出された。レイアウト解析を行わなかった場合と比較し、提案した手法によるレイアウト解析を行った場合の文字切り出しの精度が高いため、本稿で提案した手法は有効であり、4.1 節で提案したヒストグラムによる手法と比較して 4.2 節で提案した FCN による手法が有効である。

今後の課題として、文字短形作成処理について文字パーツ間の距離が大きい文字の外接短形の統合の失敗への対策が必要である。FCN による手法について、現状では FCN の学習データが不足しており、正確に文字部分、枠線部分を抽出することができない。誤認識を防ぐ方法としては、学習データの増加、入力画像のノイズとかすれの除去の適用、ネットワークのフィルタサイズの変更などで Semantic Segmentation の精度向上を目指す。

## 参考文献

- [1] 国立国会図書館  
<http://www.ndl.go.jp/> (参照 2018-8-24)
- [2] 国立国会図書館デジタルコレクション  
<http://dl.ndl.go.jp/> (参照 2018-8-24)
- [3] 中村洋治, 除村健俊, 豊川和治, 北山友.:PC 上で動く印刷文字 OCR, 情報処理学会 第 33 回(昭和 61 年後期)全国大会, pp1635-636, (1986).
- [4] Smith, R. W. (2009, July). Hybrid page layout analysis via tab-stop detection. In 2009 10th International Conference on Document Analysis and Recognition (pp. 241-245). IEEE.
- [5] Grana, C., Serra, G., Manfredi, M., Coppi, D., & Cucchiara, R. (2016). Layout analysis and content enrichment of digitized books. *Multimedia Tools and Applications*, 75(7), 3879-3900.
- [6] 帝国議会会議録検索システム  
<http://teikokugikai-i.ndl.go.jp/> (参照 2018-8-24)
- [7] LONG, Jonathan; SHELHAMER, Evan; DARRELL, Trevor. Fully convolutional networks for semantic segmentation. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015. p. 3431-3440.
- [8] THOMA, Martin. A survey of semantic segmentation. arXiv preprint arXiv:1602.06541, 2016.