

令和3年度修士論文

クラスタリングと深層学習による  
クロスカットされた文書画像の再構成

宮崎大学大学院 工学研究科 工学専攻

機械・情報系コース 情報システム工学分野

学生番号 T1903357

富永圭佑

指導教員 椋木雅之教授

令和4年1月24日

## 概要

本研究では、クラスタリングと深層学習によるクロスカットされた文書画像の再構成の手法を提案する。文書分析の分野の問題の1つとして、細断された文書画像の再構成がある。文書画像の再構成問題の中でもクロスカットされた文書画像の再構成は、「再構成のためのフラグメントの量が膨大である」「フラグメントの情報量が少ない」などの要因により複雑で困難な問題であるとされている。従来ではクラスタリングによるアプローチが適用されている。しかし、既存のクラスタリングでは精度と自動化の部分で改良の余地がある。そこで本研究では、クラスタリングの際に生成されるクラスタのサイズを同じサイズに揃えられるように制約を加えたクラスタリング手法を適用することで、クラスタリングの改善を行う。また、深層学習による再構成の手法をこれに組み合わせて再構成の精度の比較を行う。従来の手法と改良した手法で比較実験を行い、クラスタリングの比較では精度を大幅に改善できることを確認した。一方、深層学習による再構成手法についてはクロスカットされた文書画像の再構成には不十分であることを確認した。

# 目次

1. 序論 .....	3
2. クロスカットされた文書画像の再構成 .....	5
2.1. 文書画像の再構成問題.....	5
2.2. Chen ら[4]のクラスタリングを用いた従来手法 .....	7
2.2.1 クラスタリング .....	7
2.2.2 行内の再構成.....	13
2.2.3 行間の再構成.....	13
2.3. 深層学習を用いた従来手法[10] .....	15
2.3.1 サンプル抽出.....	15
2.3.2 自己教師あり学習.....	17
2.3.3 適合性評価 .....	18
2.4. 従来の問題点.....	21
3. クラスタリングと深層学習による再構成法 .....	22
3.1. 提案手法の流れ .....	22
3.2. Same-Size-K-means 法.....	24
3.3. 深層学習による行内の再構成と前処理.....	26
4. 評価実験 .....	27
4.1. クラスタリングの精度の比較実験.....	27
4.2. 行内の再構成の精度の比較 .....	36
4.3. クロスカットされた文書画像の再構成の結果 .....	39
5. 結論 .....	42
謝辞 .....	43
参考文献.....	44

## 1. 序論

近年、文書分析の分野への研究の関心が高まっている[1]。この分野における問題の1つとして、細断された文書の再構成問題がある。この問題は、例えば法科学の分野にて証拠隠滅を目的として破壊された文書を復元するために用いるなど重要な役割を果たしている[2]。細断された文書は膨大な数の断片（フラグメント）であり、これを手動で正確に再構成するのは手間と時間のかかる作業である。そこで、細断された文書画像を自動で再構成する手法が求められている。

細断された文書画像の再構成の研究では、主に3種類の細断法[3]について研究が行われている。本論文では、縦方向と横方向から機械的に滑らかに正方形あるいは短冊形に細断されたクロスカット文書画像の再構成問題について注目する。クロスカットされた文書画像の再構成手法の1つに、Chenら[4]のクラスタリングを用いた手法がある。この手法では、膨大な数のフラグメントを、クラスタリングにより文書画像の行ごとに分類した上で各々再構成を行う。これにより、フラグメント同士の組み合わせを減らし、再構成の精度を上げるとともに、処理速度も向上させている。しかし、Chenら[4]が使用していたクラスタリング手法では一部の文書画像において特定の行にフラグメントが偏る、あるいは不足するといった誤りが発生している。そのため、クラスタリングの誤りを手作業で修正を行う必要があり、再構成を自動で行えない場合が多くある。

そこで本研究では、クロスカットされた文書画像の再構成問題に対して、Chenら[4]の手法を基に、クラスタリングの精度を向上させ、手動によらない自動の再構成手法を実現することを目指す。そのために、クラスタリング手法としてSame-Size-K-means法を適用する。Same-Size-K-means法は、クラスタリングアルゴリズムの一種であるK-means法にグループ化するデータを同

じサイズに揃えるように制約を加えた手法である。これによってクラスタリングの精度を向上させる。また、クラスタリング後の行内での再構成には、従来の文書画像再構成手法において高い精度の結果を示している深層学習の手法を適用する。これにより、再構成の精度向上を目指す。

本稿の構成は、以下の通りである。2章では文書画像再構成の従来手法について述べる。3章では、Same-Size-K-means 法と再構成の手法について詳細を述べる。4章では、提案手法の有効性を示すために行った従来手法との比較実験について述べる。5章では本研究の結論と今後の課題について述べる。

## 2. クロスカットされた文書画像の再構成

### 2.1. 文書画像の再構成問題

細断された文書には、図 1 の様に手作業で破かれた文書、縦方向にのみ機械的に滑らかに細断されたスプリットカット文書、縦方向と横方向から機械的に滑らかに正方形あるいは短冊形に細断されたクロスカット文書の主に 3 種類が存在する[3, 9]。既存の研究の多くは、手作業で細断された文書画像やスプリットカットされた文書画像の再構成を行うものである。

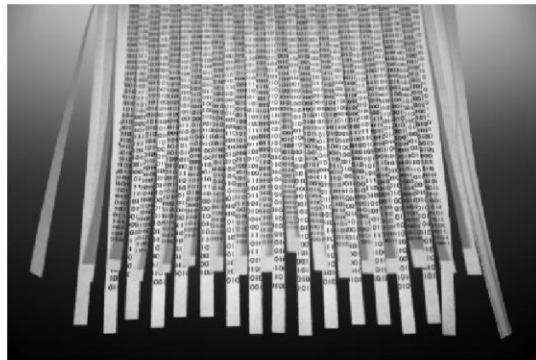
手作業で細断された境界が不規則な文書画像の再構成は、フラグメントの境界の形状の類似性に基づいて行われる[5]。しかし、シュレッダーなどで機械的に細断された文書画像のフラグメントは、境界が滑らかであるため境界が不規則な場合の再構成の方法が有効ではない。

スプリットカットされた文書の再構成では、横書きの文書を縦方向に細断したものを対象としている。そのため、フラグメントの境界部分は 1 つの文字を分割した断面を多く含む。1 つの文字内のストロークは連続しているので、2 つのフラグメントの境界の相関が高いもの同士を継ぎ合わせることで簡単に再構成できる[6]。

一方、クロスカットされた文書の再構成問題については、フラグメントの量が膨大となる、フラグメントのサイズが小さく扱える情報量が少ないなどの問題から、スプリットカットされた文書画像の再構成問題と比較して困難な問題である。この問題に関する研究は、ここ 10 年程のものがほとんどであり、蟻コロニー最適化[7]や、遺伝的アルゴリズム[8]などの様々な技術を利用した研究が行われている。近年ではクラスタリングを用いたアプローチが、クロスカットされた文書画像の再構成問題を解く鍵として注目されている。



(a)手作業



(b)スプリットカット



(c)クロスカット

図1 文書の細断法(引用[9])

## 2.2. Chen ら[4]のクラスタリングを用いた従来手法

クラスタリングによるクロスカットされた文書画像の再構成手法は、クラスタリングによって単一のクロスカットされた文書画像再構成問題を複数のスプリットカットされた文書画像の再構成問題に分割する手法である。上位問題をいくつかの下位問題にして、これらの下位問題を独立に解くという取り組みとなっている。処理は、

- ① クラスタリング
- ② 行内の再構成
- ③ 行間の再構成

の3段階からなる。手法の詳細を以下に説明する。

### 2.2.1 クラスタリング

#### 特徴量の抽出

クラスタリングにおいて、フラグメントからの特徴量の抽出は重要なステップである。文書画像のフラグメントは、白黒のものがほとんどであり特徴の差異が曖昧になりやすいからである。よって、小さくて特徴が少ないフラグメントからどれだけ特徴を抽出できるかが重要視される。

従来[9]では、図2の様にフラグメント内の行の位置情報をもとに、特徴抽出されたクラスタリングベクトルが特徴量としてクラスタリングに使用されていた。クラスタリングベクトルは、 $CV = (a_1, a_2, a_3, a_4)$ と定義される。 $a_1$ は、フラグメントの上部にある途切れた不完全な行の下辺の位置を表す（行間だった場合は0）。 $a_4$ は、フラグメントの下部にある途切れた不完全な行の上辺の位置を表す（行間だった場合は0）。 $a_2$ と $a_3$ は、フラグメントで確認できる完全な



行の上辺と下辺の位置をそれぞれ表す。しかし、まったく別のフラグメントから同じクラスタリングベクトルが抽出され、フラグメントの識別ができないという問題が発生していた。

Chen ら[4]は、クラスタリングベクトルとして一次元信号を採用した。これは、図 3 の様にフラグメントの各行の黒画素数を並べた一次元ベクトルである。ベクトルの次元数はフラグメントの縦方向の画素数と等しい。クラスタリングベクトルと比べて、フラグメントの特徴をより多く利用できるようにするためクラスタリングにおける合理的な代替である。



図 2 クラスタリングベクトルの抽出(引用[9])

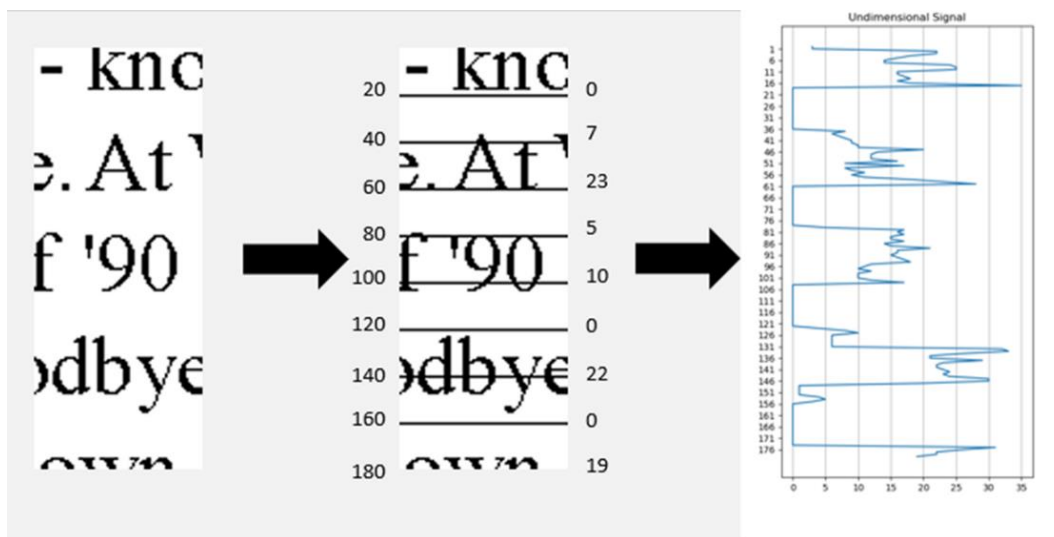


図3 一次元信号の抽出

## クラスタ中心

各行の最初のフラグメントから抽出した特徴量をクラスタ中心として定義する。クラスタ中心とは、クラスタリングを行う時に指定される重心のことである。各行ごとの最初のフラグメントは、いくつかの注目すべき特徴を持っている。例えば、図4の様にフラグメントの画像の左側は白色である。このため、各行の最初のフラグメントを発見し、その特徴量を設定することは容易である。



図4 クラスタ中心に指定できるフラグメントの例

## K-means 法

K-means 法とは、クラスタの平均を用いて任意の指定した  $K$  個のクラスタを作成する非階層型クラスタリングアルゴリズムである。

K-means 法の手順を以下に示す。

1. 各点を適当にクラスタに割り振る
2. それぞれのクラスタの重心と各点の距離を計算
3. 各点を一番近いクラスタに割り当てる
4. 2 と 3 を繰り返し、割り当てるクラスタの変更がなくなったら終了

さらに、K-means 法の手順を図で表現する。図 5 のように(a)→(b)→(c)→(d)という順序を辿ってクラスタが収束していく。(b)の段階で、まず各点が適当にクラスタへ割り振られ、その重心が計算される(重心は赤星)。(c)では、その重心との距離をもとに、再度クラスタへ割り当てられる。(新しい重心を赤星、古い重心を黄星)。この工程を繰り返して、(d)のように割り当てられるクラスタが変化しないように収束すれば、クラスタリングの完了となる。

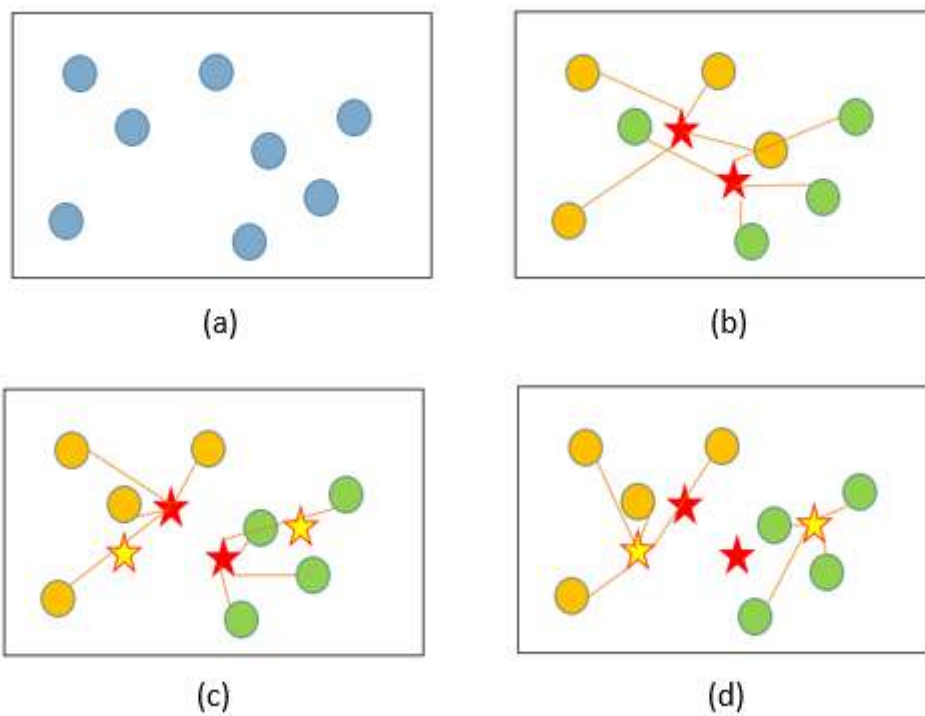


図5 K-means 法の手順

ランダムなデータを初期値としてクラスタ中心を指定するのが本来の K-means 法であるが、各行の最初のフラグメントをクラスタ中心とすることで、フラグメントを行ごとにクラスタリングする精度が向上することが確認されている[12]。よって、文献[4]では各行の最初のフラグメントをクラスタ中心とする制約付き K-means 法を使用している。

## 2.2.2 行内の再構成

クラスタリングによって行ごとに分類されたフラグメントを、それぞれ独立に再構成する。これは、複数のスプリットカットされた文書画像再構成問題と同じであると思なすことができる。ストリップカットされた文書画像再構成問題は、ほとんどの場合、巡回セールスマン問題 (TSP) として扱うことができる。具体的には、フラグメントを TSP の頂点、フラグメント間のコスト関数を辺として見なすことができる。よって、TSP を解くのに有効なアプローチが、スプリットカットされた文書画像再構成問題にも有効である。従来の手法では TSP に有効な蟻コロニー最適化を解法として使用している。しかし、クロスカットされた文書のフラグメントでは境界の情報量が限られているため理想的なコスト関数を定義することはできない。よって、Chen ら[4]は、再構成の精度を高めるためにフラグメントの数を減らす戦略を組み合わせる再構成を行っている。

## 2.2.3 行間の再構成

行内の再構成によって生成されたフラグメントを、行フラグメントと呼称する。行内の再構成が完了すると、行フラグメント同士を再構成する行間の再構成を行う。これによって、クロスカットされた文書画像全体の再構成が完了する。ここでは、図 2 のクラスタリングベクトルを利用して再構成を行う。

再構成の手順を以下に示す。

1. 各行フラグメントより、クラスタリングベクトル ( $CVC_i = [a_1, a_2, a_3, a_4]$ ) を抽出する。
2. 行フラグメントの上部が白である特徴から最初の行フラグメントを選出し、クラスタリングベクトル ( $CVC_F$ ) を決定する。現在のクラスタリングベクトル  $CVC_C = CVC_F$  とする。

3. クラスタリングベクトル  $CVC_i$  が

$$l - \delta < a_4^c + a_1^i < l + \delta \quad \text{あるいは}$$

$$nl + l' - \delta < a_3^c + a_2^i < nl + l' + \delta$$

を満たす場合、 $CVC_C = CVC_i$  として現在の行フラグメントと行フラグメント  $i$  を結合する。(  $l$  は単語行の高さ、 $l'$  は行間の高さ、 $n$  はフラグメント内の単語の行数以下の整数、 $\delta (=3)$  は誤差範囲)

4. 3 を再構成が完了するまで繰り返す。

## 2.3. 深層学習を用いた従来手法[10]

本稿でも扱う深層学習を用いた従来手法[10]について解説する。これは、スプリットカットされた文書画像の再構成の手法として扱われていた従来のアプローチの1つであり、90%を越える高い精度の結果で再構成を行っていた手法である。

手法の詳細を以下に説明する。

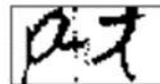
### 2.3.1 サンプル抽出

最初に、IIT-CDIP テストコレクション 1.0 の 300 dpi でスキャンされた 100 枚の文書画像（フォーム、電子メール、メモなど） [11] を縦方向に 30 に分割をする（通常の A4 シュレッダーで推定される量）。この文書画像より抽出された 32×32 サイズのサンプルによって、モデルの学習データを構成する（図 6）。

サンプルペアは、ページごとに抽出される。つまり、ペアのサンプルは同じ文書画像から得られるものである。抽出処理は、まずポジティブなサンプルペア（文書画像あたり 1,000 ペアに制限される）を収集する。その後、ネガティブなペアが収集されるが、これはポジティブなペアの数と同じ数に制限される。サンプルの抽出の際、フラグメントは上から下にスキャンされ、2 ピクセルごとにサンプルがトリミングされる。この時、空白のピクセルが 80% を超えるペアは曖昧であると見なされ破棄される。

この学習データを使って次節の学習を行い、学習モデル  $f_{left}$  と  $f_{right}$  を生成する。





(a) ネガティブなペア



(b) ポジティブなペア

図 6 サンプル抽出に使用した文書画像とサンプルの例

### 2.3.2 自己教師あり学習

自己教師あり学習とは、データ自身から独自のラベルを機械的に作り、人間によるラベル付けを必要としない学習である。サンプルペアは生成の段階において、自動的に「ポジティブ」と「ネガティブ」としてラベル付けされている。

学習データはタプル  $(X_r, X_l, y)$  で構成される。ここで  $X_r$  と  $X_l$  は、それぞれサンプル抽出において、左のフラグメントの右端のピクセルから抽出された  $r$  サンプルと、右のフラグメントの左端のピクセルから抽出された  $l$  サンプルを示し、 $y$  はグラウンドトゥールスラベルである。サンプルペアがポジティブならば  $y=1$ 、それ以外ならば  $y=0$  である。学習は Contrastive loss 関数[12] (式 (1)) で行われる。

$$L(f_{left}, f_{right}, X_l, X_r, y) = \frac{1}{2} \{ (1 - y) dist^2 + y [\max(0, m - dist)]^2 \} \quad (1)$$

この時、 $dist = \|f_{left}(X_l) - f_{right}(X_r)\|_2$  である。

図 7 に適合性評価のためのモデルの例を示す。モデルは、単語「TEXT」を構成するポジティブなサンプルペアを処理する。ポジティブ ( $y=1$ ) であるため、loss 値が低くなり、そうでないならば高くなる。

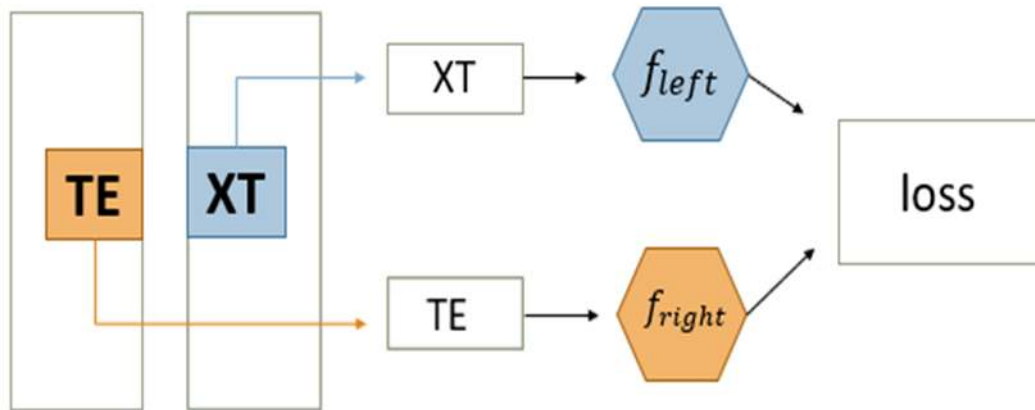


図 7 フラグメントの適合性評価のための学習モデル

### 2.3.3 適合性評価

フラグメントの適合性評価のためのアプローチとしては、まず初めに、学習したモデルを変換し別空間に埋め込む。この空間を、埋め込み空間と呼ぶ。この時、自己教師あり学習によって構築されたモデルが正しく学習されていると仮定するならば、図 8 の様にポジティブなペアは近くなるように埋め込まれ、ネガティブなペアは遠く離れるように埋め込まれる。そして、変換されて埋め込まれたモデルの埋め込み空間における距離とサンプルの正負を学習することで、フラグメントの適合性評価を行うことができる。

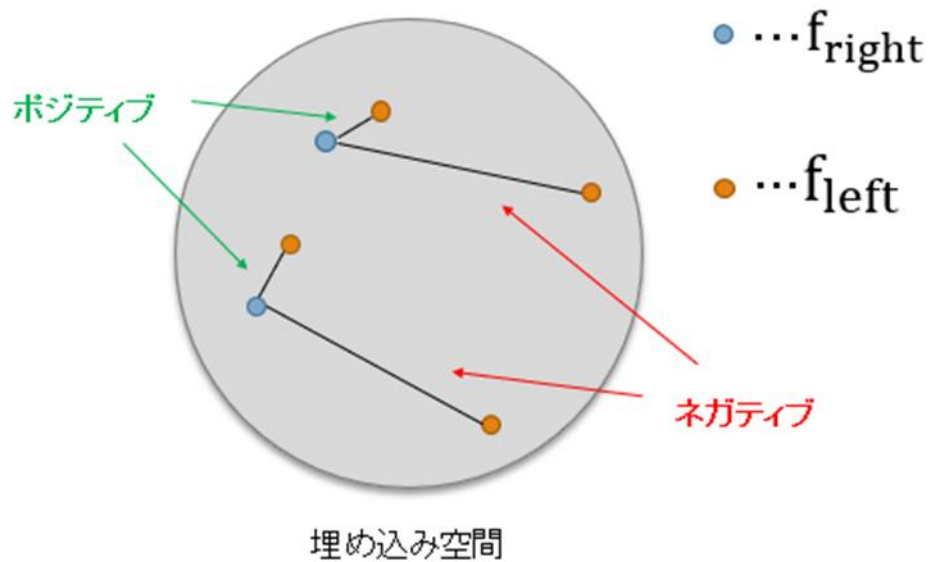


図 8 フラグメントの適合性評価のためのアプローチ

適合性評価では、細断されたフラグメントの境界から、 $h \times s$  領域を切り取る ( $s$  はサンプルのサイズ)。  $X_r$  は左のフラグメントの右端の  $s$  ピクセル、  $X_l$  は右のフラグメントの左端の  $s$  ピクセルを切り出す ( $s$  はサンプルサイズ)。そして推論モデルより、埋め込み空間に埋め込むことで  $L = f_{left}(X_l)$  (l-embeddings) と、  $R = f_{right}(X_r)$  (r-embeddings) という特徴量が生成される。これにより、  $\|R - L\|_2$  から適合性の評価を求める。

適合性評価に関して、フラグメントの埋め込みと距離計算の 2 つのステップを図 9 に示す。

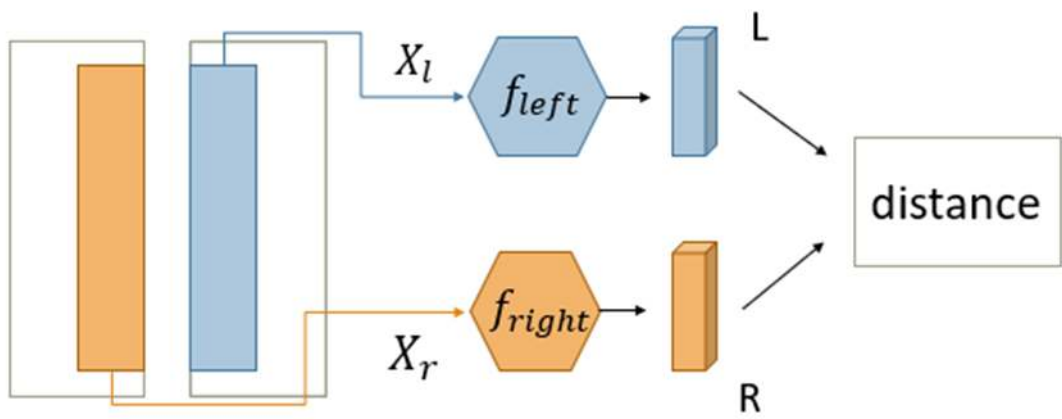
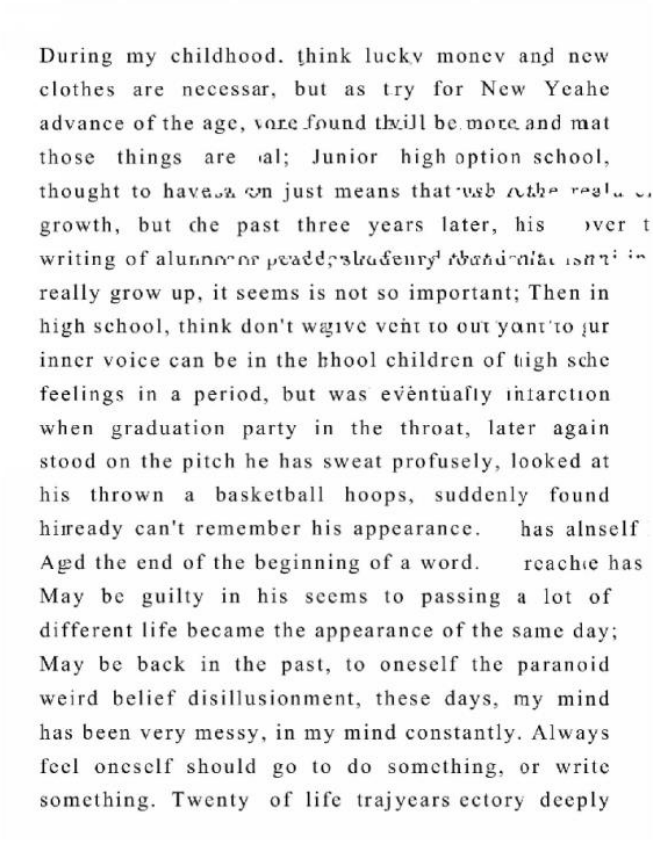


図9 フラグメントの適合性評価

## 2.4. 従来の問題点

クロスカットされた文書画像の再構成において、従来手法の問題点をあげる。クラスタリングによってフラグメントのクラス（行）分けの際に、発生するエラーの多くが特定の行にフラグメントの偏りと不足が生じてしまうエラーだった。そのまま再構成を進めると、図 10 の様に歪んだ形に再構成されてしまう。

また、深層学習による行内の再構成では、スプリットカットされた文書画像のフラグメントと違いクロスカットされた文書画像のフラグメントでは十分な情報量が足りずに精度が低下してしまう恐れがある。



During my childhood, think lucky money and new clothes are necessary, but as try for New Year's advance of the age, were found they will be more and more those things are not; Junior high school, thought to have a fun just means that with the regular growth, but the past three years later, his over the writing of a letter or reading a book is not really important; Then in high school, think don't give vent to out want to get inner voice can be in the school children of high school feelings in a period, but was eventually interaction when graduation party in the throat, later again stood on the pitch he has sweat profusely, looked at his thrown a basketball hoops, suddenly found he already can't remember his appearance. has himself And the end of the beginning of a word. reaches May be guilty in his seems to passing a lot of different life became the appearance of the same day; May be back in the past, to oneself the paranoid weird belief disillusionment, these days, my mind has been very messy, in my mind constantly. Always feel oneself should go to do something, or write something. Twenty of life trajectory deeply

図 10 クラスタリングでエラーがある再構成の例(引用[9])

## 3. クラスタリングと深層学習による再構成法

### 3.1. 提案手法の流れ

本稿でも、クロスカットされた文書画像の再構成問題を従来の手法[4]にならって、クラスタリング・行内の再構成・行間の再構成の3段階で解く。クラスタリングでは、従来のK-means法に代わってSame-size-K-means法を適用することで、クラスタリングの精度を向上させることを試みる。行内の再構成では、従来における蟻コロニー最適化に代わって、スプリットカットされた文書画像の再構成に使われた深層学習の手法[10]を、クロスカットされた文書画像の再構成に合わせた前処理と組み合わせて用いる。行間の再構成は、従来の手法をそのまま用いる。

提案する手法の流れを図11に示す。

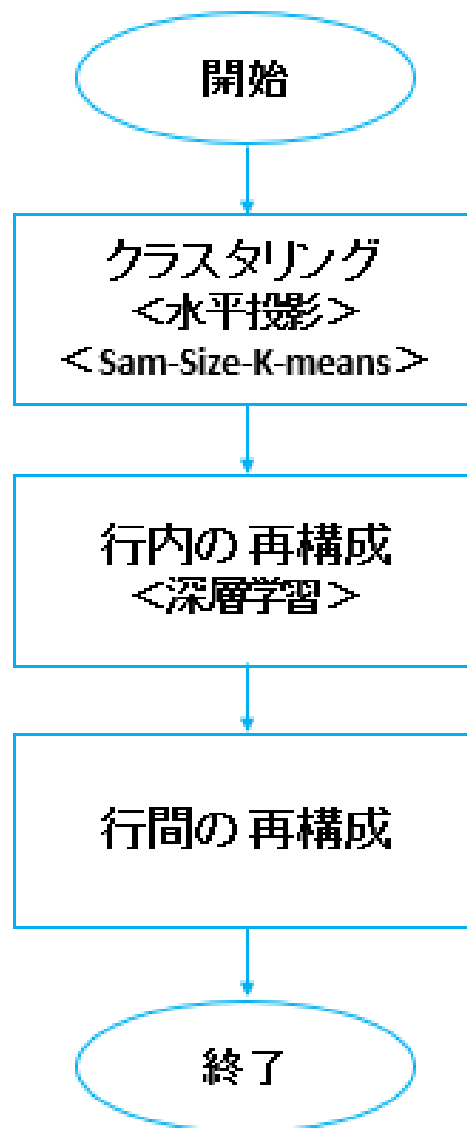


図 11 提案手法のフローチャート



## 3.2. Same-Size-K-means 法

Same-Size-K-means 法[13]は、クラスタに含まれるデータ数（クラスタサイズ）を同じ大きさに揃えるようにクラスタリングを行う K-means 法のバリエーションの一種である。

クラスタリングの手順を以下に示す。

初期化：

1. 望ましいクラスタサイズを計算する（データの総数/クラスタの数）
2. K-means 法などでクラスタリングを行いクラスタ中心の初期値を設定する
3. 最も近いクラスタ中心までの距離から、最も遠いクラスタ中心までの距離を引いた値でフラグメントを並べ替える
4. 並び替えた順序で、フラグメントを望ましいクラスタサイズ未満のもっと近いクラスタへ割り当てる。望ましいクラスタサイズになったクラスタにはそれ以上フラグメントを割り当てない。

移動候補リスト：クラスタごとに、クラスタから抜けてもよいフラグメントのリストを保持する。これにより、クラスタを移動する候補となるフラグメントを見つけやすくなる。あるクラスタに移動したいフラグメントがある場合、そのフラグメントの移動候補リスト内のフラグメントが入れ換え候補となる。

反復：

1. 現在のクラスタの平均を計算し、新たなクラスタ中心を求める
2. フラグメントごとに、クラスタ中心までの距離を計算する
3. 現在のクラスタ $C_c$ のクラスタ中心までの距離と現在のクラスタ以外で最も近いクラスタ $C_N$ のクラスタ中心までの距離の差の値でフラグメントを並べ替える
4. 並び替えた順序で各フラグメント $F_c$ について以下を行う：
  - 4.1 クラスタ $C_N$ の移動候補リスト内のフラグメント $F_N$ について、 $F_c$ とクラスタの割り当てを交換することで、クラスタ中心までの距離の和が小さくなる場合、2つのフラグメントを交換する
  - 4.2 クラスタサイズの制約を違反せずにフラグメントを移動できる場合、そのまま移動させる。
  - 4.3 フラグメント $F_c$ のクラスタが変更されていない場合、 $C_N$ の移動候補リストに $F_c$ を追加する
5. これ以上変更が行われなかった場合、クラスタリングを終了する

### 3.3. 深層学習による行内の再構成と前処理

従来の研究においても、フラグメントの数が多ければ多いほど再構成の精度は低下してしまう傾向にある[14]。よって、深層学習による再構成の前段階処理としてフラグメントの数を減らしておく手法を行うことで行内の再構成の精度の向上を試みる。

具体的には、フラグメント*i*とフラグメント*j*の距離を求める(式(3))。求められた距離が指定した閾値よりも小さい場合、2つのフラグメントは隣接関係にあると考える。この時、2つのフラグメントを接続して新しい1つのフラグメントとする(図12)。新しいフラグメントでは、フラグメント*i*の左端境界とフラグメント*j*の右端境界を新しいフラグメント自身の境界とする。このようにして、フラグメントの数を減らすことで行内の再構成の精度を向上させる。

$$d(i, j) = \|x_i - x_j\|_2 = \sqrt{\sum_{u=1}^l |x_{iu} - x_{ju}|^2} \quad (3)$$

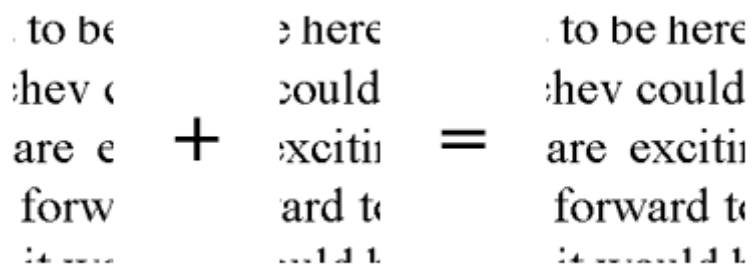


図12 フラグメントの接続

## 4. 評価実験

### 4.1. クラスタリングの精度の比較実験

この実験では、Same-Size-K-means 法（提案手法）が文書画像のクラスタリングにおいて有効であるかどうか、2.2.1.節の制約付き K-means 法（従来手法）との比較実験を行う。

実験に使用するテストデータは、解像度 1980×1224 の英文の 7 枚の文書画像 T1~T7[4]である。これを 11×17 個のフラグメントになるよう縦横に細断した。フラグメントのサイズは 180×72 画素である。文書画像の詳細を表 1 に示す。例として T1 と T6 の文書画像を図 13 と図 14 に示す。また、細断された文書のフラグメントの例を図 15 に示す。

表 1 文書画像の詳細

	行間	サイズ	フォント
T1	13	10	Times New Roman
T2	13	10	Times New Roman
T3	13	10	Times New Roman
T4	13	10	Bold Times New Roman
T5	13	10	Bold Times New Roman
T6	13	10	Bold Times New Roman
T7	13	10	Bold Times New Roman

Thank you very, very much, President Keohane. Mrs. Gorbachev, Trustees, faculty, parents, and I should say, Julia Porter, class president, and certainly my new best friend, Christine Bicknell - and, of course, the Class of 1990. I am really thrilled to be here today, and very excited, as I know all of you must be, that Mrs. Gorbachev could join us.

These are exciting times. They're exciting in Washington, and I have really looked forward to coming to Wellesley. I thought it was going to be fun. I never dreamt it would be this much fun. So, thank you for that.

More than ten years ago, when I was invited here to talk about our experiences in the People's Republic of China, I was struck by both the natural beauty of your campus and the spirit of this place.

Wellesley, you see, is not just a place but an idea - an experiment in excellence in which diversity is not just tolerated, but is embraced. The essence of this spirit was captured in a moving speech about tolerance given last year by a student body president of one of your sister colleges. She related the story by Robert Fulghum about a young pastor, finding himself in charge of some very energetic children, hits upon the game called "Giants, Wizards, and Dwarfs." "You have to decide now," the pastor instructed the children, "which you are - a giant, a wizard, or a dwarf?" At that, a small girl tugging at his pants leg, asked, "But where do the mermaids stand?" And the pastor tells her there are no mermaids. And she says, "Oh yes there are - they are. I am a mermaid."

Now this little girl knew what she was, and she was not about to give up on either her identity, or the game. She intended to take her place wherever mermaids fit into the scheme of things. "Where do the mermaids stand? All of those who are different, those who do not fit the boxes and the pigeonholes?" "Answer that question," wrote Fulghum, "And you can build a school, a nation, or a whole world." As that very wise young woman said, "Diversity, like anything worth having, requires effort - effort to learn about and respect difference, to be compassionate with one another, to cherish our own identity, and to accept unconditionally the same in others.

You should all be very proud that this is the Wellesley spirit. Now I know your first choice today was Alice Walker - guess how I know! - known for *The Color Purple*. Instead you got me - known for the color of my hair. Alice Walker's book has a special resonance here. At Wellesley, each class is known by a special color. For four years the Class of '90 has worn the color purple. Today you meet on Severance Green to say goodbye to all of that, to begin a new and a very personal journey, to search for your own true colors.

In the world that awaits you, beyond the shores of Waban - Lake Waban, no one can say what your true colors will be. But this I do know: You have a first class education from a first class school. And so you need not, probably cannot, live a "paint-by-numbers" life. Decisions are not irrevocable. Choices do come back. And as you set off from Wellesley, I hope that many of you will consider making three very special choices.

The first is to believe in something larger than yourself, to get involved in some

for moment to upset the peace of the world. It is this happy fact, now clear to the view of every public man whose thoughts do not still linger in an age that is dead and gone, which makes it possible for every nation whose purposes are consistent with justice and the peace of the world to avow nor or at any other time the objects it has in view.

We entered this war because violations of right had occurred which touched us to the quick and made the life of our own people impossible unless they were corrected and the world secure once for all against their recurrence. What we demand in this war, therefore, is nothing peculiar to ourselves. It is that the world be made fit and safe to live in; and particularly that it be made safe for every peace-loving nation which, like our own, wishes to live its own life, determine its own institutions, be assured of justice and fair dealing by the other peoples of the world as against force and selfish aggression. All the peoples of the world are in effect partners in this interest, and for our own part we see very clearly that unless justice be done to others it will not be done to us. The program of the world's peace, therefore, is our program; and that program, the only possible program, as we see it, is this:

- I. Open covenants of peace, openly arrived at, after which there shall be no private international understandings of any kind but diplomacy shall proceed always frankly and in the public view.

- II. Absolute freedom of navigation upon the seas, outside territorial waters, alike in peace and in war, except as the seas may be closed in whole or in part by international action for the enforcement of international covenants.

- III. The removal, so far as possible, of all economic barriers and the establishment of an equality of trade conditions among all the nations consenting to the peace and associating themselves for its maintenance.

- IV. Adequate guarantees given and taken that national armaments will be reduced to the lowest point consistent with domestic safety.

- V. A free, open-minded, and absolutely impartial adjustment of all colonial claims, based upon a strict observance of the principle that in determining all such questions of sovereignty the interests of the populations concerned must have equal weight with the equitable claims of the government whose title is to be determined.

- VI. The evacuation of all Russian territory and such a settlement of all

図 14 文書画像 T6



図 15 クロスカットされた文書画像のフラグメントの例

文書画像 T1~T7 に従来手法と Same-Size-K-means 法を適用してクラスタリングを行った。目視で正しいクラスタリング結果を対応付けて評価を行った。

文書画像 T2・T3・T4 においては、いずれの手法でもクラスタリングの精度 100%の結果となった。

文書画像 T1・T5・T6 については、従来手法では誤りがあったが、Same-Size-K-means 法ではすべて正しくクラスタリングが行えた。

文書画像 T1 のクラスタリングの結果を図 16 に示す。正しくは Row5 に割り当てられるはずのフラグメント 86 が従来手法では Row8 に割り当てられていた（赤字の番号はクラスタリングが誤っているフラグメントの番号を示す）。提案手法によって誤りが修正されたことが分かる。

文書画像 T6 のクラスタリングの結果を図 17 に示す。これと文書画像 T5、T7 は、従来手法で特にエラーが多い結果となった文書画像である。図より、にフラグメントが偏って割り当てられていることが分かる。また、Row7 にフラグメントが偏って割り当てられた分、Row1 のフラグメントが不足している

ことが分かる。単純に Row1 のフラグメントが Row7 に割り当てられているだけでなく、Row7 のフラグメントが Row1 にも誤って割り当てられていることも分かる。これは、Row1 と Row7 のクラスタ中心として指定したフラグメントの特徴が似ていたからだと考えられる。

文書画像 T7 については、提案手法によって従来手法から精度の改善には成功しているが、完全に正しくクラスタリングを行うことはできなかった。文書画像 T7 のクラスタリングの結果を図 18 に示す。クラスタ内のフラグメント数の偏りや不足は発生していないが、3 つのフラグメントでズレが発生している。本来は Row11 のフラグメント 87 が Row4 に、Row4 のフラグメント 53 が Row10 に、Row10 のフラグメント 100 が Row10 に割り当てられてしまっている。このような誤りの修正が、提案手法のクラスタリングにおける今後の課題であると考えられる。

最終的に全体のクラスタリングの比較実験の結果を表 2 に示す。7 枚の文書画像中 4 枚の文書画像において、Same-Size-K-means 法によりクラスタリングの精度が向上した。7 枚の文書画像の内、3 枚の文書画像は従来手法でも 100%であり、Same-Size-K-means 法でも精度を落とさずにクラスタリングを行えている。以上より、クロスカットされた文書画像において Same-Size-K-means 法によるクラスタリングが有効であることが示された。



Row1 (17) [0, 10, 47, 52, 59, 73, 77, 79, 93, 96, 132, 148, 153, 155, 166, 170, 180]  
 Row2 (17) [19, 21, 26, 36, 62, 80, 112, 117, 136, 145, 149, 152, 168, 171, 173, 182, 184]  
 Row3 (17) [2, 13, 29, 31, 44, 48, 58, 60, 63, 71, 92, 95, 105, 115, 144, 160, 179]  
 Row4 (17) [5, 6, 38, 51, 54, 55, 65, 91, 97, 109, 123, 134, 137, 140, 157, 175, 186]  
 Row5 (16) [1, 16, 22, 72, 76, 82, 85, 89, 104, 111, 114, 118, 119, 122, 125, 147]  
 Row6 (17) [12, 28, 61, 70, 75, 94, 100, 103, 107, 113, 135, 141, 143, 151, 163, 181, 185]  
 Row7 (17) [8, 18, 32, 34, 41, 42, 43, 49, 50, 66, 88, 124, 128, 146, 150, 161, 178]  
 Row8 (18) [11, 17, 20, 24, 40, 67, 68, 81, 84, **86**, 87, 102, 106, 108, 129, 133, 142, 183]  
 Row9 (17) [7, 15, 30, 33, 39, 45, 74, 83, 90, 98, 121, 126, 138, 139, 154, 164, 165]  
 Row10(17) [9, 14, 27, 35, 37, 57, 78, 101, 127, 131, 156, 159, 162, 167, 172, 174, 177]  
 Row11(17) [3, 4, 23, 25, 46, 53, 56, 64, 69, 99, 110, 116, 120, 130, 158, 169, 176]

(a)制約付き K-means 法

Row1 (17) [0, 10, 47, 52, 59, 73, 77, 79, 93, 96, 132, 148, 153, 155, 166, 170, 180]  
 Row2 (17) [19, 21, 26, 36, 62, 80, 112, 117, 136, 145, 149, 152, 168, 171, 173, 182, 184]  
 Row3 (17) [2, 13, 29, 31, 44, 48, 58, 60, 63, 71, 92, 95, 105, 115, 144, 160, 179]  
 Row4 (17) [5, 6, 38, 51, 54, 55, 65, 91, 97, 109, 123, 134, 137, 140, 157, 175, 186]  
 Row5 (17) [1, 16, 22, 72, 76, 82, 85, **86**, 89, 104, 111, 114, 118, 119, 122, 125, 147]  
 Row6 (17) [12, 28, 61, 70, 75, 94, 100, 103, 107, 113, 135, 141, 143, 151, 163, 181, 185]  
 Row7 (17) [8, 18, 32, 34, 41, 42, 43, 49, 50, 66, 88, 124, 128, 146, 150, 161, 178]  
 Row8 (17) [11, 17, 20, 24, 40, 67, 68, 81, 84, 87, 102, 106, 108, 129, 133, 142, 183]  
 Row9 (17) [7, 15, 30, 33, 39, 45, 74, 83, 90, 98, 121, 126, 138, 139, 154, 164, 165]  
 Row10(17) [9, 14, 27, 35, 37, 57, 78, 101, 127, 131, 156, 159, 162, 167, 172, 174, 177]  
 Row11(17) [3, 4, 23, 25, 46, 53, 56, 64, 69, 99, 110, 116, 120, 130, 158, 169, 176]

(b) Same-Size-K-means 法

図 16 T1 のクラスタリング結果

Row1 (10) [2, 33, 42, 61, **74**, 79, **130**, 142, 148, 178]  
 Row2 (17) [0, 3, 13, 18, 40, 55, 56, 85, 89, 91, 92, 102, 120, 122, 128, 132, 182]  
 Row3 (17) [11, 15, 22, 30, 36, 39, 60, 72, 76, 84, 93, 99, 106, 108, 118, 165, 179]  
 Row4 (17) [4, 16, 19, 44, 45, 50, 51, 65, 66, 96, 103, 140, 146, 152, 159, 166, 171]  
 Row5 (17) [28, 29, 32, 46, 52, 54, 62, 64, 70, 98, 112, 119, 126, 141, 147, 157, 168]  
 Row6 (18) [7, 9, 14, 23, 24, 35, 47, 48, 49, 67, **69**, 75, 111, 116, 123, 145, 156, 183]  
 Row7 (24) [**10**, **17**, **27**, 41, **43**, 53, 77, **82**, 83, **88**, 104, **107**, 121, **127**, 131, 133, 134, 135, 143, 150, 153, 162, 164, **170**]  
 Row8 (17) [6, 8, 21, 31, 63, 80, 86, 87, 94, 105, 110, 138, 139, 144, 154, 169, 184]  
 Row9 (17) [34, 78, 90, 97, 100, 109, 113, 114, 124, 155, 158, 160, 163, 175, 176, 180, 181]  
 Row10(17) [12, 25, 26, 38, 57, 59, 71, 81, 95, 101, 115, 125, 129, 151, 161, 172, 186]  
 Row11(16) [1, 5, 20, 37, 58, 68, 73, 117, 136, 137, 149, 167, 173, 174, 177, 185]

(a)制約付き K-means 法

Row1 (17) [2, 10, 17, 27, 33, 42, 43, 61, 79, 82, 88, 107, 127, 142, 148, 170, 178]  
 Row2 (17) [0, 3, 13, 18, 40, 55, 56, 85, 89, 91, 92, 102, 120, 122, 128, 132, 182]  
 Row3 (17) [11, 15, 22, 30, 36, 39, 60, 72, 76, 84, 93, 99, 106, 108, 118, 165, 179]  
 Row4 (17) [4, 16, 19, 44, 45, 50, 51, 65, 66, 96, 103, 140, 146, 152, 159, 166, 171]  
 Row5 (17) [28, 29, 32, 46, 52, 54, 62, 64, 70, 98, 112, 119, 126, 141, 147, 157, 168]  
 Row6 (17) [7, 9, 14, 23, 24, 35, 47, 48, 49, 67, 75, 111, 116, 123, 145, 156, 183]  
 Row7 (17) [41, 53, 74, 77, 83, 104, 121, 130, 131, 133, 134, 135, 143, 150, 153, 162, 164]  
 Row8 (17) [6, 8, 21, 31, 63, 80, 86, 87, 94, 105, 110, 138, 139, 144, 154, 169, 184]  
 Row9 (17) [34, 78, 90, 97, 100, 109, 113, 114, 124, 155, 158, 160, 163, 175, 176, 180, 181]  
 Row10(17) [12, 25, 26, 38, 57, 59, 71, 81, 95, 101, 115, 125, 129, 151, 161, 172, 186]  
 Row11(17) [1, 5, 20, 37, 58, 68, 69, 73, 117, 136, 137, 149, 167, 173, 174, 177, 185]

(b) Same-Size-K-means 法

図 17 T6 のクラスタリング結果

Row1 (17) [8, 14, 15, 20, 27, 33, 39, 72, 81, 89, 125, 127, 129, 162, 181, 185, 186]  
 Row2 (17) [18, 22, 28, 30, 42, 70, 80, 103, 104, 112, 121, 130, 144, 146, 148, 149, 150]  
 Row3 (06) [**1, 38, 117, 131, 135, 184**]  
 Row4 (17) [2, 10, 31, 32, 43, 52, 53, 59, 62, 74, 107, 133, 145, 151, 159, 160, 178]  
 Row5 (17) [13, 17, 36, 40, 54, 56, 58, 64, 78, 84, 97, 122, 123, 158, 168, 169, 182]  
 Row6 (12) [0, 21, **23**, 61, 67, 82, 110, 115, 134, 137, 154, 156]  
 Row7 (17) [3, 11, 34, 44, 46, 68, 75, 83, 85, 106, 108, 126, 141, 161, 175, 176, 180]  
 Row8 (16) [16, 37, 49, 51, 60, 65, 69, 88, 93, 95, 102, 124, 136, 140, 177, 179]  
 Row9 (34) [**6, 7, 9, 19, 26**, 29, **50**, 55, 66, 71, **76**, 77, **79**, 86, **91**, 92, 94, **99**, 101, 105, **119**, 132, 139,  
 142, 143, **153**, 155, 157, 163, **164, 166**, 172, **173, 174**]  
 Row10(16) [4, 5, 25, 41, 45, 48, 57, 98, 109, 114, 120, 138, 147, 165, 170, 183]  
 Row11(18) [12, 24, 35, 47, 63, 73, 87, 90, 96, **100**, 111, 113, 116, 118, 128, 152, 167, 171]

(a)制約付き K-means 法

Row1 (17) [8, 14, 15, 20, 27, 33, 39, 72, 81, 89, 125, 127, 129, 162, 181, 185, 186]  
 Row2 (17) [18, 22, 28, 30, 42, 70, 80, 103, 104, 112, 121, 130, 144, 146, 148, 149, 150]  
 Row3 (17) [6, 7, 9, 19, 26, 50, 76, 79, 91, 99, 119, 143, 153, 164, 166, 173, 174]  
 Row4 (17) [2, 10, 31, 32, 43, 52, 59, 62, 74, **87**, 107, 133, 145, 151, 159, 160, 178]  
 Row5 (17) [13, 17, 36, 40, 54, 56, 58, 64, 78, 84, 97, 122, 123, 158, 168, 169, 182]  
 Row6 (17) [0, 1, 21, 38, 61, 67, 82, 110, 115, 117, 131, 134, 135, 137, 154, 156, 184]  
 Row7 (17) [3, 11, 34, 44, 46, 68, 75, 83, 85, 106, 108, 126, 141, 161, 175, 176, 180]  
 Row8 (17) [16, 23, 37, 49, 51, 60, 65, 69, 88, 93, 95, 102, 124, 136, 140, 177, 179]  
 Row9 (17) [29, 55, 66, 71, 77, 86, 92, 94, 101, 105, 132, 139, 142, 155, 157, 163, 172]  
 Row10(17) [4, 5, 25, 41, 45, 48, **53**, 57, 98, 109, 114, 120, 138, 147, 165, 170, 183]  
 Row11(17) [12, 24, 35, 47, 63, 73, 90, 96, **100**, 111, 113, 116, 118, 128, 152, 167, 171]

(b) Same-Size-K-means 法

図 18 T7 のクラスタリング結果

表2 クラスタリングの精度の比較実験

	制約付き K-means 法		Same-Size-K-means 法	
	エラー(個)	精度(%)	エラー(個)	精度(%)
T1	1	99	0	100
T2	0	100	0	100
T3	0	100	0	100
T4	0	100	0	100
T5	18	90	0	100
T6	12	93	0	100
T7	24	87	3	98

## 4.2. 行内の再構成の精度の比較

この実験では、行内の再構成の精度の向上を検討する。まず初めに、深層学習による従来手法の単体による再構成を行った場合と、前処理でフラグメントを接続してから再構成を行った場合との精度の比較を行う。文書画像 T1 における 11 行でそれぞれ再構成を行い、フラグメント同士で誤った接続が行われた数の合計を比較する。次に、従来手法[4]の再構成と深層学習に前処理を組み合わせた再構成の手法で精度の比較を行う。行内には全部で 17 組の組み合わせがある。再構成の精度の比較には以下の式(4)を用いる。

$$\text{Accuracy} = 1 - \frac{\text{the total number of error pairs in rows}}{\text{the number of pairs in rows}}$$
$$1 - \frac{\text{the total number of error pairs in rows}}{11 \times 17} \quad (4)$$

再構成で誤った接続の数を表 3 に示す。合計で 73 個の誤った接続が 43 個に減らせることが分かった。再構成の前処理としてフラグメントを接続して数を減らすと、行内の再構成における誤ったフラグメントの接続が減ることを確認できた。これにより、フラグメントの数を減らす前処理が深層学習による再構成の手法と組み合わせても機能することが分かった (図 19)。しかし、表より Row3 だけは誤ったフラグメントの組み合わせの数が増えていることが分かる。Row3 では、図 20 の様に前処理の段階で誤った組み合わせのフラグメントにも関わらず高い類似性があったためそのまま接続されてしまっていた。このように、クロスカットされた文書のフラグメントでは、境界面の情報量が限られてくるため、フラグメント同士が強い隣接関係にあっても必ずしも正しい組み合わせではないというリスクがあることが分かる。

従来手法[4]と提案手法の行内の再構成の比較の結果を表 3 に示す。深層学習による再構成では、前処理と組み合わせても再構成の精度の結果では劣ることが分かった。スプリットカットされた文書画像では高い精度を示している手法

でも、クロスカットされた文書では多少精度が落ちることは予想していたが、クロスカットされた文書画像の再構成が困難であることが分かった。

表 2 行内の再構成の比較

	深層学習	深層学習+前処理
Row1	5	5
Row2	5	3
Row3	7	8
Row4	6	5
Row5	3	3
Row6	5	0
Row7	8	5
Row8	6	0
Row9	8	5
Row10	10	6
Row11	11	3
合計	74	43

compassionate with one another, to cherish our own identity, and to accept unconditionally the same in others.

You should all be very proud that this is the Wellesley spirit. Now I know your first choice today was Alice Walker -- guess how I know! - known for The Color Purple.



(a) 深層学習のみ

compassionate with one another, to cherish our own identity, and to accept unconditionally the same in others.

You should all be very proud that this is the Wellesley spirit. Now I know your first choice today was Alice Walker - guess how I know! - known for The Color Purple.

(b) 前処理あり

図 19 Row7 の再構成の比較

the un  
 ars a  
 republ  
 pirit  
 e is

+

k yo  
 ted h  
 uck l  
 an id

=

the un k yo  
 ars ated h  
 republ uck l  
 pirit  
 e is an id

図 20 前処理による誤った接続の例

表 3 2つの手法による行内の再構成精度の比較

	提案手法	従来手法[4]
精度(%)	77.0	94.7

### 4.3. クロスカットされた文書画像の再構成の結果

クロスカットされた文書画像の再構成を、最初から最後まで全体を通して行った結果について考察する。

図 21 に、文書画像 T1 における再構成の結果を元の文書画像 T1 とともに示す。クラスタリングについては、提案手法により正しく行えていることが確認できる。また、行間の再構成に関しては、図で見れば明らかに行がかみ合っていないように見える箇所がある。しかし、行間の再構成は 2 章で説明した通り、クラスタリングベクトルを抽出して行っていたので、行内の再構成の結果には左右されずに再構成を行うことができる。よって、クラスタリングが正しく行えていれば、行間の再構成も正しく行えていることが確認できる。

以上より、提案手法におけるクロスカットされた文書画像の再構成が上手くいっていない部分の原因は、行内の再構成であることが分かる。全体を通して見ると、フラグメントの上部、あるいは下部で文字の行が途切れている場合は、再構成が上手くいっていないものが多い。これは、フラグメントに含まれる情報量の単純な大小よりも、フラグメント内に文字の行がしっかりと含まれている場合の方が、行内の再構成を正しく行いやすくなるのだと考えられる。



Thank you very, very much, President Keohane. Mrs. Gorbachev, Trustees, faculty, parents, and I should say, Julia Porter, class president, and certainly my new best friend, Christine Bicknell - and, of course, the Class of 1990. I am really thrilled to be here today, and very excited, as I know all of you must be, that Mrs. Gorbachev could join us.

These are exciting times. They're exciting in Washington, and I have really looked forward to coming to Wellesley. I thought it was going to be fun. I never dreamt it would be this much fun. So, thank you for that.

More than ten years ago, when I was invited here to talk about our experiences in the People's Republic of China, I was struck by both the natural beauty of your campus and the spirit of this place.

Wellesley, you see, is not just a place but an idea - an experiment in excellence in which diversity is not just tolerated, but is embraced. The essence of this spirit was captured in a moving speech about tolerance given last year by a student body president of one of your sister colleges. She related the story by Robert Fulghum about a young pastor, finding himself in charge of some very energetic children, hits upon the game called "Giants, Wizards, and Dwarfs." "You have to decide now," the pastor instructed the children, "which you are - a giant, a wizard, or a dwarf?" At that, a small girl tugging at his pants leg, asked, "But where do the mermaids stand?" And the pastor tells her there are no mermaids. And she says, "Oh yes there are - they are. I am a mermaid."

Now this little girl knew what she was, and she was not about to give up on either her identity, or the game. She intended to take her place wherever mermaids fit into the scheme of things. "Where do the mermaids stand? All of those who are different, those who do not fit the boxes and the pigeonholes?" "Answer that question," wrote Fulghum, "And you can build a school, a nation, or a whole world." As that very wise young woman said, "Diversity, like anything worth having, requires effort - effort to learn about and respect difference, to be compassionate with one another, to cherish our own identity, and to accept unconditionally the same in others.

You should all be very proud that this is the Wellesley spirit. Now I know your first choice today was Alice Walker - guess how I know! - known for *The Color Purple*. Instead you got me - known for the color of my hair. Alice Walker's book has a special resonance here. At Wellesley, each class is known by a special color. For four years the Class of '90 has worn the color purple. Today you meet on Severance Green to say goodbye to all of that, to begin a new and a very personal journey, to search for your own true colors.

In the world that awaits you, beyond the shores of Waban - Lake Waban, no one can say what your true colors will be. But this I do know: You have a first class education from a first class school. And so you need not, probably cannot, live a "paint-by-numbers" life. Decisions are not irrevocable. Choices do come back. And as you set off from Wellesley, I hope that many of you will consider making three very special choices.

The first is to believe in something larger than yourself, to get involved in some

Trust me, thank you very, very much, President Gorbachev, Keohane, Mrs. es, family, parents, and I should say, Julia Pident, and certainly, class president, my new best friend, Christine Bicknell - and, class of 1990. I am of course, the Clilly thrilled to be here today, and now all of us I know very excited, must be, that Mrs. Gorbachev could join us.

These are exciting times. Washington, they're exciting and I have really looked forward to coming to it was going to Wellesley. It is to be fun. I never dreamt it would be this much.

More than ten years ago, when I was invited to talk to our international experience abroad in the People's Republic of China, I was struck by the beauty of our campus and the spirit of this place.

Wellesley, but diversity is embraced. The essence of this spirit was captured in a message given last year speaking by a student body president of one of our members. She related the story of my by Robert Fulghum about himself in charge of some very energetic children. "Giants, Wizards, and Dwarfs." "You, hits upon the game call have to decide the children, "which you are - a giant, now," the pastor instructs wizard, or a dwarf tugging at his pants leg, asked, "But warf?" At that, a small girl here do the pastor tells her there are no mermaids. "maids stand?" And she says, "Oh yes there are - they are. I am a mermaid."

Now this little girl knew what she was, and she was not giving up on either her identity, or the game. She intended to take her place with the mermaids. fit into the scheme of things. "Where do the mermaids live? Those who stand different, the holes?" "Answer the poxes and the big goose who do cover that question," well, a nation, or. "And a man can build a whole world." As my, like anything you, "Dionysus said that very wise man worth having, respect difference, to earn about to do requires effort - to be compassionate with one another, to cherish our own identity, and to accept unconditionally the same in others.

You should all be very proud that this is the Wellesley spirit. Now I know your first choice today was Alice Walker - guess how I know! - known for The Color Purple. I know for the color. Alice Walker instead you got my book. At Wellesley, each is a special class is color. The '90 has worn the color you saw our years the Color Purple meet on. See and by to all of that, and a very good Green to see begin a personal journey.

In the words of W. Laban, no wonder the spirit that awaits, beyond shores of Babylon can say could have. You are first will be. But your truth is what this I do know educationally cannot, live. And so from a first school for you need not, be a "painter. The colors are numbers." And as you set off Wellesley will know you, I hope that many make it to the very special.

The first is to believe something get involved, and loved than yours in sorrow me

図 21 (b)提案手法で再構成した文書画像 T1

## 5. 結論

本研究では、クロスカットされた文書画像の再構成における手法の改善の検討を行った。まず、クラスタリングでは、従来手法ではクラスタリングの際にクラスタに割り当てられるフラグメントの数が偏ってしまったり不足してしまったりする誤りが多かった。これに対して、提案手法では新しいクラスタリングのアプローチとしてクラスタのサイズが同じになるようにクラスタリングを行う **Same-Size-K-means** 法を適用してクラスタリングを行い、比較実験によってクラスタリングの精度を改善できることを示した。次に、行内の再構成では、横方向の並べ替えのみを考えればよいことからストリップカットされた文書画像の再構成の手法において高い精度を残していた深層学習による再構成の手法を適用して精度の向上を行えるか検討した。しかし、クロスカットされた文書画像では著しく再構成の精度が下がってしまうことが分かった。また、前処理によりフラグメントの数を減らすことで再構成の精度を向上させた上でも、従来手法と比較しても再構成の精度は及ばなかった。この結果から、スプリットカットされた文書画像再構成に比べてクロスカットされた文書画像の再構成が困難であることが分かった。

今後の課題としては、行内の再構成の精度の向上と、再構成対象となる文書画像の複雑化が挙げられる。本研究でテストデータとして使用した文書画像は行の最初のフラグメントの認識が容易であることや、行間やフォントサイズが統一されているなどの様々な条件が整えられている。例えば、クラスタリングの際に、行の最初のフラグメントをクラスタ中心に指定できない場合などで、クラスタリングの精度を維持できるようにする方法を考える必要がある。

## 謝辞

本研究を行うにあたり、多くの方々に協力をしていただきました。大変お世話になりましたことを、この場を借りてお礼申し上げます。

指導教員である椋木雅之教授には研究や論文作成に関して様々のご指導を頂きました。本当にありがとうございました。

坂本教授、井上教授には、副査を務めていただきました。お忙しい中、貴重な時間を割いて下さりましてありがとうございます。

椋木研究室の皆様には、研究を進める中で様々な助言をいただきました。充実した研究室生活を過ごすことができたと思います。大変感謝しています。

## 参考文献

- [1] Faure, C., & Vincent, N. “Document image analysis for active reading,” Proceedings of the International Workshop On Semantically Aware Document Processing And Indexing, 2007
- [2] Ukovich, Anna, et al. “Shredded document reconstruction using MPEG-7 standard descriptors,” Proceedings of the Fourth IEEE International Symposium on Signal Processing and Information Technology, 2004.
- [3] C. Schauer, M. Prandtstetter, and G. R. Raidl, “A memetic algorithm for reconstructing cross-cut shredded text documents,” Hybrid Metaheuristics, vol. 6373 of Lecture Notes in Computer Science, pp. 103–117, Springer,2010.
- [4] Chen, Junhua, et al. “A solution to reconstruct cross-cut shredded text documents based on constrained seed K-means algorithm and ant colony algorithm,” Expert Systems with Applications 127: 35-46,2019.
- [5] Justino, Edson, Luiz S. Oliveira, and Cinthia Freitas. “Reconstructing shredded documents through feature matching,” Forensic science international 160.2-3 : 140-147,2006.
- [6] H. Y. Lin and W. C. Fan-Chiang, “Reconstruction of shredded document based on image feature matching,” Expert Systems with Applications, vol. 39, no. 3, pp. 3324–3332, 2012.
- [7] B. Biesinger, C. Schauer, B. Hu, and G. R. Raidl, “Enhancing a genetic algorithm with a solution archive to reconstruct cross cut shredded text document,” Computer Aided Systems Theory-EUROCAST, pp. 380–387,2013.

- [8] M. Prandtstetter and G. R. Raidl, “Meta-heuristics for reconstructing cross cut shredded text documents,” Proceedings of the ACM 11th Annual Genetic and Evolutionary Computation Conference (GECCO '09), pp. 349–356, 2009.
- [9] H. Xu, J. Zheng, Z. Zhuang, S. Fan “A solution to reconstruct cross-cut shredded text documents based on character recognition and genetic algorithm,” Abstract and applied analysis, 2014.
- [10] Paixao, Thiago M., et al. “A deep learning-based compatibility score for reconstruction of strip-shredded text documents,” Conference on Graphics, Patterns and Images (SIBGRAPI), 2018.
- [11] D. Lewis, G. Agam, S. Argamon, O. Frieder, D. Grossman, and J. Heard. “Building a Test Collection for Complex Document Information Processing,” Conf. on Res. and Develop. in Inf. Retrieval, pp. 665–666, 2006.
- [12] S. Chopra, R. Hadsell, and Y. LeCun. “Learning a similarity metric discriminatively, with application to face verification,” Conf. on Comput. Vision and Pattern Recognit. (CVPR), pp. 539–546, 2005
- [13] Same-size k-Means Variation - ELKI Data Mining, [https://elki-project.github.io/tutorial/same-size\\_k\\_means](https://elki-project.github.io/tutorial/same-size_k_means)
- [14] J. Chen, D. Ke, Z. Wang, Y. Liu “A high splicing accuracy solution to reconstruction of cross-cut shredded text document problem,” Multimedia Tools and Applications, 77 (15), pp. 19281-1930, 2018