

修士論文

深層学習を用いた
漫画中のオノマトペの検出

令和 1 年度修了

三重大学大学院 工学研究科 情報工学専攻
ヒューマンインターフェース研究室

加藤ななみ

はじめに

漫画は日本を代表するメディアコンテンツの1つであり，世界中で広く親しまれている．日本文化を世界に発信していくために漫画にも注目が集まっており，漫画を様々な言語で出版するために漫画翻訳の需要が高まっている．漫画は，擬態語や擬声語を表現する手書き文字であるオノマトペを含む．オノマトペは，その効果で聴覚のイメージに訴えかけ，臨場感を高め，漫画の表現を広げる重要な要素である．漫画翻訳は，紙媒体をスキャンして電子化し，手動で文字を消して翻訳したセリフを貼り付けたり，描き直して行われるため，人手や時間がかかる作業である．また，電子書籍市場の発展により，電子漫画の需要も増加している [1] [2]．そこではオノマトペに動きをつけるなど，電子漫画ならではの付加価値を持たせる試みが行われている [3] [4]．そのため，もしオノマトペを自動で検出できれば，翻訳の手間・費用を削減することが可能である．また，オノマトペを別レイヤーに配置して漫画画像に階層構造を持たせることで，利便性の向上が期待できる．近年では，漫画解析の研究が増え，コマや吹き出しの自動検出の研究が進み，実際にサービスが提供されているものもある [5] [6]．しかしオノマトペの検出は限定的となっているものがほとんどである．さらに，オノマトペの書き方は作者や場面によってさまざまであり，従来の文字認識手法では認識が難しい．

先行研究として，橋本ら [7]，劉ら [8]，石橋ら [9] の研究がある．しかし，[7][8] は太い黒文字のオノマトペ以外は検出不可能である．また，文献 [9] は太い黒文字のオノマトペや，枠線の幅が平均的な白抜き文字のみの検出となる．

そこで本研究では，深層学習を用いた漫画画像中のオノマトペ検出手法を提案し，検出可能な形状を限定しないオノマトペの検出を実現する．本手法は，漫画画像を入力し，黒画素と白画素がそれぞれオノマトペと背景を表す2値画像を出力する．この手法は主にU-Netによるオノマトペ検出，検出領域の膨張，ノイズ削減の3ステップからなる．まずU-Netにより漫画画像からオノマトペを画素単位でセグメンテーションし，2値画像を出力する．次に検出領域の欠落を補うため検出領域を膨張する．最後に検出された2値画像内の連結成分ごとにオノマトペであるかどうかをConvolutional Neural Networksにより判定しオノマトペでない連結成分を棄却することでノイズを削減する．また，公開されて

いる漫画画像中のオノマトペのアノテーションデータがないため、5 作品の漫画に対しオノマトペのアノテーションを付与した。

評価実験は、東京大学による漫画データセット Manga109 [10] からオノマトペの書き方が異なる 3 種類の漫画を選択し、評価には再現率、適合率、F 値を用いた。各漫画それぞれ、学習に 90 画像、検証に 10 画像、評価に 20 画像を使用した。実験の結果、最大で 73.6%の F 値が得られ、再現率 80.7%、適合率 67.9%であった。黒文字・白抜き文字・小さな文字のいずれのオノマトペも検出できた。

目次

| | |
|----------------------------------|-----------|
| はじめに | i |
| 第 1 章 序論 | 1 |
| 1.1 研究背景 | 1 |
| 1.2 関連研究 | 2 |
| 1.3 本研究の概要 | 3 |
| 第 2 章 本研究に関連する技術 | 4 |
| 2.1 深層学習 | 4 |
| 第 3 章 提案手法 | 11 |
| 3.1 概要 | 11 |
| 3.2 アノテーション | 14 |
| 3.3 U-Net による検出 | 15 |
| 3.4 オノマトペ領域の膨張 | 18 |
| 3.5 CNN による非オノマトペの消去 | 19 |
| 第 4 章 実験 | 21 |
| 4.1 実験条件 | 21 |
| 4.2 パラメータ選択のための予備実験 | 27 |
| 4.3 提案するオノマトペ検出手法性能評価実験 | 29 |
| 第 5 章 結言 | 38 |
| 5.1 本研究のまとめ | 38 |
| 5.2 今後の展望 | 38 |
| 付録 A 研究で用いたデータの参照場所, 実行環境 | 41 |
| A.1 プログラム, 実験データ | 41 |
| A.2 実行環境 | 41 |

| | |
|-----------|----|
| 目次 | iv |
| 付録 B 発表資料 | 42 |
| 謝辞 | 45 |

第 1 章

序論

1.1 研究背景

漫画は日本を代表するメディアコンテンツの一つであり，世界中で広く親しまれている．漫画やアニメといったコンテンツは政府によるクールジャパン戦略の一部に含まれており [11]，日本の魅力を世界に発信していくために漫画にも注目が集まっている．そのため，漫画を様々な言語で出版しようと，漫画翻訳の需要が高まっている．漫画の翻訳は，紙媒体をスキャンして電子化し，手作業で文字を消して翻訳したセリフを貼り付けたり，描き直して行われるため，人手や時間がかかる．そのため，擬態語や擬音語を表現する手書き文字であるオノマトペは翻訳されないまま漫画が出版されることもある．そこで，オノマトペを自動で検出できれば，翻訳の手間・費用の削減が期待できる．

また，電子書籍市場の発展により，電子漫画の需要も増加している．紙媒体が低迷する中，電子書籍の国内市場規模は成長を続けており，2018 年度は 2826 億円と推計され，2017 年度から 585 億円 (26.1 %) 増加している．さらに漫画は市場での占有率を高め，電子書籍市場の 84.5% を占める．現在紙媒体でのみ出版されているコンテンツの電子化を進める動きもあり [1][2]，これからも，電子漫画の需要はますます高まっていくと考えられる．

電子漫画では，CM などで見られるモーションコミックのように，漫画の構成要素を動かすといった電子漫画ならではの効果をつけるなど，付加価値を持たせる試みが行われている [3][4][12]．しかし，既存の漫画は単純な画像データであることがほとんどである．そこで，オノマトペを自動検出し，ページ上に重畳した別レイヤーに配置して漫画画像に階層構造を持たせることができれば，利便性の向上が期待できる．また，漫画の要素を構造的にデータ化することで，電子書籍を提供するサイト等での検索の利便性やレコメンドの高性能化が期待できる．

漫画中のオノマトペは，その効果で聴覚のイメージに訴えかけ，臨場感を高め，漫画の

表現を広げている重要な要素である。しかし、オノマトペの自動検出には以下のような困難がある。第一にオノマトペは背景と一体化していることがある、第二にページの端で文字が途切れていることがある、第三に書き方が作者や場面によってさまざまである。これらが原因となり、従来の文字認識手法では検出と認識が難しい。

1.2 関連研究

1.2.1 漫画画像に関連する研究

近年では、漫画解析の研究が増え、コマや吹き出し検出の研究が進み、実際にサービスとして提供されているものもある [5][6]。

石井は漫画のコマ検出、キャラクター情報抽出からなる漫画画像解析研究を行い、2値線画像解析のための画像特徴量として、2値細線の方向成分をとらえる HBOP 特徴量を提案している [13]。柳澤らは Fast R-CNN が漫画のコマと吹き出しの検出に、Faster R-CNN がキャラクターの顔とテキストの検出に有効であることを示している [14]。

また、田中らは AdaBoost によりセリフ検出を行うことで吹き出し候補を検出し、SVM を用いて吹き出しの分類および検出を行う手法を提案し、約 86% の精度を出している [17]。Dubray らは、U-Net を用いて、主に欧米で出版された漫画の吹き出し検出を行い、94% を超える F 値を得ている [16]。

オノマトペ検出は、主にオノマトペの形状に着目したハンドクラフトな特徴量を元に検出を試みる研究がほとんどであり、まだ実用的なレベルには至っていない。橋本らによる研究 [7] では、オノマトペのほとんどが同一の画素値で構成されているとし、連結成分の面積と縦横の大きさを用いて検出している。しかし正抽出数よりも過検出・未検出数が多い。石橋らによる研究 [9] では、オノマトペが一定線幅の線分で構成される点に着目して、黒文字にはモルフォロジー演算、白抜き文字には領域分割手法と、二種類の手法を用いて抽出している。しかし、黒文字抽出手法では黒画素を膨張・収縮させることで細かい線の背景を除去してオノマトペを抽出するため、小さなオノマトペが検出できない上、黒領域の大きな背景が含まれてしまう。白文字抽出手法は漫画画像を小さな領域に分割し領域内の輝度値の分散値によって検出するが、分割領域の大きさをオノマトペの黒枠の太さの平均から決めているため、黒枠の太さが平均から大きく外れたオノマトペを検出できない。劉らによる研究 [8] では、オノマトペは均一な線幅で書かれると想定し、ドロネー三角形分割を利用した幅の均一なりボン状領域の抽出に加え、同一文字は類似した色を持つという特徴を合わせて利用してオノマトペを抽出する。しかしこの手法では文字が小さすぎると、同色の三角形連結領域を取得するために必要なドロネー三角形分割で得た三角形の個数が足りないため、小さな文字を検出できない。

1.3 本研究の概要

本研究では、深層学習を用いた漫画画像中のオノマトペ検出手法を提案し、検出可能な形状を限定せずオノマトペを検出することを目標とする（図 1.1）。

本手法は、漫画画像を入力し、黒画素と白画素がそれぞれオノマトペと背景を表す 2 値画像を出力する。この手法は主に U-Net によるオノマトペ検出、検出領域の膨張、ノイズ削減の 3 ステップからなる。また、公開されている漫画画像中のアノテーションデータにはオノマトペのアノテーションが含まれていないため、5 作品の漫画に対し独自にオノマトペのアノテーションを付与して用いる。

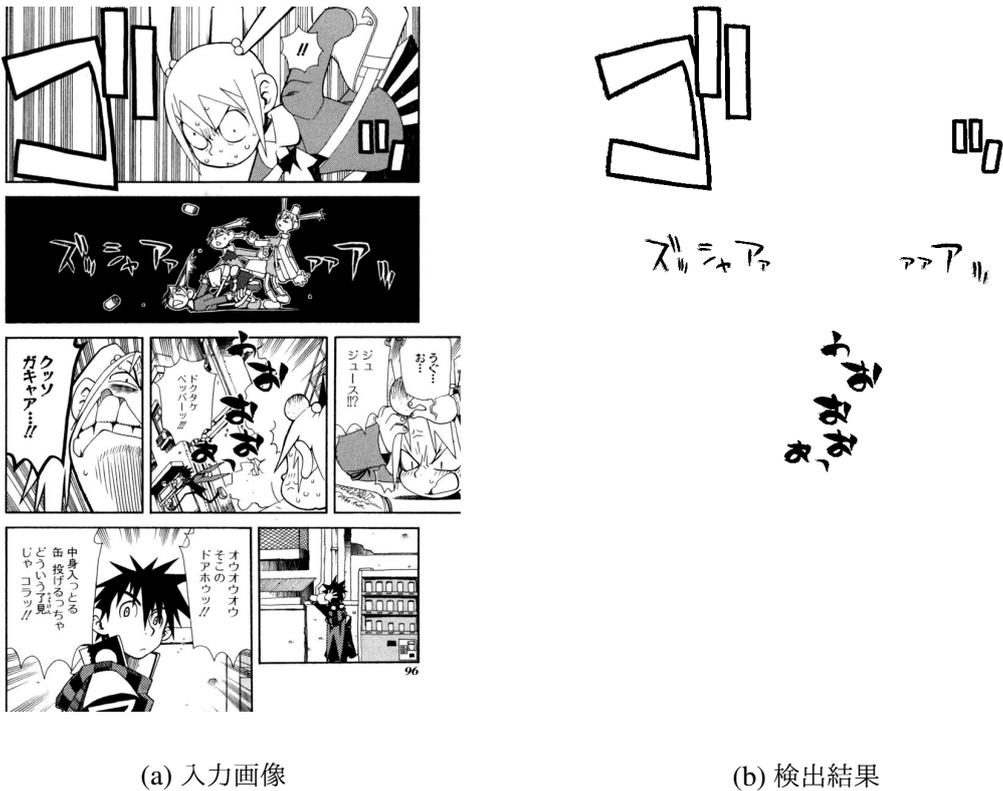


図 1.1: 研究の目標

第 2 章

本研究に関連する技術

本章では，提案手法で用いる要素技術について記す．

2.1 深層学習

深層学習（ディープラーニング）は機械学習手法の一つであり，ディープニューラルネットワークとも呼ばれる．近年，多層のネットワークを接続したディープニューラルネットワーク技術の進展により，画像認識や音声認識等の性能が向上し，様々な認識問題に対して高い精度を出している．画像認識においては畳み込みニューラルネットワーク (Convolutional Neural Networks (CNN)) が主に用いられている．

2.1.1 Convolutional Neural Network (CNN)

CNN は，複数の畳み込み層とプーリング層，全結合層，出力層で構成される．畳み込み処理 (Convolution) と活性化関数によって画像から特徴マップを生成し，プーリングにより縮小する．畳み込みとプーリングを交互に繰り返し，活性を伝播させることで，多チャンネル特徴マップが得られる．得られた特徴マップを全結合層に受け渡し，全結合層・出力層でモデルが推定した事後確率ベクトルを出力する．

畳み込み層

畳み込み (convolution) は，入力された画像の一部と任意の大きさのフィルターの内積の計算を，フィルターをスライドさせながら入力画像全体に対して行う処理である．畳み込み処理の概略図を図 2.1 に示す．

l 層目のフィルタサイズ $W^l \times H^l$ への入力 x^l に対しての畳み込み処理は式 (2.1) で定

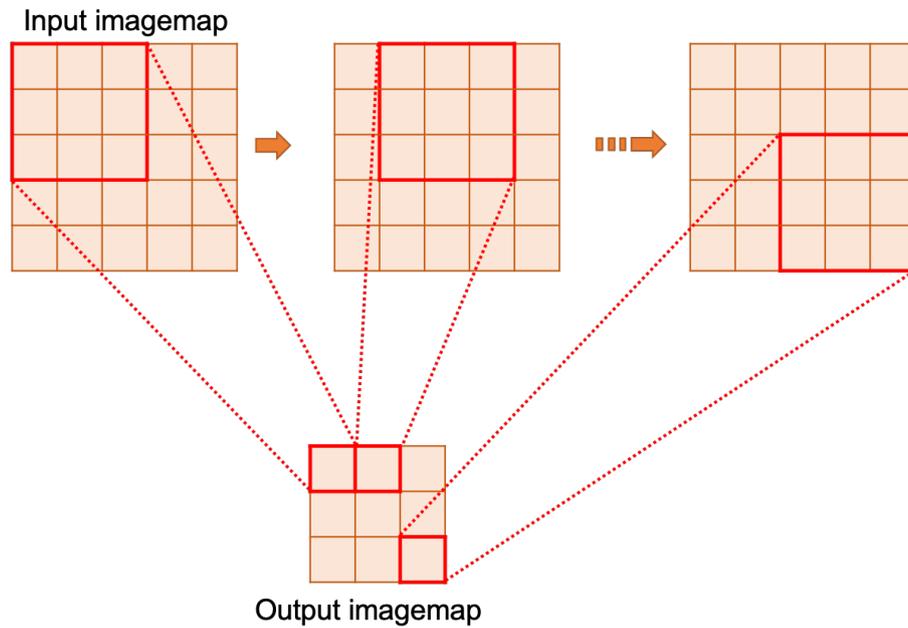


図 2.1: 畳み込み処理

義できる.

$$y_{i,j}^{l+1} = \sum_{p=-H^l/2}^{H^l/2} \sum_{q=-W^l/2}^{W^l/2} f_{p,q} \times x_{i+p,j+q}^l \quad (2.1)$$

ここで、 $y_{i,j}^{l+1}$ は $l+1$ 層目に出力された特徴マップ y の i 行 j 列目の要素、 $f_{p,q}$ はフィルタの p 行 q 列目の要素を表す。3 階のテンソル型が入力の場合も同様に定義できる。学習段階では、画像中の特徴を見つけた場合に活性化するようにフィルタを最適化する。

畳み込みで得られた特徴マップを活性化関数に通し、非線形変換を行う。代表的な活性化関数である ReLU を式 (2.2) に示す。概略図を図 2.2 に示す。

$$z_{i,j}^l = \max\{0, y_{i,j}^l\} \quad (2.2)$$

畳み込み、活性化により画像の持つ特徴を得たのち、位置ずれに対して頑健にするため、任意の大きさの局所領域内の最大値を取り、特徴マップのサイズを圧縮する最大値プーリングを行う。特徴マップサイズが 4×4 、局所領域サイズが 2×2 の場合の例を図 2.2 に示す。

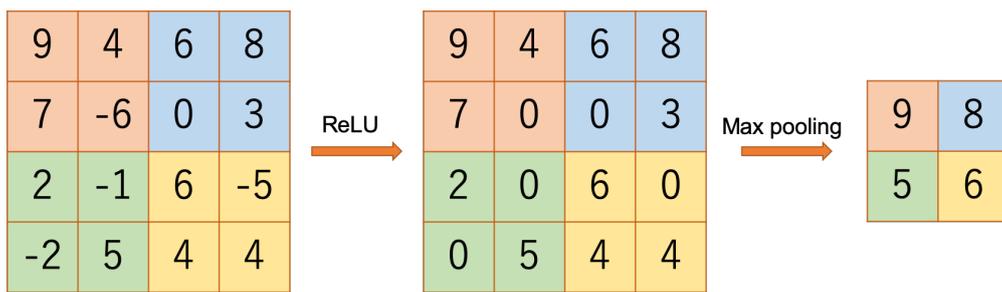


圖 2.2: ReLU, Max pooling

2.1.2 Fully Convolutional Networks (FCN)

Fully Convolutional Networks (FCN) [18] は、全て畳み込み層で構成されている CNN である。ネットワーク構成の例を図 2.3 に示す。FCN は画像を画素単位でクラス分類し画像内を領域分割する、セマンティックセグメンテーションなどに用いられる。

一般的な CNN では全結合層を用いるが、FCN では全結合層と同様の働きをする畳み込み層のみを用いる。最後に逆畳み込み演算を行い入力画像と同じ大きさまで画像を拡大する。よって任意の入力画像に対応した大きさの画像を出力できるため、教師データとして画像を用い画像対画像の関係を学習できる。

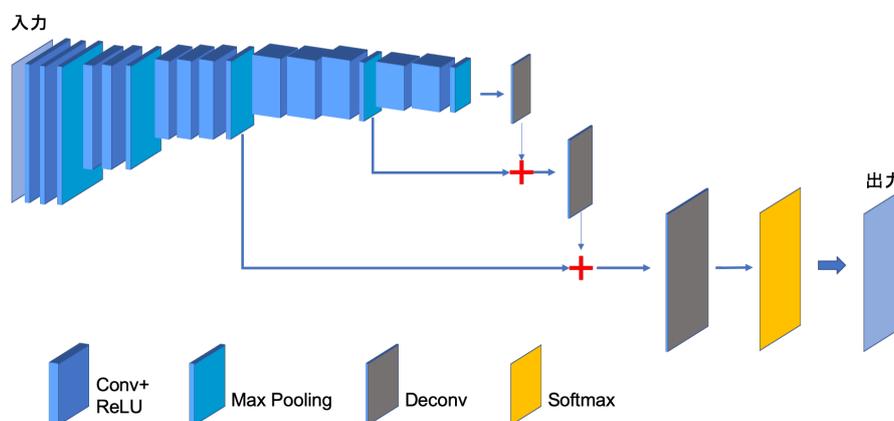


図 2.3: Fully Convolutional Networks (FCN)

逆畳み込み層

逆畳み込み (Deconvolution) は、前層から入力された特徴マップに対して画素間への空白の追加等を施しマップサイズを拡大し、拡大したものにフィルタを適用して畳み込みを行う。通常の畳み込みを行うとマップサイズは縮小されるが、畳み込みの前に任意のサイズの拡大を行うことで、入力より大きいサイズの特徴マップを出力できる。入力マップサイズが 2×2 、フィルタサイズが 2×2 の場合の例を図 2.4 に示す。

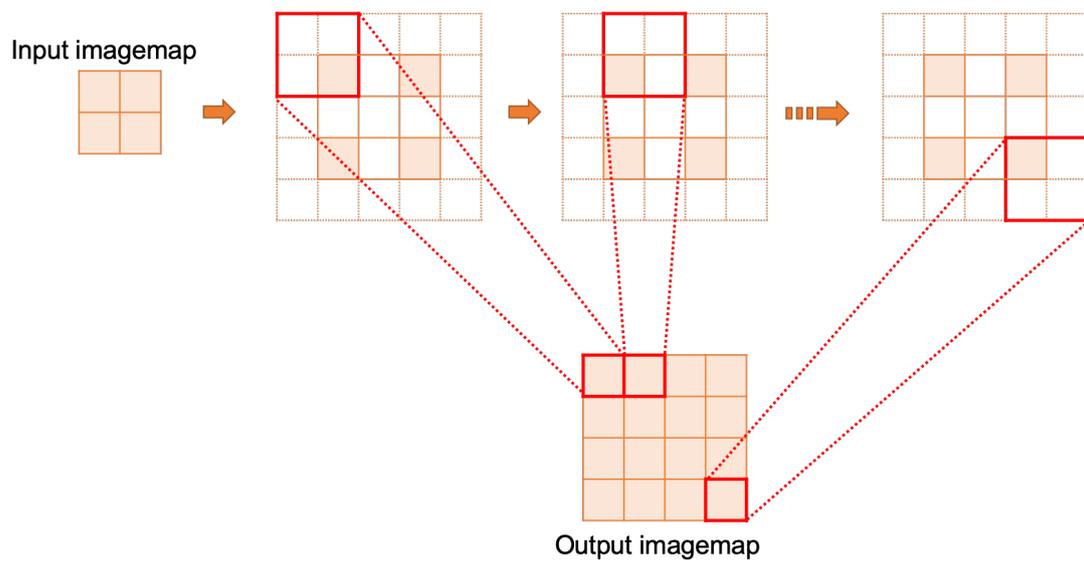


図 2.4: 逆畳み込み処理

2.1.3 U-Net

U-Net [19] は医用画像のセグメンテーションを目的に開発された，セマンティックセグメンテーション手法の一つである．図 2.5 にネットワーク構成の例を示す．他の分野での応用もされており，セマンティックセグメンテーション手法として高い精度を誇る．FCN の 1 種で，全結合層を持たず，畳み込み層で構成される．左右対象のエンコーダー・デコーダー構造で，エンコーダー（図 2.5 で左半分）で畳み込まれた特徴マップを，デコーダー（図 2.5 で右半分）で逆畳み込みし，元の解像度となる．

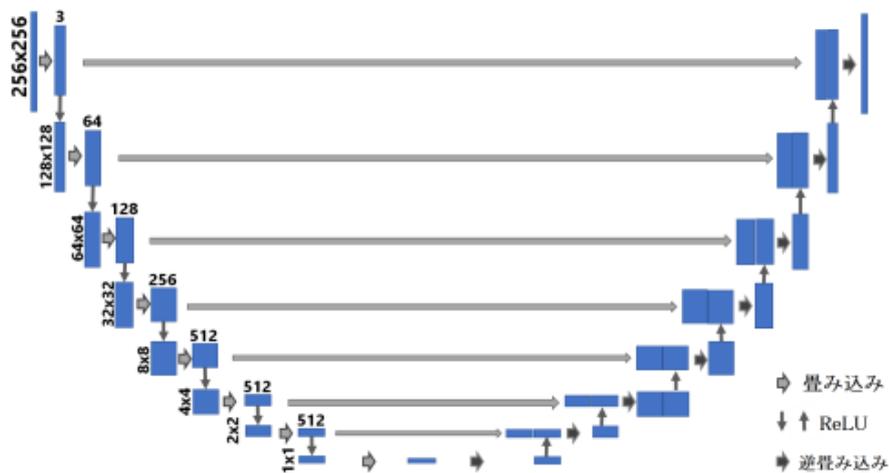


図 2.5: U-Net

スキップ接続

エンコーダーで畳み込んでいくと特徴マップが小さくなり細部の情報が失われてしまう．デコーダーの逆畳み込み層の出力マップと，同じサイズの対応するエンコーダーの出力マップを連結することで，細部の認識精度を高めている．スキップ接続の概要を図 2.6 に示す．

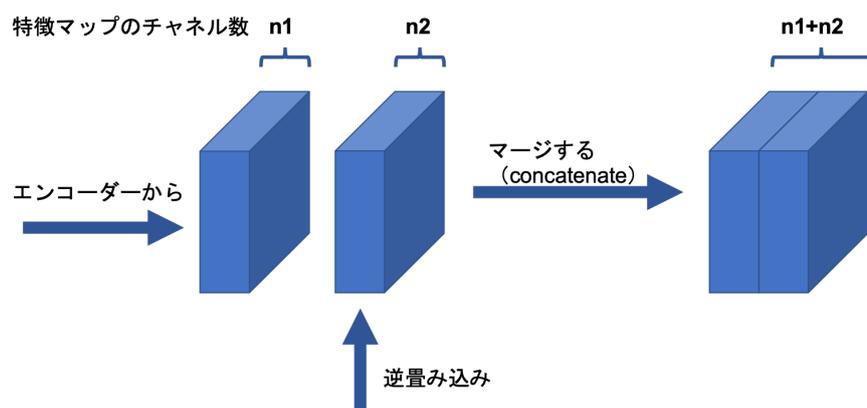


図 2.6: スキップ接続での特徴マップの連結

第 3 章

提案手法

3.1 概要

本研究で提案する検出手法では，漫画画像を入力とし，オノマトペ領域が抽出された画像を出力とする．入力された漫画画像に対し，U-Net を用いてオノマトペのみを残す画像変換を行う．入力は画像 1 つにつき漫画 1 ページが描かれたグレースケール画像，出力はオノマトペが黒色，オノマトペ以外が白色である 2 値画像とする．入出力の例を図 3.1 に示す．

提案手法によるオノマトペの検出は主に、U-Netによる検出、検出領域の膨張、ノイズ削減の3つの処理で構成される。流れを以下に示す。

1. U-Netによる検出

(a) 前処理

入力された漫画画像を白黒反転し、 256×256 pixelの画像に分割する。

(b) U-Netによる検出

前処理をした画像をU-Netに入力し、オノマトペであると予測した領域を白色、それ以外を黒色で表した画像を出力する。

(c) 後処理

U-Netの出力画像に対し、前処理と逆の要領で画像をつなぎ合わせ元の大きさに戻し、白黒反転する。

2. オノマトペ領域の膨張

オノマトペであるとして出力された領域を膨張する。

(a) U-Netの後処理後の出力画像から、オノマトペと予測された領域のうち、面積が一定値以下のものをノイズとして消去する。

(b) 入力画像と出力画像を比較し、どちらも黒画素である部分を黒、それ以外を白とする。

(c) 以下の(i), (ii)の処理を適切な回数繰り返す。

i. 残った黒画素を膨張する。

ii. 入力画像と出力画像を比較し、どちらも黒画素である部分を黒、それ以外を白とする。

3. CNNによる非オノマトペの消去

抽出された領域が真にオノマトペであるか判断する。オノマトペでないと判断された領域はノイズとして棄却する。

(a) 前処理

検出領域膨張後の出力画像から連結成分を切り出す。また、元の漫画画像(入力画像)から同じ部分を切り出す。

(b) 連結成分単位での分類

CNNにより連結成分がオノマトペか否かで2値分類する。オノマトペでないと分類されたものは消去する。

3.2 アノテーション

後述の Manga109 [10] にはオノマトペのアノテーションがつけられていないため、漫画画像に対しオノマトペ部分を黒、オノマトペでない部分を白で塗りつぶした画像を作成する。このアノテーションを正解画像として U-Net の学習を行う。

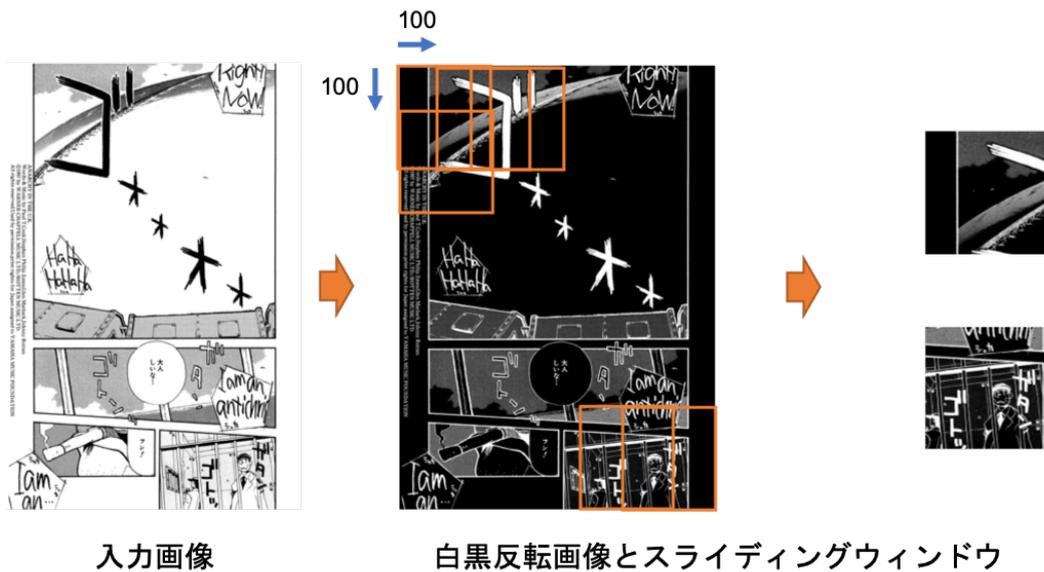


図 3.2: 前処理

3.3 U-Net による検出

3.3.1 前処理

まず入力画像を白黒反転し、次にスライディングウィンドウ形式で 256×256 pixel の小画像に分割する。ウィンドウサイズは 256×256 pixel, スライド幅は縦横どちらも 100 pixel である。流れを図 3.2 に示す。

3.3.2 U-Net による検出

学習済の U-Net によりオノマトペを検出する。前処理後の 256×256 pixel の小画像を入力すると、オノマトペ部分を白、オノマトペでない部分を黒で塗りつぶした 256×256 pixel の小画像を出力する。本手法で用いる U-Net の構成を図 3.3 に示す。

3.3.3 後処理

U-Net の出力した小画像をつなぎ合わせて入力画像と同じサイズに戻し、白黒反転する。オノマトペ部分が黒、オノマトペでない部分が白となる。つなぎ合わせる際に重複した画像の部分は画素値を足し合わせ、255 を超えた場合は 255 とする (式 3.1)。 $I(i, j)$ を

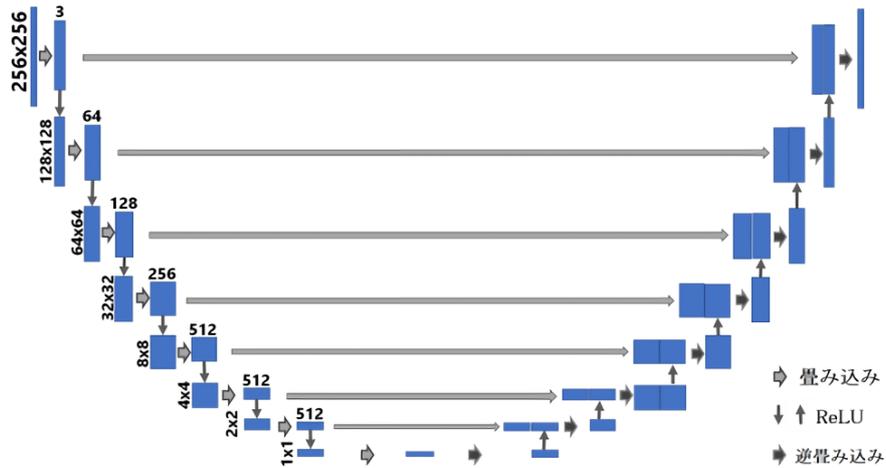


図 3.3: 提案手法の U-Net の構成

後処理後の画像の i 行 j 列目の要素, $P_k(i_p, j_p)$ を小画像の i 行 j 列目の要素とする.

$$I(i, j) = \min \left(\sum_k P_k(i_p, j_p), 255 \right) \quad (3.1)$$

これにより, 複数の小画像で切り出され U-Net による抽出が行なわれているオノマトペは, 図 3.4 のように少なくとも 1 つの画像で抽出されていればオノマトペとして検出される.

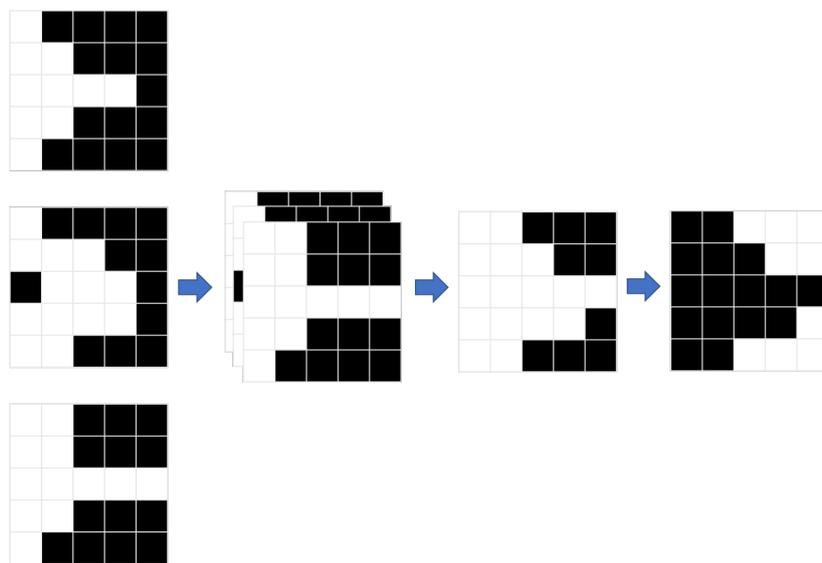


図 3.4: 後処理

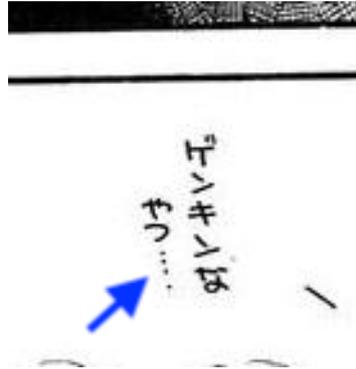


図 3.5: 極小領域

3.4 オノマトペ領域の膨張

まず U-Net の後処理後の出力画像から，オノマトペと予測された領域のうち，面積が一定値以下のものをノイズとして消去する．この際，検出対象となっている手書き文字のうち図 3.5 のように「・・・」が非常に小さいため除去されてしまうものがあるが，翻訳やエフェクト付与の対象になることは少ないと考えて問題なしとする．次に，入力漫画画像（オリジナルの漫画画像）とノイズ除去後の出力画像を比較し，どちらも黒画素である部分を黒，そうでない部分を白とする（式 3.2）．式 3.2 において， $I'_d(i, j)$ を比較処理後画像の i 行 j 列目の画素， $I_d(i, j)$ を出力画像の i 行 j 列目の画素， $I_o(i, j)$ を入力漫画画像の i 行 j 列目の画素とし，黒画素を 1，白画素を 0 とする．

$$I'_d(i, j) = I_d(i, j) \times I_o(i, j) \quad (3.2)$$

次に，抽出された領域を膨張する．以下の 2 つの処理をワンセットとし，適切な回数繰り返す．

1. 残った黒画素に対しフィルターサイズ 3×3 で膨張処理する．
2. 入力漫画画像と 1. の出力画像を比較し，どちらも黒画素である部分を黒，それ以外を白とする（式 3.2）．

これにより，入力漫画画像を参照しつつ連続する黒領域を広げることで，一部が欠けた状態で抽出されたオノマトペの欠けた部分を埋めていく．

3.5 CNN による非オノマトペの消去

膨張処理により，誤検出された領域も肥大してしまうため，学習済の CNN により連結成分がオノマトペであるかどうか分類する．オノマトペでないと判断された領域はノイズとして消去する．

3.5.1 前処理

膨張処理後の画像を白黒反転し，ここから，連結成分を外接矩形で切り出す．連結成分がひらがなの小さな一部である場合などには，オノマトペであるか判別が難しいため，外接矩形の縦または横が 28 pixel 未満である場合，連結成分の重心が中央になるように，28 pixel 未満の辺を 28 pixel の大きさに変更して切り出す．処理例を図 3.6 に示す．また，入力漫画画像 (オリジナルの漫画画像) も白黒反転し同一範囲を切り出す．

3.5.2 CNN での分類

学習済の CNN により連結成分がオノマトペであるかどうか判別する．前処理で切り出した，連結成分の画像および同範囲の入力漫画画像のペアを 28×28 pixel に正規化し CNN に入力する．出力として，連結成分がオノマトペであれば 1，オノマトペでなければ 0 を得る．本手法で用いる CNN の構成を図 3.7 に示す．

3.5.3 非オノマトペ消去

CNN によりオノマトペでないとされた連結成分を背景色 (白) で塗り替え，消去する．オノマトペが黒，オノマトペでない部分が白である画像を出力する．



図 3.6: 非オノマトペ消去の前処理

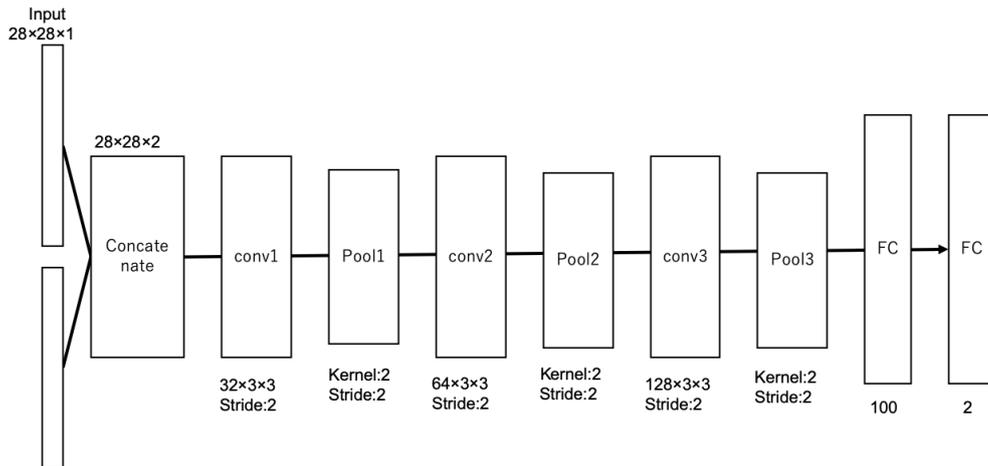


図 3.7: 本手法で用いる CNN の構成

第 4 章

実験

4.1 実験条件

4.1.1 データセット

Manga109

日本の漫画を収録した代表的なデータセットとして，東京大学による漫画画像データセット Manga109 [10] がある．このデータセットには 1970 年代～2010 年代に公開されたプロの漫画家による様々なジャンルの漫画が，109 作品 2114 ページ含まれている．漫画の見開き 2 ページ分を，サイズ 1654×1170 pixel のグレースケール画像ファイル 1 個として収録している．Manga109 に含まれるすべての漫画にはアノテーションが付与されており，アノテーションデータは漫画 1 作品に対し 1 個の xml 形式ファイルが対応している．アノテーションはキャラクターの顔，キャラクターの全身，コマ枠，テキストの 4 種類がつけられており，全てに id と矩形領域 (xmin, ymin, xmax, ymax) が付与されている．種類別の固有情報として，キャラクターの顔とキャラクターの全身には漫画内で統一されたキャラクター id が，テキストにはテキストの内容が与えられている．コマ枠は固有情報を持たない．テキストはセリフ，モノログ，状況説明を表す活字と一部の手書き文字に対して付与されており，オノマトペには付与されていない．

アノテーション

本研究では，Manga109 でアノテーションが付与されていないものも含む全ての手書き文字に対してアノテーションを付与する (図 4.1)．漫画中の手書き文字には本来の擬音語，擬態語以外も含むが，本研究では以下の理由によりアノテーションの対象とした．

- 図 4.2 (a), (b) に含まれるセリフとオノマトペのように，類似した書体で描写される場合が多い．
- 図 4.2 (c) のように吹き出し内のセリフが手書き書体で描画されている場合がある．
- 図 4.2 (d) のように，作者のコラムがコマ内に手書きにより記述される場合がある．

形状特徴のみを用いる検出の段階において，これらを切り分けることは困難である．

アノテーション付与作業は目視により手作業で行った．

使用したデータ

本実験では，Manga109 からオノマトペの書き方が異なる 3 種類の漫画，4 コマ漫画 [20](図 4.1(a)(d))，バトル漫画 [21](図 4.1(b)(e))，恋愛漫画 [22](図 4.1(c)(f)) を選択し，各漫画 120 ページを実験に用いた．また，公平に性能を評価するため 4 分割交差検証を行う．図 4.3 のように各漫画を 4 分割し，そのうち 3 セットを学習に，残り 1 セットのうち 10 ページを検証用，20 ページを評価用に用いた．これを 4 通りの組み合わせで行い，実験結果の評価は 4 通りの平均を用いる．



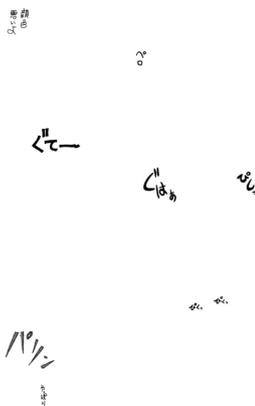
(a) 4コマ漫画



(b) バトル漫画



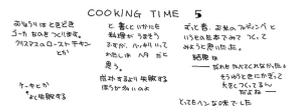
(c) 恋愛漫画



(d) 4コマ漫画の
アノテーション



(e) バトル漫画の
アノテーション



(f) 恋愛漫画の
アノテーション

図 4.1: Manga109

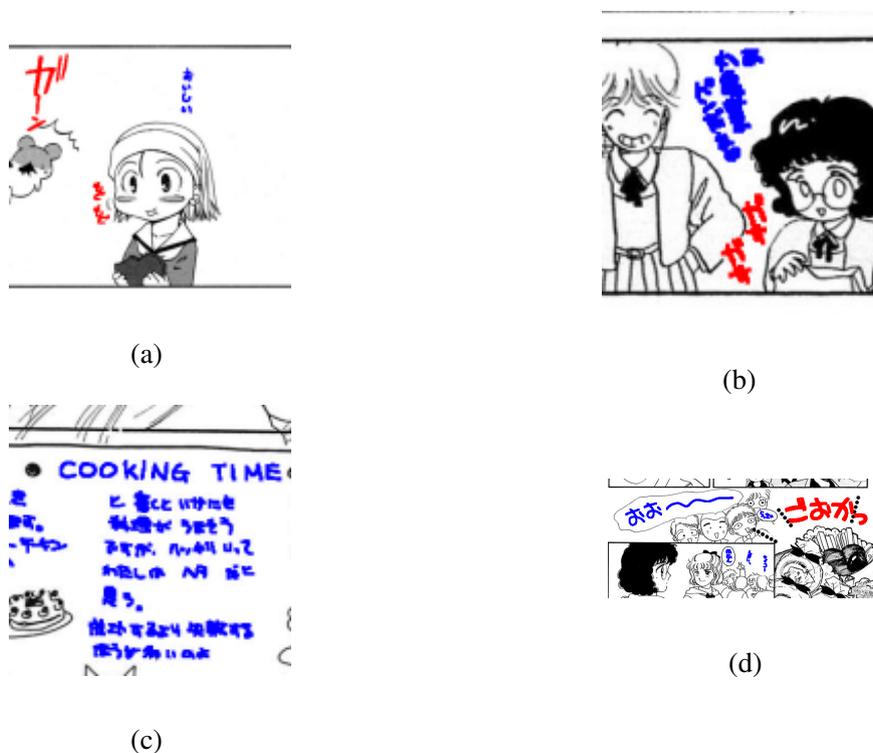


図 4.2: オノマトペ (赤字) に類似した手書き文字 (青字) の例

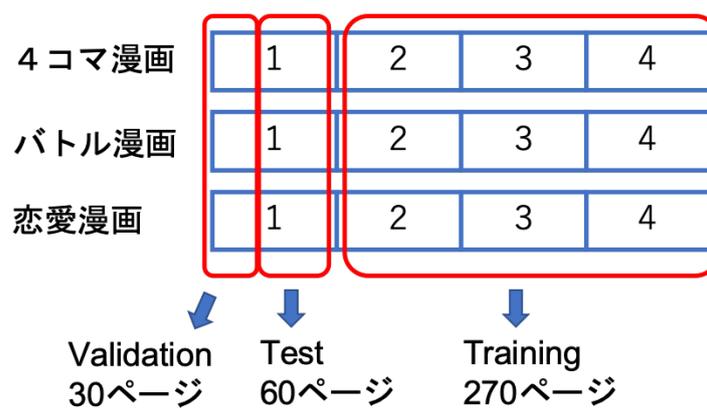


図 4.3: 4分割交差検証

4.1.2 評価基準

オノマトペ検出精度の評価の指標として、画素単位で適合率 (Precision), 再現率 (Recall), F 値 (F-measure) を用いる。これらは以下の式 (4.1), (4.2), (4.3) で定義する。ここで、適合率とはオノマトペとして検出された画素のうち実際にオノマトペであるものの割合、再現率は実際にオノマトペである画素のうちオノマトペとして検出されたものの割合を示し、F 値は適合率と再現率の調和平均である。

テストデータ内の全画素について TP, FP, FN の画素数を数え、適合率, 再現率, F 値を計算する。

$$Precision = \frac{TP}{TP + FP} \quad (4.1)$$

$$Recall = \frac{TP}{TP + FN} \quad (4.2)$$

$$F - measure = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (4.3)$$

- TP : オノマトペである画素をオノマトペとして検出した画素数
- FP : 非オノマトペ画素をオノマトペとして検出した画素数
- FN : オノマトペである画素を非オノマトペとして検出した画素数

4.1.3 オノマトペ検出のための U-Net の学習

学習用データセットと検証用データセット、および対応するアノテーション画像に前処理（入力画像を白黒反転し、 256×256 pixel の小画像に分割する）を施した後、U-Net に学習データ、検証データとして入力し後述のデータ拡張を行う。学習時の損失関数には 1-ダイス係数、活性化関数には ReLU、最終層には sigmoid を用いた。ダイス係数は 2 つの集合の類似度を表す。それぞれ式 (4.4), (2.2), (4.5) で表される。 $Y_{predicted}$ は予測、 Y_{true} は正解、 a は関数への入力を表す。学習はバッチサイズ 16、エポック数 300 で行い、検証用データで最も小さい損失値であったエポックでのモデル（学習した重みパラメータ）を保存し、評価に用いる。

$$loss = 1 - \frac{2 \cdot |Y_{predicted} \cap Y_{true}|}{|Y_{predicted}| + |Y_{true}|} \quad (4.4)$$

$$sigmoid(a) = \frac{1}{1 + \exp(-a)} \quad (4.5)$$

4.1.4 非オノマトペ判別のための CNN の学習

まず、U-Net 学習用データセットの漫画画像を学習済 U-Net に入力し、オノマトペ領域を予測した画像を得る。これらの画像にランダムでノイズを付与し、膨張処理を行う。これらの画像を CNN の学習用画像とする。CNN の学習用画像に、CNN の前処理（連結成分の切り出し）を行い、得られた連結成分画像および同範囲の漫画画像、連結成分のラベル（オノマトペか否か）を CNN に与え、学習する。学習時の損失関数にはバイナリクロスエントロピー、活性化関数には ReLU、最終層には softmax を用いた。それぞれ式 (4.6), (2.2), (4.7) で表される。softmax の出力ベクトル y は、入力が各クラスに属する確率を表しており、 y_i は i 番目のクラスである確率を表す。 y は予測値、 t は正解の値、 a は関数への入力を表す。学習はバッチサイズ 32、エポック数 50 で行い、検証用データで最も小さい損失値であったエポックでのモデル（学習した重みパラメータ）を保存し、評価に用いる。

$$E = - \sum_n^N \sum_i^D t_{ni} \log y_i \quad (4.6)$$

$$y_i = \frac{\exp(a_i)}{\sum_j^D \exp(a_j)} \quad (4.7)$$

U-Net と CNN の学習と検出の画像の内訳の流れを図 4.4 に示す。

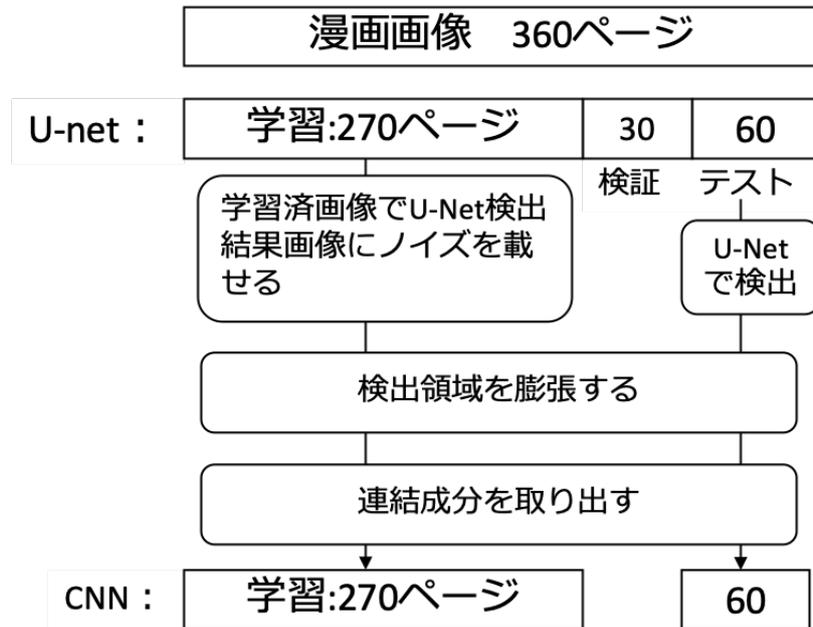


図 4.4: U-Net と CNN の学習と検出の画像の内訳の流れ

4.2 パラメータ選択のための予備実験

学習時のパラメータ選択のための予備実験として、先に述べた 4 分割されたデータのうち 1, 2, 3 のデータを学習に、4 のデータを検証と評価に用いて下記の実験を行った。

4.2.1 データ拡張

学習画像のバリエーションを増やすために、U-Net 学習時に入力画像に対し以下の変換をランダムで行い、どの変換を行うときに最も有効かを調べた。表 4.1 に実験結果を記す。

実験より、水平・垂直反転を行ってデータ拡張した場合の F 値が最も高かった。拡大縮小や回転処理を施した際には、変換により画像サイズが変化し余った部分を黒で埋めたためノイズとして悪影響を及ぼし、水平・垂直反転よりも精度が上がらなかったと考えられる。よって学習時には入力画像にランダムに水平・垂直反転を行うこととする。

4.2.2 前処理のウィンドウサイズ

U-Net の前処理で小画像をスライディングウィンドウ形式で切り出す際のウィンドウサイズを、 256×256 pixel と 512×512 pixel の 2 通りで変更して行い、U-Net によるオノマ

表 4.1: データ拡張実験結果 (%)

| 変換内容 | 適合率 | 再現率 | F 値 |
|---------------------|------|------|-------------|
| 変換なし | 48.8 | 78.6 | 60.2 |
| 水平反転 | 52.9 | 85.0 | 65.2 |
| 垂直反転 | 54.9 | 81.2 | 65.5 |
| 水平・垂直反転 | 54.2 | 89.1 | 67.4 |
| 拡大・縮小 | 42.8 | 85.4 | 57.0 |
| 0-180 度の間でランダムな角度回転 | 47.4 | 88.1 | 61.6 |
| 0-90 度の間でランダムな角度回転 | 52.9 | 87.1 | 65.8 |
| 水平垂直反転+回転 (0-180°) | 44.5 | 82.5 | 57.9 |
| 水平垂直反転+拡大縮小 | 50.4 | 86.3 | 63.6 |

表 4.2: 前処理ウィンドウサイズ変更実験結果 (%)

| ウィンドウサイズ | 適合率 | 再現率 | F 値 |
|----------|------|------|-------------|
| 256×256 | 54.2 | 89.1 | 67.4 |
| 512×512 | 41.0 | 92.5 | 56.8 |

トベ検出の精度を評価する。今回用いる U-Net の構造上、入力画像サイズは 1 辺が 256 の倍数となるためこの値を用いる。表 4.2 に実験結果を記す。

実験より、ウィンドウサイズ 256×256 pixel とした場合、512×512 pixel より F 値が約 10.3% 上回った。よって前処理はウィンドウサイズ 256×256 pixel で行うこととする。

4.3 提案するオノマトペ検出手法性能評価実験

4.3.1 実験概要

各処理の有効性を検証するため、前章で説明した、以下の処理をそれぞれ行った後の検出精度を評価する。

- U-Net で検出
- U-Net で検出 + 膨張
- U-Net で検出 + 膨張 + ノイズ除去

4.3.2 U-net で検出

結果を表 4.3 に示す。検出結果画像を図 4.5, 4.6, 4.7 に示す。一部の欠けやかすれがあるが、3種類の漫画で黒太文字, 白抜き文字, 小さな手書き文字いずれも検出している。しかし一部の絵や活字の誤検出が見られる。検出が不十分だったオノマトペの補完と、誤検出の削減を行う必要がある。

表 4.3: 実験結果 U-Net による検出 (%)

| | 適合率 | 再現率 | F 値 |
|-------------|------|------|------|
| U-Net による検出 | 63.9 | 82.9 | 71.9 |



(a) 入力画像

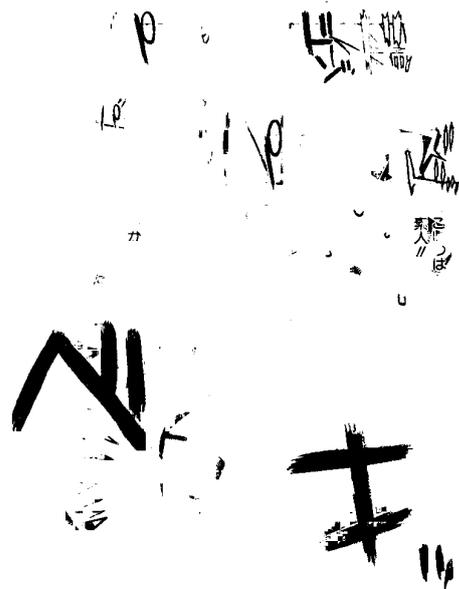


(b) U-Net 検出後

図 4.5: U-Net によるオノマトペ検出結果画像 -4 コマ漫画

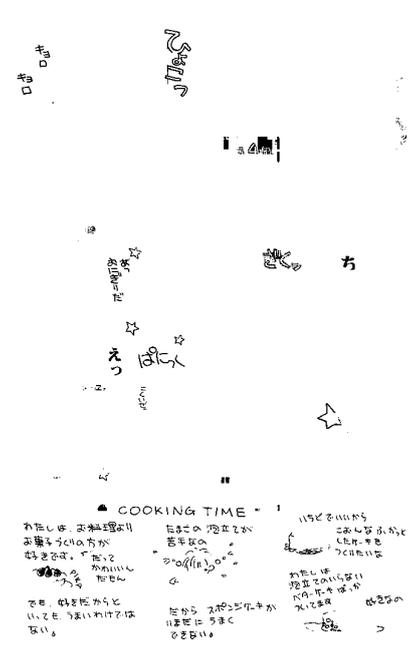


(a) 入力画像



(b) U-Net 検出後

図 4.6: U-Net によるオノマトペ検出結果画像 -バトル漫画



(a) 入力画像

(b) U-Net 検出後

図 4.7: U-Net によるオノマトペ検出結果画像 -恋愛漫画

4.3.3 検出領域の膨張

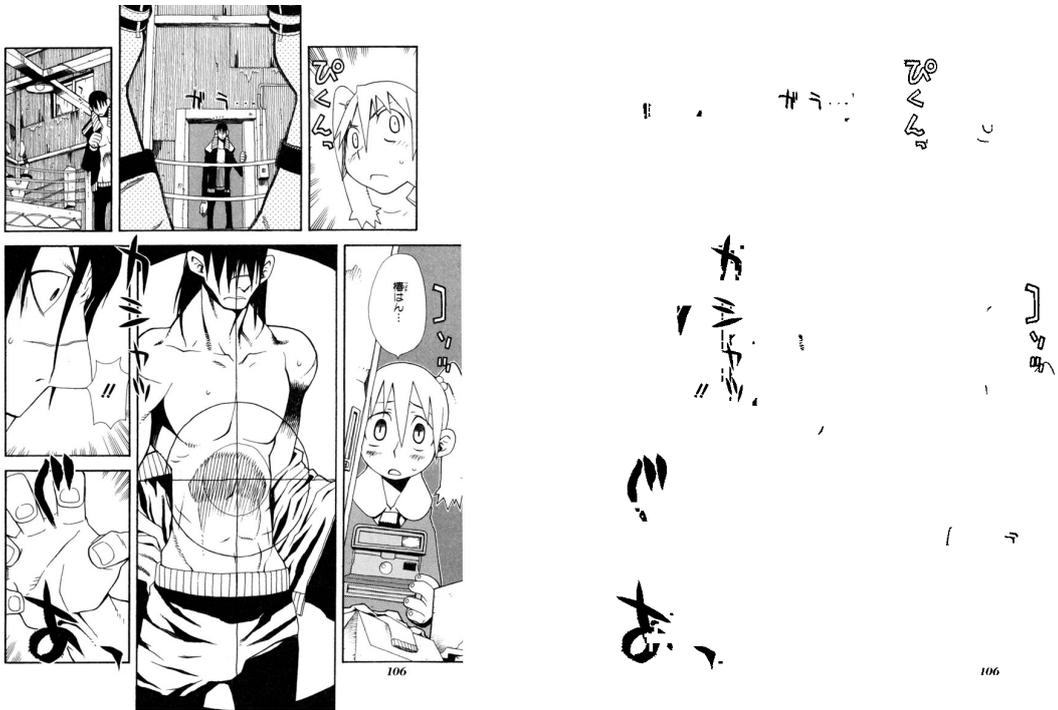
部分的に検出されたオノマトペが存在するため、オノマトペの膨張処理を行う。まず面積が画素数 5 pixel 以下の小さな領域を削除する。膨張処理を行う回数を 1 回から 5 回に変化させ、それぞれの検出精度を評価する。結果を表 4.4 に示す。

表 4.4: 実験結果 領域膨張 (%)

| | 適合率 | 再現率 | F 値 |
|--------------------------|-------------|-------------|-------------|
| U-Net で検出 + 小領域削除 | 64.9 | 82.3 | 72.3 |
| U-Net + 小領域削除 + 膨張 (1 回) | 61.0 | 84.4 | 70.5 |
| U-Net + 小領域削除 + 膨張 (2 回) | 57.5 | 86.0 | 68.7 |
| U-Net + 小領域削除 + 膨張 (3 回) | 55.2 | 86.8 | 67.2 |
| U-Net + 小領域削除 + 膨張 (4 回) | 53.0 | 87.4 | 65.8 |
| U-Net + 小領域削除 + 膨張 (5 回) | 51.1 | 87.9 | 64.4 |

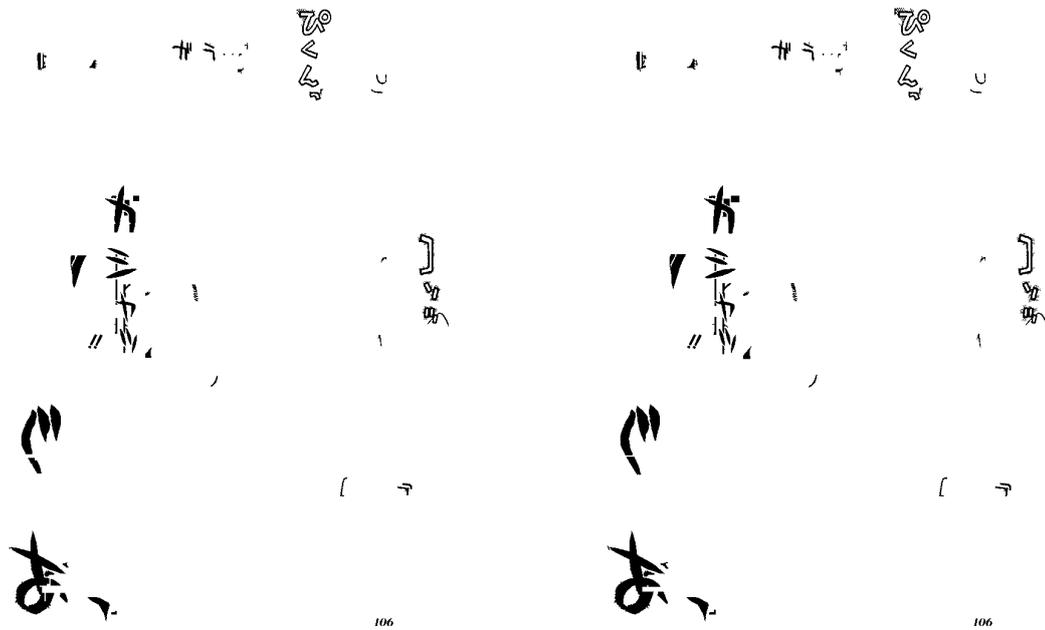
5 pixel 以下の細かなノイズを除去することで適合率が上昇し、非常に小さな手書き文字の一部などが失われることで再現率が低下したものの、U-Net 検出後と比較し F 値が 0.4% 上昇した。膨張処理後は、1 回の膨張処理を行なった場合が最も高い F 値となったが、U-Net 検出後と比較し F 値が 1.4% 低下した。膨張を繰り返すほど、未検出だった領域が検出されていくため再現率が上昇するが、誤検出部分も広がるため適合率が低下し、F 値としては全体で下がってしまう結果となった。しかし 1 回の膨張では部分的に検出されたオノマトペの欠損部分を十分に埋められないことが多い。また、検出できたオノマトペと連結する画素を少しずつ広げていくため、全く検出されなかったオノマトペは膨張処理では補完することができなかった。図 4.8 に膨張処理回数による検出結果の変化の例を示す。画像左下のオノマトペ“ぐあっ”に注目すると、膨張していくことで欠損部分が埋まっていくことがわかる。

膨張処理後、拡大してしまった誤検出領域には、検出すべきオノマトペと連結しているものと独立しているものがある。検出すべきオノマトペとは独立している誤検出連結成分について、連結成分単位で除去できれば適合率の低下を抑えられる。



(a) 入力画像

(b) 1回膨張



(c) 3回膨張

(d) 5回膨張

図 4.8: オノマトペ検出結果画像の膨張処理回数による差

4.3.4 CNN による非オノマトペ消去

検出領域の膨張を行うと誤検出された領域も膨張され適合率が低下した。膨張処理後の画像に対し連結成分ごとにオノマトペか否かを判定し、オノマトペでないとされた領域を削除するノイズ領域削減を行う。結果を表 4.5 に示す。

表 4.5: 実験結果 ノイズ削減 (%)

| | 適合率 | 再現率 | F 値 |
|----------------------------------|-------------|-------------|-------------|
| U-Net + 小領域削除 + 膨張 (1 回) + ノイズ削減 | 67.9 | 80.7 | 73.6 |
| U-Net + 小領域削除 + 膨張 (2 回) + ノイズ削減 | 65.3 | 81.3 | 72.0 |
| U-Net + 小領域削除 + 膨張 (3 回) + ノイズ削減 | 63.7 | 82.3 | 71.7 |
| U-Net + 小領域削除 + 膨張 (4 回) + ノイズ削減 | 60.9 | 81.9 | 69.6 |
| U-Net + 小領域削除 + 膨張 (5 回) + ノイズ削減 | 59.2 | 80.9 | 68.2 |

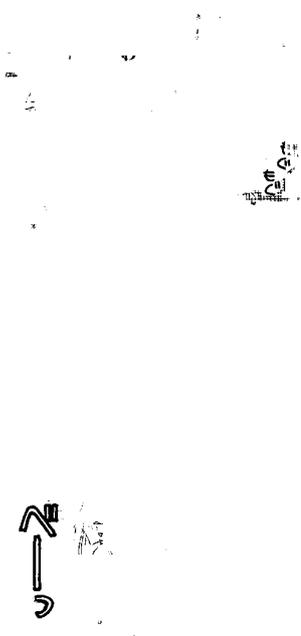
ノイズを削減したことで適合率が上昇したが、誤って削除されてしまったオノマトペも存在するため再現率が低下した。最も高い F 値である 1 回の膨張処理をした画像にノイズ除去を行った場合、適合率の上昇により F 値は U-Net による検出後と比較し 1.7% 上昇、領域膨張後と比較して 3.1% 上昇した。

恋愛漫画の画像に対し、U-Net による検出、5 回の領域膨張、ノイズ除去を順に行った結果画像を図 4.9 に示す。太く書かれた白抜き文字のオノマトペ、黒文字のオノマトペともに概ね検出されている。

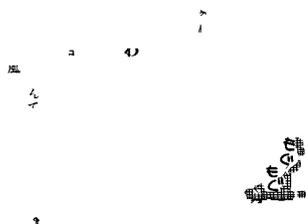
U-Net 検出後 (図 4.9(b)) と膨張後 (図 4.9(c)) を比較すると、一部が欠けていたオノマトペ “もぐもぐ” の欠けていた部分が埋められていることがわかる。しかし、オノマトペ周辺のスクリーントーンが余分に検出されてしまった。また、膨張後 (図 4.9(c)) とノイズ除去後 (図 4.9(d)) を比較すると、“もぐもぐ” の周辺の誤検出が除去されたが、繋がってしまっていたオノマトペも除去されてしまった。



(a) 入力画像



(b) U-Net 検出後



(c) 膨張後



(d) ノイズ除去後

図 4.9: オノマトペ検出結果画像 (5 回膨張)

4コマ漫画の画像に対し，U-Netによる検出，3回の領域膨張，ノイズ除去を順に行った結果画像を図4.10に示す．太く書かれたオノマトペ，小さな手書き文字ともに概ね検出されている．

U-Net検出後(図4.10(b))と膨張後(図4.10(c))を比較すると，一部が欠けていたオノマトペ“カッ”の欠けていた部分が埋められていることがわかる．しかし，オノマトペ周辺の黒い背景が余分に検出され，他のノイズも鮮明に出てしまった．また，膨張後(図4.10(c))とノイズ除去後(図4.10(d))を比較すると，複数のノイズが除去されていることがわかる．しかし，連結成分単位でのノイズ除去であるため，オノマトペと同一連結成分に含まれるノイズは除去できなかった．また検出していたオノマトペの一部が誤って除去されてしまった．オノマトペの文字の一画が誤って消されている場合が多く，近くに存在する複数の連結成分がオノマトペであった場合にはノイズとして除去しないなどの対処が必要である．しかし絵の中に複雑に書かれているオノマトペも多数存在するため，オノマトペと非オノマトペの連結成分が複雑に配置された場合も考慮しなければならない．

第 5 章

結言

5.1 本研究のまとめ

本研究では、深層学習により漫画画像からオノマトペを検出し 2 値画像として出力する手法を提案した。提案手法は以下の手順で構成される。

1. 学習した U-Net を用いてオノマトペ領域を検出する。
2. 検出された領域を膨張する。
3. 検出された領域を連結成分単位で分類しノイズを削減する。

漫画画像データセット Manga109 から 3 種類の漫画作品を用いて提案手法の評価実験を行った。U-Net での検出はオノマトペの形状や大きさを問わず検出でき、検出精度は F 値で 71.9% となった。

U-Net での検出ではオノマトペの部分的な欠けが見られたため、領域の膨張による補完を行なったところ未検出領域は減少したが過検出・誤検出が増大してしまった。そこで CNN によるノイズ除去を行ったところ、一部オノマトペも消去されてしまうが、誤検出を削減できた。提案手法でノイズ除去を行った結果、F 値は最大で 73.6% となった。

5.2 今後の展望

オノマトペのアノテーションとして 2 値画像を作成したが、Manga109 に付属のアノテーションは xml 形式となっている。また 2 値画像のアノテーションには文字情報は含まれていない。今後の研究のために、文字情報を含む Manga109 と同一の形式のオノマトペアノテーションの作成が必要である。また、本研究でアノテーションを付与できた漫画は Manga109 に収録された 109 作品中 5 作品である。オノマトペのアノテーションを増量し、より幅広い漫画での検証が必要である。

提案手法ではオノマトペ検出後のノイズ除去を行ったが、連結成分単位での除去のため検出したオノマトペと同一連結成分に含まれてしまったノイズは除去できない。また、そもそもオノマトペと背景が繋がって書かれている場合も多数ある。このような場合にオノマトペを分離する方法を検討する必要がある。

本研究ではオノマトペの検出を行ったが、翻訳のためには検出後、文字認識を行わなければならない。また、翻訳のためにはひとつのオノマトペを構成する文字群をクラスタリングしなくてはならない。しかし、漫画中のオノマトペの特徴として、図 5.1 の赤枠内のオノマトペのようにひとつのオノマトペを構成する文字が常に隣接しているとは限らない点がある。さらに、オノマトペは同一の漫画内でも様々な大きさ、形状をとる。そのため通常の文字の認識とは異なる方法を検討しなければならない。



図 5.1: 構成文字が離れているオノマトペの例

付録 A

研究で用いたデータの参照場所, 実行環境

A.1 プログラム, 実験データ

本研究に関するプログラム, データをヒューマンインタフェース研究室サーバー内の以下のディレクトリに置く.

```
/net/xserve0/users/nanami/nanami_M/Master
```

A.2 実行環境

本研究の実験は GPU を用いて以下の環境で行った.

- Ubuntu 16.04
- Python 3.7.3
- Tensor Flow 1.14.0
- Keras 2.2.4

付録 B

発表資料

修士論文発表で用いたプレゼンテーション資料を以下に掲載する。

**深層学習を用いた
漫画中のオノマトペの検出**
Extraction of Onomatopoeia from Comic Images
Using DNN

三重大学 工学研究科 情報工学専攻
ヒューマンインターフェース研究室
加藤ななみ

オノマトペとは

オノマトペとは

▶擬態語、擬声語を表す言葉

漫画のオノマトペ

▶手書き文字

▶聴覚に訴え、読者をひきつける効果



背景

日本の漫画は世界中で読まれている.[1]

→漫画翻訳の需要の増加

スマホやタブレットにより

電子書籍が普及している.[2]

→電子コミックの増加



自動でオノマトペを検出し
手間・費用を削減



[1]内閣府文化庁文化政策課文化政策室「クールジャパン戦略」について、2013
[2]読者サービスセンター「電子書籍の普及状況」について、2016
[3]読者サービスセンター「電子書籍の普及状況」について、2016

研究目的

漫画画像中のオノマトペの検出

・形状を限定せずオノマトペを検出する。



先行研究と問題点

石橋ら[3]:
マンガ画像における手書きオノマトペ検出
 >太黒文字, 平均的な白抜き文字のみ検出可

劉ら[4]:
漫画画像からの手書き擬音文字抽出
 >太黒文字のみ検出可



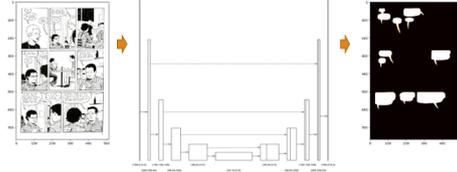
形状を限定しないオノマトペの検出

[3] 石橋 隆典, 渡辺 誠, “マンガ画像における手書きオノマトペ検出”, 京大大学院学術論文, 2015.
 [4] 劉 宇, 李 宇, 王 宇, 李 宇, 李 宇, “漫画画像からの手書き擬音文字抽出”, 機械学習 & データ分析, 25(4), 2014.

先行研究 吹き出し

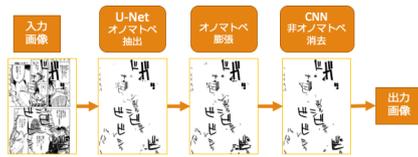
U-Net セグメンテーションCNN

Deep CNN-Based Speech Balloon Detection and Segmentation for Comic Books[5]

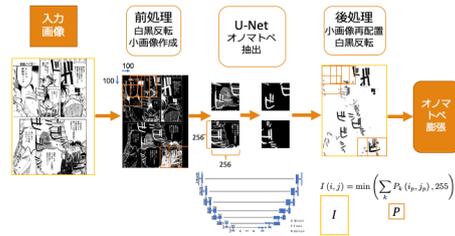


[5] D. DURAY AND J. LAURIOUX, “DEEP CNN-BASED SPEECH BALLOON DETECTION AND SEGMENTATION FOR COMIC BOOKS”, 2019 INTERNATIONAL CONFERENCE ON DOCUMENT ANALYSIS AND RECOGNITION (ICDAR), SYDNEY, AUSTRALIA, 2019, PP. 1227-1234.

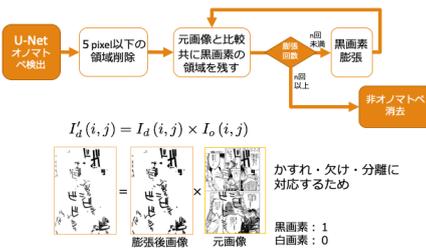
提案手法—概要



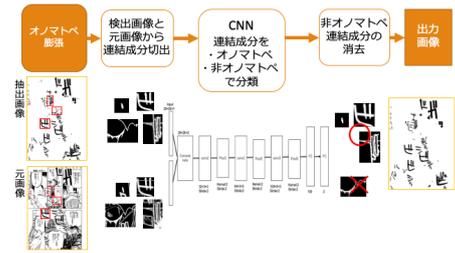
U-Netによるオノマトペ抽出



オノマトペ膨張



CNNによる非オノマトペの消去



データセット 1/3

Manga109[6]

- 東京大学による漫画データセット
- 109冊, 2114ページ分
- 1654×1170(pixel)



[6] MANGA109, <http://www.manga109.org/>

データセット 2/3

- Manga109に含まれる漫画作品に対するオノマトペのアノテーションを作成した。
- 正解データとして用いる。

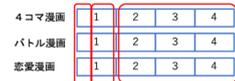
作成したアノテーション

白: 背景 黒: オノマトペ



データセット 3/3

Manga109から3作品
各120ページ全360ページ用いる。
各漫画均等に4分割し、



Validation 30ページ
Test 60ページ
Training 270ページ

4分割交差検証を行い4回の平均で評価する。



実験—評価方法

画素単位の適合率と再現率, F値を求める。

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F\text{-measure} = \frac{2Recall \cdot Precision}{Recall + Precision}$$

| | | | |
|------|--------|-------|--------|
| | | 真の結果 | |
| | | オノマトベ | 非オノマトベ |
| 予測結果 | オノマトベ | TP | FP |
| | 非オノマトベ | FN | TN |

実験結果

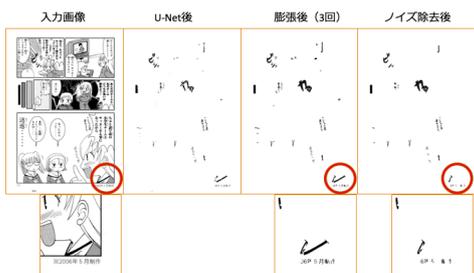
各処理実行後の適合率,再現率を求めF値を算出

| | 膨張回数 | 適合率 | 再現率 | F値 |
|-----------|------|-------|-------|-------|
| U-Net抽出後 | | 63.9% | 82.9% | 71.9% |
| オノマトベ膨張後 | 1回 | 61.0% | 84.4% | 70.5% |
| | 3回 | 55.2% | 86.8% | 67.2% |
| 非オノマトベ消去後 | 1回 | 67.9% | 80.7% | 73.6% |
| | 3回 | 63.7% | 82.3% | 71.7% |
| | 5回 | 59.2% | 80.9% | 68.2% |

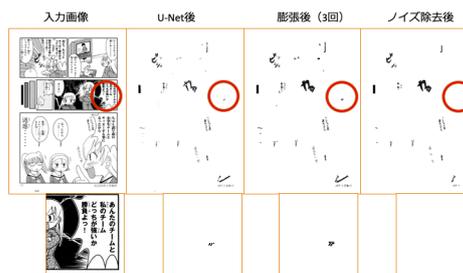
実験結果例



実験結果例



実験結果例



まとめ

- まとめ
 - ・漫画画像から形状を問わずオノマトベを検出する手法を提案した。
 - ・実験の結果, 約70%のF値で検出できた。
- 今後の展望
 - ・オノマトベと繋がったノイズの分離
 - ・検出後に文字認識を行う。
 - ・アノテーション増量

謝辞

研究を進めるにあたり，多くの助言をいただきました若林哲史教授，盛田健人助教，白井伸宙助教，大山航埼玉工業大学教授，毎回のディスカッションで助言をいただいた三宅康二名誉教授に深く感謝いたします。さらに，ともに切磋琢磨した同期の皆さまにも感謝いたします。最後に，6年間大学生活をいつも支えてくれた両親，家族に今一度深く感謝の意を表し，本論文の締めといたします。

参考文献

- [1] インプレス総合研究所, “電子書籍ビジネス調査報告書 2019” , 2019
- [2] 全国出版協会, “出版月報 2020 年 1 月号”, 2020
- [3] 松下光範, 今岡夏海, “デジタルコミック制作のための動的な音喩表現生成システム”, 人工知能学会, 1C1-OS4a-3, 2011.
- [4] 佐藤剣太, 中村聡史, 鈴木正明, “電子コミックの表現を豊かにする手書き文字アニメーション生成手法”, 第 30 回人工知能学会全国大会 (JSAI2016), vol. 4L4-4in2, p. 1-4, 2016. MI2009-34), pp.187-192, May 2009.
- [5] ふきだしズーム, <https://japan.googleblog.com/2017/01/google-play.html>
- [6] Mantra, <https://mntr.jp/index.html> 参照 Jan. 29, 2020.
- [7] 橋本直樹, 佐藤貴明, 澤野弘明, 鈴木裕利, 堀田政二, “漫画のコマ画像からの手書きオノマトペの抽出とその映像効果付与手法の提案”, 情報処理学会第 80 回全国大会, 2018.
- [8] 劉向萍, 東海林健二, 森博志, 外山史, 宮道壽一, “漫画画像からの手書き擬音文字抽出”, 映像情報メディア学会, 25-4, 2012.
- [9] 石橋延高, 渡辺裕, “マンガ画像における手書きオノマトペ抽出”, 早稲田大学卒業論文, 2015.
- [10] Manga109, <http://www.manga109.org/>
- [11] 内閣府知的財産戦略推進事務局 クールジャパン戦略, “クールジャパン戦略について”, 2019. https://www.cao.go.jp/cool_japan/about/about.html 参照 Jan. 29, 2020.
- [12] 佐々木茂, 福原明彦, 田中誠一, 山脇奈緒美, “HTML5 の canvas 要素を用いた モーションコミックコンテンツの開発”, 第 78 回全国大会講演論文集 2016(1), 635-636, 2016.
- [13] 石井大祐, “高機能なマンガコンテンツのための画像解析手法に関する研究”, 早稲田大学大学院, 博士論文, 2014.
- [14] H. Yanagisawa, T. Yamashita and H. Watanabe, ”A study on object detection method from manga images using CNN,” 2018 International Workshop on Advanced Image

- Technology (IWAIT), Chiang Mai, 2018, pp. 1-4.
- [15] 石井大祐, 河村圭, 渡辺裕, “コミック画像のコマ分割処理における制御パラメータに関する検討”, 電子情報通信学会パターン認識・メディア理解研究会 PRMU2009-34.
- [16] 田中孝昌, 外山史, 宮道壽一, 東海林健二, 漫画画像の吹き出し検出と分類, 映像情報メディア学会誌, vol.64, No.12, 2010.
- [17] D. Dubray and J. Laubrock, “Deep CNN-Based Speech Balloon Detection and Segmentation for Comic Books,” 2019 International Conference on Document Analysis and Recognition (ICDAR), Sydney, Australia, pp. 1237-1243, 2019.
- [18] Jonathan Long, Evan Shelhamer, Trevor Darrell, “Fully Convolutional Networks for Semantic Segmentation”, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015.
- [19] Olaf Ronneberger, Philipp Fischer, Thomas Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation”, MICCAI (Medical Image Computing and Computer-Assisted Intervention), 2015.
- [20] 新居さとし, あくはむ, 第1巻, 講談社, 2005.
- [21] 猪原大介, 学園ノイズ, 第1巻, 東京三世社, 2003.
- [22] くりた陸, ゆめ色クッキング, 第1巻, 講談社, 1988.
- [23] Augereau, Olivier Augereau, Motoi Iwata, Koichi Kise, “A survey of comics research in computer science”,