

# ユーズド衣料品の再利用促進のための分類システムの構築

## A Construction of Classification System for Encouragement to Recycle Used Clothes.

奥村紀之<sup>\*1</sup>  
Noriyuki OKUMURA

水沼千枝<sup>\*1</sup>  
Chie MIZUNUMA

田中駿<sup>\*2</sup>  
Shun TANAKA

浅谷晋作<sup>\*2</sup>  
Shinsaku ASATANI

上泰<sup>\*3</sup>  
Yasushi KAMI

<sup>\*1</sup>神戸松蔭女子学院大学  
Kobe Shoin Women's University

<sup>\*2</sup>明石工業高等専門学校  
National Institute of Technology, Akashi College

<sup>\*3</sup>広島工業大学  
Hiroshima Institute of Technology

The fashion industry is often cited as the world's second-largest polluting sector, necessitating the development of systems for clothing reuse. This study aims to construct a system capable of determining the reusability of clothing items based on images. The judgment system utilizes classification models from machine learning, with parameters set through 5-fold cross-validation. Experimental results demonstrated that a model fine-tuned using Vision Transformer (ViT) achieved an F1 score of 0.98. Consequently, this establishes a method for straightforwardly determining clothing reusability from pictures.

### 1. はじめに

ファッション業界は世界第2位の汚染産業と言われる<sup>\*1</sup> <sup>\*2</sup>ほどに、深刻な環境汚染を引き起こしているとされている。そのため、SDGsの観点からも、ファッション業界において再資源化して活用するリサイクルだけではなく、むしろそのまま再活用するリユースの促進が重要な取り組みであるという認識が広がっている。リサイクルにおける再資源化では、大量の水を使い、CO<sub>2</sub>の排出量も増加も大きな問題となるため、利用可能な衣料品は衣料品のまま循環させる仕組みが重要となる。

本研究では、衣料品のリユース促進のための取り組みとして、個々人が所持している衣料品がリユースに適しているものかそうでないものかを写真から判定するシステムを構築し、家庭内で廃棄されてしまう衣料品をリユース品として提供しやすい環境を提供することを目指す。衣料品のリユースに関して身近なものを考えると、リサイクルショップのようなものを想定する。しかし、リサイクルショップに持ち込む前段階として、そもそも手持ちの衣料品が買取可能なものであるかの判断が付かず、まとめて廃棄されてしまっているという現状もある。

別の観点からも考えてみよう。リユース可能な衣料品を循環させる仕組みは、日常生活において様々なファッションを簡単に楽しむことができるという点においても有益であることは分かる。しかし、2024年1月1日に発生した能登半島地震や2011年3月11日に発生した東日本大震災<sup>\*3</sup>、1995年1月17日に発生した阪神・淡路大震災など、大規模な災害においても、衣料品を提供する仕組みは非常に重要となる。一方で、たとえ善意からであったとしても、提供される衣料品には、およそ利用には適さないもの（破れている、汚れている、臭いがあるなど）も多く含まれており、被災地において選別しなければならないという問題が発生している<sup>\*4</sup>。同時に、利用不可

能な衣料品を廃棄しなければならない問題も発生する。

そこで、本研究では、衣料品を撮影した画像から簡易にリユースの可否を判定するための分類器（衣料品分類システム）を構築し、各家庭、事業所などで廃棄される衣料品を削減し、リユースを促進するための仕組みを構築する。衣料品分類システムに利用するデータセットはリサイクル業者と共同で作成している[水沼 24]。分類器には、ベースラインとしてSVM (Support Vector Machine) による分類を実施し、F1スコアで0.85となった。また、TransformerをベースとしたViT (Vision Transformer)、畳み込みモデルであるVGGNet、GoogleNetをそれぞれにファインチューニングを行い、ViTモデルにおいてF1スコアで0.98となった。

### 2. 衣料品リユースに関する取り組み

衣料品リユースの必要性として、衣料品大量廃棄問題に関する論文を紹介する。また、衣料品を循環させる取り組みとしては、一般的にはリサイクルショップへの持ち込み、または衣料品専用の回収ボックスへの投函となる。そのため、リユース促進のための機械学習モデルというものはあまり例がなく、企業のとしての取り組み事例としてZOZOTOWNの取り組みを紹介する。

#### 2.1 衣料品の大量廃棄問題

北村の報告では、衣料品は傷みや汚れによって着られなくなったもの以外にも、流行の影響によって、時代遅れと感じてしまうことでまだまだ着られる服でも大量に廃棄されてしまう問題を取り上げている[北村 23]。ファストファッションの台頭や売れ残りの問題、需要と共有のバランスの崩壊などから衣料品大量廃棄の問題について述べており、本研究で対象としているリユース促進の根幹をなす報告となっている。特に、売れ残りに関しては新品であることから、適切な循環モデルの構築が必要となると考えられる。

#### 2.2 ZOZOTOWNでの取り組み

ファッション関連でのECサイトとして有名なZOZOTOWNでは、過去にZOZOTOWNを通じて購入した衣料品をAIによって判定し、再販価格を予測するモデルを構築し運用してい

連絡先: 奥村紀之, 神戸松蔭女子学院大学, 神戸市灘区篠原伯母野山町1-2-1, 078-882-6122, noriyuki@shoin.ac.jp

<sup>\*1</sup> [https://www.unic.or.jp/news\\_press/features\\_backgrounders/32952/](https://www.unic.or.jp/news_press/features_backgrounders/32952/)

<sup>\*2</sup> <https://unctad.org/news/sustainable-african-design-focus-womens-day>

<sup>\*3</sup> <https://pgd.tohoku.ac.jp/lgs/gsafty/common/download/Student-activity-report-of-self-planning-2018.2.pdf>

<sup>\*4</sup> 内閣府阪神・淡路大震災教訓情報資料集: [https://www.bousai.go.jp/kyoiku/kyokun/hanshin\\_awaji/data/detail/2-2-](https://www.bousai.go.jp/kyoiku/kyokun/hanshin_awaji/data/detail/2-2-1.html)

1.html

る\*5\*6。自社で取り扱っている商品を循環させる仕組みとして成功しており、AIを活用することで買取価格が3割ほど向上したとの報告があり、ユーザにとっても利用しやすいシステムとなっている。

### 3. リユース可否の判定システム

衣料品のリセール（再販）可否を判定するサービスは、一部の企業が提供している。しかし、可否を判定するための学習データ（衣料品画像）は非公開であるため、簡単に再現することが難しい。したがって、リユース可否の判定システムを構築するためには、1) データセットの構築、2) 分類器の構築が必要となる。以下、データセットの準備と前処理、および、分類器の構成について述べる。

#### 3.1 データセットの準備

衣料品の分類を行うためのアノテーション済みオープンデータは著者が調査した範囲では見当たらず、まずデータセットを構築する必要がある。データセットの構築には、リサイクル業者である株式会社 Kurokawa、およびリサイクル工房あづまの協力を得て、各所に持ち込まれる衣料品を仕分けるためのマニュアルに即してアノテーションを実施している。特に、利用可能か不可かの判断が曖昧になるような部分については、店頭に並べるとすればという観点から再評価を実施し、リユース可否の判定と統合して評価している。詳細は文献[水沼 24]にて報告している。

本研究においては、衣料品画像の収集が必須であるため、可能な限り背景の影響を抑制するために、ベニヤ板などの素材に衣料品を上げ、撮影を実施している(図 3.1)。このように広げた状態の衣料品の表と裏を複数のカメラで撮影している。スマートフォンのカメラを利用することが一般的と考えられる。しかし、メーカーや世代の違いによりカメラ性能も異なるため、本研究では、デジタル一眼レフカメラ、複数の世代の iPhone、複数メーカーの Android 端末を使用することで、カメラ性能に依存しない分類器の構築を目標としている。

このようにして撮影した画像は延べ 14,105 枚である。



図 1: 撮影した衣料品画像の例

#### 3.2 画像データの前処理

図 3.1 に示したとおり、撮影した画像には背景となるベニヤ板や工場内の床などが同時に映り込んでいる。分類器の構築にはこのような背景は不要であるため、Python 実装である

rembg\*7 を用いて全ての画像に対し背景除去を行った。図 3.2 は図 3.1 に対し rembg による背景除去を行った例である。



図 2: rembg による背景除去を行った例

全ての画像に対し、確実に背景が除去されることを期待するが、実際には図 3.2 に示すような抽出ミスが発生する。上の例では、T シャツの背景を除去しようとしたところ、T シャツにプリントされているロゴのみが抽出されてしまっている例を示し、下の例では、背景の一部分のみが残ってしまっている例を示している。



図 3: rembg による背景除去の失敗例

以上のように前処理として背景除去を実施した結果、14,105 枚の画像のうち、11,231 枚が適切に除去されている。本節以降の実験では、この 11,231 枚の画像を利用している。なお、本研究では、再販可能かどうかという点ではなく、再販可能である含めリユース可能であるかどうかの 2 値分類を行う。11,231 枚の画像のうち、再販可能なもの、リユース可能なものを合わせて 6,185 枚、資源化対象となるものが 5,046 枚であり、ほぼ均等な割合でリユース可否の画像が準備されている。画像サイズは  $224 \times 224$  または  $384 \times 384$  の正方形へリサイズして利用している。この際、縦横比の維持は考慮していない。

#### 3.3 分類器の構築

リユースの可否を分類するための分類器として、SVM (Support Vector Machine) をベースラインとして学習させる。これに対し、Vision Transformer のモデルである ViT、畳み込みモデルである GoogleNet、VGGNet をそれぞれファインチューニングしたモデルを学習させ比較検討する。

##### 3.3.1 グリッドサーチによる SVM の学習

SVM (Support Vector Machine) は古典的分類モデルの一つであり、線形分類を行うモデルである。しかし、線形分離が

\*5 <https://digital-shift.jp/ai/230216#item65160>

\*6 <https://netshop.impress.co.jp/node/8150>

\*7 <https://github.com/danielgatis/rembg>

困難であるデータに対しては、カーネル関数を利用し、超空間での分離を行うことで分離平面を構築することが可能である [Cortes 95].

本研究では、カーネル関数として線形カーネル、多項式カーネル、RBF カーネル、シグモイド関数を検証している。各種パラメータの詳細を表 1 に示す。これらのパラメータに対し、5 分割交差検定によるグリッドサーチを利用し最適な組み合わせを抽出している。

表 1: グリッドサーチに用いたパラメータ

| カーネル    | C               | gamma         | degree  |
|---------|-----------------|---------------|---------|
| linear  | 1, 10           |               |         |
| poly    | 1,10, 100, 1000 | 0.001, 0.0001 | 2, 3, 4 |
| rbf     | 1,10, 100, 1000 | 0.001, 0.0001 |         |
| sigmoid | 1,10, 100, 1000 | 0.001, 0.0001 |         |

### 3.3.2 ViT, VGGNet, GoogleNet の学習済みモデルのファインチューニング

近年の画像処理においては、言語処理（機械翻訳）の流れで開発された Transformer を画像に適用した Vision Transformer (ViT) が広く活用されている [Dosovitskiy 21]. また、ViT が提案されるまで主流であった CNN（畳み込みニューラルネットワーク）に基づくモデルとして、GoogleNet [Szegedy 15], VGGNet [Simonyan 15] などが提案されている。これらのモデルをファインチューニングすることにより、本研究で目標としているリユースの可否判定（2 値分類）を実現する。

上述の 3 種類のモデルを利用した理由として、手持ちの開発環境で現実的な時間で学習（ファインチューニング）が完了するモデルであることが挙げられる。開発環境は、Ubuntu 22.04.3 LTS, CPU は AMD Ryzen 7 5700g, 学習用の GPU として GeForce RTX 3060 が 2 枚である。表 2 に実験に利用した学習済みモデルを記載する。

表 2: 実験に利用した ViT, VGGNet, GoogleNet のモデル

| ViT                   | VGGNet   | GoogleNet    |
|-----------------------|----------|--------------|
| vit_tiny_patch16_224  | vgg11_bn | inception_v1 |
| vit_tiny_patch16_384  | vgg13_bn | inception_v3 |
| vit_small_patch16_224 | vgg16    | inception_v4 |
| vit_small_patch16_384 | vgg16_bn |              |
| vit_base_patch16_224  | vgg19_bn |              |
| vit_base_patch16_384  |          |              |
| vit_large_patch16_224 |          |              |

ファインチューニングにおいては、エポック数を 10 としている。それぞれのモデルに対し 5 分割交差検定を行い、F1 スコアの平均により評価する。

### 3.4 評価方法

本研究では、構築したモデルの評価指標として F1 スコアを採用している（式 1）。Precision は精度、Recall は再現率を示しており、F1 は精度と再現率の調和平均となっている。5 分割交差検定による評価を実施しているため、各モデルの評価は 5 つの評価値の平均としている。

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (1)$$

また、ユーズド衣料品のリユース可否という評価は非常に曖昧であり、境界が明確でないことが多い。そのため、ROC 曲線（Receiver Operating Characteristic curve）における AUC（Area Under Curve）を算出し、F1 スコアとは別の観点から分類性能を評価している。AUC は 0.5 の時が最も性能が悪く、1.0 の時に最高性能となる評価指標である。

## 4. 実験結果

本節では分類実験の結果を示す。なお、実験結果についてはそれぞれのモデルで F1 スコアが最高になったもののみを示している。

### 4.1 SVM のグリッドサーチの結果

SVM のグリッドサーチでは、C を 10, gamma を 0.0001, カーネルを RBF としたときに最も良い評価値を得た。この際、平均の F1 スコアは 0.85, 平均の AUC は 0.89 となっている。この評価値をベースラインとする。

### 4.2 ファインチューニングの結果

各ニューラルネットワークのモデルをファインチューニングした結果を、前節で述べた SVM の評価結果をベースラインとして表 3 に示す。

表 3: ファインチューニングの結果

| モデル       | 最高性能         | F1 スコア      | AUC         |
|-----------|--------------|-------------|-------------|
| SVM       |              | 0.85        | 0.89        |
| ViT       | base_16_384  | <b>0.98</b> | <b>0.98</b> |
| VGGNet    | vgg16_bn     | 0.92        | 0.91        |
| GoogleNet | inception_v1 | 0.93        | 0.93        |

## 5. 考察

前節の結果から、Vision Transformer が最も良好な評価結果となっていることが分かる。紙面の都合上全ての評価結果を示していないが、いずれのモデルにおいても 10 エポックのファインチューニングを行った後の F1 スコアは 0.2 程度の差であり、いずれのモデルもベースラインの SVM の評価よりも良い結果となっている。

図 5. に示す画像は、SVM では分類に失敗しているが、ViT では適切に分類できている例である。

図 5. のボトムスは、リユース可能という人間の判断であるが、SVM ではリユース不可と判定されている。その理由として考えられることは、膝や股の部分に光沢が見られ、この部分が色あせと判定されてしまっている可能性である。撮影時のカメラの角度、光の当たり方によって、色あせと光沢の違いが判定しづらくなっている。ViT の場合は、こういった画像については適切に判定することができており、光の成分による影響を回避できているものと考えられる。

SVM が正解しており、ViT では適切に判定できなかった画像は見られなかったため、SVM でなければ判定できない衣料品は本実験においては存在しなかった。したがって、リユース可否の判定には、ViT をファインチューニングしたモデルが最適であると判断される。GoogleNet や VGGNet も SVM と比較すれば良好な結果となっているが、Transformer を適用した ViT が従来の CNN モデルよりも分類に適していると言える。

また、本稿では AUC による評価も行っている。AUC による評価は、分類が曖昧な境界線付近における分類性能を表現す



図 4: SVM では失敗するが, ViT で適切に分類される衣料品

る上で重要な評価値であり, 1 に近づくほど良い分類器が構築できていると言える. 本実験においてはベースラインである SVM における AUC が 0.89 であることにに対し, F1 スコアで最も性能が良かった ViT のモデルにおいて 0.98 であることから, ViT のモデルの分類性能が極めて高いということが示されている.

以上のことから, 本研究における分類性能は ViT のモデルが最も良好であり, 実用に耐えうる性能を有していると言える.

## 6. おわりに

本稿では, ユーズド衣料品のリユース可否を判定するための分類モデルを構築した. Vision Transformer, GoogleNet, VGGNet の比較実験を行い, Vision Transformer をファインチューニングしたモデルが最高性能となることを示した. また, ベースラインとして実装した SVM に対し, SVM では分類できなかった衣料品画像を ViT では適切に分類できていることを示し, 従来の機械学習手法よりも Vision Transformer の方が良好な結果が得られることを確認した.

今後の課題として, 衣料品画像の前処理への対応が挙げられる. 分類器の学習のため, 衣料品画像以外の背景画像を除去しているが, 適切に除去できていない画像が 3,000 枚程度存在している. これらの画像は学習には使用されていないため, より適切な背景除去手法を検討する必要がある.

一方で, 本研究はアーバンイノベーション神戸の助成により実施しており, その目的は, 家庭から廃棄されてしまう衣料品をリユースできる環境を構築することにある. そのための社会実装として Web サービス等を検討しているが, 背景除去の前処理がうまく実行できなかった場合は, ユーザに写真の撮り直しを依頼すれば十分であるという見方もある. そのため, 適切に背景除去ができていない 3,000 枚の画像を活用し, 前処理が成功しているかいないかを判定する分類モデルを作成することでも対応が可能であると考えている.

## 謝辞

本研究は, 2023 年大学発アーバンイノベーション神戸 (物価高騰等対策・募集テーマ指定型) において採択された「衣料品リユース活性化を支援する AI システムの開発と社会応用に関する研究」の助成により実施した<sup>\*8</sup>. また, 本研究の実施に際し, 株式会社 Kurokawa<sup>\*9</sup>, リサイクル工房あづま<sup>\*10</sup> の専門家からのコメントを基にデータセットの構築を行った. ここに感謝申し上げる.

## 参考文献

- [水沼 24] 水沼 千枝, 池谷 知子, 黒崎 優美, 野手 紀花, 奥村 紀之, 上 泰, 花田 美和子: ユーズド衣料品分類のための画像データセットの構築 (2024), 2024 年度 人工知能学会全国大会 (第 38 回)
- [北村 23] 北村 真琴: ファッションの大量廃棄問題とマス・カスタマイゼーション研究の課題 (2023), 東京経大会誌 (経営学), 東京経済大学経営学会, 318 巻, pp.137-154
- [Cortes 95] Corinna Cortes, Vladimir Vapnik: Support-Vector networks(1995), Machine Learning, Springer, Vol.20, pp.273-297.
- [Dosovitskiy 21] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, Neil Houlsby: An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale(2021), International Conference on Learning Representations.
- [Szegedy 15] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, Andrew Rabinovich: Going deeper with convolutions(2015), IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Vol.1, No.1.
- [Simonyan 15] Karen Simonyan, Andrew Zisserman: Very Deep Convolutional Networks for Large-Scale Image Recognition(2015), International Conference on Learning Representations.

\*8 [https://www.city.kobe.lg.jp/a05822/daigakurenkei/uik\\_bukkakoutou.html#syuin](https://www.city.kobe.lg.jp/a05822/daigakurenkei/uik_bukkakoutou.html#syuin)

\*9 <http://www.kurokawa1953.com/>

\*10 <https://www.city.kobe.lg.jp/a04164/kurashi/recycle/gomi/shisetsu/recyclingkobo/azuma.html>