

衣服に適したスニーカー画像を深層学習を用いて 推奨するシステムの提案

2018SE002 伴和真

指導教員：沢田篤史

1 はじめに

近年、個人の好みに合った情報や商品を提案する推奨システムが普及してきた。衣服や靴などのファッション製品を販売するサイトも一般化しており、多くの人々が利用している。EC サイトやアプリケーションの普及に伴い、大量の商品がインターネットを通じて販売されている。

衣服や履物、アクセサリなど身に着ける商品の組み合わせを推奨することは難しい問題である。ファッションは個人の感性に依存することで正解が一つでないことが難しい要因である。また、膨大な商品がある中でコーディネートと相性の良いアイテムを探すことは容易でなく、見落とししてしまうことが多い。

特にスニーカーは種類が豊富で、他の商品と適したものを推奨することが難しい。背景として、スニーカー市場の成長やスポーツブランドのウェア・シューズをコーディネートに組み込むスタイルの流行などが要因の一つである [1]。

本研究の目的は、ユーザのコーディネート画像と相性の良いスニーカー画像を推奨するシステムの提案である。推奨システムは、ユーザのコーディネート画像を入力として、そのコーディネートと相性の良いスニーカー画像を出力する。

提案する推奨システムは Python で実装した。コーディネート画像と推奨するスニーカー画像の特徴に注目し、画像入力のみで推奨するために深層学習を用いた。画像の特徴を捉えることに優れた CNN (Convolutional Neural Network) を用いる。必要なデータはスクレイピングを行い収集した。

実装した推奨システムと既存のシステムとの比較を行い、妥当性や優位性、問題点について考察した。

2 ファッションコーディネート推奨システムとその課題

衣服や履物、アクセサリなど身に着ける商品の組み合わせを推奨することは個人の好みとの相性や、ファッションに正解がないなど難しい問題がある。また、膨大な商品が存在する中で自分好みの商品を探し出すことは難しい。

堀ら [2] は、過去の研究 [3] をもとにこの問題を解決するためにコーディネートの重要な特徴の一つであるカテゴリを判定しシステムの構築・評価を行っている。服の画像とそのカテゴリを取得し、CNN を用いてカテゴリを分類している。誤認識を防ぐためにカテゴリの階層性を用いて分類し、精度が向上することを確認している。カテゴリに注目しているため、本研究では色に注目する。

山口 [4] は、これまで取り組んできた事例を含め研究課題を解説し、今後の展望についてまとめている。コーディネートに対して良い悪いを判断することは難しいが、二枚の画像の比較で判断がつくことがあると述べている。この判別方法を発展させるとシステムの精度向上につながる。

3 深層学習を用いたスニーカー推奨システム

入力したコーディネート画像に対して相性の良いスニーカー画像を推奨するシステムの仕組みについて述べる。

特徴量はコーディネート画像から衣服領域の色を抽出し特徴量とする。ファッションに使われる白、黒、赤、青、黄の五色に分類されたラベルを作り、ラベルごとに応じたスニーカー画像を紐づけする。

3.1 システムの概要

図 1 にシステムの構成を示す。図下部の入力をコーディネート画像とし、出力をコーディネートと相性の良いスニーカー画像とする。図上部のコーディネート画像、スニーカー画像のデータセットを画像整形し、CNN を用いて学習させた。学習させたデータをモデルとし、コーディネート画像を入力することにより相性の良いスニーカー画像が出力される。推奨するスニーカー画像の決定は Web サイト上のコーディネート画像を正解としてコーディネート画像に写るスニーカーを選び、色ラベルごとに 40 種類ずつ用意した。入力されたコーディネート画像と正解データを比較し、判別された色の中からランダムでスニーカーを出力する。

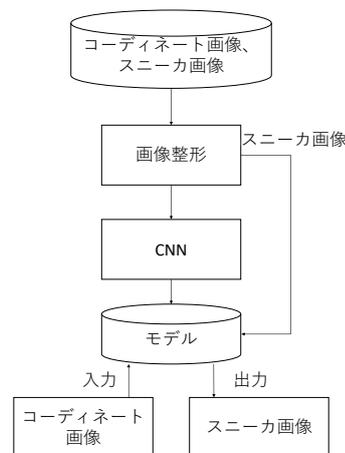


図 1 システムの構成

3.2 学習に用いるデータ

学習に用いるデータは、スクレイピングツール Octoparse^{*1}を用いて Web サイトよりコーディネート画像 4243 枚、推奨するスニーカー画像を 200 枚取得した。取得したデータセットの中には、全身のコーディネート画像や上半身下半身のみの画像があるので画像整形、画像選別を行った。取得したコーディネート画像から学習に使用できる画像を白、黒、赤、青、黄に選別した。

3.3 分類手法

推奨システムは、決定的なアルゴリズムに基づいて実現することが難しいことから深層学習を用いる。分類方法には画像の特徴を捉えられる CNN を用いることが適していると考えた。コーディネートが持つ色に注目し、その中でも主に使用される白、黒、赤、青、黄の五種類に分類する。収集した画像数が少ないことから学習済みモデルである VGG-16 モデルを使用した。

3.4 画像整形

画像内の背景と衣服を識別するために、衣服領域の抽出を行う必要がある。背景処理には、背景処理ツール ClippingMagic^{*2}を用いた。衣服などの必要となる要素以外を背景とみなし背景を白色に統一し、VGG-16 モデルの入力サイズに合わせ画像サイズを 224 × 224 に統一した少ない画像はデータ拡張の一つである反転を用いて画像を増やした。

4 スニーカー推奨システムの実装と評価

既存の画像識別システムと作成したシステムの評価を行った。提案したシステムを Python で実装し、コーディネート画像を色ごとの五種類にラベル分けした。ラベル分けした五種類の色に対し相性の良いスニーカーを各ラベルと紐付けた。学習にはコーディネート画像が少ないので学習済みモデルである VGG-16 モデルを使用した。図 2 で示されたように学習は epoch 数 100 で頭打ちとなった。精度は平均 0.90 ほどでテスト画像 10 枚のうち安定して 8, 9 枚の正解を出した。一方、急激なロス減少もみられた。

5 考察

学習結果は、0.90 ほどまで安定して上がったので色における分類に成功したといえる。学習に用いるコーディネート画像は 1926 枚と少なかったことから、画像データを多く収集し、ラベルの色だけでなく衣服の形やテクスチャなど細分化すること、学習に用いる CNN を適切にチューニングすることでより高い精度、学習時間短縮が見込める。データ拡張は精度向上に寄与することが多く過学習が起こりにくいのでコントラストや明るさを調節することでより正確な結果が得られる可能性がある。

^{*1} <https://www.octoparse.jp/>

^{*2} <https://ja.clippingmagic.com/>

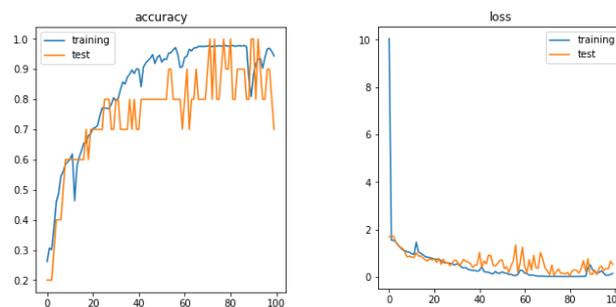


図 2 epoch 数 100 までの学習結果

6 おわりに

近年、個人の好みに合った情報や品物を推奨するシステムが普及している中で、本研究では、コーディネートと相性の良いスニーカーを推奨するシステムを提案した。衣服や履物など身につける商品を推奨することは難しく、商品が多く見落としてしまう問題がある中で、本研究では、ユーザのコーディネート画像と相性の良いスニーカー画像を推奨するシステムの提案をした。ユーザが入力するコーディネート画像に対して相性が良いだけでなく、幅広いスニーカーを用意することでユーザ自身では探すことの難しいスニーカーを提案することを可能なシステムを目指した。

コーディネート画像とスニーカー画像の特徴に着目し、入力画像のみでスニーカーを推奨するために深層学習を用いて Python で実装した。コーディネート画像から五種類の色ラベルに判別し、色ラベルのスニーカー群からランダムに推奨することで単に相性が良いだけでなくユーザが探し出すことの難しいスニーカーを提案することが可能となった。

より精度の高い推奨システムを目指すには、衣服の形やテクスチャなどの特徴にも着目する必要がある。判別に用いる CNN を適切にチューニングする必要がある。データセットやラベルを増やすことで判別も細分化できる。

参考文献

- [1] 株式会社 矢野経済研究所, “スポーツシューズ市場に関する調査を実施 (2020 年)”, https://www.yano.co.jp/press-release/show/press_id/2496, 2020. (Accessed 2021.9.28)
- [2] 堀和紀, 坂原誠, 岡田将吾, 新田克己, “深層学習を用いたファッションコーディネート因子の抽出”, 人工知能学会全国大会論文集, 2016.
- [3] 堀和紀, 岡田将吾, 新田克己, “オンラインファッションカタログを利用した画像とテキストからの組み合わせ推薦”, 人工知能学会全国大会論文集, 2015.
- [4] 山口光太, “Computer Vision meets Fashion”, 映像情報メディア学会誌, Vol. 71, No. 6, pp. 760-763, 2017.