入力画像の特徴に基づいた WordPress テーマの推薦
Recommendation of WordPress Themes Based on Input Image Features

田中　大地
Daichi Tanaka
法政大学情報科学部ディジタルメディア学科
E-mail: daichi.tanaka.4m@stu.hosei.ac.jp

Abstract

There are several ways for beginners to have their own web sites. A widely-used content management system called WordPress allows them to create web sites using their own themes without the specialized knowledge of HTML and CSS. A theme is a collection of web page designs and functions, and there are about 7000 themes on WordPress.org. However, it is difficult for beginners to find suitable themes. This paper proposes an intuitive method for finding a theme by image search and processing. It uses thumbnail images of themes as a database for image search. To simplify text and photograph elements in the thumbnail images, Tesseract OCR and OpenCV are used for character recognition, binarization, and edge extraction. It also processes images in the database using a feature analyzer named MobileNetV3. The processed images are compared with the query image, and the images with the closest Euclidean distances to the query image are output as a search result. As an example of its application, a web application using the proposed method was created. The web application allows users to draw their own images and search for the desired themes close to the drawn images.

1. はじめに

近年，テクノロジーの進歩にともなうインターネットサービスの普及により，ウェブサイトを利用する人や，ビジネスや趣味，教育のために自身のウェブサイトを作成する人が増えている．Facebook，Twitter，Instagram などの SNS の台頭により，個人が自身の日々の記録のためや趣味のためにウェブサイトを持つことは 10 年ほど前と比べると少なくなっている．しかし，依然として人々が人や会社について興味をもった時調べる媒体は Google や Yahoo などの主要な検索エンジンであり，各種 SNS と異なり自分についての情報のみを多く記載できるウェブサイトを個人で持つことの重要性は変わらない．

ウェブサイトを個人で持つために自作する方法の 1 つに，CMS と呼ばれるコンテンツ管理システムを用いて製作する方法があり，代表的なものに WordPress，Joomla，Drupal などがある．特に WordPress は初心者向けかつ豊富なテーマとプラグイン，デザインや機能の拡張が手軽に行える点や，公式フォーラムにサポートやチュートリアルが充実している点など，様々な優れた特徴を持っている．テーマとは，ウェブサイト全体のデザインを決めるテンプレートのことであり，ユーザーがデザインや機能，カスタマイズ性などを考慮して選べる．現在，テーマの数が約 7000 種類存在し，初心者が自分の希望に最も沿うものを選ぶことは難しい．

本研究では，ユーザーのニーズや興味に合わせた WordPress のテーマを容易に見つけられるようにするツールの開発を目的とする．そのために，ユーザーの入力をキーワードではなく画像で受け取る機能を構築する．画像検索はキーワードのみの検索と比べて情報量の多い媒体をベースに検索を行えるため，よりニーズに近いものを得られるメリットがある．この機能は，テーマのサムネイルを収集し，その画像の特徴量と入力画像の類似度を測る．本研究では画像の類似度で計算を行うとその内容によって精度に影響する恐れがあるため，コンテンツの内容ではなく配置でその類似度を計算する必要がある．このため，あらかじめデータベース内の画像をグレースケールに変換し，OpenCV を利用した文字と写真の認識によってコンテンツの位置を強調する加工を行う．さらに上記の加工を行ったデータに対し，検索精度を測る実験を行った．複数の画像に対し，その画像を入力したときの検索結果の正解データをあらかじめ定めておき，それを入力した際のシステムの出した結果の適合率と再現率，F 値を計測した．また，それらのシステムを利用したウェブサービス，WordDepth を開発した．WordDepth では，ユーザーがウェブ上のキャンバスを用いてシステムの入力画像を描写し，検索結果を昇順に提示する．

2. 関連研究

廣池ら [1]は，画像を対象とした高速な類似画像検索技術 EnraEnra を開発した．類似画像検索技術とは，画像の問い合わせに用いて似ている画像を探す技術であり，検索対象の画像から数百次元のベクトルを画像特徴量として抽出し，ベクトル同士の距離の比較によって検索結果を得る．しかし，本研究で行うテーマとクエリの類似度比較では元データの内容がノイズとなり上手く働かない．増永ら [2]は，情報検索のためのベンチマークを構築し，日本語文書での情報検索システムのためにデータベースを用意して検索実験を行った．本研究で扱うのは画像データについての検索だが，実験ではここで用いられていたものを参考にした．

Supervisor: Prof. Hiroshi Hosobe

3. 準備

3.1. HDBSCAN

HDBSCAN とは，データの密度に基づいてクラスタを形成するアルゴリズムである [3]．ノイズや外れ値を自動的に除外できる点が特徴となる．K-means アルゴリズムや DBSCAN のようにクラスタ形成に必要なパラメータを事前に入力する必要はなく，今回のような大規模データを扱ううえで役立つ．

3.2. 特徴量

特徴量とは，機械学習やデータ分析でデータの中から有用な情報を取り出しモデルに与えるための属性や指標のことである．画像データの場合，画像の各画素のピクセル値，特定の形状やエッジの有無，色のヒストグラムや領域の面積など画像内の物体やパターンを区別するために用いられる．本研究では，画像から特徴量を得るためのエンジンとして，MobileNetV3 を用いる．この特徴化器は軽量かつ高効率でありながら，高い認識精度を持つ．MobileNetV3 は Inverted Residual Block や Squeeze-and-Excitation モジュールなどの技術を採用し，計算量を抑えて重要な特徴を強調できる．

4. 提案手法

WordPress.org の Theme Directory では様々なテーマを検索できるが，いずれもキーワードやタグによる検索であり，画像入力による検索は行えない．本論文では，WordPress のテーマについて，画像検索の方法を用いて利用者の求めるテーマを調べて示す方法を提案する．各画像に対して 2 種の加工を施しコンテンツの内容に関係なく位置のみを特徴として扱えるようにする．

4.1. 画像の収集

テーマ推薦には情報検索システム，中でも画像検索の手法を用いる．そのため，データベースとなる画像約 7500 枚を WordPress.org 内の Theme Directory からスクレイピングによって収集する．これらは，各テーマのダウンロードページのサムネイルであり，そのテーマの持つ特徴を表す．

4.2. 画像の加工

収集した画像は，そのままの状態では画像の詳細が検索精度に大きく影響するため不適である．ここで，画像に対して 2 種の加工を施す．1 つ目の加工では，画像内の背景を白色にする．画像内の画素をカウントしていき，最頻色を持つものを白色に変更する．2 つ目の加工では，画像内に存在する写真の要素と文字の要素をそれぞれ青いボックスと赤いボックスで塗りつぶす．画像加工前と加工後のイメージは図 1 である．画像内の写真要素を抽出するために OpenCV で二値化とエッジ検出を行って輪郭を抽出し，文字要素の抽出には，文字認識エンジン Tesseract OCR を用いる．このエンジンは画像内に存在する文字要素を設定で単語や文節，文章ごとに囲めるため，それを利用して各単語を小さな赤いボックスで囲う．しかし，このエンジンでは少ない行のテキストブロックを認識できない．そのため，HDBSCAN アルゴリズムである距離内に存在する小さな赤いボックスを結合していくクラスタリングを行い，画像内の各テキストブロックをまとめて大きなボックスで囲む．これらの処理を示したのが図 2 である．

図 1　加工前と加工後のイメージ

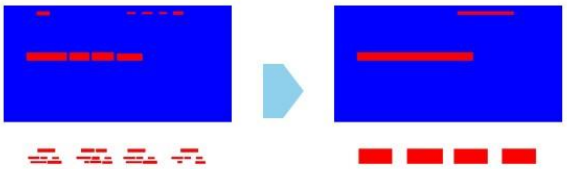


図 2：文字要素のクラスタリング処理

4.3. 画像の特徴化

加工後の画像をデータベースとして，すべてのデータベース内の画像に対し特徴化器を用いた特徴化を行い複数次元のベクトルに加工する．このベクトルを用いて画像検索を行う．検索を行う際には，入力画像に対してデータベースに行った処理と同様の特徴化を行う．特徴化した入力画像とデータベース内の画像とのユークリッド距離を測り，最も近いものを検索結果として提示する．なお，類似度を測定する方法としてユークリッド距離を測るほかにコサイン類似度を測る方法があるが，こちらはベクトルの角度を重視しスケールの影響などは無視されるため，各色のボックス配置を重視する今回の画像検索では不適と判断した．

5. 実装

WordPress.org 内のテーマページから各テーマのサムネイル画像を取得する必要があるため，Python の BeautifulSoup ライブラリを用いてスクレイピングを行った．ファイルの読み込みや画像の加工，検索には Python を使用する．スクレイピングによって得られた画像に対して，OpenCV で画像内の文字と写真の強調加工を行う．ここでは文字の認識に Tesseract OCR を用い，得た文字要素の中心点の座標を利用して HDBSCAN アルゴリズムでクラスタリングを行う．これにより，Tesseract OCR では不可能なテキストコンテンツの囲い込みを行える．写真の認識には画像を二値化した後に Canny エッジ検出を用いてエッジを抽出し，それによって得た輪郭情報を利用する．また，それらの処理の下処理として背景画像をすべて白にする加工も行う．さらに，特徴化器を用いて加工された画像を特徴化し，入力画像との類似度測定を行えるようにする．特徴化器には MobileNetV3 を利用する．

データベースは WordPress 公式サイトからスクレイピングした 6588 枚のサムネイル画像とする．

6. 実験

提案手法の性能を評価するため，ベンチマークセットを作成して評価実験を行った．評価指標として，正解データを設定したうえで再現率(Recall)，適合率(Precision)，F 値を使用した．

6.1. ベンチマークセット

実験には増永ら [2]の先行研究と同様にベンチマークを作成して用いる．本研究では，実際に使用するデータセットのうちから 480 件分の画像をランダムに選出したものを用いてテスト用のデータセットを作成した．評価指標に用いる正解データとして，著者が事前にペイントツールで作成した赤と青のボックスからなるウェブデザインらしい画像を元とし，その画像に似ていると著者が判断したコンテンツ配置を持つ画像数件をベンチマークの正解として選定した．複数画像に対し実験を行ったため，ベンチマークデータの正解データの画像数はそれぞれ 6〜10 件と一定ではなかった．

6.2. 評価指標

本実験では，通常の情報検索システムと同様に評価指標として再現率(Recall)，適合率(Precision)，F 値(F-Score)を使用した．正解データのうちツールが正しいと予測したものの割合が再現率であり，次式で求められる．

$$\mathrm{Recall} = \frac{\mathrm{TruePositive}}{\mathrm{TruePositive} + \mathrm{FalseNegative}}$$

正しいと予測したもののうち正解データの含まれる割合が適合率となり，次式で求められる．

$$\mathrm{Precision} = \frac{\mathrm{TruePositive}}{\mathrm{TruePositive} + \mathrm{FalsePositive}}$$

それらを用いて調和平均として計算されるのが F 値であり，次式で求められる．

$$\text{F-Score} = \frac{2 \times \mathrm{Precision} \times \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}}$$

これらの式で，TruePositive はシステムが正解と判断したもののうち，正解データに含まれるもの，FalseNegative は正解データであるがシステムが誤って不正解と判断したもの，FalsePositive は正解データに含まれないが，システムが正解と判断したものを示す．

情報検索システムでは，再現率が高い場合そのシステムは情報の検出漏れを防ぎ，検索結果が正しいデータを漏らさず取得する．一方，得られる結果にはノイズも少なからず含まれる．これは，情報の見落としを防ぐため医療診断や防犯システムで重視される．適合率が高い場合，そのシステムの検索結果はほとんど正しいことを示す．これは，ユーザーの求める情報を正しく表示する必要がある情報推薦システムやフィルタリングで重視される．F 値が高いことは，システムは再現率も適合率も両方優れていることを表す．これは，システムの信頼性が高くノイズなどは少ないことを示し，ユーザーが求める正しい結果を漏れなく無駄なく提供できていることを意味する．

6.3. 実験結果

本実験では，ある画像に対してそれを入力とした際，システムが出力すべき正解の画像データを任意の数だけ設定し，実際にプログラムを動かし検索結果にどの程度の数の正解が含まれているかを計測する．例えば，正解データが 5 個の時，検索結果の 7 件中 4 件が正解データの画像だった場合，再現率は 0.80，適合率は 0.57 となる．本実験では画像に対して検索結果が何件の場合，各評価指標が高くなるかを計測するため，3, 5, 7 件の場合について実験を行った．また，画像は A〜H までの 8 件に対して行い，それぞれ正解データの数は 6 件が 4 枚，7, 8, 10 件がそれぞれ 1 枚ずつであった．

画像 A に対して得られた結果が

表 1 である．ここでは，検索結果の件数を増やすごとに再現率や適合率が上がるわけではなく，5 件の時が最も高くなっている．しかし，F 値に注目すると 3 件の時に最もスコアが高くなっている．また，すべての画像に対し同様の実験を行った結果の各評価指標の平均値を示すものが表 2 である．結果によると F 値の値こそ検索結果数によって大きな変化はないが，再現率，適合率を見ると件数が 5 件の時に最も高いスコアを示している．これは，再現率と適合率が検索結果の量に依存し，それらが少なすぎたり多すぎたりする場合，0.1 を下回るほどの低スコアとなることがあったため，3 件と 7 件の時の評価指標の揺れが大きかったことに起因していると考える．

表 1：画像 A に対する実験結果

検索結果件数	再現率	適合率	F 値
3 件	0.33	0.50	0.50
5 件	0.66	0.60	0.43
7 件	0.44	0.54	0.46

表 2：すべての画像に対する評価指標の平均

検索結果件数	再現率	適合率	F 値
3 件	0.33	0.44	0.53
5 件	0.75	0.60	0.52
7 件	0.46	0.50	0.52

7. 応用例

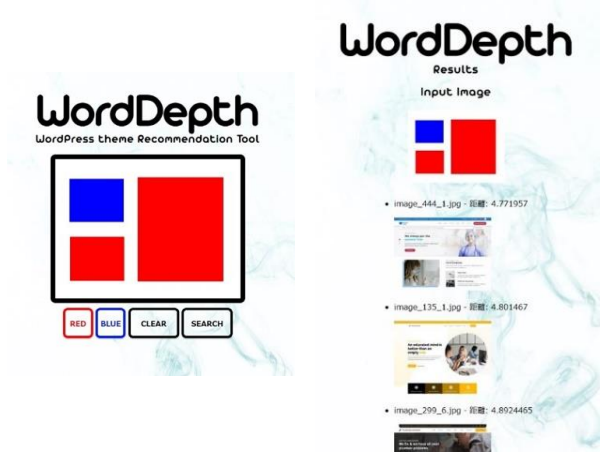
提案手法を利用したシステムの応用例の 1 つとして，ユーザーの入力をもとに，WordPress テーマの検索を支援するウェブサービスである WordDepth を開発した．このサイトのサーバーとウェブアプリの開発には Python 用のウェブアプリフレームワークである Flask を用いた．

本サービスでは，ウェブ上でユーザーに画像を描画してもらい，それを入力画像（クエリ）としてプログラムを動作させる．入力画像には写真や図，イラストなどの画像コンテンツを置きたい位置に青いボックスを，文章やタイトルなどの文字コンテンツを置きたい位置に赤いボックスを，サイト中央のキャンバスにユーザーに描画してもらったものを使用する（図 3(a)中央）．キャンバスでは，RED ボタンと BLUE ボタンで描写するボックスの色の変更を行えるようにし，CLEAR ボタンをクリックすることでキャンバス内に描画中の入力画像がリセット

される．ユーザーが入力画像を描写し終えたとき，右下の SEARCH ボタンを押すことで画像をもとに検索を行う．

検索結果として，当初の目的通り WordPress テーマのサムネイル画像と，0 を最低値とした入力画像と結果画像とのユークリッド距離を類似度として提示するようにした（図 3(b)）．検索結果の件数は，実験結果をもとに再現率と適合率の最高値が見込める 5 件とした．

それぞれ HTML と CSS を用いてユーザーが直感的に操作できるデザインを目指した．入力画面では，最低限の操作を行えるボタンのみ配置し，最も重要なキャンバスが最大になるようにデザインされている．また，結果画面では最も上に入力した画像を表示し，その下に結果が続いていくようにした．



(a) (b)

図 3 : (a)ウェブアプリの入力画面，(b)ウェブアプリの検索結果画面

8. 議論

8.1. 実験結果

実験の結果，正解データに対し検索結果の件数が小さい場合は再現率と適合率は低い値をとることが多かった，しかし，結果の件数が大きい場合でも同様に小さい値や，最大値をとることも多く，結果は安定しなかった．これらの値の揺れが中程度の検索結果の件数をとる場合は少なくなり，高い平均スコアを得られた．しかし，どの場合でも F 値はほぼ一定の値を示している．これは，本システムが検索結果の件数が変わっても全体的なバランスを維持しており，特定の評価指標に偏っていないことを示す．

適合率が再現率より 25%も高い数値を示していることは本システムの検索結果の網羅性の欠如を表す．必要な情報を漏らす可能性や，重要な情報が結果に含まれない問題や，フィルタリングにより特定の画像が優先され多様性が欠ける問題などにより，ユーザー満足度が低下する恐れもある．一般的に情報探索では適合率よりも再現率が優先されるため，再現率を上昇させる手段を検討する必要がある．具体的には，現在設定している検索精度の閾値を見直し，ノイズを許容することでより多くの結果をポジティブと判断するように変更することや，特徴化に用いるモデルの変更などがある．

8.2. 応用例

応用例では，本システムを利用したウェブサービスを開発した．提案手法は，検索システムの入力情報をユーザーによって作成された画像としているため，ウェブ上で動作するキャンバスと相性が良い．キャンバスによって画像を作成することでユーザーが画像を事前に作成しアップロードする手間を軽減できる．本サービスで，これからウェブページを作成したい人が WordPress 上の膨大な数のテーマの中から自分の求める最適なものを探す手間や煩わしさを軽減できる可能性が見込める．

9. おわりに

本論文では，クラスタリング，文字認識，特徴化器での検索を用いて，WordPress テーマの推薦を行う手法を提案した．この手法では，画像の単純化加工とそれを利用した画像検索アルゴリズムによって，ユーザーが求める WordPress テーマを推薦した．実験では，提案手法に基づいて加工した画像を用いた画像検索を行い，実用的と言える範囲の検索精度を得た．また，応用例では同検索システムを用いたウェブサービスである WordDepth を開発した．このサービスでは検索システムを動作させるための入力画像をサイト上のキャンバスのみで作成できるようにする．これにより，自前で画像を用意する必要が無くなり画像作成，アップロード，検索までの作業をサイト上で完結させることができる．

今後の課題として，検索精度の向上させることがあげられる．ユーザーが望む WordPress テーマをより確実に推薦するために，文字認識エンジンやエッジ抽出など各画像加工の見直しや，使用する特徴化器 MobileNetV3 から変更し，このような単純画像に特化したものを作成すればより検索精度が上がる可能性がある．また，応用例で作成したウェブサービス WordDepth の UI の改良や，具体的には動作サンプルの GIF 画像を表示しユーザーが困惑することなく操作できるようにすることなどがある．

文　献

[1] 廣池敦, "類似画像検索システム「EnraEnra」," 人工知能, vol. 29, no. 5, pp. 430-438, 2014.

[2] 小川泰嗣, 木本晴夫, 田中智博, 石川徹也, 増永良文, 芥子育雄, 豊浦潤, 福島俊一, 宮内忠信, "日本語情報検索システムのためのベンチマークの構築," 情報処理学会研究会報告：DBS, no. 100, pp. 145-152, 1994.

[3] R. J. G. B. Campello, D. Moulavi, A. Zimek and J. Sander, "Hierarchical Density Estimates for Data Clustering, Visualization, and Outlier Detection," *ACM Transactions on Knowledge Discovery from Data*, vol. 10, no. 1, pp. 5:1-51, 2015.

[4] A. J. Onumanyi, D. N. Molokomme, S. J. .Isaac and A. M. Abu-Mahfouz, "AutoElbow: An Automatic Elbow Detection Method for Estimating the Number of Clusters

in a Dataset," *Applied Sciences,* vol. 12, no. 15, pp. 7515:1-17, 2022.