

テーマパークの楽しみ方推薦を目的とした 画像付きツイートの自動分類の試み

土筆 勇都[†] 福原 知宏^{†‡} 山田 剛一[†] 増田 英孝[†]
東京電機大学[†] マルティススープ株式会社[‡]

1. はじめに

若い女性を中心に、SNS で共有することを目的とした写真撮影が流行している。「インスタ映え」という新語もでき、SNS 上で話題となっている人気の場所を調べ、そこへ行って見栄えのよい写真を撮影する SNS 利用者も多い。しかし、SNS 上で見つけた写真の場所を探す場合、写真からその撮影場所を特定することは容易ではない。本研究では、Twitter から収集した画像付きツイートをを用いて、テーマパーク来場者の楽しみ方を分析し、それをを用いてお薦めスポットを推薦することを目的としている。本論文ではテーマパーク内の場所やキャラクタを対象とし、収集した画像に対し、畳み込みニューラルネットワーク (CNN) を用いた画像の自動分類を行った。

2. 関連研究

橋本ら [1] は、SNS における炎上検知と炎上フィルタリングを目的に炎上画像の分類を行っている。その手法は、画像に写っている物体を手で抽出し、抽出された物体を特徴量として SVM によって分類を行うというものである。その結果 92.6% の精度で分類することができており、炎上検知率も 72.2% と良い結果が得られている。

本研究では、CNN を用いてテーマパーク内で多くの人が撮影している場所やキャラクタの画像を学習させることによって、収集した画像から自動的に対象物を特定し、分類を行う。橋本らとは炎上画像の代わりにテーマパーク内で撮影された画像を特定し、分類を行うという点で類似している。

3. 本研究の目的と提案システムの概要

本研究ではテーマパークを楽しもうとしているユーザに対し、お薦めの撮影スポットを推薦し、そこまでのルート案内を行うことで、撮影スポットへ容易に辿りつけることを可能にし、テーマパークをより楽しんでもらうことを目的としている。この時、ユーザはスマートフォンを用いてシステムを利用する。提案するシステ

ムの典型的な利用シナリオを以下に示す(図 1)。

システムは 2 で CNN を用いて事前に画像分類を行い、3 でユーザが撮影スポットのある領域内に入ったことを検出し、4 でユーザに該当する撮影スポットを推薦し、ユーザが選択した撮影スポットへ案内を開始する。

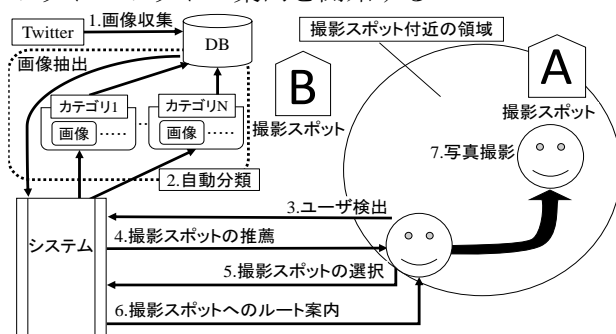


図 1. 提案する典型的なシステムの利用シナリオ

4. 実現方法の検討

4.1 概要

上記を実現するためには、テーマパーク内で既に撮影された画像と被写体の情報として被写体の名前と位置情報を特定し、予め分類しておく必要がある。テーマパーク内で既に撮影された画像については、テーマパークに関連する単語をキーワードとして Twitter から画像付きツイートを取得可能である。しかし、収集した画像全てがテーマパーク内で撮影されたものでないことやツイート内容と全く関係のない画像である場合が多いため、これを手作業で分類することは困難である。

そこで、本研究では、Twitter からテーマパーク内で撮影された画像付きツイートを大量に収集し、これに対して CNN を用いた自動分類を試みた。今回対象としたテーマパークは東京ディズニーランドである。手始めに、収集した画像を目視した際に多く保存されていたアトラクションや建造物とキャラクタを含めた「ミッキー、ミニ、シンデレラ城、プーさんのハニーハント、モンスターズインクのロゴがある壁、その他」の 6 カテゴリ分類を試みる。

4.2 ツイートの収集方法

TwitterAPI の SearchAPI を用いてディズニーランドに関連するキーワードを指定し、2017 年 8 月 14 日から 12 月 25 日の間画像付きツイート

Identification method of popular spots in a theme park from Twitter texts and images
[†] Hayato Tsukushi, ^{†‡} Tomohiro Fukuhara, [†] Koichi Yamada, [†] Hidetaka Masuda
 Tokyo Denki University([†]), MULTISOUP([‡])

を対象に約 10 万ツイートを集めた。ただし、重複した画像を集めてしまうことを防ぐため、リツイートされたツイートや Bot などの Web サービスを利用して投稿されたツイートは収集対象外とした。キーワードは、ディズニーランド内のアトラクション名、キャラクタ名、イベント名である。

4. 3 訓練データ

Google 検索や Instagram から収集した画像を縦 28 ピクセル横 28 ピクセルにリサイズを行ったものと縦 233 ピクセル横 233 ピクセルにリサイズを行ったものを用意した (表 1)。この時、キャラクタの画像については、顔を矩形領域で切り取り、建造物の画像については、建造物自体を矩形領域で切り取り、それぞれ画面一杯に写るように領域を抽出した。以下紙面の都合上モンスターズインクのロゴがある壁はロゴ壁、プーさんのハニーハントはハニーハントと略して表に記述する。

表 1. カテゴリ毎の訓練データの枚数

ミッキー	ミニー	シンデレラ城	ハニーハント	ロゴ壁	その他
223枚	155枚	173枚	112枚	70枚	37枚

4. 4 学習方法

LeNet のような小さな構造の CNN に 28×28 にリサイズした訓練データを用いた学習 (以下 CNN とする) と 233×233 にリサイズした訓練データを用いて ImageNet の大規模画像セットで事前学習済みのモデルである VGG16 に対して Fine-tuning による転移学習を行った。転移学習を行う理由は、少ない訓練データでも高い精度を出すことが見込めることができるからである。

5. 実験

転移学習を行ったモデルは、12 層目までの重みを固定し、VGG16 モデルの全結合層を全て削除し、代わりに 256 ユニット 1 層、カテゴリ分類用の 6 ユニット 1 層を全結合層として用い、13 層目から全結合層まで Fine-tuning を行った (以下 VGG16 とする)。最適化アルゴリズムとして確率的勾配降下法を用い、学習減衰率は 10^{-6} 、モーメンタムは 0.9 とした。以下に CNN と VGG16 のパラメータをまとめた表を示す (表 2)。そして、CNN と VGG16 から生成された分類器を用いて、Twitter から収集した画像の一部に対し、4.1 で述べたカテゴリへの分類を試みる (表 3)。

6. 結果と考察

CNN に関する全体の精度は、0.58、各カテゴリの精度の平均は、0.50、VGG16 に関する全体の精度は、0.65、各カテゴリの精度の平均は、0.69 となった。

CNN の精度は 58% と低く、VGG16 は CNN よりも

7 ポイント高い 65% の精度を得ることができたため、少ない訓練データでも精度を向上させることに成功している。しかし、65% では大量の画像分類を試みるためには不十分な精度である。これらの結果から、今回の分類対象では、訓練データを増やし転移学習を行うことで精度の向上が見込めるということ、訓練データとして用いた画像サイズが適切でなかったために正しく識別できていないということが考えられる。

表 2. CNN と VGG16 に用いたパラメータ

パラメータ	CNN	VGG16
画像サイズ	28px × 28px	233px × 233px
学習ステップ	500	1000
バッチサイズ	50	50
学習率	10^{-3}	10^{-6}
活性化関数	ReLU	ReLU
出力関数	ソフトマックス関数	ソフトマックス関数

表 3. 分類結果に対する評価

分類器の種類	適合率		再現率		F値	
	CNN	VGG16	CNN	VGG16	CNN	VGG16
ミッキー	0.54	0.86	0.52	0.24	0.53	0.38
ミニー	0.33	0.67	0.13	0.53	0.19	0.59
ロゴ壁	0.71	0.94	0.94	1.00	0.81	0.97
ハニーハント	1.00	0.40	1.00	1.00	1.00	0.57
シンデレラ城	0.83	1.00	0.69	0.69	0.75	0.82
その他	0.43	0.47	0.56	0.71	0.49	0.56
平均	0.64	0.72	0.64	0.69	0.63	0.65

7. おわりに

今回は、3. で述べたことを実現するために必要な事前処理である被写体を特定し、分類を行うという処理を実現するために Twitter から収集したテーマパーク内で撮影された画像に対し CNN を用いて 6 カテゴリへ画像の自動分類を試みた。その結果 VGG16 のモデルを Fine-tuning して生成された分類器の方が高い精度を得ることができた。

今後は、訓練データと様々な種類の画像を分類可能にするためにカテゴリ数を増やしつつ、今回得られた知見を元に Fine-tuning を再度行い、精度を向上させていく。また、今回は 1 つの分類器で分類を行ったが、今後、大分類から小分類について段階的に分類を行う方式を検討する。

参考文献

- [1] 橋本ほか, "機械学習を利用した SNS における炎上画像の推定", 情報処理学会第 78 回全国大会, 5S-02 (2016).