

# Web 画像を用いた物体認識実験

1510151 柳 裕太

2018 年 2 月 8 日

# 目次

<b>第 1 章</b>	<b>2 クラス物体分類実験</b>	<b>2</b>
1.1	課題内容 . . . . .	2
1.2	設計方針 . . . . .	2
1.3	プログラムの説明 . . . . .	4
1.4	実験 . . . . .	7
1.5	考察 . . . . .	7
1.6	感想 . . . . .	7
<b>第 2 章</b>	<b>Web 画像検索リランキング実験</b>	<b>10</b>
2.1	課題内容 . . . . .	10
2.2	設計方針 . . . . .	10
2.3	プログラムの説明 . . . . .	10
2.4	実験 . . . . .	10
2.5	考察 . . . . .	10
2.6	感想 . . . . .	10
<b>付録 A</b>	<b>プログラムリスト</b>	<b>11</b>
A.1	レポート課題 1 . . . . .	11
A.2	レポート課題 2 . . . . .	11
<b>参考文献</b>		<b>12</b>

# 第 1 章

## 2 クラス物体分類実験

### 1.1 課題内容

異なる 2 クラス各 200 枚の画像データセットをそれぞれポジティブ画像・ネガティブ画像として分類し、それらを対象に 2 クラス画像分類を行った。今回対象としては以下のパターンである。

- ポジティブ: タスマニアデビル
- ネガティブ: カピバラ
- ポジティブ: タスマニアデビル
- ネガティブ: それ以外

なお”それ以外”とは、`/usr/local/class/object/bgimg` に存在する 900 枚の画像から無作為に選出した 600 枚 (ポジティブ画像数の 3 倍) のことを指す。

分類は以下の 3 種類の方法を行い、分類精度の比較を行った。

- カラーヒストグラムと最近傍分類
- BoF ベクトルと非線形 SVM による分類
- MatConvnet の標準ネットワーク (AlexNet) による DCNN 特徴量と線形 SVM

なお、評価は 5-fold cross validation 方式を採用した。

### 1.2 設計方針

#### 1.2.1 画像収集方法

実験に先立ち、画像の収集を行った。Flicker から画像を収集し、横幅 320px にリサイズされた画像への URL を表示するサイト (<https://goo.gl/v8wbsR>) を使用した。検索ワードは

- タスマニアデビル: Tasmanian Devil
- カピバラ: Capybara

とした。得られた 300 枚の画像から明らかに当該動物が映っていない画像をここから除外し、さらに全体の画像が 200 枚になるように調整した。

### 1.2.2 codebook/filelist 作成

予め、後の処理で使用する全ポジティブ・ネガティブ画像の SIFT 特徴が記されたコードブックを作成した。また同時に分析対象のファイルの path が記された filelist も作成した。該当スクリプトと出力結果ファイルは

- タスマニアデビル/カピバラ (以下、t/c と表記)
  - codebook: `mk_codebook_tc.m` → `codebook-tc.mat`
  - filelist: `flist_tc.m` → `filelist-tc.mat`
- タスマニアデビル/それ以外 (以下、other と表記)
  - codebook: `mk_codebook_other.m` → `codebook-other.mat`
  - filelist: `flist_other.m` → `filelist-other.mat`

である。

### 1.2.3 カラーヒストグラムと最近傍分類

評価を行う `capybara_hist.m`(カピバラ相手) と `others_hist.m`(それ以外相手)、2 者に共通して画像のカラーヒストグラムを返す関数 `mk_hist.m` の 2 ファイルを作成した。最終的にどれだけの割合で正しくクラス分類できたか確率を返すようになっている。

### 1.2.4 BoF ベクトルと非線形 SVM による分類

評価を行う部分と、入力された filelist 記載の画像を BoF ベクトル化して返す部分を関数として実装した。該当するスクリプト・出力結果ファイルは以下の通りである。

- t/c
  - 評価: `bof_svm_tc.m`
  - BoF 化: `mk_code_tc.m` → `all_bovw-tc.mat`
- other
  - 評価: `bof_svm_others.m`
  - BoF 化: `mk_code_others.m` → `all_bovw-others.mat`

最終的にどれだけの割合で正しくクラス分類できたか確率を返すようにした。

### 1.2.5 MatConvnet の標準ネットワークによる DCNN 特徴量と線形 SVM

評価を行う部分と、入力された filelist 記載の画像の DCNN 特徴量を返す部分を関数として実装した。該当するスクリプト・出力結果ファイルは以下の通りである。

- t/c
  - 評価: `dcnn_svm_tc.m`
  - DCNN: `mk_dcnnlist.m` → `dcnn-tc.mat`
- other

- 評価: dcnn\_svm\_others.m
- DCNN: mk\_dcnnlist.m → dcnn-others.mat

最終的にどれだけの割合で正しくクラス分類できたか確率を返すようにした。

## 1.3 プログラムの説明

### 1.3.1 カラーヒストグラムと最近傍分類

#### 学習準備

- 該当スクリプト
  - t/c: capybara\_hist.m #1-22
  - other: others\_hist.m #1-22

まず予め作成した各画像への path が記された list 変数を、ポジティブ/ネガティブ画像へ分割した (いずれもインデックス番号 200 が境目となっている)。その後 5-fold cross validation のためポジ/ネガ双方を 5 分割して、内 4 つをポジ/ネガで統合して学習用 (train) とし、残りを学習後の評価に用いる (eval) ことにした。なお、正誤判定のために学習・評価ともに該当画像がポジなら 1, ネガなら -1 を収録した eval\_label, train\_label 配列も定義した。

なお学習・評価が終わると、5 分割中の割り当てをローテーションして再度学習・評価を行う。この繰り返しはすべての画像が評価対象となるまで、つまり 5 回行われる。

#### 学習

- 該当スクリプト
  - t/c: capybara\_hist.m #23-28
  - other: others\_hist.m #23-28
  - mk\_hist.m

train の全ての画像を mk\_hist.m に渡すことで、カラーヒストグラムの取得を行う。mk\_hist.m 内では、受け取った画像をモノクロ化してから各ドットの濃度を 64 分割中のヒストグラムに割り当て、正規化した配列を返している。各画像の結果は配列 db の要素として逐次追加される。

#### 評価

- 該当スクリプト
  - t/c: capybara\_hist.m #29-
  - t/c: others\_hist.m #29
  - mk\_hist.m

今度は train ではなく eval の全画像を対象にカラーヒストグラムを取得し、各画像ごとに最もカラーヒストグラムのインターセクションが少ない画像を探索する。そのインターセクションが最も近い画像のクラスが最も濃淡に近い画像と考え、該当するクラスに分類する。最後に、分類するクラスが正しいかどうか判定し、適合率を ac 配列に収める。

5-fold cross validation のため、評価は 5 回行われる。各回の適合率 ac は accuracy 配列に収

められ、最終的な適合率を算出する。

### 1.3.2 BoF ベクトルと非線形 SVM による分類

#### 画像の BoF ベクトル化

- 該当スクリプト: `mk_code.m`

こちらでは、予め解析対象となる全画像 (t/c: 400 枚, other:800 枚) を BoF ベクトル化させて、結果を `*.mat` として保存している。

まず、予め codebook 化されたファイルを読み込み、それを元に各画像の SIFT 特徴点を取得している。その結果を `code` 配列の要素として追加し、すべての画像の BoF ベクトル化が終了したら配列を保存している。

この BoF ベクトル化された配列は、この後学習や評価に使われることとなる。

#### 学習準備

- 該当スクリプト
  - t/c: `bof_svm_tc.m` #1-22
  - other: `bof_svm_other.m` #1-22

大まかな流れは第 1.3.1 節と同じだが、違う点は予め `mk_code()` 関数によって既に全解析対象画像が BoF ベクトル化されている点である。

#### 学習

- 該当スクリプト
  - t/c: `bof_svm_tc.m` #23-29
  - other: `bof_svm_other.m` #23-29

`train` 配列が学習対象画像の BoF ベクトルを持っているため、`train_label` と共に非線形 SVM モデル (`model`) を作成する。

#### 評価

- 該当スクリプト
  - t/c: `bof_svm_tc.m` #30-
  - other: `bof_svm_other.m` #30-

同じく `eval` 配列が評価対象画像の BoF ベクトルを持っているため、`model` と共に分類関数に入力し、分類結果を得ている。

その後、分類結果と正解クラスとの照合を行い、適合率を出力している。最後は、両クラスにおいて誤りだった画像を出力するための処理が下続く。

### 1.3.3 MatConvnet の標準ネットワークによる DCNN 特徴量と線形 SVM

#### DCNN 特徴量取得

- 該当スクリプト: mk\_dcnnlist.m

こちらでは、予め解析対象となる全画像 (t/c: 400 枚, other:800 枚) の DCNN 特徴量を取得し、結果を\*.mat として保存している。

まず学習済みモデルを読み込み、それを元に各画像の DCNN 特徴点を取得している。その結果を dcnn\_list 配列の要素として追加し、すべての画像の DCNN 特徴量の取得が終了したら配列を保存している。

この DCNN 特徴量配列は、この後学習や評価に使われることとなる。

#### 学習準備

- 該当スクリプト
  - t/c: dcnn\_svm\_tc.m #1-33
  - other: dcnn\_svm\_other.m #1-33

大まかな流れは第 1.3.2 節と同じだが、違う点は予め mk\_dcnnlist() 関数によって既に全解析対象画像の DCNN 特徴量が取得されている点である。

また、この後 DCNN 特徴点を取得するためのセットアップが#1-12 にて行われている。

#### 学習

- 該当スクリプト
  - t/c: dcnn\_svm\_tc.m #34-40
  - other: dcnn\_svm\_other.m #34-40

第 1.3.2 節と同様に、train 配列が学習対象画像の DCNN 特徴量を持っているため、train\_label と共に線形 SVM モデル (model) を作成する。

#### 評価

- 該当スクリプト
  - t/c: dcnn\_svm\_tc.m #41-
  - other: dcnn\_svm\_other.m #41-

同じく eval 配列が評価対象画像の DCNN 特徴量を持っているため、model と共に分類関数に入力し、分類結果を得ている。

その後、分類結果と正解クラスとの照合を行い、適合率を出力している。最後は、両クラスにおいて誤りだった画像を出力するための処理が下にく。

## 1.4 実験

### 1.4.1 適合率

各方法における適合率は以下の表の通りとなった。

表 1.1: 各分類クラス・分類方法における適合率

Classification method	t/b	other
Color histogram	0.345000	0.460000
BoF with rbf SVM	0.935000	0.742500
DCNN with linear SVM	0.980000	0.987500

### 1.4.2 誤答画像

実際に誤った分類が行われた画像は以下の図 1.1, 1.2 の通りである。

## 1.5 考察

**カラーヒストグラムと最近傍分類** 3 方法の中では、カラーヒストグラムが最も不安定な分類を行っていたことが読み取れた。また、図??, 1.2 のカラーヒストグラムによる誤答画像を肉眼で見ても、あまり類似性を見ることができなかった。これはカラーヒストグラムによる最近傍分類は、あまり精度の高いものではないことを強く示していると考えられた。

**BoF ベクトルと非線形 SVM による分類** t/c, other 双方の適合率を見ると、カラーヒストグラムよりはいずれも高かった。特に t/c においては DCNN と比べても遜色ない精度となっていたものの、other においては明らかに精度が落ちていることが表 1.1 から読み取れた。

これは、t/c においてはポジティブ/ネガティブ共に特定の種類の動物に限定していたたことにより、分離平面を定めることが容易であったものの、other ではネガティブが特定の動物に限定していなかった影響で分離が難しくなったことが原因として考えられる。

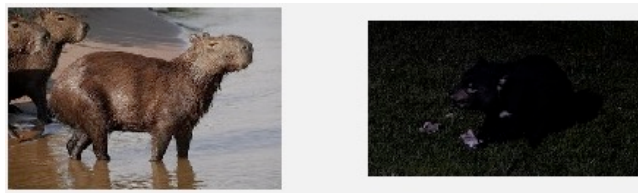
**MatConvnet の標準ネットワークによる DCNN 特徴量と線形 SVM** t/c, other とともに適合率が約 98% と非常に安定した数値となっていた。

詳しく見ると other の適合率が t/c より 0.75% 優れた結果となっているが、これは other の全体のサンプル数が t/c の二倍であったことが影響していると考えられる。

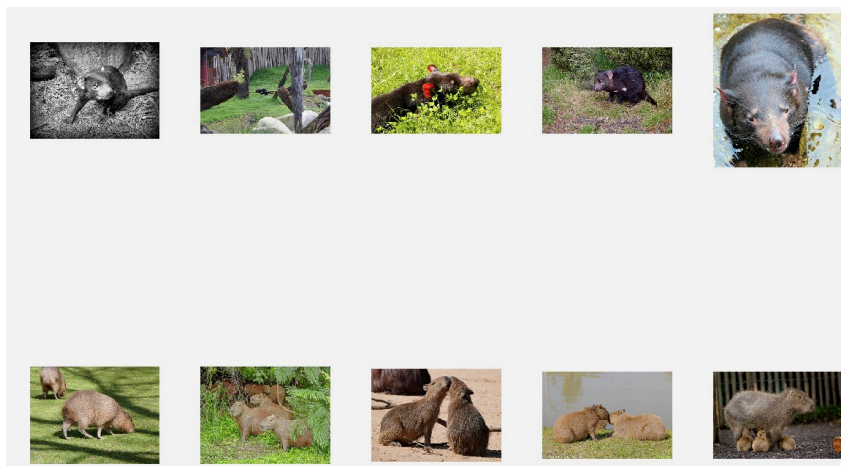
## 1.6 感想

3 種の画像認識による分類方法を実践することで、三者の認識方法の違いを知ることができた。

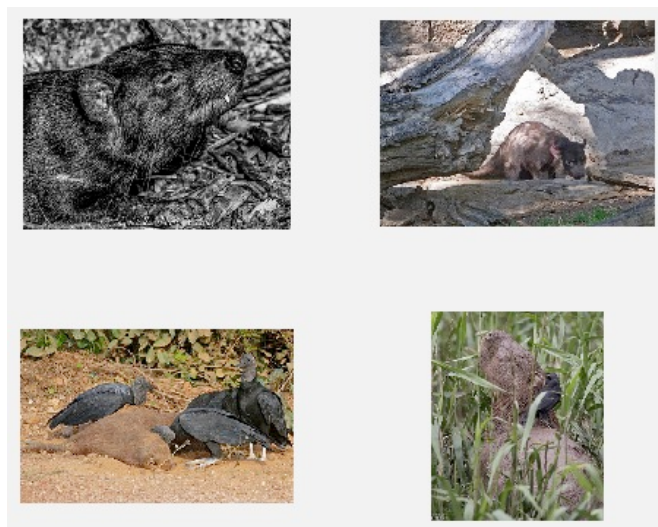




(a) Color histogram(左:判定対象画像 右:最近傍画像)

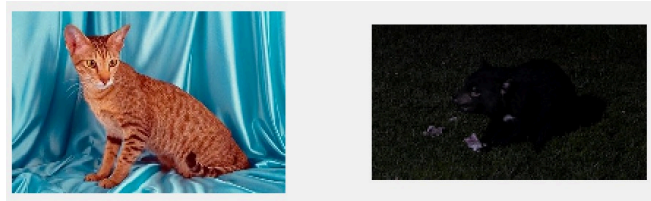


(b) BoF with rbf SVM(上:ネガティブ判定された画像 下:ポジティブ判定された画像)

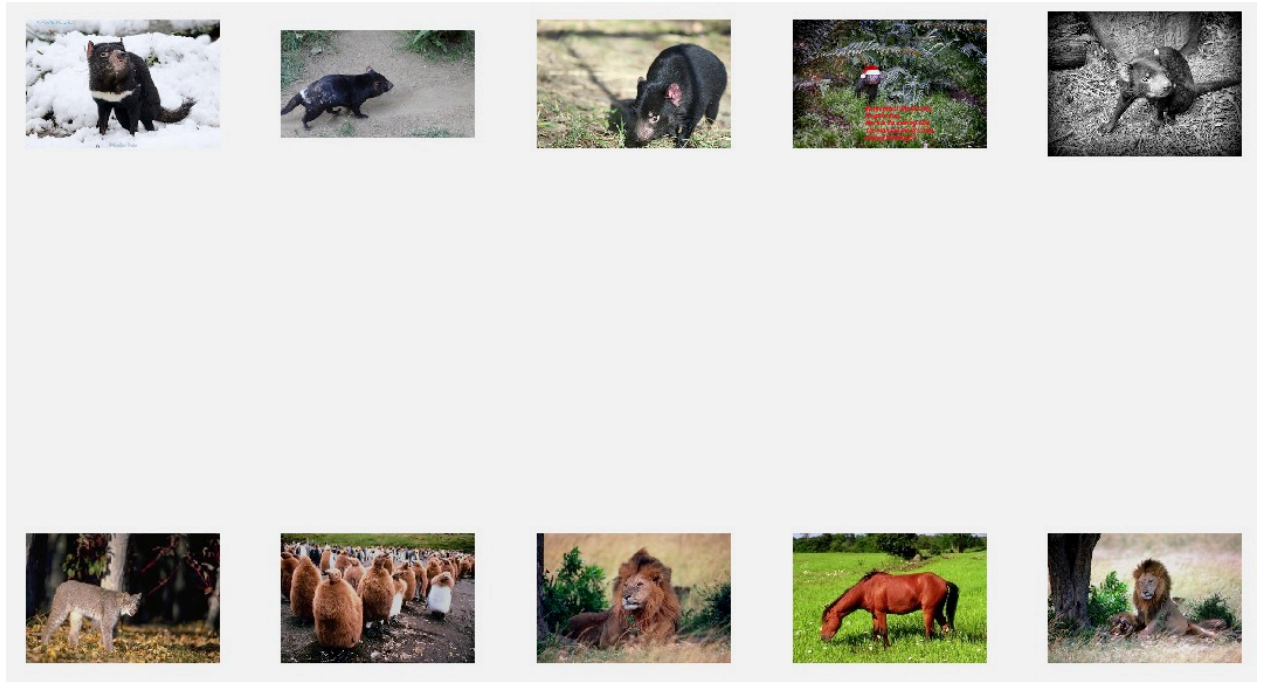


(c) DCNN with linear SVM(b と同様)

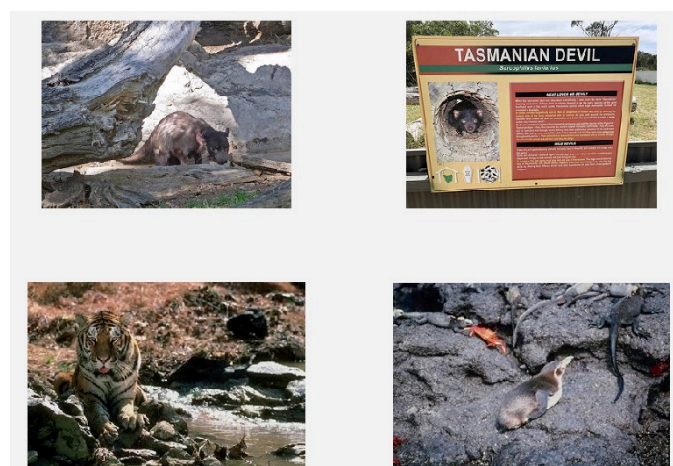
図 1.1: t/b 誤答画像一覧



(a) Color histogram(左:判定対象画像 右:最近傍画像)



(b) BoF with rbf SVM(上:ネガティブ判定された画像 下:ポジティブ判定された画像)



(c) DCNN with linear SVM(b と同様)

図 1.2: other 誤答画像一覧

## 第 2 章

# Web 画像検索リランキング実験

- 2.1 課題内容
- 2.2 設計方針
- 2.3 プログラムの説明
- 2.4 実験
- 2.5 考察
- 2.6 感想

## 付録 A

# プログラムリスト

A.1 レポート課題 1

A.2 レポート課題 2

# 参考文献

- [1] K.Yanai, "物体認識論 演習 レポート課題", the-UEC(Last modified: 27-Jan-2018)