

# OpenCV を用いたコマーシャルの同定について

## Identifying CMs by Neural Networks with OpenCV

小浦方大悟, 森和好

Daigo Korakata, Kazuyoshi Mori

会津大学

The University of Aizu

キーワード : テレビコマーシャル (Television Commercial), 機械学習 (Machine Learning),  
ニューラルネットワーク (Neural Network), 深層学習 (Deep Learning),  
パターン認識 (Pattern recognition)

連絡先 : 〒 965-0005 福島県会津若松市一箕町大字鶴賀字上居合 90  
会津大学 コンピュータ理工学部 コンピュータ理工学科 システム解析学講座  
小浦方大悟, Tel: 0242(37)2615, Fax: 0242(37)2747, E-mail: s1210195@u-aizu.ac.jp

## 1. はじめに

### 1.1 背景

テレビは一般的に広く普及しており, その中でもテレビコマーシャル (CM) は社会に広く浸透していることから, 当時の社会や流行を反映していることも多い<sup>1)</sup>. コマーシャルを調査することによって, 社会情勢を知ることができる他, 視聴者が惹きつけられるコマーシャルの傾向がわかれば, 広告事業にも貢献ができる<sup>2, 3, 4)</sup>. その上で, 事前に収集したデータを幾つかのカテゴリに分類することができれば, より効率的に調査を進められると考えられる.

しかしながら, 分類を手動で行うのは手間がかかり非効率である. よって自動でコマーシャルを分類するシステムの構築が必要である. その際に機械学習を導入し, ニューラルネットワークによるパターン認識を用いることで, 分類の自動化を図ることにした.

動画を分割して静止画像として扱うことで簡単化した後, より詳細な入力方法を調べるために, 今回は画像の大きさや学習効率及び精度の関係に焦点を置いた実験を行うことにした.

### 1.2 ニューラルネットワーク

ニューラルネットワーク<sup>5)</sup> は人間の脳の神経回路を模した数学モデルで, 文字認識や音声認識などのパターン認識に有用である. 多種多様なコマーシャルから特徴量を発見し手動で設定するのは困難であると考え, 今回は 2 つの中間層を持つモデルを用いた深層学習 (ディープラーニング)<sup>6, 7, 8, 9)</sup> によって自動で計算することにする.

### 1.3 OpenCV と TensorFlow

Open Source Computer Vision Library (OpenCV)<sup>10)</sup> とは, Intel が開発及び公開した

オープンソースのコンピュータビジョン向けライブラリである。本研究では主として動画処理に用いる。

TensorFlow(テンソルフロー)<sup>6)</sup>とは、Googleが開発しオープンソースとして公開した人工知能のソフトウェアライブラリである。深層学習にも対応しており、本研究ではニューラルネットワークの構築に利用する。

## 2. 実験

### 2.1 準備

事前に250個のコマーシャルを収集し、動画から3フレーム毎について静止画像として取り出し、商品別に“食料品”、“車”、“化粧品”、...のようにラベル(教師信号)を付けた。ラベリングの際は、その1枚の画像から予測できる範囲でカテゴリを決定した。この作業によって食料品10036枚、車5305枚、化粧品4221枚の計19562枚の学習用データ画像が得られた。その後、新たに30個のコマーシャルを収集し同様の手法でラベルを付け、得られた食料品2085枚、車703枚、化粧品1294枚の計4082枚の試験用データ画像は学習中の精度確認に用いられる。

### 2.2 概要

4層パーセプトロンのニューラルネットワークを用いてコマーシャルを3種類(“食料品”、“車”、“化粧品”)のいずれかに分類する。その際に入力として与えられる画像サイズを適宜変更して実験を行い学習時間と精度に及ぼす効果を検証する。現在はコマーシャルは16:9のアスペクト比で放映されている。縮小の際の画像の歪みによる特徴量への影響を考え、画像サイズを決定する際はアスペクト比を保つように配慮した。今回は40x22, 53x30, 80x45, 160x90, 320x180の5種類の画像サイズで実験を行う。

Table 1 画像サイズに対する学習時間と回数

画像サイズ	学習回数	学習時間
40x22	10000	03:37:06
53x30	10000	04:46:27
80x45	7853	09:05:20
160x90	2856	09:54:47
320x180	1328	13:41:25

Table 2 画像サイズに対する各予測精度

画像サイズ	学習精度	試験精度	動画精度
40x22	0.887384	0.579863	53%
53x30	0.918107	0.592847	56%
80x45	0.960025	0.653601	70%
160x90	0.960025	0.618569	63%
320x180	0.960738	0.613915	66%

ニューラルネットワークのパラメータは、入力層のユニットの数は画像サイズに応じて適宜設定し、1つ目の中間層は64、2つ目の中間層は24、出力層は3とする。教師信号として、学習データ画像を与える。ミニバッチを用いて1回の学習に256個のサンプルを使用し、最大10000回繰り返す。また、学習データに対する予測精度が96%以上になると打ち切る。

学習が終了した後、かかった時間を計測する。そして出来上がったモデルを用いて試験用データを動画として分類し、予測精度を求める。各々の得られた値を比較して、より学習効率の良い入力方法について考察をする。今回は動画として分類する仕組みとして、動画の全フレームに対して学習済みモデルによる分類を行い、最も分類率が高かったカテゴリを採用することにした。

### 2.3 結果と考察

Figs. 1-10に各実験の学習回数と学習データ画像及び試験データ画像に対する予測精度の関係を示し、Figs. 11, 12はこれらの結果をまとめた図となっている。各図では縦軸が予測精度、横軸が学習回数を表す。また入力する画像サイズに対する学習時間と回数をTable 1に、画

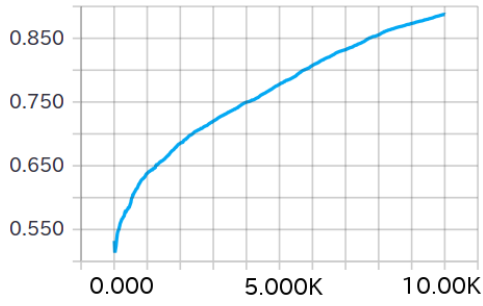


Fig. 1 学習精度と学習回数 (40x22)

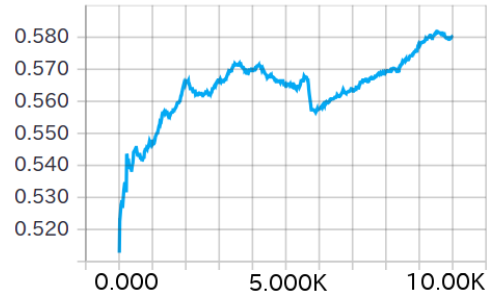


Fig. 6 試験精度と学習回数 (40x22)

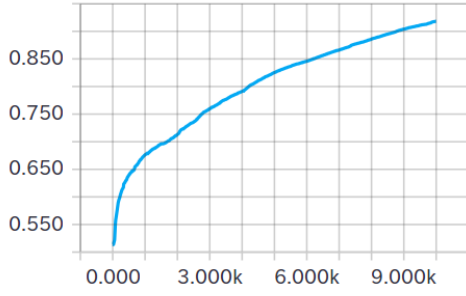


Fig. 2 学習精度と学習回数 (53x30)

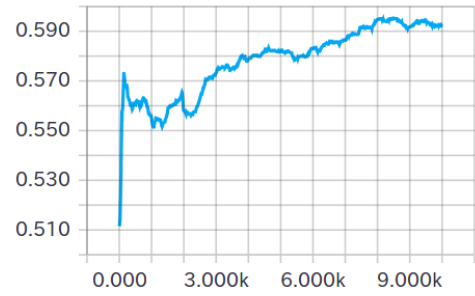


Fig. 7 試験精度と学習回数 (53x30)

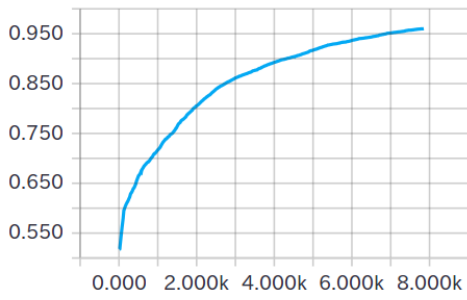


Fig. 3 学習精度と学習回数 (80x45)

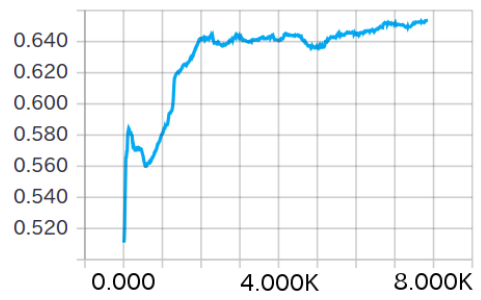


Fig. 8 試験精度と学習回数 (80x45)

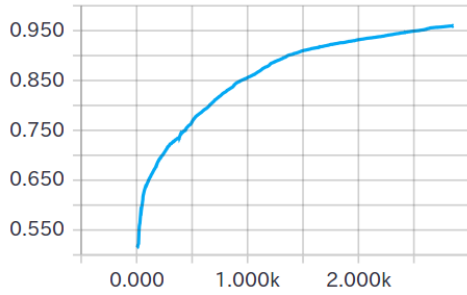


Fig. 4 学習精度と学習回数 (160x90)

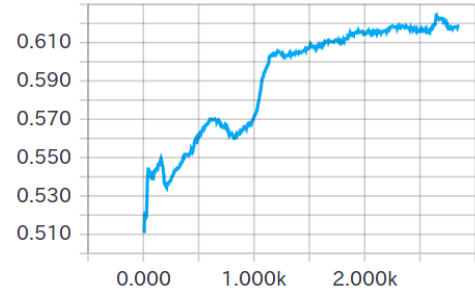


Fig. 9 試験精度と学習回数 (160x90)

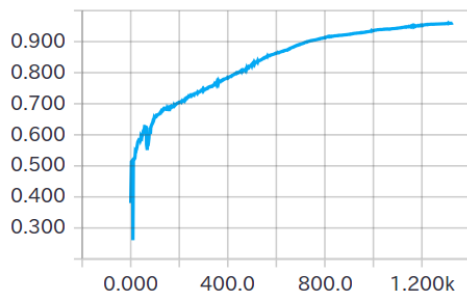


Fig. 5 学習精度と学習回数 (320x180)

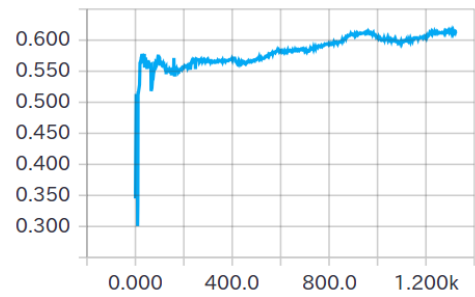


Fig. 10 試験精度と学習回数 (320x180)

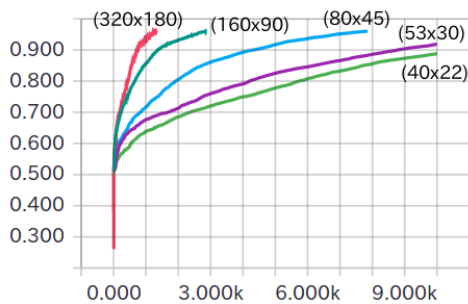


Fig. 11 学習精度と学習回数 (全て)

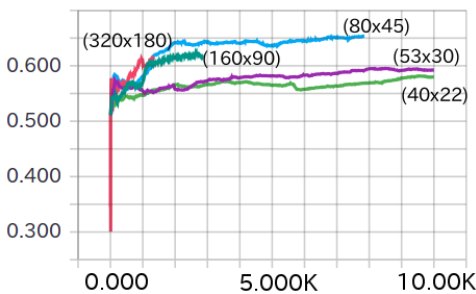


Fig. 12 試験精度と学習回数 (全て)

像サイズに対する各予測精度を Table 2 に示す。Fig. 11 と Table 1 から画像が小さいと学習時間は短く済むが学習データに対する予測精度は低く、逆に大きいと精度は高いが時間が多くかかることがわかる。Fig. 12 と Table 2 を見ると画像が小さいとき試験データに対する精度も低いことから、画像の縮小によって重要な特徴が失われ学習に影響を及ぼしたと考えられる。画像を大きくすると試験データに対する精度も向上したが、160x90 以上になると僅かに低下したことから、余分または不要な特徴が含まれていることによる学習への影響が考えられる。以上の結果から、画像サイズは特徴量を残しつつ適度に縮小することで学習をより効率的に進めることができると考えられる。

### 3. まとめと今後の課題

今回の実験によって、入力する画像サイズの変更による学習効率と予測精度の関係について検証できた。しかし、試験画像データに対する精度は最高で 65%程度であり、また動画に対する精度は最高で 70%だったが、現状の多数決の

仕組みだと典型的なコマーシャルにしか対応できないこともわかり、入力の工夫の他にも更なる改善が必要であると考えられる。

コマーシャルは開始直後と終了直前に商品に関する重要な情報が含まれる場合が多いため、それらに重みを付けることによって、改善できていると考えている。それ以外にも今後の課題として、カテゴリ数と分類精度の向上のために、畳み込みニューラルネットワークの利用、音声データや物体検出によるロゴの認識などを用いたシステムの改善を図る。

### 参考文献

- 1) 山田奨治, テレビ・コマーシャルと文化研究, 日本研究 35, 527/536 (2007)
- 2) 嶺田明美, 長澤輝世, 広告の表現について (1)-テレビコマーシャルの表現形式と文末表現-, 學苑 862, 13/23 (2012)
- 3) 嶺田明美, 長澤輝世, 広告の表現について (2) テレビコマーシャルにおける業種と表現形式を中心に, 學苑 864, 20/29 (2012)
- 4) 片柳伊佐, 猪狩, 良介 テレビCMクリエイティブの分類と広告効果の関係性: ブランド浸透度別の比較, 日経広告研究所報 48(4), 18/25 (2014)
- 5) 平野廣美 (著): Cでつくるニューラルネットワーク, パーソナルメディア株式会社 (1991)
- 6) MNIST For ML Beginners, <https://www.tensorflow.org/versions/r0.11/tutorials/mnist/beginners/index.html#the-mnist-data>
- 7) 中井悦司 (著): TensorFlow で学ぶディープラーニング入門 畳み込みニューラルネットワーク徹底解説, マイナビ出版 (2016)
- 8) 岡谷貴之 (著): 深層学習 (機械学習プロフェッショナルシリーズ), 講談社 (2015)
- 9) 神鳥敏弘 (編): 深層学習 Deep Learning (監修: 人工知能学会), 近代科学社 (2015)
- 10) OpenCV documentation Index, <http://docs.opencv.org/>