

2018年1月19日
株式会社インプレスR&D
<https://nextpublishing.jp/>

大好評第三弾！
『TensorFlow はじめました 3 Object Detection — 物体検出』発行
Google の機械学習を使いこなす！

インプレスグループで電子出版事業を手がける株式会社インプレス R&D は、『TensorFlow はじめました 3 Object Detection — 物体検出』(著者:有山圭二)を発行いたします。

『TensorFlowはじめました3 Object Detection — 物体検出』
<https://nextpublishing.jp/isbn/9784844398080>



著者:有山圭二

小売希望価格:電子書籍版 1000円(税別)／印刷書籍版 1500円(税別)

電子書籍版フォーマット:EPUB3／Kindle Format8

印刷書籍版仕様:B5判／カラー／本文78ページ

ISBN:978-4-8443-9808-0

発行:インプレス R&D

<<内容紹介>>

【大好評「TensorFlow はじめました」第三弾！】

本書は、TensorFlow がオープンソースで公開されるまで「機械学習」に触れたことがなかった筆者が、TensorFlow を通じて機械学習に挑戦して、七転八倒した成果をまとめた「TensorFlow はじめました」シリーズの第三弾です。

今回は画像の中から物体(イラストなら「顔」の部分など)を検出する「物体検出」を題材に、畳み込みニューラルネットワークモデルの学習と評価・検証を行っています。

(本書は、次世代出版メソッド「NextPublishing」を使用し、出版されています。)

画像の中に人間の「顔」があるイラストをテーマに機械学習でそれを検出する方法を解説

第2章 グリッドベースの物体検出

2.1 物体検出とは

物体検出 (Object Detection) は、画像の中に含まれる「物体」を検出するタスクです。

図2.1: 人物、カンバス (イーゼル)、絵の具など、さまざまな物体が含まれている



以前は、画像から物体らしいものが写っている候補領域 (Region Proposal) を見つけて、それらについて分類 (Classification) を実行するのが一般的でした。

現在はYOLO (You Only Look Once)やSSD (Single Shot MultiBox Detector)とといった手法が発表されていて、検出と分類を1つのネットワークで行うことができます。本書では、TensorFlowを使ってイラスト画像から顔を検出する機械学習のモデルを作成します。

モデルは、次の手順で作成します。

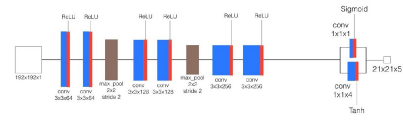
1. モデルの定義
2. データセットの作成
3. 学習
4. 検証

2.2 モデルの定義

リスト21のprepare_layersとoutput_layersは、本書で使う物体検出のモデル (図22) を構築します。

prepare_layersは、縦横192px (INPUT_IMAGEで定義)の画像を入力すると、畳み込みとプーリングを経て [21, 21, 256] の中間結果を返します。output_layersは、prepare_layersで得た中間結果を元に [21, 21, 5] を出力します。

図2.2: モデル構造



リスト21: tfbook_model/model0.py

```
import tensorflow as tf

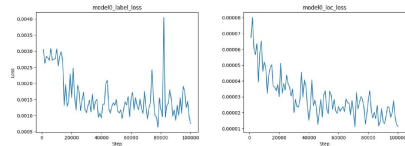
NAME = 'model0'
NUM_CLASSES = 1

INPUT_IMAGE = 192

OUTPUT_SHAPES = [
    [21, 21],
```

豊富なグラフとサンプルコードも掲載

図2.11: 誤差の変化



2つの誤差は、概ね順調に下がっているように見えます。特に座標の誤差 (model0_loc_loss) は、確信度誤差 (model0_label_loss) と比較して誤差が小さく、ステップ数を重ねるにつれて下がっています。

次に、この学習済みモデルを使って、実際にイラスト画像から顔領域を検出します。

2.5 検証

リスト214は、画像を学習済みモデルに入力して結果を出力するプログラムです。

リスト214: eval_ssd.py

```
import json
import os

import numpy as np
import tensorflow as tf
from PIL import Image, ImageDraw

import box_util
from tfbook_model import model0 as base_model

# 使用するモデル
from tfbook_model import model0 as model

FLAGS = tf.app.flags.FLAGS
tf.app.flags.DEFINE_string('image_dir', None,
    "処理対象のディレクトリ")
tf.app.flags.DEFINE_string('output_dir', None,
    "出力するディレクトリ")
tf.app.flags.DEFINE_string('train_path', None,
    "訓練結果のファイルパス")
```

```
def _load_image(image_path, image_size):
    image = Image.open(image_path).convert('L')
    resized_image = image.resize((image_size, image_size))
    image = np.array(resized_image.getdata()) \
        .reshape(image_size, image_size, 1) \
        .astype(np.float32)
    return image

def _write_regions(boxes, image_file, output_path):
    image = Image.open(image_file).convert('RGB')
    draw = ImageDraw.Draw(image)
    for box in boxes:
        rect = box.rect()
        region = [rect[0] + image.width,
            rect[1] + image.height,
            rect[2] + image.width,
            rect[3] + image.height]
        draw.rectangle(region, outline=0xFF0000)
    image.save(output_path, format='JPEG')

def _get_result_path(image_path, dest_dir):
    name, ext = os.path.splitext(os.path.basename(image_path))
    path = os.path.join(dest_dir, '%s-result.jpg' % name)
    return path

def _create_boxes(flatten_logits_value):
    boxes = []
    for shape in model.OUTPUT_SHAPES:
        boxes.extend(box_util.build_boxes(shape))

    for index, logits in enumerate(flatten_logits_value):
        loc = logits[:4]
```

グリッドベースの物体検出についてグラフィカルに説明

座標の差分 (offset)

第2章で学習用データセットを作成したとき、顔領域とグリッドのboxとの対応は、顔領域の中心座標を基準に決定しました。またboxには、顔領域の座標との差分 (offset) を設定しました。

図3.3: 顔領域の中心座標があるboxに割り当てる

Box(4, 10)
left : 0.4761
top : 0.1986
right : 0.5237
bottom : 0.2258
confidence: 1.0
left : -0.0464
top : -0.0188
right : 0.0258
bottom : 0.0518

Box(17, 11)
left : 0.0895
top : 0.0000
right : 0.0000
bottom : 0.0000
confidence: 0.0
left : 0.0000
top : 0.0000
right : 0.0000
bottom : 0.0000

Box(6, 14)
left : 0.6666
top : 0.2857
right : 0.7124
bottom : 0.3333
confidence: 1.0
left : -0.0583
top : -0.0827
right : 0.0485
bottom : 0.0299

Box(12, 17)
left : 0.0895
top : 0.0952
right : 0.0571
bottom : 0.1428
confidence: 1.0
left : -0.0722
top : -0.0885
right : 0.0325
bottom : 0.0388

Box(14, 10)
left : 0.4761
top : 0.6666
right : 0.5238
bottom : 0.7142
confidence: 1.0
left : 7
top : 7
right : 7
bottom : 7

1024x1024

しかし、顔の領域にはさまざまな大きさがあります。たとえば、図3.4の場合、顔の中心座標となるboxと実際の顔の領域の大きさが乖離していて、単純に差分をとると非常に大きな値になります。

図3.4: boxに対して顔の領域が大きすぎる

画像を単純に縦横21劃のグリッドに分割するのは限界があります。では、画像に含まれるさまざまな大きさの物体を認識するには、どのようにすれば良いのでしょうか。そのために、モデルに使われている「畳み込み層 (Convolution Layer)」に注目します。

3.2 畳み込み層

畳み込み層は、入力データから特徴マップを作成します (図3.5)。

図3.5: 畳み込み層の動作

56 | 第3章 物体認識奮闘記

第3章 物体認識奮闘記 | 57

<<目次>>

第1章 TensorFlow の基礎

- 1.1 TensorFlow とは
- 1.2 データフローグラフ
- 1.3 テンソル (Tensor)
- 1.4 変数とプレースホルダー
- 1.5 演算子のオーバーロード
- 1.6 ブロードキャスト

第2章 グリッドベースの物体検出

- 2.1 物体検出とは
- 2.2 モデルの定義
- 2.3 データセットの作成
- 2.4 学習 (訓練)
- 2.5 検証

第3章 物体認識奮闘記

- 3.1 確信度と座標
- 3.2 畳み込み層
- 3.3 モデルの定義
- 3.4 データセットの再作成
- 3.5 学習・訓練
- 3.6 検証

付録 参考資料

<< 著者紹介 >>

有山 圭二(ありやま けいじ)

有限会社シーリス代表。Android アプリケーションの受託開発や、Android に関するコンサルティング業務の傍ら、技術系月刊誌への記事執筆、AOSP(Android Open Source Project)へコントリビュートを行っている。著書に『TensorFlow はじめました』シリーズ(インプレス R&D 刊)、共著に『TensorFlow 活用ガイド[機械学習アプリケーション開発入門]』(技術評論社刊)。

<< 販売ストア >>

電子書籍:

Amazon Kindle ストア、楽天 kobo イーブックストア、Apple iBookstore、紀伊國屋書店 Kinopyy、Google Play Store、honto 電子書籍ストア、Sony Reader Store、BookLive!、BOOK☆WALKER

印刷書籍:

Amazon.co.jp、三省堂書店オンデマンド、honto ネットストア、楽天ブックス

※ 各ストアでの販売は準備が整いしだい開始されます。

※ 全国の一般書店からもご注文いただけます。

【株式会社インプレス R&D】 <https://nextpublishing.jp/>

株式会社インプレス R&D(本社: 東京都千代田区、代表取締役社長: 井芹昌信)は、デジタルファーストの次世代型電子出版プラットフォーム「NextPublishing」を運営する企業です。また自らも、NextPublishing を使った「インターネット白書」の出版など IT 関連メディア事業を展開しています。

※NextPublishing は、インプレス R&D が開発した電子出版プラットフォーム(またはメソッド)の名称です。電子書籍と印刷書籍の同時制作、プリント・オンデマンド(POD)による品切れ解消などの伝統的出版の課題を解決しています。これにより、伝統的出版では経済的に困難な多品種少部数の出版を可能にし、優秀な個人や組織が持つ多様な知の流通を目指しています。

【インプレスグループ】 <https://www.impressholdings.com/>



株式会社インプレスホールディングス(本社: 東京都千代田区、代表取締役: 唐島夏生、証券コード: 東証1部9479)を持株会社とするメディアグループ。「IT」「音楽」「デザイン」「山岳・自然」「モバイルサービス」を主要テーマに専門性の高いコンテンツ+サービスを提供するメディア事業を展開しています。2017年4月1日に創設25周年を迎えました。

【お問い合わせ先】

株式会社インプレス R&D NextPublishing センター

〒101-0051 東京都千代田区神田神保町 1-105

TEL 03-6837-4820

電子メール: np-info@impress.co.jp