

ディープラーニングを用いたリアルタイム動画画像認識

D-2 Real Time Video Recognition using Deep Learning

佐久間 龍樹[†] 黒木 啓之[†]Ryuki Sakuma[†] Takashi Kuroki[†][†] 東京都立産業技術高等専門学校[†] Tokyo Metropolitan College of Industrial Technology

1 はじめに

今日において人工知能はその優秀さからデータの予測、データの分析など様々な分野で用いられている。この技術は我々が生活している場面でも多く用いられている。

特に人工知能は画像解析、画像認識の分野で利用されており、車載カメラや無人コンビニになどが開発され競争が激化している。

しかしリアルタイムの動画画像認識に用いられる人工知能は普及の段階に達しておらずまだ発展途上段階にある。そこで本研究は人工知能の多くに用いられているディープラーニングを使用して、リアルタイムで動画画像の認識を行うことを目的とする。

2 研究方法

本研究では0~9のハンドサインをディープラーニングを用いてリアルタイムでの動画画像認識を以下の行程で行う。

1. 静止画の0~9のハンドサインのデータセットを取得してディープラーニングで学習を行う
2. 学習を行ったモデルをJSON形式で出力する
3. 学習済みモデルを使って動画画像認識を行うため、サーバを立ててWebカメラを用いて動画をキャプチャする
4. キャプチャした動画を学習済みモデルに入力して、最後にキャプチャした動画画像が0~9のどのハンドサインなのかを認識させる

3 研究結果

計測を行う際に以下の条件の元で計測する。

- ハンドサインの背景を白にする
- 肌色に近い色のものを写りこまないようにする
- 画面の $\frac{1}{3}$ 以上の大きさにハンドサインを作る

上記の条件を元に図1のハンドサインを行い、計測を行った結果を表1に示す。ただしnumberはハンドサインに相当する数字を、valueは認識結果の出力値を表す。

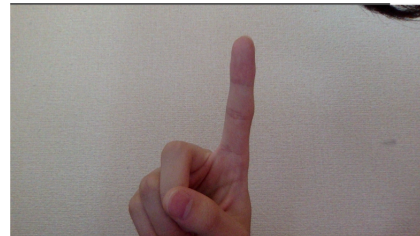


図1 条件を元に行ったハンドサイン

表1 認識の出力結果

number	value
one	0.946170
zero	0.023130
eight	0.012399

この結果より学習したデータセットに基づいて条件を与えてハンドサインをキャプチャした結果、定義されたハンドサインと同じ数字として認識した。

4 おわりに

本研究を通してディープラーニングを用いたリアルタイムの動画画像認識を行うことができた。しかし偏ったデータセットは正しい認識を行うことができないことがわかった。

今後の課題として、様々な入力パターンに対応していないため、モデルの学習方法を変更したり学習するためのデータに偏りなく様々なデータを入れる必要がある。