

Python を用いた光学文字認識

岡田 義弘

黒井 大夢

1. 研究動機

- ・先端技術である AI に興味があり、自分たちで作りたいと思ったため。
- ・OCR を作ることで、事務処理での人の負担を減らすことができると考えたから。

2. 研究内容

- (1) OpenCV を使用して Python で画像の加工を行う。
- (2) Tesseract OCR を使用して選択した画像内の文字を認識し、その認識した文字データをメモ帳に出力する。
- (3) Python を使って自作 OCR を作成する。

3. 使用機器・ソフトウェア

(i) Python

1991 年に開発されたオープンソースのプログラミング言語。人工知能開発によく用いられる。
簡潔なプログラムが書けることが特徴。

(ii) OpenCV

画像処理・画像解析および機械学習などの機能が用意されているオープンソースのライブラリ。
Python に限らずさまざまな言語で使用できる。

(iii) Tesseract OCR

Google が提供しているオープンソースの OCR (光学文字認識) エンジン。
画像の中の文字をテキストファイルとして読み込むことができる。

Python から実行する場合、PyOCR というライブラリをインストールする必要がある。

4. 研究の成果

(1) OpenCV をインストールし、画像のグレースケール化とエッジ加工を行う。

- ・グレースケール…白と黒の中間色であるグレーの濃淡を、254 階調で表したものの。

白黒よりも豊かな表現が可能。

- ・エッジ加工…画像内の画素値の変化が大きい箇所(エッジ)を検出する処理。
エッジ検出を行うことで物体の輪郭を抽出することができる。



取り込む画像①



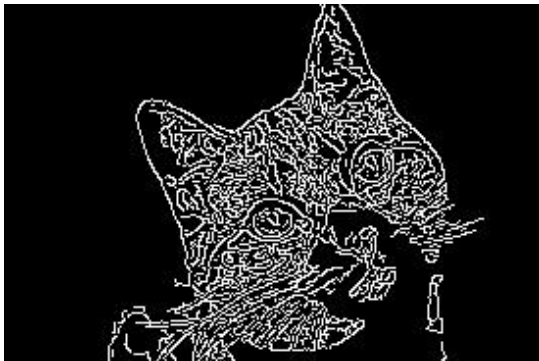
実行結果①-1 (グレースケール化)

Python を用いた光学文字認識

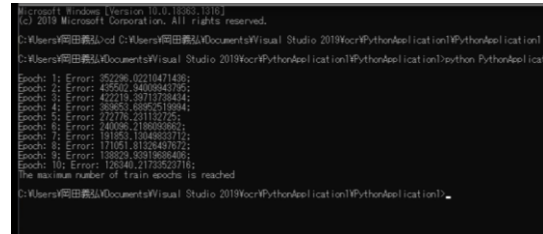
岡田 義弘

黒井 大夢

材をインプットしてニューラルネットワークに入れて訓練させ、ネットワークを形成するパラメータを最適化する。



実行結果①-2 (エッジ加工)



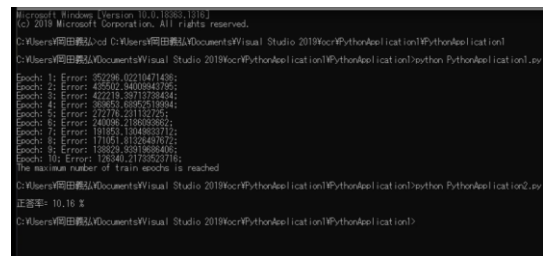
ニューラルネットワーク学習

(2) Tesseract OCR をインストールし、性能テストをする。

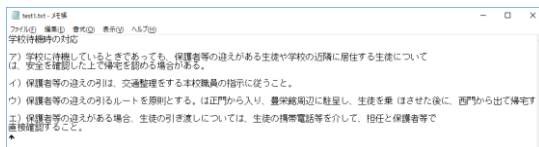
学校待機時の対応

- ア) 学校に待機しているときであっても、保護者等の迎えがある生徒や学校の近隣に居住する生徒については、安全を確認した上で帰宅を認める場合がある。
- イ) 保護者等の迎えの車は、交通整理をする本校職員の指示に従うこと。
- ウ) 保護者等の迎えの車は正門から入り、豊栄館周辺に駐車し、生徒を乗車させた後に、西門から出て帰宅するルート为原则とする。
- エ) 保護者等の迎えがある場合、生徒の引き渡しについては、生徒の携帯電話等を介して、担任と保護者等で直接確認すること。

取り込む画像②

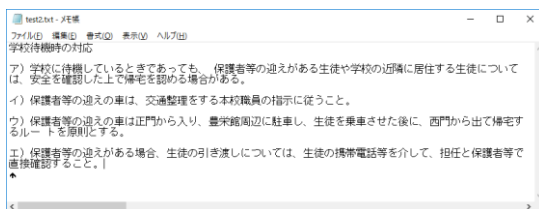


学習の結果



実行結果②-1

取り込む画像の画素数を多くすることで解像度が上がり、精度が向上する。



実行結果②-2

(3) Python を使用して OCR プログラムを作成する。

解答付きの学習素材を解答付きの学習素

5. まとめ・反省

・画像加工、画像認識に関して

OpenCV による画像の加工では、OCR の精度を上げるためにグレースケールや、エッジ加工などを学ぶことができた。エッジ加工では、キャニー法を学んだが、他にも加工方法が多々あったのでそれらについても学んでいきたい。

・OCR に関して

光学文字認識の仕組みを一から学ぶことができた。自分たちで OCR を作ったが、Tesseract OCR よりも認識精度がよくなかった。ニューラルネットワークの学習方法を理解するのに時間がかかってしまったのが理由としてあるので、今後知識を深め OCR の精度が上げられるようにしたい。