

# データマイニング 3班

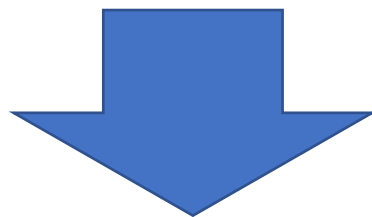
184528D 下地 剛史

185761E 多和田 真都

185767D 藤渕 はな

# 目的・目標

画像に描かれている文字を一つ一つ手で打ち込むのは、少ない文字であれば簡単であるが、**文字数が増えれば仕事量は増加**してしまう



**画像に描かれている文字を認識できるツール**を作成したい！！

# アプローチの全体像

## 1.文字を画像に描く

- ・画像にjpg形式で文字を描いて保存

## 2.プログラムで読み込み、判断

- ・ツール自体をPython(keras)で実装
- ・使用したデータベースは  
"mnist"(数字認識のテストに利用),  
"手書き教育漢字データベースETL8"
- ・モデルに畳み込み層を用いる事で  
認識率を高める
- ・モデルの最適化の手法としてAdamを  
用いる事でETLの認識率を高める

## 3.判断した結果を出力

最終的にETLを用いて文字認識の精度を  
99%まで高める事ができた

# 予定していた実験計画

1.文字を画像に描く

・文章が描かれた画像を読み込む  
(未実装)

2.プログラムで読み込み、判断

・文字を一文字ずつでなく、複数文字の画像を読み込み、判断をする(未実装)  
→**mser**を利用し、文字領域の抽出を行うが、ひらがなの認識が難しく断念

3.判断した結果を出力

# データセットの構築方法

ひらがな文字に関するデータを用意し、画像に変換する



MNISTと同じデータ型になるように画像データを修正する

# ETL文字データベース

ひらがな、カタカナ、漢字など、約120万の文字画像データが収集されている





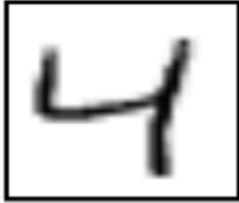

ETL8...教育漢字881種、ひらがな75種



1文字につき161個の画像データ ✖ 75種類 = 12075個

# mnistのデータの内容

70000個の手書き数字画像データ { 60000個の学習用データ  
10000個の検証用データ

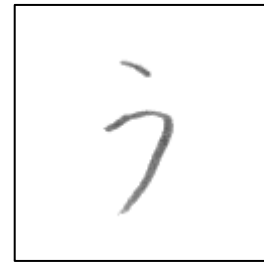
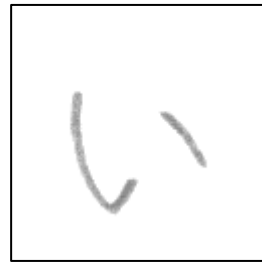
画像データ				
ラベルデータ	5	0	4	1

それぞれの画像に正解を示すラベルデータがある



12075個の手書きひらがな画像データ { 11325個の学習用データ  
750個の検証用データ

画像データ



||

||

||

||

ラベルデータ

"あ"

"い"

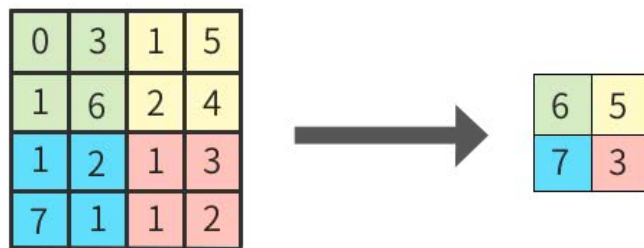
"う"

"え"

# 機械学習の進め方

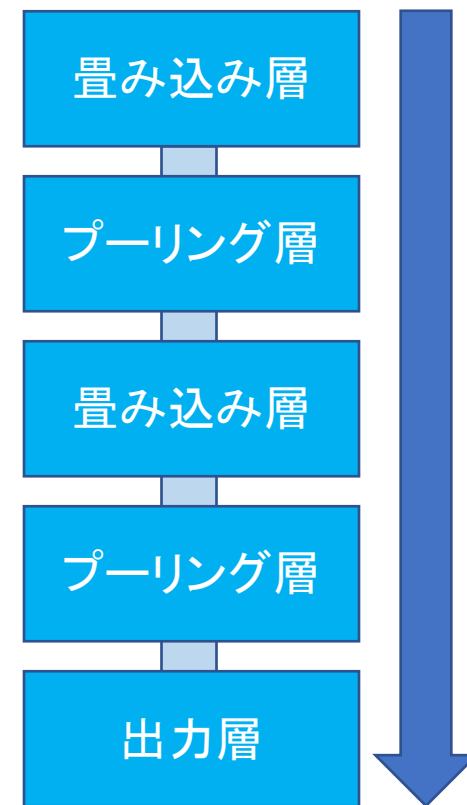
## 使用した学習機

- CNN
- LeNet(プーリング層、畳み込み層が二つずつ)



Max Pooling

引用:[https://deepage.net/deep\\_learning/2016/11/07/convolutional\\_neural\\_network.html](https://deepage.net/deep_learning/2016/11/07/convolutional_neural_network.html)



# 機械学習の進め方

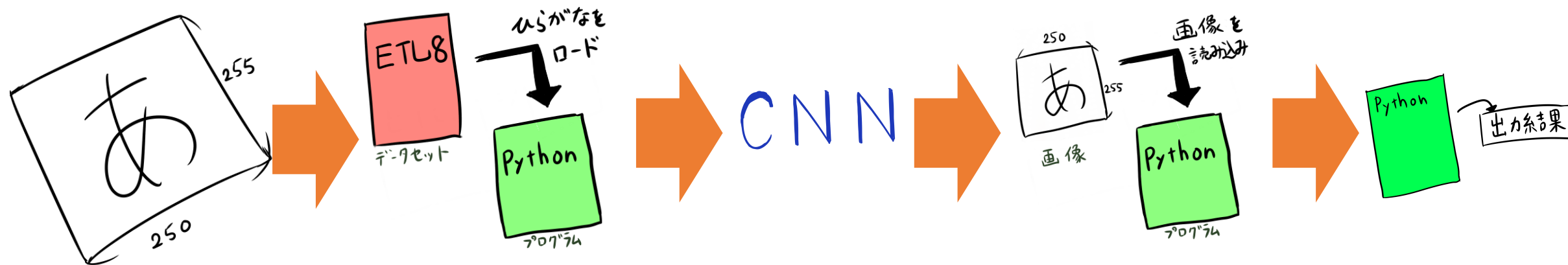
## パラメータの調整

- 数字認識からひらがな認識に移行した際に、入力画像サイズを28x28から40x40に増加させた。
- 畳み込み層のカーネルサイズは3x3。
- プーリング層のプーリングサイズはそれぞれ2x2と4x4。

# 実験

- 実験設計

1. ひらがな1文字を1つの画像に描く ( $250 \times 255$ )
2. ETL8をロードする
3. 畳み込みニューラルネットワーク (CNN) を構築
4. Python上で画像を読み込む
5. 画像を判断する



# 実験結果

あ

予測文字	あ	ね
確率	99.9%	0.01%

い

予測文字	い	の
確率	約100%	0.01%以下

う

予測文字	う	の
確率	99.6%	0.21%

え

予測文字	え	み
確率	77.9%	11.9%

お

予測文字	お	わ
確率	98.9%	0.85%

へ

予測文字	べ	ぺ
確率	97.3%	1.9%

へ

予測文字	べ	ぺ
確率	99.5%	0.4%

ば

予測文字	ば	ぱ
確率	94.7%	5.2%

ば

予測文字	ば	ぱ
確率	53.9%	46.0%

# 考察

半濁音を濁音と誤認識する割合が高い。

学習データに似せた半濁音文字は正しく認識することから、学習データの数が少ない事や多様性が低い事が考えられる。

入力データの前処理(2値化、サイズ)によって学習速度や予測精度に大きな影響が出たため、モデルの構成の他にも入力データの質の高さが予測精度を上げる要因となっている。