第 6 回　知能システム学特論レポート

15344203　有田 裕太
15344206　緒形 裕太
15344209　株丹 亮
12104125　宮本 和

西田研究室，計算力学研究室

2015 年 7 月 6 日

進捗状況

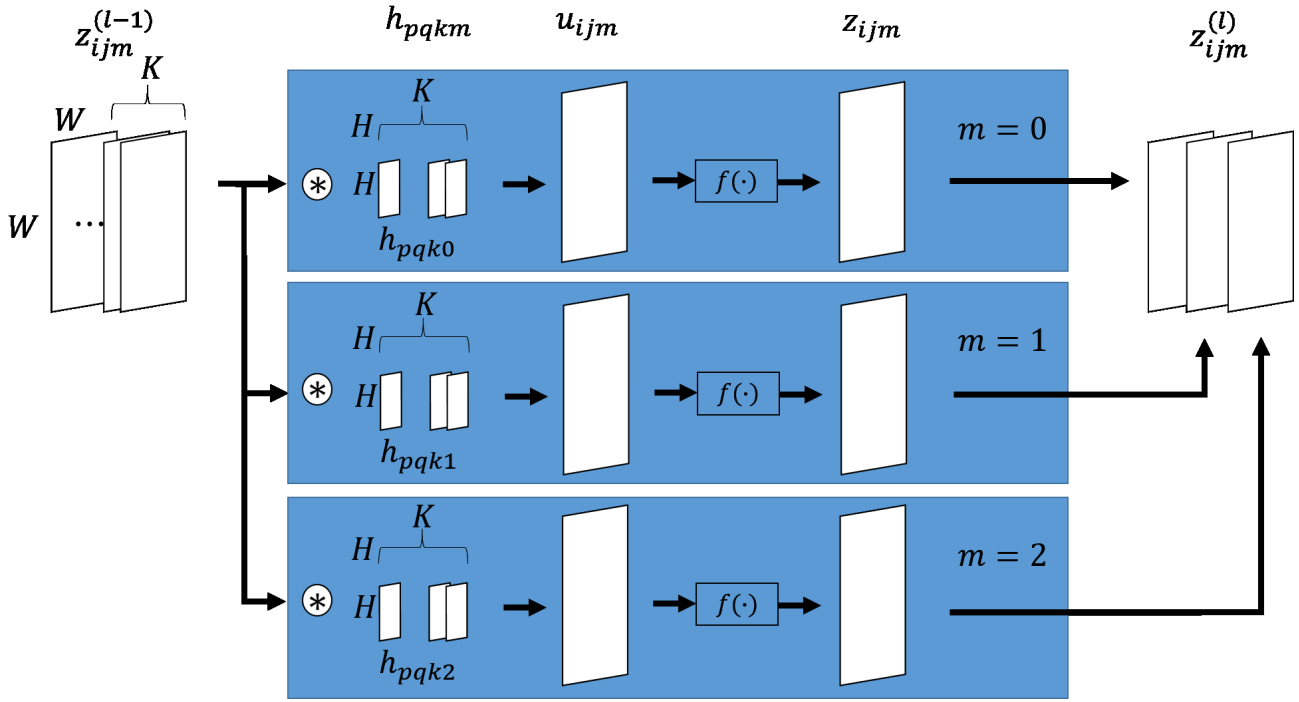
理論研究の進捗

畳込みニューラルネットワークの理論について

プログラミングの進捗

プログラム実行環境の見直し
データセット作成

畳み込み層



- 各フィルタについて並列に計算され,u_{ijm} が出力される．各チャンネルについて並列に画像とフィルタの畳み込みを行い全チャンネルにわたり加算する．

$$u_{ijm} = \sum_{p=0}^{K-1} \sum_{q=0}^{H-1} \sum_{q=0}^{H-1} z_{i+p,j+q,k}^{(l-1)} h_{pqkm} + b_{ijm} \tag{1}$$

畳み込み層

活性化関数への適応

$$z_{ijm} = f(u_{ijm}) \tag{2}$$

式 (1), (2) は，層間に特別な構造を持つ単層ネットワークとして表現できる.

- 入力のユニット数：$W \times W \times K$
- 出力のユニット数：$W \times W \times M$

畳み込みの計算の局所性により，出力層のユニット 1 つは入力層の $W \times W \times K$ 個のユニットとのみ結合する.

- その時の重みがフィルタ係数 h_{pqkm}
- 同一チャネルの全ユニットで共有されるため
  重み共有 (weigh sharing, weight tying) と呼ばれる

プーリング

プーリング

プーリング層は通常，畳み込み層の直後に設置され，プーリング層は畳み込み層で抽出された特徴の位置感度を低下させる働きがある．

サイズ $W \times W \times K$ の入力画像上で画素 (i, j) を中心とする $H \times H$ の正方領域とり，この中に含まれる画素の集合を P_{ij} で表す．P_{ij} 内の画素についてチャネル k ごとに独立に，H^2 個ある画素値を使って 1 つの画素値 u_{ijk} を求める．

- 最大プーリング (max pooling)

$$u_{ijk} = \max_{(p,q) \in P_{ij}} z_{pqk} \tag{3}$$

- 平均プーリング (average pooling)

$$u_{ijk} = \frac{1}{H^2} \sum_{(p,q) \in P_{ij}} z_{pq} \tag{4}$$

プーリング

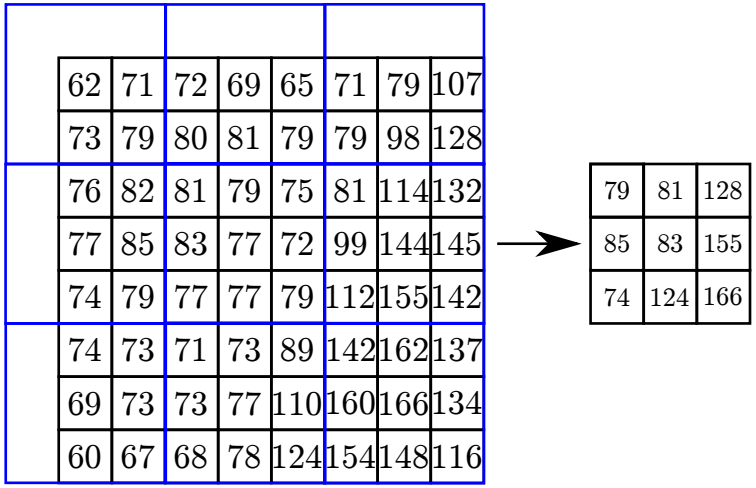


Figure : プーリングの例．プーリングサイズ 3×3，ストライド $s = 3$，ゼロパディングで最大プーリングを行った場合

平均プーリングや最大プーリングなどのプーリングを含む一般性を持った表記として，次の Lp プーリング (Lp pooling) がある．

$$u_{ijk} = \left(\frac{1}{H^2} \sum_{(p,q) \in P_{ij}} z_{pqk}^{P} \right)^{\frac{1}{P}} \tag{5}$$

$P = 1$ で平均プーリング，$P = \infty$ で最大プーリングが表現できる．

GPU を用いた学習実行時の演算処理高速化

- caffe がサポートしている NVIDIA の CUDA を使う（導入済み）
- Deep Learning 用の CUDA ライブラリ cuDNN を使用う
  （デベロッパー登録申請認可）

GPU による並列計算

- Ubuntu 14.04 のマシンにライブラリをインストールし Caffe をコンパイルし直し成功
- 正確な計測は行っていないが，明らかに学習速度の向上が見られた.

データセット作成のためのプログラムの作成

- 独自のデータセットを作成する
- 写真や動画から人の顔を切り出し，データセットを作成，そして学習という一連の操作を行う計画.
- まずデータセットを作成するためには大量の人物が一人ずつ写った写真データを用意する必要がある.

顔検出プログラムの作成

- OpenCV に実装されている顔検出アルゴリズムを用いて，人物が写った部分を自動的に切り出すプログラムを作成した.
- 大量にある画像データや動画から切り出す.

画像データからの切り出し

実行は以下のようにソースとなる画像が含まれたディレクトリと，出力先のディレクトリを指定する.

```
$ python facedetect.py src output
```

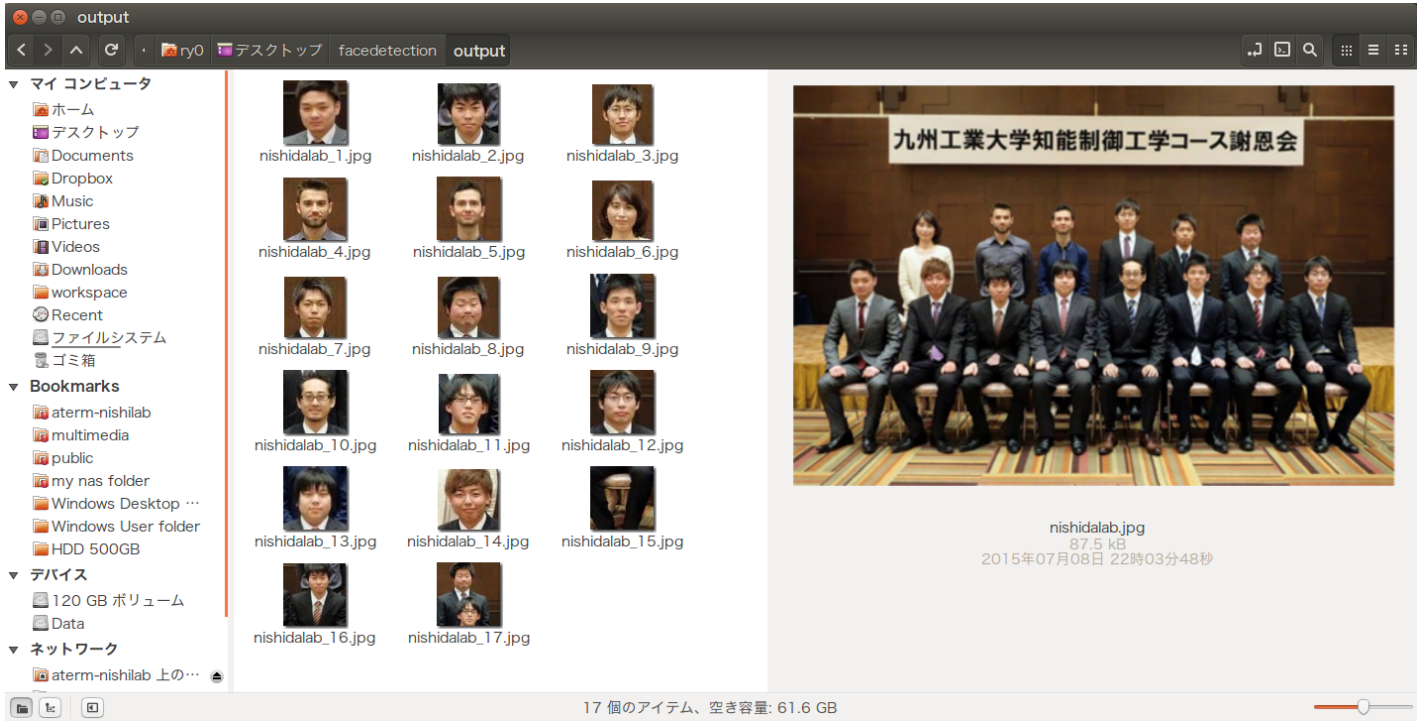
Python で実装した機能を以下に示す.

- ディレクトリを指定したらそのディレクトリに入っているすべての画像ファイルを顔認識して切り取る.
- 指定したディレクトリの中に含まれているサブディレクトリの中もすべて探索してすべて取り込む.
- プログレスバーを設置して進捗を可視化

```
66% (925 of 1522) |#################### | Elapsed
     Time: 0:00:07 ETA:  0:04:26
```

- 壊れた画像ファイルを読み込まれた場合でも例外処理.

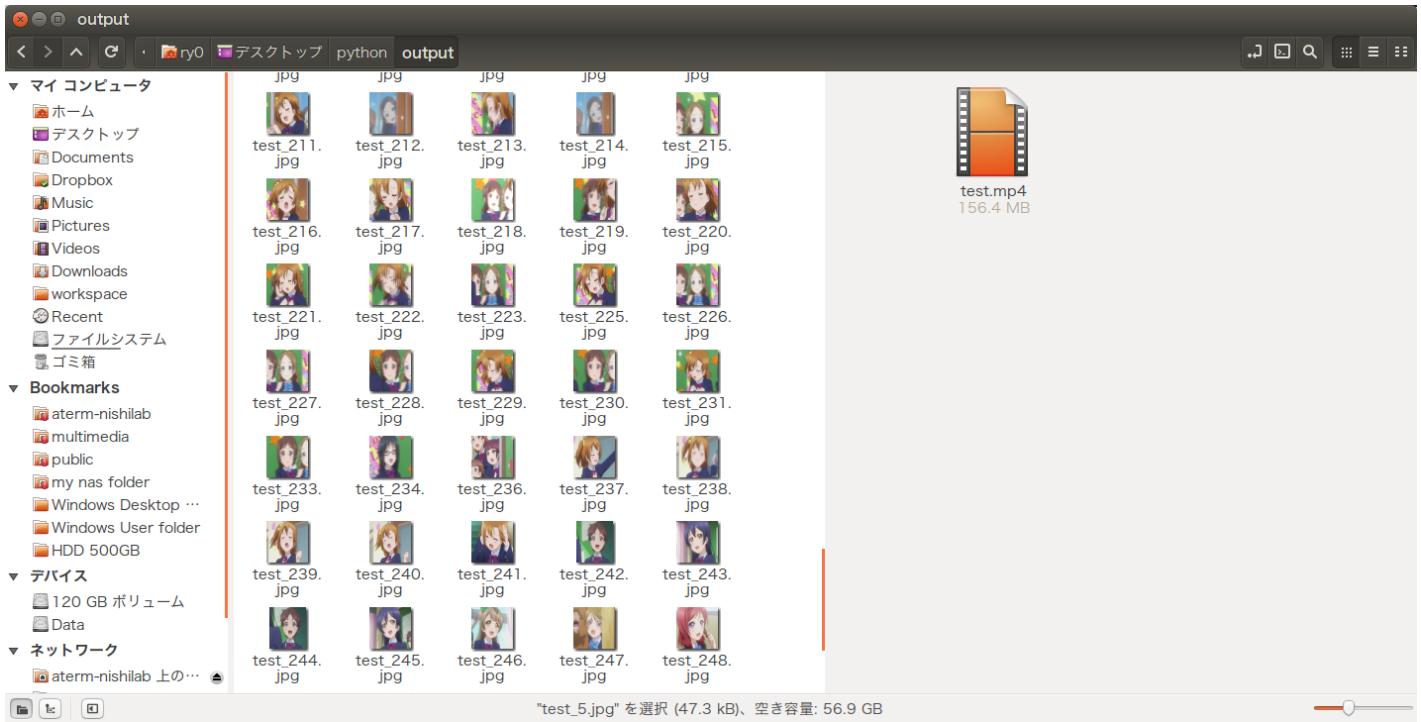
画像データからの切り出し



動画からの切り出しプログラムも画像場合とほぼ同様の機能を有しており，ソースは動画ファイルを直接指定する．

```
1 $ python facedetect_video.py src/test.mp4 output
```

動画データからの切り出し



データセットの作成支援

このプログラムを用いて，アニメのキャラクターやアイドルグループのメンバーなどの写真から顔の部分のみを切り出すことが可能となる．
しかし切り出したあとは自分で顔を識別してフォルダ分けする作業が必要である．

今後の課題

理論研究

CNN の詳細な調査
プーリングの理論

プログラミング

データセットの作成