

知能システム学特論レポート

(DL2 班) Caffe on Ubuntu

2015 年 6 月 29 日

1 報告者

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

2 進行状況

- 理論研究
- 順伝播型ネットワークについて

3 理論研究

3.1 単純型細胞と複雑型細胞

単純型細胞と複雑型細胞のモデルを Fig. 1 に示す. (a),(b) は 4×4 の入力層が中間層のそれぞれと結合していることを示しており, (c),(d) は入力パターンの位置変化に伴う中間層および出力層の状態の変化を示している.

中間層は単純型細胞, 出力層は複雑型細胞のモデルとなっており, 中間層は位置選択性が厳密であるが, 出力層は入力パターンが少しずれても活性化する. 全体への入力 (c) から (d) のように変化すると, 中間層でも図のように活性化するユニットが変化するが, 出力層のユニットは活性化したままである. これは出力層は中間層のユニットが一つでも活性化されていれば活性化するためである.

この2つの細胞をモデル化した二層構造のペアを繰り返す構造が畳み込みニューラルネットに用いられている. 物体カテゴリ認識は長年コンピュータには難しいとされてきたが, 近年畳み込みニューラルネットによってそれも解決されつつある. 神経科学の分野では, 多層の畳み込みニューラルネットが霊長類の脳の高次視覚野と似た振る舞いを示すことが示されている.

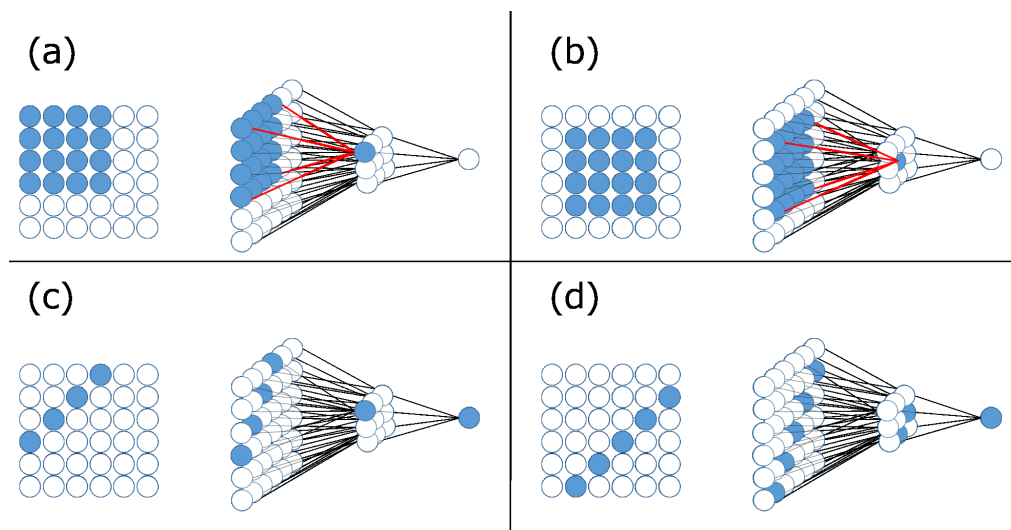


Fig.1 単純型細胞と複雑型細胞のモデル

3.2 全体の構造

4 プログラム

4.1 caffeNet の構造

caffe では訓練済みのニューラルネットワークをダウンロードし、使用することができる。先日、示したサンプルの画像認識に用いた CNN の構造を Fig.2 に示す。赤色は畳込み層、オレンジ色はプーリング層、青色は正規化層、紫色は全結合層を表している。また、活性化関数にはソフトマックス関数を用いている。すなわち、この CNN は 5 つの畳込み層、3 つのプーリング層、2 つの正規化層、そして 3 つの全結合層からできている。

Fig.2 を出力するには、caffe のディレクトリ直下で、次のコマンドを実行する。

```
python/draw_net.py models/bvlc_reference_caffenet/train_val.prototxt caffeNet.png --
rankdir BT
```

4.2 中間層の出力

畳込み 1 層目のフィルタと Fig. に示す入力画像を入力したときの出力を得ることができた。1 層目のフィルタを Fig.4, その出力を Fig.5 に示す。畳込み 1 層目は全部で 96 個の 11×11 のフィルタから構成されており、ストライドが 4 である。従って、出力される画像のサイズは 55×55 であり、96 個の画像が出力される。

5 今後の課題

- 理論研究を進める。
- Caffe を使いこなす

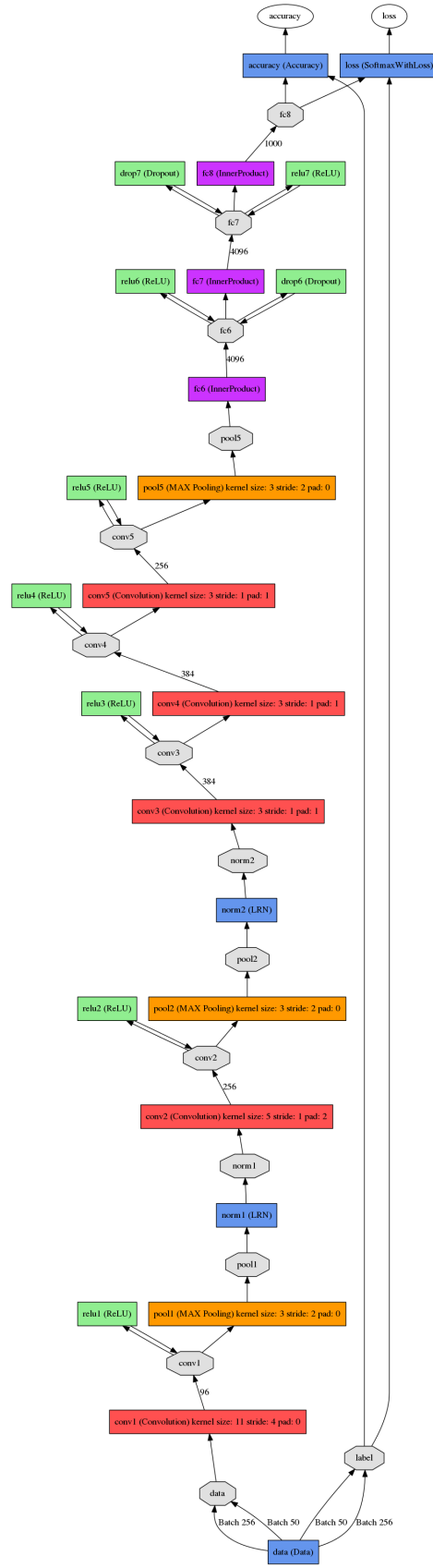


Fig.2 caffeNet

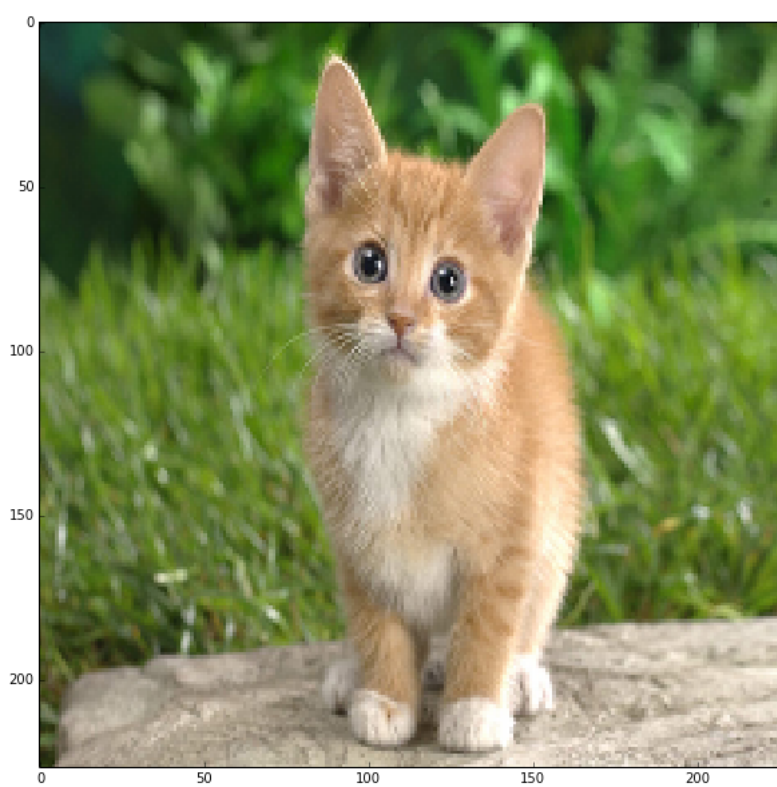


Fig.3 入力画像

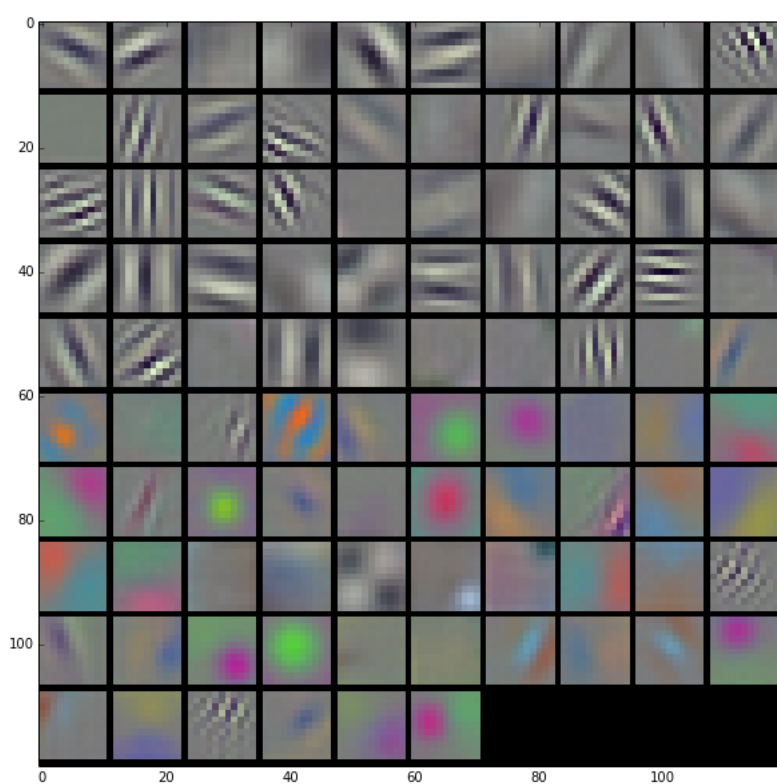


Fig.4 畳込み層 1 層目のフィルタ

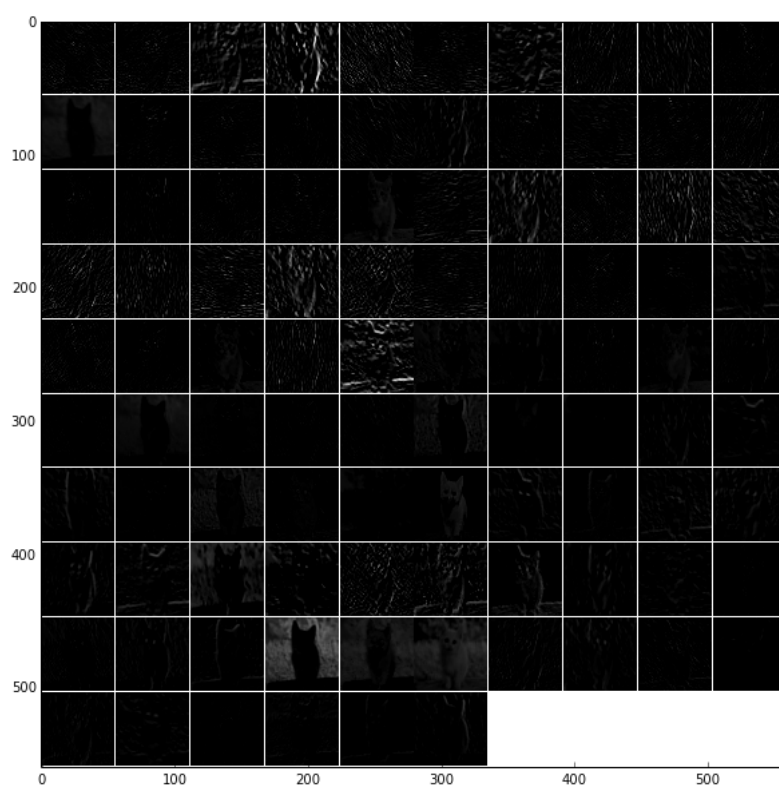


Fig.5 畳込み層 1 層目のフィルタ出力