

知能システム学特論レポート

(DL2 班) Caffe on Ubuntu

2015 年 7 月 13 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

- 畳み込みネットワークと正規化層の理論について
- データセットの作成準備

3 理論研究

3.1 畳込み層

3.2 単一チャネルの正規化

単一チャネルの画像 x_{ij} に対し，プーリングと同様，画素 (i, j) を中心とする $H \times H$ の正方領域 P_{ij} を考える．減算正規化とは，入力画像の各画素濃淡から， P_{ij} に含まれる画素の濃淡の平均，つまり $\bar{x}_{ij} = \sum_{(p,q) \in P_{ij}} x_{i+p,j+q}$ を差し引く．

$$z_{ij} = x_{ij} - \bar{x}_{ij} \quad (3.1)$$

ここで差し引く \bar{x}_{ij} には，重み付き平均

$$\bar{x}_{ij} = \sum_{(p,q) \in P_{ij}} w_{pq} x_{i+p,j+q} \quad (3.2)$$

を使う場合もある．その場合 w_{pq} は

$$\sum_{(p,q) \in P_{ij}} w_{pq} = \sum_{q=0}^{H-1} \sum_{q=0}^{H-1} w_{pq} = 1 \quad (3.3)$$

であり、領域の中央で最大値をとり、周辺部へ向けて低下するようなものとする。領域の中央部をより重視し、周辺部の相対的な影響を少なくするためである。

同じ領域内で、さらに画素値の分散を抑える操作が除算正規化である。 P_{ij} 内の画素値の分散は

$$\omega_{ij}^2 = \sum_{(p,q) \in P_{ij}} w_{pq} (x_{i+p,j+q} - \bar{x}_{ij})^2 \quad (3.4)$$

となるが、減算正規化を施した入力画像をこの標準偏差で割る。

$$z_{ij} = \frac{x_{ij} - \bar{x}_{ij}}{\omega_{ij}} \quad (3.5)$$

この計算をそのまま行くと、濃淡変化が少ない局所領域ほど濃淡変化が増幅され、ノイズが強調される。そこで、入力画像のコントラストが大きい部分にのみ適用するために、ある定数 c を設定し、濃淡の標準偏差がこれを下回る ($\omega_{ij} < c$) で除算する

$$z_{ij} = \frac{x_{ij} - \bar{x}_{ij}}{\max(\omega_{ij}, c)} \quad (3.6)$$

や、同様の効果が ω_{ij} に応じて連続的に変化する

$$z_{ij} = \frac{x_{ij} - \bar{x}_{ij}}{\sqrt{\omega_{ij} + c}} \quad (3.7)$$

を使う。減算正規化および式(3.6)による除算正規化の計算例を図3.1に示す。これらの正規化の結果、画像の画素値は負の値をとり得るため、画素値の最大値と最小値が $[0, 255]$ の範囲に収まるように画素値を線形変換している。

3.3 プーリング

4 プログラミング

4.1 画像を学習させて分類器を作るまでの流れ

データセットを作成し、学習を行ってモデルを作成するまでの流れを Fig. 1 に示す。Fig. 1 より (1) の大量の写真データまたは動画からの切り出しに関しては

5 今後の課題

- 理論研究を進める。
- データセットの作成，学習の実行

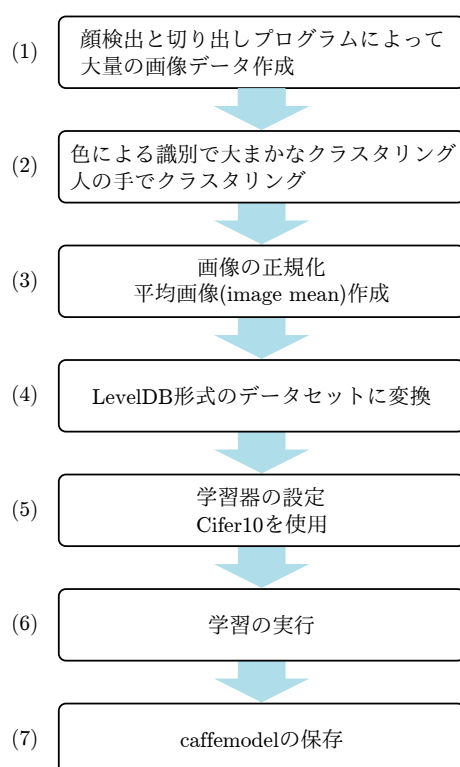


Fig.1 モデルを作成するまでの流れ