

1 課題 1

ネットワークの構造を変更し，認識精度の変化を確認する．

1.1 結果

実行結果を表 1 に示す．バッチサイズは 64，エポック数は 10，学習率は 0.01，最適化手法はモメンタム SGD とする．

表 1：ネットワーク構造の変更

| | 認識精度 | 中間層のユニット数 | 層数 | 活性化関数 | カーネルサイズ |
|----|--------|------------------------------|----|-----------|---------|
| 1 | 0.6761 | 1024 | 5 | ReLU | 3×3 |
| 2 | 0.6763 | 1024, 1024 | 6 | ReLU | 3×3 |
| 3 | 0.6784 | 2048, 2048 | 6 | ReLU | 3×3 |
| 4 | 0.6662 | 2048, 2048 | 6 | LeakyReLU | 3×3 |
| 5 | 0.7141 | 2048, 2048 | 6 | ELU | 3×3 |
| 6 | 0.7145 | 2048, 2048 | 6 | ELU | 5×5 |
| 7 | 0.6599 | 2048, 2048 | 6 | ReLU | 5×5 |
| 8 | 0.67 | 2048, 2048, 2048 | 7 | ReLU | 5×5 |
| 9 | 0.6755 | 1024, 1024, 1024 | 7 | ReLU | 5×5 |
| 10 | 0.6678 | 1024, 1024, 1024, 1024 | 8 | ReLU | 5×5 |
| 11 | 0.6594 | 1024, 1024, 1024, 1024 | 8 | LeakyReLU | 5×5 |
| 12 | 0.6698 | 1024, 1024, 1024, 1024 | 8 | LeakyReLU | 3×3 |
| 13 | 0.6712 | 512, 512, 512, 512 | 8 | LeakyReLU | 3×3 |
| 14 | 0.6714 | 512, 512, 512, 512, 512 | 9 | LeakyReLU | 3×3 |
| 15 | 0.7049 | 512, 512, 512, 512, 512 | 9 | ELU | 3×3 |
| 16 | 0.7113 | 1024, 1024, 1024, 1024, 1024 | 9 | ELU | 3×3 |

1.2 考察

活性化関数以外の構造を同じにして ReLU と LeakyReLU の精度を比較すると，3 行目と 4 列目では 0.00122, 10 行目と 11 行目では 0.0084 となり，ReLU と LeakyReLU の性能に大きな差はないと考える．また，同じ ReLU の発展形である ELU は LeakyReLU と比べより高い認識精度を出している．

2 課題 2

学習設定を変更し、認識精度の変化を確認。

2.1 結果

実行結果を表 2 に示す。層数は 5 層として、中間層のユニット数は 1024、活性化関数は ReLU、カーネルサイズは 3×3 で固定する。

表 2：学習設定の変更

| 認識精度 | バッチサイズ | 学習回数 | 学習率 | 最適化手法 |
|--------|--------|------|------|-----------|
| 0.6732 | 64 | 10 | 0.01 | SGD |
| 0.6878 | 64 | 20 | 0.01 | SGD |
| 0.6952 | 64 | 30 | 0.01 | SGD |
| 0.7091 | 128 | 40 | 0.01 | SGD |
| 0.7076 | 256 | 50 | 0.01 | SGD |
| 0.4899 | 256 | 50 | 0.01 | Adadelata |
| 0.701 | 256 | 50 | 0.01 | Adagrad |
| 0.5323 | 256 | 50 | 0.01 | Adam |

2.2

モーメンタムを使用しない Adadelata や Adagrad は SGD と比べて認識精度が低下している。

3 課題 3

認識精度が向上するようにパラメータを変更。

3.2 結果

表 3：学習パラメータとネットワーク構造の変更

| 認識精度 | 中間層のユニット数 | 層数 | 活性化関数 | カーネルサイズ | バッチサイズ | 学習回数 | 最適化手法 |
|--------|------------------|----|-------|--------------|--------|------|-------|
| 0.6795 | 1024 | 5 | ReLU | 3×3 | 256 | 30 | SGD |
| 0.716 | 1024 | 5 | ELU | 3×3 | 256 | 30 | SGD |
| 0.6965 | 1024, 1024, 1024 | 7 | ELU | 3×3 | 256 | 30 | SGD |
| 0.7247 | 2048, 2048 | 6 | ELU | 3×3 | 128 | 30 | SGD |

3.3 考察

層数を増やすと認識精度は向上するが上限があり，ある一定以上層数を増やしても認識精度は向上しなくなる．学習回数も同様に，一定以上増やしても認識精度は向上せず，逆に認識精度が低下する．これらは，過学習や勾配消失によるものであり，層数と学習回数は正確に見極める必要がある．