

8 ニューラルネットワークを用いた数字認識の研究

植野 凱虎

指導教員 飯坂 覚

1 はじめに

本研究では機械学習やニューラルネットワークについて知識を身に着け、手書き数字の認識方法を通してニューラルネットワークのプログラムを作成します。

2 数字認識のニューラルネットワーク

図 1 のとおりニューラルネットワークは、入力層から入った数字がノードをつなぐエッジを通して伝搬し、中間層である隠れ層を介して出力層に伝わる仕組みになっています。これに、エッジやノードの重みを調節する誤差逆伝搬法¹⁾を用いて、数字認識をさせています。図 1 では、手書き数字データ「7」を学習させており、ニューラルネットワークを通した結果が「1」と間違えて出力された例です。そこで誤差逆伝搬法を用い、データが通ってきたノードとエッジの重みを大きくし、再度「7」の画像を入力したとき重みが大きいノードとエッジは通らず他のところに行くという仕組みになっています²⁾。これを「7」と認識されるまで繰り返します。

3 開発環境

次のような環境のもと C#で数字認識プログラムを作成しました。手書き数字の入力にはペンタブレットを使用します。

開発環境 : Windows10

開発言語 : C#

使用 IDE : Visual Studio

使用機材 : ペンタブレット (WACOM : CTL-490)

4 数字認識プログラム

数字認識プログラムは学習プログラムと認識プログラムから構成されています。

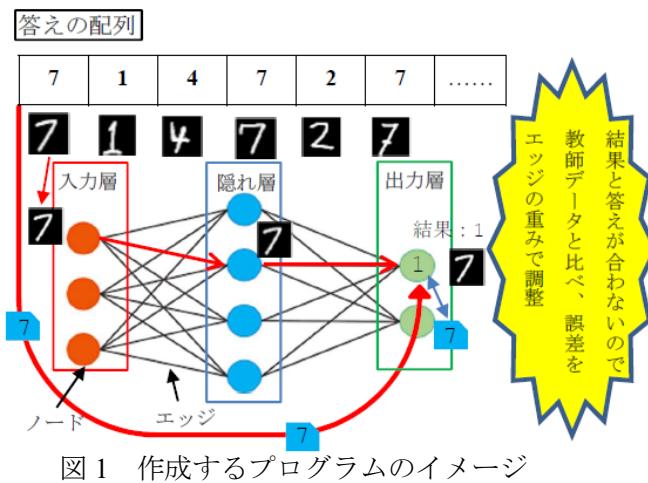
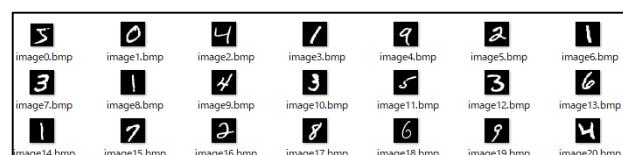


図 1 作成するプログラムのイメージ

4.1 学習プログラム

学習プログラムでは、あらかじめ手書き数字を多く学習させる必用があります。それには一般に公開されている「0~9」の画像 60,000 枚手書き数字データ「MNIST」を用いました。手書き数字データの目視化した様子を図 2 に示します。これに対し、図 2 と同じ順番の正解数字データセット(図 3)も用意します。このうち先頭から 50,000 枚が学習データで残り 10,000 枚がテストデータとなっています。

図 4 は隠れ層のノード数を 100 としたときの学習の様子で、x 軸が学習した数字の枚数、y 軸が

図 2 「MNIST」による手書き数字画像
を可視化したもの

| |
|--|
| +0 +1 +2 +3 +4 +5 +6 +7 +8 +9 +A +B +C +D +E |
| 000000 00 00 08 01 00 00 FA 60 05 00 04 01 09 02 01 03 |
| 000010 01 04 03 05 03 06 01 07-02 08 06 09 04 00 09 01 |

正解データセット

図 3 手書き数字画像と同じ順番の
正解数字データセット

正答率を表しています。数字を読み込むと徐々に正答率が上がっています。図 5 はこの学習プログラムにテストデータ 10,000 枚を入力したときの結果で、96.65% で正しく数字を認識したという結果です。なお、このエッジとノードの情報は保存できます。

4.2 認識プログラム

右の余白に数字を書いて認識を行い、結果は左に表示します。図 6 ではタブレットで 2 と書いたところ、左に 2 と認識されている様子です。

5 考察

友人にも協力してもらい手書き認識を行ったところ、書いた数字と同じ値が出る確率が 44% (サンプル 100 回) になりました。この結果を踏まえ、正答率を上げるには隠れ層を増やせば良いと考えました。

表 1 に隠れ層のノードの数を 100、200、400、700、1000 の 5 つに変えて実験を行った結果を示

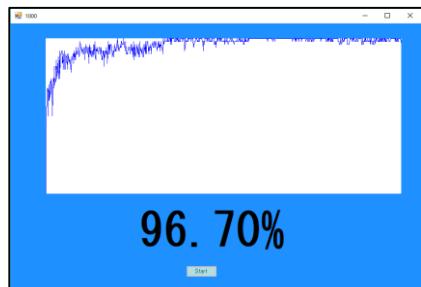


図 4 学習中の様子

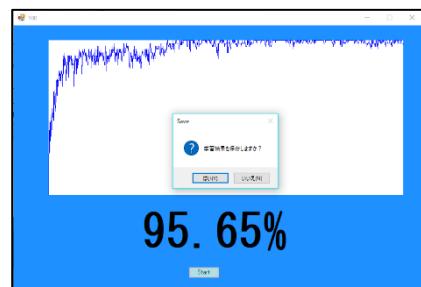


図 5 テスト後の結果表示

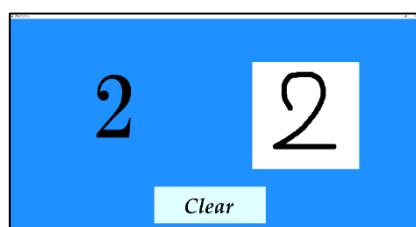


図 6 認識の様子

表 1 実験結果

| ノード数 | 学習時間 | テスト結果(%) | 実験結果(%) |
|------|---------|----------|---------|
| 100 | 3 : 48 | 95.26 | 44 |
| 200 | 9 : 29 | 96.07 | 48 |
| 400 | 22 : 06 | 96.62 | 51 |
| 700 | 30 : 30 | 96.69 | 56 |
| 1000 | 39 : 50 | 96.70 | 58 |

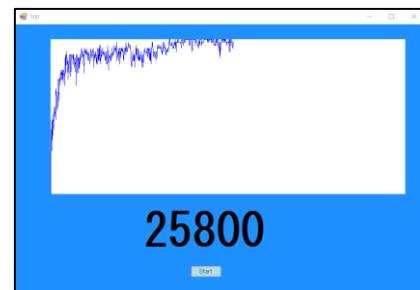


図 7 隠れ層のノード数 1,000 個のテスト結果

します。学習時間は要しているもののテスト結果、並びに実験結果は上昇していることがわかります。また、図 7 から早い段階で数字の正答率が上昇していることがわかります。

認識は予想どおりでしたが、全体的に精度が向上しないという結果が出ました。その原因としては、認識プログラムの余白部分に 500×500 ピクセルで数字を書いてもらうのですが、それを 28×28 ピクセルに圧縮する過程で数字の特徴に欠落が出ている可能性があります。よって、直接 28×28 ピクセルで記入するか、圧縮プログラムを工夫する必要があります。

6. 終わりに

卒業研究でニューラルネットワークについて研究することにより人工知能で使われている機械学習がどのようなものか知ることが出来ました。私の研究目標は機械学習やニューラルネットワークについて学習することでしたので、目標を達成することができました。

【参考文献】

- 誤差逆伝播法を宇宙一わかりやすく解説してみる
<https://www.yukisako.xyz/entry/backpropagation>
- “日経ソフトウェア 2018 年 9 月号” ,
日経 BP 社 (2018)