

誤差逆伝播法を用いた画像のカラー化に関する研究

研究者：下吉 賢信 指導教員：濱川 恭央

1. 目的

誤差逆伝播法（BackPropagation, BP 法）を用いて、白黒画像のカラー化を行うことを目的とする。

2. 理論

2.1 ニューラルネットワーク (Neural Network, NN)

人間の脳は100億から140億個の神経細胞が互いに繋がり、巨大なシステムを構築している。このような神経細胞を構成素子とし、計算機上でシミュレートすることを目指した回路網をニューラルネットワークと呼ぶ。ニューラルネットワークには様々なモデルがあるが、そのうちのひとつとして、パーセプトロンが挙げられる。

2.1.1 単層パーセプトロンとは

単層パーセプトロンとは、NNのひとつであり、複数の信号を入力として受け取り、ひとつの信号を出力するものである。その出力は1か0の2値であり、1を出力することを、「ニューロンが発火する」と表現することがある。図1に2つの信号を入力として受け取るパーセプトロンを示す。 x_1, x_2 は入力信号、 y は出力信号、 w_1, w_2 は重みと呼ばれ、式1のような挙動を示す。図1における丸(○)は、ニューロンと呼ばれる。ここで、 θ は閾値を示す。

$$y = \begin{cases} 0 & (w_1x_1 + w_2x_2 \leq \theta) \\ 1 & (w_1x_1 + w_2x_2 > \theta) \end{cases} \quad (1)$$

例えば、 $(w_1, w_2, \theta) = (0.5, 0.5, 0.8)$ のようにパラメータを設定すると、AND ゲートと同じ挙動を期待できる。

2.1.2 バイアス

式1の θ を $-b$ として、式2に変換できる。

$$y = \begin{cases} 0 & (b + w_1x_1 + w_2x_2 \leq 0) \\ 1 & (b + w_1x_1 + w_2x_2 > 0) \end{cases} \quad (2)$$

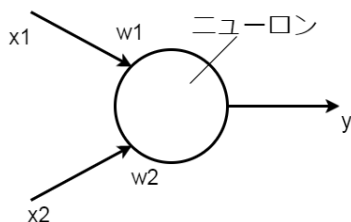


図1 パーセプトロン

ここで b をバイアスと呼ぶ。式2で示されるように、パーセプトロンでは入力信号に重みが乗算された値とバイアスの和が計算され、その値が0を上回ると1を出力し、そうでなければ0を出力する。このことから、バイアスはニューロンの発火のしやすさの度合いであるといえる。

2.1.3 順伝播型ニューラルネットワーク

単層パーセプトロンでは、線形非分離な問題を解くことができない。だが、パーセプトロンを多層にし、出力を実数値にすることにより、これを解決した。こうして出来たモデルを順伝播型ニューラルネットワーク (Feedforward neural network, FFNN) と言う。図2にFFNNの例を示す。ここで、一番左の層を入力層、一番右の層を出力層、それぞれに挟まれた層を中間層または隠れ層と呼ぶ。中間層は数に特に決まりは無く、それぞれの層のニューロンの数にも制限は無い。

それぞれのニューロンは活性化関数を内包しており、それにより発火の仕方が決定される。

2.1.4 活性化関数

活性化関数を $h()$ で表すとすると、式2は以下の式に書き換えられる。

$$a = b + w_1x_1 + w_2x_2 \quad (3)$$

$$y = h(a) \quad (4)$$

$$h(x) = \begin{cases} 0 & (x \leq 0) \\ 1 & (x > 0) \end{cases} \quad (5)$$

これを明示的に示すとすると、図3のようになる。

図に示され通りこれまでのニューロンの丸の中に、活性化関数によるプロセスがある。つまり、重みつき信号の和が a というノードになり、活性化関数 $h()$ によって y というノードに変換されることが示されている。ちなみに式5は、ステップ関数と呼ばれる活性化関数である。

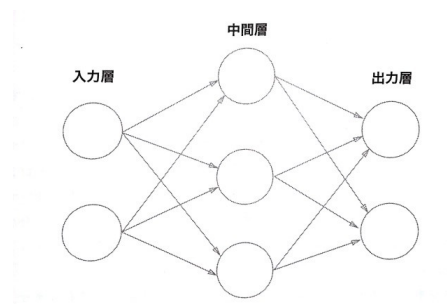


図2 順伝播型ニューラルネットワーク

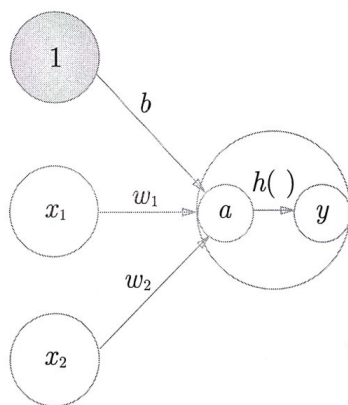


図3 活性化のプロセス

2.1.5 活性化関数

本研究で利用した活性化関数を示す。

- シグモイド関数

ニューラルネットワークでよく用いられる活性化関数のひとつであり、式6で表される。出力が0から1の実数値である。

$$h(x) = \frac{1}{1 + \exp(-x)} \quad (6)$$

- ランプ関数 (ReLU)

式7で表される。後述する勾配消失問題を解決する関数として近年になって利用頻度が上がった。

$$h(x) = \max(0, x) \quad (7)$$

2.2 学習

機械学習の問題では、訓練データとテストデータの2つに分けて、学習や実験などを行うのが一般的である。その場合、まず訓練データのみを用いて学習を行い、最適なパラメータを探索する。そして、テストデータを使って、その訓練したモデルの実力を評価する。訓練データとテストデータを分けるのは、モデルの汎用的な能力を評価するためである。

2.2.1 損失関数

損失関数はニューラルネットワークの性能の悪さを示す指標である。現在のニューラルネットワークが教師データに対してどれだけ適合していないかと言うことを表す。損失関数にマイナスをかけた値は、どれだけ性能がよいかと言う指標として解釈できるため、どちらを指標としたとしても行うことは本質的に同じである。損失関数として用いられる関数をいくつか以下に示す。入力するデータ点が x で学習させたい関数は $f(w, x)$ 、教師データを t とする。

- 二乗損失

式8に示す。出力データが教師データからずれる

にしたがって、そのずれの二乗の損失を与える関数である。負のズレも、正のズレと同じ扱いである。

$$L = (t - f(w, x))^2 \quad (8)$$

- Huber 損失

式9に示す。この損失関数は、ズレがある範囲内ならば二乗損失を、それより外なら直線状に増加する損失を与える。学習データにノイズが入った場合、それに引っ張られる学習を抑制することができる。

$$L = \begin{cases} (t - f(w, x))^2 & (f \in [t - \delta, t + \delta]) \\ 2\delta(|t - f| - \frac{\delta}{2}) & (otherwise) \end{cases} \quad (9)$$

- ϵ -許容損失

式10に示す。ズレが $\pm\epsilon$ 以内であれば、損失を与えない関数である。うまく回帰が出来ていないデータに対してのみ重点的に学習を進めていくことになり、細かい部分は学習を行わない。

$$L = \max(|t - f(w, x)| - \epsilon, 0) \quad (10)$$

2.2.2 勾配法

機械学習の問題の多くは、学習の際に最適なパラメータを探索することである。ニューラルネットワークも同様に最適なパラメータを学習時に見つけなければならない。最適なパラメータというのは、損失関数が最小値をとるときのパラメータの値である。喪失関数は、最小値まで勾配を持つ。しかし、一般的に損失関数は複雑で、パラメータ空間は広大であるため、どこに最小値をとるのか、見当がつかない。そこで、勾配をうまく利用して関数の最小値を探すことを目指したものが勾配法である。

勾配法では、現在の場所から勾配方向に一定の距離だけ進む。そして、移動した先でも同様に勾配を求め、また、その勾配方向へ進むと言うように、繰り返し勾配方向へ移動する。このように勾配方向へ進むことを繰り返すことにより、関数の値を徐々に減らすのが勾配法である。ニューラルネットワークの学習では勾配法がよく用いられる。例えば、式11についての勾配法を数式で表すと式12のように書くことが出来る。

$$f(x_0, x_1) = x_0^2 + x_1^2 \quad (11)$$

$$\begin{aligned} x_0 &= x_0 - \eta \frac{\partial f}{\partial x_0} \\ x_1 &= x_1 - \eta \frac{\partial f}{\partial x_1} \end{aligned} \quad (12)$$

式12の η は学習率と呼ばれ、一度の学習でどれだけパラメータを更新するかを決める。

式12は一回の更新式を示しており、このステップを繰り返し行う。つまり、ステップごとに変数の値を更新していき、そのステップを何度か繰り返すことによって

徐々に誤差関数の値を減らしていく。ここでは、変数が二つの場合を示しているが、変数の数が増えても、同じような式によって更新される。

勾配法は最小値を探す場合を勾配降下法、最大値を探す場合を勾配上昇法と言う。

2.2.3 誤差逆伝播

誤差が出力とは逆に伝播していくため BP 法と名付けられた。図 4 のように望ましい出力との誤差を逆伝播させることにより次第に結合係数を変化させ、最終的に正しい解を得られるようにする手法である。

誤差逆伝播を行う為には、活性化関数が微分可能であることが必要不可欠である。ステップ関数は微分不可能であったため、それに近い出力を持ち、且つ微分可能であるシグモイド関数が考案された。シグモイド関数の微分は式 13 に表されるように、単純である。

$$f(x)' = \left(\frac{1}{1 + \exp(-x)} \right)' = f(x)(1 - f(x)) \quad (13)$$

2.2.4 勾配消失問題

誤差逆伝播では出力値の微分が学習の係数にかかっている。シグモイド関数が活性化関数だった場合、入力が大きくなればなる程微分の値は小さくなっていくため、学習が滞ってしまうことがある。これを、勾配消失問題と呼ぶ。複数の層でこれが起こった場合、誤差逆伝播で誤差を伝えていくうちに、誤差が正しく伝播されなくなる。これを解決する為の関数が ReLU である。ReLU は入力値が正であれば、入力値をそのまま出力する。このとき微分は必ず 1 の値を取り、出力の値が学習を滞らせることが一切無くなる。

2.2.5 スパース化

ReLU のもうひとつの特徴は、ネットワークのスパース化に貢献すると言う点である。ニューラルネットワークがスパースであるとは、出力が 0 であるユニットが多数存在することを指す。あるユニットへのい入力が 0 未満だった場合、ReLU の微分は 0 であるため、これが学習の係数に掛かり、そのユニットは全く学習が行われな

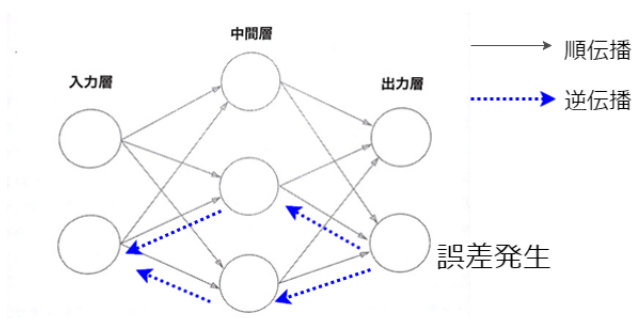


図 4 誤差逆伝播

表 1 実行環境

OS	Windows 7 Enterprise
Processor	Intel Core i7-4820k CPU @ 3.70GHz
RAM	32.0GB
GPU	GeForce GTX 1080Ti
GPU Memory	11.0GB

くなる。その後、そのユニットに対する学習は行われなくなり、残ったユニットだけで学習が行われる。ニューラルネットワークは本来冗長であるため、学習するパラメータを厳選するという点においても ReLU は貢献している。

3. 実験

実行環境を表 1 に示す。

3.1 研究手法

画像は、ピクセルごとに RGB のパラメータを持つ。RGB とは、赤、緑、青、の 3 つの原色を混ぜて幅広い色を再現する手法である。RGB それぞれを FFNN に学習させてカラー化を行うことは冗長であり、計算機の資源不足に陥ることも考えられる。そのため、RGB を HSV 色空間へ変換してから学習させる手法を提案する。HSV 空間とは、色相、彩度、明度で色を表現する空間であり、明度 V はカラー化の前後で変化しない。そのため、学習させる必要のあるものは、H,S の 2 つのみとなり、より簡易で高性能なニューラルネットワークを構成できると考えた。この手法を基に、2 つの実験を行った。

3.2 最適なモデルの考察

まず、図 5 のようなモデルを考案した。青色の四角い枠で囲まれた小さな NN を PNN(a Partial Neural Network) と呼ぶ。この PNN をつなぎ合わせていくのだが、その途中に図 5 に示されるように、結合を無視し、PNN の出力との和をとる層がある。これを skip connection と呼ぶ。FFNN は、層が深くなるにつれ入力

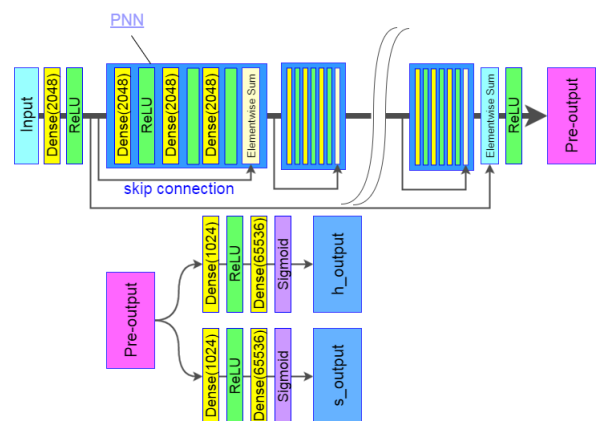


図 5 モデル

データの詳細な部分が曖昧になっていく。その解決を図るため、skip connection を用いて、浅い層の情報を加味した NN を構成することを目指し、実装した。本実験の目的は最適な PNN の数を検討することである。出力画像の評価指標は SSIM を用いた。SSIM は、輝度、コントラスト、周囲のピクセル平均、分散、共分散をとることで、ピクセル単体でなく周囲との相関を取り込んだ指標であり、人間の資格特性を反映した評価の手法である。

- ワープロ等の DTP ソフトで作成すること。
- 用紙は A4 縦で、本文は 2 段組の横書きとする。
- 上余白 25mm, 下余白 20mm, 左右余白 20mm とする。すべてのページで同じ余白とする。
- 文字の大きさは、10pt を基本とする。1 ページあたりの行数は 50 行とし、1 段の 1 行あたりの文字数は 25 文字とする。
- 1 ページ目から「参考文献」のページまでで、6 ページ以上 15 ページ以内であること。
- 1 ページ目のみページ上部にタイトル、研究者名、指導教員名、あらまし、キーワードを記入する。この部分のみ、左右余白を 30mm とする。
- 本文のページ番号は各ページの下部中央に、所定の形式で番号を付加すること。付録のページ番号は本文とは独立に付与すること。
- 図と表にはそれぞれ通し番号を付与し、本文中ではその番号を参照して説明すること。

余白の長さや行数、文字数は標準値を示した。若干の変更は可能とする。本資料は卒業研究論文の書式に合わせたサンプルである。

4. 文字フォントについて

論文等を記述するにあたって、文字フォントを揃えて記述すると読みやすく、また、綺麗に見える。文字フォントについては、当文書内では明朝もしくはゴシックという表現で指定する。

全体として日本語フォント 2 種類、欧文フォント 2 種類の使用を標準とする。図表内での使用フォントも同じ物を使用することを標準とする。例外としてプログラムリスト等の表示には、等幅のフォントを使用すると見やすくなる。

タイトル：

14pt 明朝，中央寄せ，行送り 15pt

サブタイトル：

12pt 明朝，中央寄せ，行送り 1.5 行

研究者名，指導教員名：

10pt 明朝，中央寄せ

あらまし：

9pt 明朝，あらましの部分のみ 9pt ゴシック，行送り 12pt

キーワード：

9pt 明朝，キーワードの部分のみ 9pt ゴシック，行送り 12pt

本文：

10pt 明朝，両端揃え，各段落の先頭で 1 文字分の字下げ

明朝は、縦が太く横が細い、ウロコ（飾り）のあるフォントを使用する。例えば、日本語フォントでは、IPA P 明朝や MS P 明朝が該当する。同時に使用する英語等のフォントは、セリフ（飾り）のある Century や Times Roman などが該当する。ゴシックは、縦横などの線の太さが一定で飾りのないフォントを使用する。日本語で使用するフォントとしては、IPA P ゴシックや MS P ゴシックが該当する。同時に使用する英語等のフォントは、Arial や Sans-serif などが該当する [1]。

5. 章タイトルの形式（11pt ゴシック）

章のタイトルは、11pt ゴシックとする。タイトル行の前に 0.5 行の改行を付加する。番号は、〈章番号〉. とする。

5.1 節タイトルの形式（10pt ゴシック）

節のタイトルは、10pt ゴシックとする。タイトル行の前に 0.5 行の改行を付加する。番号は、〈章番号〉. 〈節番号〉とする。

5.1.1 小節タイトルの形式（10pt ゴシック）

小節のタイトルは、10pt ゴシックとする。タイトル行の前に 0.5 行の改行を付加する。番号は、〈章番号〉. 〈節番号〉. 〈小節番号〉とする。

6. 内容について

卒業研究は実験型研究と開発型研究に大きく分けられる。それぞれに対する論文の標準的な構成は以下のとおりである。

• 実験型研究の標準的な例

1. 研究題目，研究者名，指導教員名
2. あらまし等
3. 研究の背景，目的，概要など
4. 理論
5. 実験の概要，使用した機器や設備
6. 実験結果
7. 考察
8. 結論（あとがき）
9. 謝辞
10. 参考文献
11. 付録

• 開発型研究の標準的な例

1. 研究題目，研究者名，指導教員名
2. あらまし等

表 2 総点の分布

総点	人
500 ～ 699	2
700 ～ 899	11
900 ～ 1099	19
1100 ～ 1299	5

3. 研究の背景, 目的, 概要など
4. 方法論
5. 各構成要素の詳細
6. 動作結果
7. 評価
8. 結論 (あとがき)
9. 謝辞
10. 参考文献
11. 付録

7. 図や表の挿入について

図の下には図番号とタイトルを添える。表には上に表番号とタイトルを添える。写真については、図と同様に下に写真番号とタイトルを添える。これらのキャプションは 9pt ゴシックとする。掲示した図表については、必ず本文中で説明する。

図や表は、段の幅に収まるように作成することが、基本である。収まりきらない場合は、左右の段をまたぐことができる。いずれの場合でも、ページの上部もしくは下部に挿入する。段の途中で文章を分ける配置は好ましくない。また、図や写真の縦横比はオリジナルから変更することのないように心がける。

図??は <http://jeena.net/rp-hw-button> に掲載されていたビットマップ形式 (png 形式) の図である。図中の文字は拡大縮小すると画素が目立ち美しくない。図??はベクトルデータ形式 (Windows の EMF 形式を EPS 形式に変換した) の図である。拡大しても画素が目立つことはない。可能な限り、ベクトルデータ形式で図を取り込むこと。表 2 は前期末の総点の分布である。

8. 参考文献について

参考文献は、研究で参考した文献や Web ページ等で、卒業論文中で参考している順に、以下のようにリストアップしておく。参考文献に挙げたものを参考している部分に、必ずこのように番号を付加する [2]。参考文献の見出しには、番号を振らないことが慣例である。

[2] 以降の参考文献のデータは、和歌山大学システム工学部坂間千秋教授の「卒業論文の書き方」<http://www.wakayama-u.ac.jp/~sakama/sotsuron/>からお借りした。これらは実際には利用されていないが、リストの書式のサンプルとしてあげてある。実際のリストの書式は、指導教員の指示に従い記述すること。

8.1 参考文献の書式

参考文献の見出しは、章見出しと同じ形式で番号を振らない。参考文献のリストは、9pt 明朝とする。

9. その他

- 本文は「～である」調を用いること。
- 句読点は、全角のコンマ、とピリオド。を用いること。
- 著作権法を順守し、他人の著作物 (文章や図表) を自分の論文中に流用してはならない。必要場合は引用の慣行に準ずること。すなわち、必要最小限の分量であること、改変をしないこと、出典を明示することである。
- ページ番号に使用する記号番号 (X1 の部分) は、後日連絡します。

参考文献

- [1] 奥村晴彦 他, [改訂第 6 版] L^AT_EX 2_ε 美文書作成入門, 技術評論社, 2013
- [2] 長尾真, 知識と推論, 岩波講座ソフトウェア科学 14, 1988
- [3] 実近憲昭, ゲームと AI, 人工知能学会誌 vol.5, pp.527–537, 1990
- [4] 人工知能の歴史, <http://www.jinkouchino-no-rekishu.com>, 2010

10. 数式の書き方

数式には文中に $y = ax^2 + bx + c$ と数式を書く場合と、次のように別行立てで書く場合がある。

$$y = ax^2 + bx + c$$

別行立ての数式には、`equation` 環境を用いて番号をつけることができる。

$$y = ax^2 + bx + c \tag{14}$$

式 14 は、2 次関数の式である。

11. パッケージの利用方法

`\documentclass[twocolumn,fleqn]{jsarticle}` とヘッドラインに記述する。twocolumn は 2 段組みにするため、fleqn は数式を中央ではなく、左端をそろえるオプションである。

論文を作成するにあたり、プリアンブル部に以下の記述を勧める。

```
\usepackage[dvipdfmx]{graphicx}
\usepackage{amsmath}
\usepackage{newtxtext,newtxmath}
\usepackage{float}
\usepackage{nidanfloat}
\usepackage{ascmac}
\usepackage{gthesis}
```

上から順に、`graphicx` は図を取り込むために必要、

`amsmath` は数式の AMS 拡張機能を使う, `newtxtext` は Times と Helvetica を使用する, `newtxmath` は数式のフォントをそろえるため, `float`, `nidanfloat` は図表の配置を適切にするため, `ascmac` は上記のような枠などを作るため, 最後の `gthesis` は情報工学科用の論文スタイルである.

`\begin{document}`以降に`\title` 命令でタイトルを指定, `\subtitle` 命令でサブタイトルを指定, サブタイトルがない場合は`\subtitle` 命令を使用しない. `\author` で作者の氏名を `supervisor` で指導教員の氏名を指定する. `abstract` 環境にあらましを, `keyword` 環境にキーワードを書く. これらを指定したのち, `\maketitle` 命令を書いておくと 1 ページ目の上部が完成するので, 以降に本文を記述する.

pL^AT_EX で `jsarticle` を用いて作成した場合, Word に比べて文字が小さめになるが, そのままでかまわない.

目次

1.	目的	1
2.	理論	1
2.1	ニューラルネットワーク (Neural Network, NN)	1
2.2	学習	2
3.	実験	3
3.1	研究手法	3
3.2	最適なモデルの考察	3
4.	文字フォントについて	4
5.	章タイトルの形式 (11pt ゴシック)	4
5.1	節タイトルの形式 (10pt ゴシック)	4
6.	内容について	4
7.	図や表の挿入について	5
8.	参考文献について	5
8.1	参考文献の書式	5
9.	その他	5
10.	数式の書き方	5
11.	パッケージの利用方法	5

図目次

1	パーセプトロン	1
2	順伝播型ニューラルネットワーク	1
3	活性化のプロセス	2
4	誤差逆伝播	3
5	モデル	3

表目次

1	実行環境	3
2	総点の分布	5