

AI研究のための手書き動物イラストデータセットの開発

山下 明 博

The Development of a Dataset of Animal Illustrations for AI Research

Akihiro YAMASHITA

造形デザイン学科, 家政学部,
安田女子大学

要 旨

AI研究の画像認識分野において必要とされる学習用画像データセット「安田女子大学・手書き動物イラストデータセット」を開発した。

これは、安田女子大学の学生420名に、10種類の動物のイラストの描画を依頼し、そこで収集した画像から開発したデータセットである。

次に、データセットの有効性を確認するために、ニューラルネットワーク、および、8種類の畳み込みニューラルネットワークによって、与えられたイラスト画像が、10種類の動物のうちどれかを分類させる画像認識プログラムを作成した。また、過学習の傾向があったため、データ拡張により、67,200件までデータセットを拡張した。

その結果、最高96%の確率で分類を行うことができた。

キーワード：AI、動物イラスト、畳み込みニューラルネットワーク、深層学習

はじめに

AI¹⁾ 研究において欠かせないのが、AIによる学習を進めるためのデータの集合であるデータセットである。

従来から、AIに画像認識²⁾ を学習させるときに使用するデータセットは、MNISTやCIFAR-10など、様々なものが提供されており、これらのデータセットを使用した研究が多く行われてきた。

しかし、安田女子大学でAIを学ぶ学生にとって、他者から提供されたデータセットを利用するのではなく、自ら開発したデータセットを使用して学んだ方が、学習効果や理解度が高まると考えた。

そこで、安田女子大学の学生420名に、動物のイラストの描画を依頼し、そこで収集した画像認識用のデータセットを開発することにした。

I. 画像認識用データセットの開発

この章では、画像認識用データセットについて、従来から提供されているMNISTデータセット、CIFAR-10データセットと、今回開発した「安田女子大学・手書き動物イラストデータセット」の比較を行う。

1. MNISTデータセット

画像認識の学習において、最も多く利用されているのが、MNISTデータセットである。MNISTデータセットは、0から9までの数字画像の集合であり、 28×28 画素の白黒画像となっている。また、訓練画像60,000枚、テスト画像10,000枚から構成されており、60,000枚の訓練画像で、AIによる画像認識の学習を行い、その後、10,000枚のテスト画像で、学習効果の測定が行われるといった使い方がなされる。図1に、MNISTデータセットの先頭10件のデータを示す。

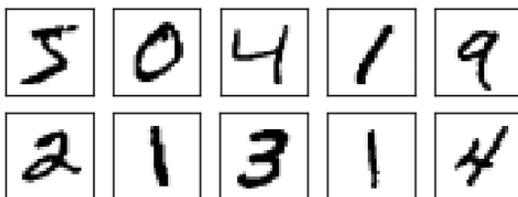


図1 MNISTデータセットの先頭10件のデータ

2. CIFAR-10データセット

CIFAR-10データセットは、カラー画像認識の学習において、よく利用されている。CIFAR-10



図2 CIFAR-10データセットの例

データセットは、airplane、automobileといった、10種類の画像の集合である。32×32画素のRGB画像となっており、訓練画像50,000枚、テスト画像10,000枚から構成されている。図2に、CIFAR-10データセットの例を示す。

3. 安田女子大学・手書き動物イラストデータセットの開発

2019年6月に、安田女子大学の学生420名に、「イヌ」「ウマ」「ヒト」「トラ」「キツネ」「ネコ」「ウシ」「サル」「ライオン」「タヌキ」という10種類の動物のイラストの描画を依頼した。そして、それらのデータを、AIによる画像認識のための学習に利用するために、64×64画素の白黒画像として切り出し、10種類の動物を示す数字ラベル（イヌ=0、ウマ=1、ヒト=2…、タヌキ=9）と組み合わせた、合計4,200枚のデータセットを開発した。

図3に、安田女子大学の学生による、10種類の手書き動物イラストの例を示す。

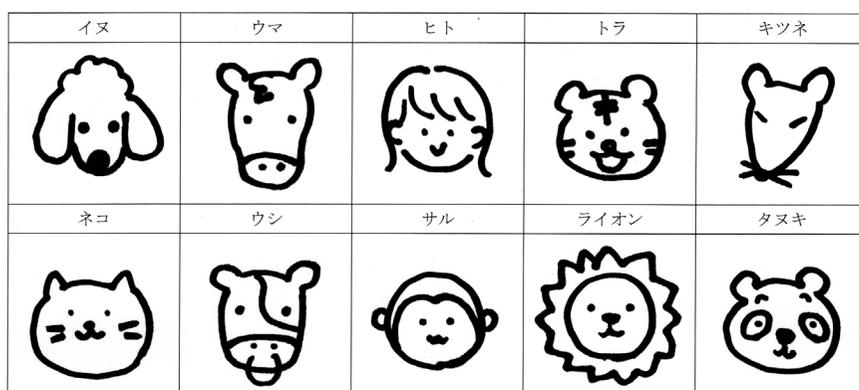


図3 安田女子大学の学生による手書き動物イラストの例

II. ニューラルネットワークによる手書き動物イラストの分類

はじめに、安田女子大学で開発した手書き動物イラスト画像4,200件を、訓練画像4,000件とテスト画像200件に分け、図4に示すニューラルネットワーク³⁾により分類を学習させた。

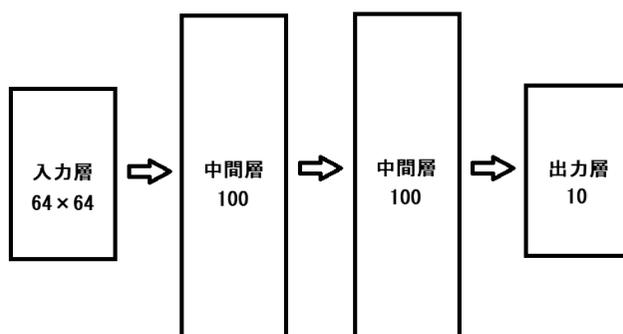


図4 ニューラルネットワーク

その結果、図5のように、訓練時は分類精度が100%になったにもかかわらずテスト時には分類精度が約50%に留まった。これは、イラストの分類テストを100回試行したときに、正解を50回しか導けなかったことを示す。

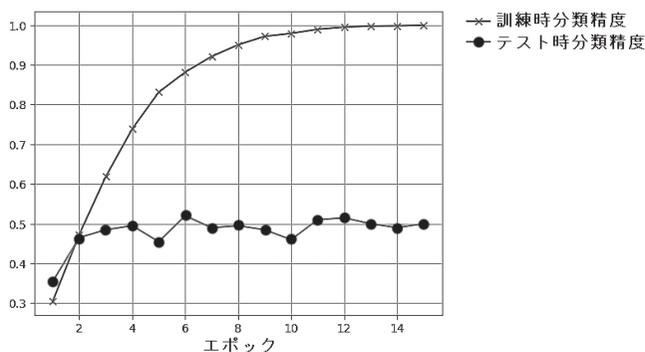


図5 ニューラルネットワークによる4,000件の訓練と200件のテスト結果

分類の学習においては、OSはWindows10、開発環境はMicrosoft Visual Studio Express 2017、開発言語はPython3.7.0を使用し、AI学習のためのオープンソースライブラリChainerを使用した。

図5において、テスト時分類精度が低い原因としては、データセットの件数が少なく、過学習に陥っていることが考えられる。

Ⅲ. 深層学習による手書き動物イラストの分類

前章で行った、ニューラルネットワークによる、手書き動物イラストの分類の学習を進展させ、深層学習（ディープラーニング）⁴⁾を行うことにより、分類精度の向上を目指す。

図6に、今回作成した8種類の畳み込みニューラルネットワークを示す。

これら8種類の畳み込みニューラルネットワークを使用し、安田女子大学で開発した手書き動物イラスト画像2,100件を、訓練画像2,000件とテスト画像100件に分け、分類を学習させた結果を図7に示す。

これより、畳み込み処理（畳み込み+パディング+プーリング）は、4回が良いと考えられる。畳み込み処理を5回行った場合、プーリング処理により画像のサイズが小さくなりすぎるため、学習効果が下がってしまう。また、4回の畳み込み処理のうち、どこかのプーリングを省略し、画像のサイズを小さくしない方法で調べたところ、4回目のプーリング処理を抜いた場合に最も成績が良くなった。そこで、4回の畳み込み処理（4回目のプーリング処理抜き）のニューラルネットワークを中心に、他のニューラルネットワークも含めて検証を行うことにした。

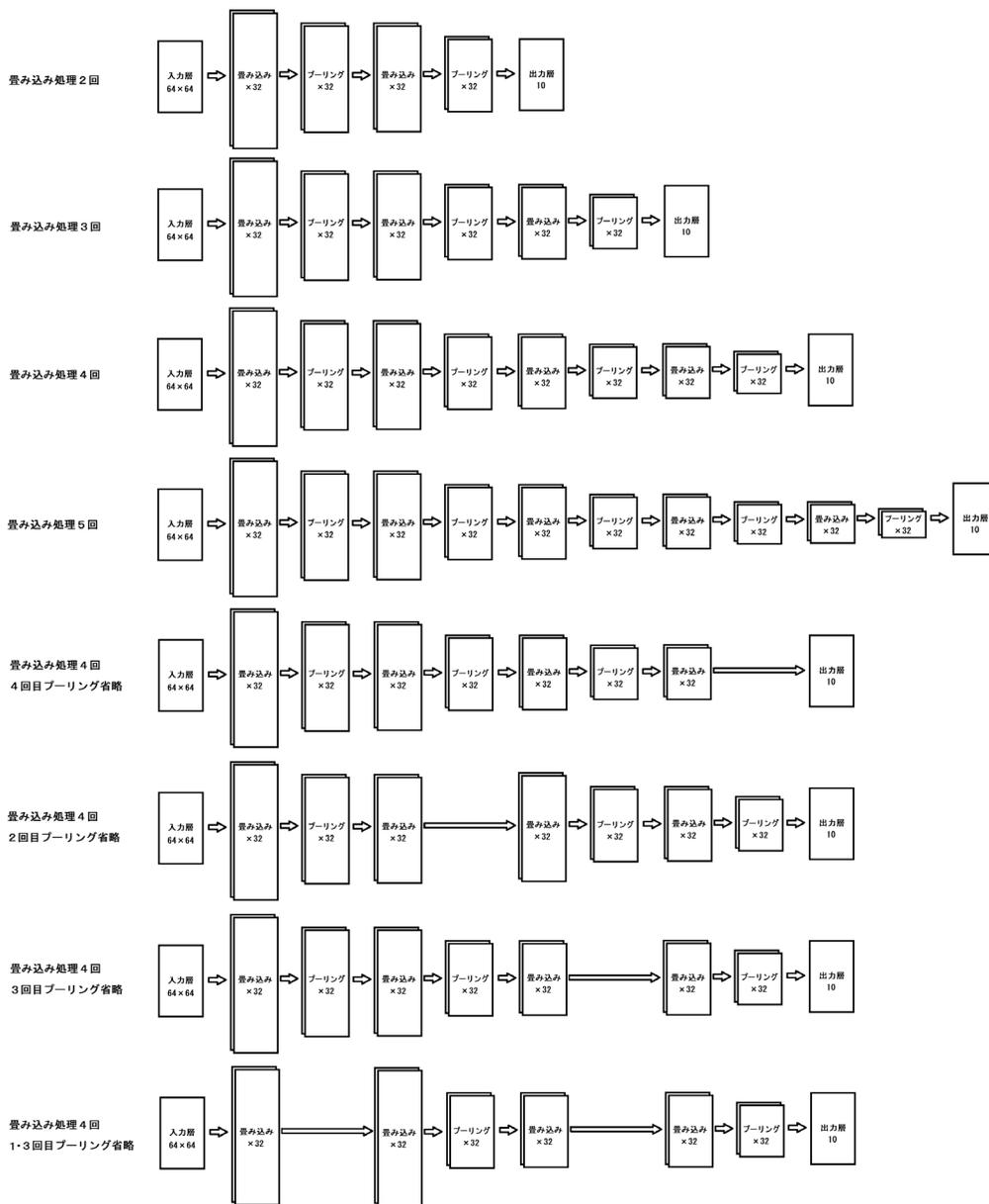


図6 8種類の畳み込みニューラルネットワーク

IV. 畳み込みニューラルネットワークにおける精度向上のためのデータ拡張

安田女子大学で開発した、手書き動物イラスト画像データセットは、4,200件である。学習精度を向上させるには、データ件数が少ないという問題がある。そこで、もともとのイラスト画像に対し、左右反転、上下反転、画像の回転といったデータ拡張を行うことにより、訓練画像

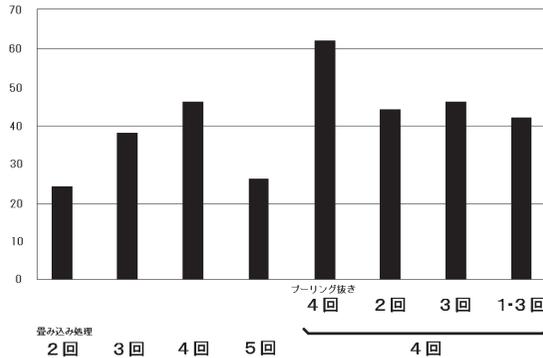


図7 8種類の畳み込みニューラルネットワークによる学習結果

4,200件を水増しし、畳み込みニューラルネットワークにより分類を学習させることにする。図8に、基本となる動物イラストと、そのデータを拡張したイラストを示す。

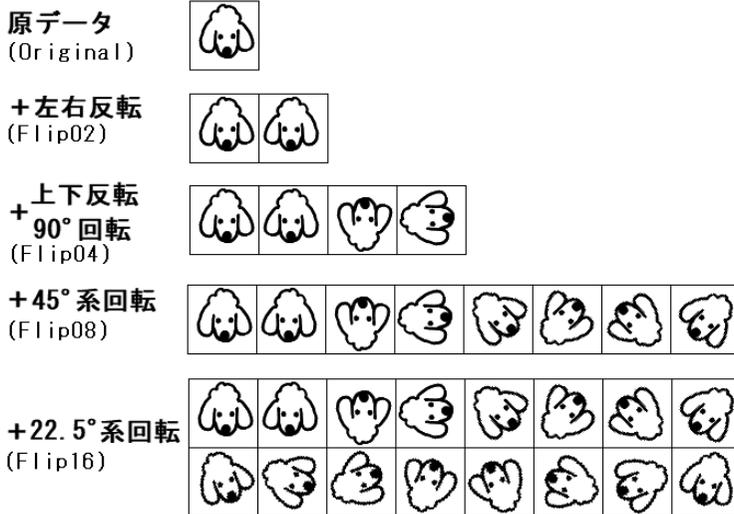


図8 原データとデータ拡張の例

これらのデータ拡張により、データセットは、4,200件から、8,400件、16,800件、33,600件、67,200件まで増やすことができる。

そこで、前章で用意した、8種類の畳み込みニューラルネットワークを使用し、データ拡張した画像データの件数を変えながら、深層学習を行った。

V. データセット読み込み速度の高速化

手書き動物イラスト画像データセットの分類を行うための学習において、当初、1,000件程度のデータを読み込むときにも、10分程度の時間がかかるという問題が生じた。

これは、イラスト画像データセットをCSV形式のファイルから読み込み、PythonのNumpy拡張モジュールの配列に格納する箇所において、非常に時間がかかることが原因であった。

具体的には、Numpyのasarray関数でNumpy配列を作成し、Numpyのappend関数によるデータ挿入を、例えばデータが1,000件あれば1,000回繰り返して行っていた。

これを改良し、最初に、Pythonのネイティブな機能であるリスト配列を作成し、そこにPythonのappend関数でデータ挿入を行い、全データ挿入完了後、Numpyのasarray関数で、Pythonのリスト配列をNumpy配列に変換するように、プログラムを変更した。

その結果、データセット読み込み速度を、約10分から約0.12秒に短縮できた。これは、約5,000倍高速化することができたことを意味する。

この、Numpy配列に直接appendするのではなく、Pythonのリスト配列にappendし、その後Numpy配列に変換するという高速化は、他でも効果を上げており、800件のデータで、587.8秒が0.074秒と、約7900倍高速になった例も報告されている⁵⁾。

VI. 畳み込みニューラルネットワークによる学習結果

様々な畳み込みニューラルネットワークを使用し、データセットの件数を変更して、AIに手書き動物イラストの分類を学習させた場合、最高の正答率を示したのは、67,200件までデータ拡張を行ったうち、訓練画像2,000件、テスト画像100件に絞ったデータセットを使用し、4回の畳み込み処理（2回目のプーリング処理抜き）のニューラルネットワークの場合で、50回テストを行った結果であった。この場合、48回の分類成功、すなわち、96%の確率で分類を行うことができた。図9に、訓練時の分類精度と、テスト時の分類精度のエポックごとのグラフを示す。

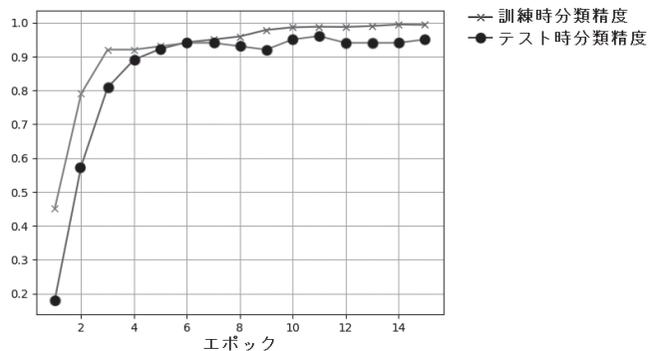


図9 畳み込みニューラルネットワークによる2,000件の訓練と100件のテスト結果

また、分類に失敗したのは、以下の2つの画像であった。図10に、分類に失敗したデータを示す。

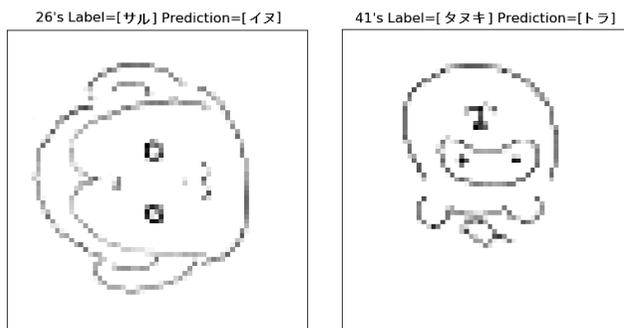


図10 分類に失敗したデータ（サルをイヌと認識、タヌキをトラと認識）

左側の、サルのイラスト（右90°回転）をイヌに分類した理由は不明である。一方、右側のタヌキのイラスト（上下反転）をサルに分類した理由は、タヌキの口の形が、トラの特徴である額の模様と誤って認識されたからではないかと考えられる。

結 論

本稿では、安田女子大学の学生420名が描画した、「イヌ」「ウマ」「ヒト」「トラ」「キツネ」「ネコ」「ウシ」「サル」「ライオン」「タヌキ」という10種類の動物のイラストから、AIによる画像認識のための学習に利用するための合計4,200枚のデータセットを開発したことを述べた。

MNISTのような手書き文字画像や、CIFAR-10のようなカラー画像といった、他者から提供されるデータセットでも、AIによる画像認識の学習を行うことはできる。また、最近では、学習済みモデルを使い、新たに少ないデータと学習量でモデルを作る「転移学習」という手法も使用されることがある。しかし、AIの基礎を学ぶ段階においては、自らデータセットを開発するという手間をかけ、データセット作成方法を学ぶことも、学習効果の向上に資すると考える。

今後、今回開発した手書き動物イラストデータセットに対し、画像イメージのサイズを統一する変換等を加えることにより、認識率がどのように変化するかについて検証を行いたい。

注

1. AIは、Artificial Intelligenceの略称であり、人工知能のことである。
2. 画像認識は、画像および動画から文字や顔などのオブジェクトや特徴を認識し検出するパターン認識技術である。手書き数字認識などが実用化されている。斎藤康毅（2016）、『ゼロから作るDeep Learning：Pythonで学ぶディープラーニングの理論と実装』、東京：オライリー・ジャパン社、P.72参照。
3. ニューラルネットワークは、神経細胞をモデル化した計算素子である人工ニューロンを組み合わせたもので学習を行うもので、機械学習の範疇に入る。小高知宏（2016）、『機械学習と深層学習：C言語によるシミュレーション』、東京：オーム社、P.116参照。
4. 畳み込みニューラルネットワーク（Convolutional Neural Network）は、CNNとも略され、畳み込み層とプーリング層を組み合わせて学習を行うものである。小高知宏2016、P.205参照。
5. PythonのリストとNumpy配列に対するappend速度の比較は、<http://kurora-shumpei.hatenablog.com/entry/2019/01/12/python%E3%81%AE%E3%83%AA%E3%82%B9%E3%83%88%E3%81%A8numpy%E3%81%AEappend%E3%81%AE%E9%80%9F%E5%BA%A6%E6%AF%94%E8%BC%83>

参照 (2019.09.08確認)。

[2019. 9. 26 受理]

コントリビューター：染岡 慎一 教授 (造形デザイン学科)

