

ディープラーニングを用いての写真と絵画,イラストの識別

大崎 一輝† 田村 仁† 檜山 正樹†† 入江 俊†† 仲田 仁††

日本工業大学工学部 創造システム工学科† 機械システム工学専攻††

1. 研究背景

近年,スマートフォンなどの写真撮影や,検索エンジンの画像検索での画像収集などで,図1, 図2, 図3のように,容易に写真や絵画,イラストなどの画像データを入手することができる.それにより大量の様々な画像データをユーザが有することになる.そして,その画像をユーザが手動で画像を1つ1つ振り分けることになる際,時間がかかり,ユーザへの負担になる.そこで,写真と絵画,イラストを自動識別し振り分けることにより,ユーザへの負担を大幅に軽減できるだろうと考えた.



図1 写真の例



図2 絵画の例



図3 イラストの例

2. 研究概要

先行研究として参考文献[1]の描線に着目した写真とイラストの識別を行っていた.この研究では,描線(イラストの輪郭を構成する主要な線)の有無を見て写真,イラストを識別していた.

本研究では先行研究のように手動で特徴を取り出すのではなく,以下の手法を用いて,写真,イラスト,絵画それぞれの画像を学習させ特徴を抽出し識別する学習器を作成する.

また,従来の画像識別では画像の特徴抽出が困難だったが,ディープラーニングでは画像の特徴をより明細に抽出できる.そのため,今回はディ-

プラーニングの手法の一つである二次元の画像から特徴抽出ができる,畳み込みニューラルネットワークを用いる.

2.1. 畳み込みニューラルネットワーク

この畳み込みニューラルネットワークは,人の視覚野の動きをモデル化した構成になっている.また,畳み込みニューラルネットワークとはいくつかの層があり,図4のようにそれぞれ入力層,畳み込み層,プーリング層,全結合層,出力層があり,この中の畳み込み層とプーリング層を複数回繰り返した構造をしている.

畳み込み層では入力に対して重みフィルタの内積を計算する.1層目の畳み込み層で入力サンプルに対して畳み込み処理を行い,特徴を取得する.

プーリング層では,畳み込み層から出力された特徴マップを縮小し,着目する領域を設定してその領域の特徴マップの値から新たな特徴マップの値を求める.

全結合層では重み付き結合を計算し,活性化関数によりユニットの値を求める.活性化関数には,シグモイド関数,tanh関数,ReLUなどを用いる.出力層では,尤度関数を用いて,各クラスの尤度を求める.[2]

また,畳み込みニューラルネットワークをすることでエッジやグラデーションが抽出される.

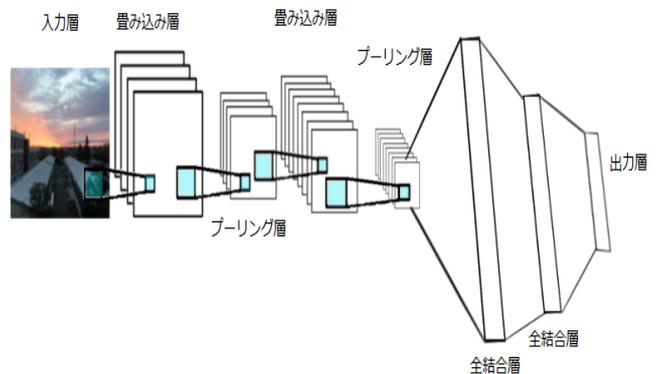


図4 畳み込みニューラルネットワーク

Identification of photos, paintings, and illustrations using deep learning

†Kazuki Osaki, Hitoshi Tamura

††Masaki Hiyama, Suguru Irie, Hitoshi Nakada

†Nippon Institute of Technology

††Mechanical Systems Engineering Major

3. 実験手順

学習に必要な写真, 絵画, イラストの画像をBing APIを用いて写真, 絵画を約8千枚, イラストを約2千枚収集した.^[1] また, 収集した画像には何枚か同一の内容の画像があり, それをあらかじめ削除しておき,^[4] 画像群は手動で分類分けを済ませておいた. これを先ほど説明した畳み込みニューラルネットワークに入力して分類器を作成した.^[5] また学習した分類器を用いて画像特徴を出力した.

4. 実験結果

写真, 絵画, イラストそれぞれすべての画像を, 約3万枚を学習させた. 学習させた回数は50回であり, また, 80%を学習に, 残り20%を評価に用いた. その結果, 図5, 図6のようになった. lossとaccはそれぞれ, 誤差と認識精度である. val_lossとval_accはそれぞれ評価に用いた誤差と認識精度である.

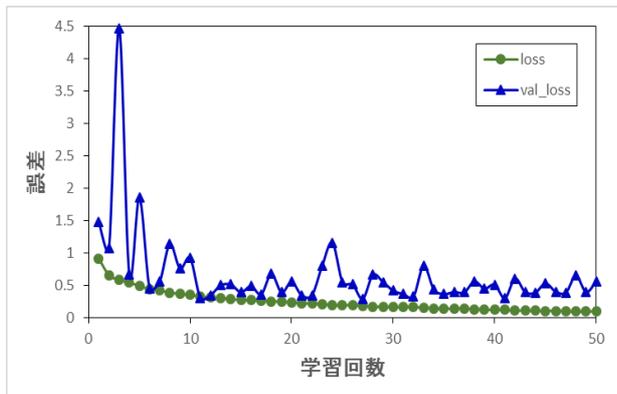


図5 学習誤差

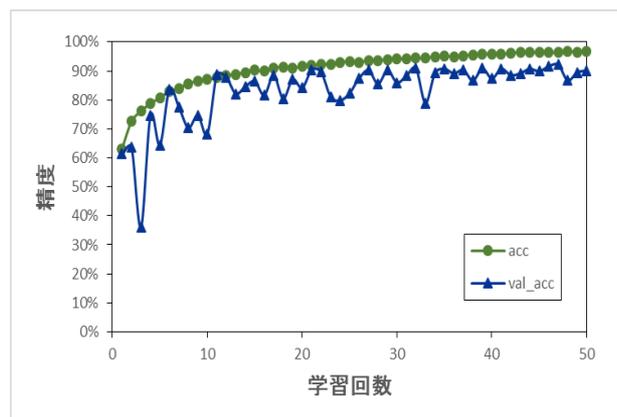


図6 認識精度の推移

図5から, 学習回数が増えるごとに誤差が小さくなっていくのを確認した. また, 図6のように学習の数を重ねるごとに認識精度が上がっていくのを確認した. 今回は約90%の認識精度を得ることができた.

表1の判別精度の比較では絵画の精度はないが, 先行研究^[1]の描線に着目した写真とイラストの識別の実験結果である, 描線画像判別手法と, 色相画像判別手法の特徴値を合わせた全特徴値の約74%の認識精度を超えることができた.

また, 次の図7のような8×8画素の特徴を32枚抽出することができた. 図7を見ると人間の視覚では理解できない特徴だった.

表1 判別精度の比較

	先行研究 ^[1]	本研究
写真	0.634	
イラスト	0.854	
絵画		
全画像	0.744	0.901

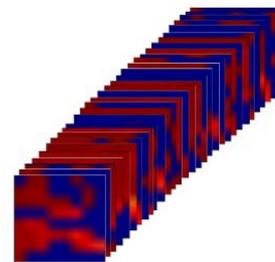


図7 抽出された特徴(8×8画素)

5. まとめ

図5から, lossは低くなっていくが, val_lossは少々高くなっている部分もあり, 過学習が生じている. この場合でも学習回数が増えるごとに誤差が小さくなっていくのが分かった.

図6からは, 表1の先行研究の認識精度を超えることができた.

また, 先行研究の描線に着目した写真とイラストの識別では, 人間が特徴を選出していたが機械学習に手法を用いることで, 人の手による特徴抽出を行う必要なく分類ができ, ユーザの負担を軽減できると考える.

参考文献

- [1] 河野 倫充, 久保田 康之 描線に着目した写真とイラストの識別 日本工業大学
- [2] 山下隆義 著 (2016) イラストで学ぶディープラーニング pp36-47
- [3] PythonでBing Search APIを使って画像データを集める方法 <http://urx3.nu/AsU>, (参照 2016-06-02)
- [4] sim-g's junk space <http://hp.vector.co.jp/authors/VA032597/Software/index.html>, (参照 2016-08-04)
- [5] Keras Documentation <https://keras.io/ja/layers/convolutional/>, (参照 2016-10-15)