

出土石器の分類器の作成及びデータ拡張の比較検討

戸次道宏* 山口文彦†

長崎県立大学情報セキュリティ学科‡

1 はじめに

近年、機械学習が様々な分野で活用されている。考古学においても例外ではなく、古代文字の解読などに活用されている例がある。機械学習は対象を分類するために用いられる。出土石器の分類は専門家の手により行われているが、人の手による分類は人的コストが大きく作業に専門的な知識が必要となる。機械学習による自動分類システムを利用することで人の手による作業量が削減され、専門的な知識を持たない人でも作業が可能になると期待される。そこで機械学習の手法の一つである CNN を用いて出土石器の画像を入力とする分類学習器を作成することを考える。

機械学習を行うには予め大量の偏りのないデータを用意することが理想であるが、出土石器の場合、出土状況に左右されるなどの理由で偏りのないデータを用意することが困難である。機械学習においてデータの偏りを軽減する手法としてデータ拡張が挙げられる。データ拡張とは、元のデータに対して変形やノイズを加えることで類似した違うデータを作り出すことである。本研究では出土石器の分類に適したデータ拡張の方法を検討する実験を行った。

2 研究にあたって

機械学習の様々なアルゴリズムが様々なところで利用されている。本研究は画像分類の分野に該当する。画像分類は近年、深層学習の成果があり普及した。考古学分野においてもくずし文字の解読などの研究が行われている [1]。しかし、出土物の分類や解析については 3D データを用いたパターンマッチングなどの手法は研究が行われている一方、機械学習による自動分類システムを作成、実装した研究例や論文は少ない。データ拡張方法は訓練データにノイズや変形を施すもので様々な機械学習の研究で利用されている。拡張手法を取得する研究もなされている [2]。

3 実験準備

実験に使用するために「福島県文化財センター白河館」のホームページである「まほろんのホームページ」で公開されている文化財データベース [2] から出土石器の画像データを収集した。分類に際して各クラス名のラベル付けはこのデータベースで用いられているものに準じた。クラス分類の観点は用途、形状、制作過程が中心である。収集した画像をそれぞれのクラスごとに訓練データ、検証データに分割した。分割前のデータに拡張を行うと、分割した際に拡張元が同一のデータが訓練データと検証データに存在することになり、汎化性能が正しく評価できない。そこで、分割した後の訓練データを拡張する。クラスによってはデータが 1 つしかないものもあったが、分割できないため本研究では除外することとした。収集した画像データは映っている石器の外接矩形を半自動で検出し、トリミングした。

4 実験

データを拡張には多くの手法があるが、本実験では以下の拡張手法を比較する。また以降では各拡張方法を括弧内の略称で呼称する。

- i. 画像のコントラスト調整 (Cn)
- ii. ペッパーノイズ (Pe)
- iii. ソルトノイズ (Sa)
- iv. ガウス分布によるノイズ (Ga)
- v. カラーチャンネル調整 (CC)
- vi. せん断処理 (Sh)

これに加えて比較用の手法として、同一画像を無加工でコピーして増やしたのみの処理 (Cp) を行って拡張した合わせて 7 つの拡張方法で実験を行った。

Cn は画像のコントラストを低減もしくは強調する。強調する場合は一定の数値 A 以下の低輝度画素を 0 に、一定数値 B 以上の高輝度画素を 255 に、その中間の輝度の画素を調整する。A を 40~50、B を 195~205 からランダムに決定した。低減する場合は全体的な輝度の幅を小さくするように処理を行った。Pe 及び Sa はそれぞれ黒と白の点をランダムに画像全体の 0.1% から 0.4% に描いた。Ga は画像の各画素に標準偏差 σ の正規

*Michihiro Hetsugi

†Fumihiko Yamaguchi

‡Department of Information Security, University of Nagasaki

分布乱数を加えることで輝度を変化させる。 σ は10~15の範囲でランダムに決定した。CCは画像のカラーチャンネルを-50~50の間でランダムにシフトさせた。Shは座標(x,y)で表される点を(x+my,y)に移すような変換処理を行う。mは-20~20の範囲でランダムに決定した。

各拡張方法について、1クラス500枚まで拡張した76クラスの拡張訓練データを用意し、学習を行いその精度を比較した。学習器としてVGG16[3]のImageNet学習済みモデルを用い、転移学習を行った。モデルのうち、15~18層の再学習を行った。全結合層はノード数4096を持つものを2つ、出力用のものを1つ用意した。データの入力サイズは128px × 128px、ミニバッチサイズは128、エポック数は50で学習を行った。また学習時のバッチ入力時に回転及び上下左右方向への水平移動処理をランダムに適用している。学習時、検証データに対する精度であるval_accが最高値になる度にその時のモデルの状態を保存するようにしている。

図1はShで拡張した訓練データの学習グラフである。縦軸が精度、横軸がエポック数である。表1はval_accの最高値及び最高値が出た時の訓練データに対する精度accを拡張方法毎にまとめたものである。

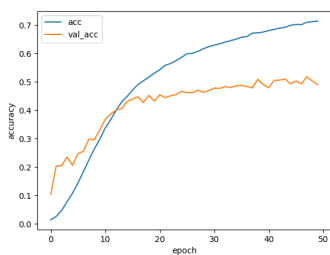


図1: Shの学習グラフ

表1: 単一拡張方法での精度比較

手法	val_acc	acc
Cn	0.3587	0.6252
Sa	0.4429	0.7059
Pe	0.4503	0.6936
Ga	0.4981	0.7032
CC	0.2855	0.7025
Sh	0.5187	0.7099
Cp	0.2432	0.7251

7つの拡張方法の中で最も精度が優れている拡張方法はShであると分かった。次にそれぞれ精度が良かった拡張方法を順に組み合わせて学習を行い、その精度を確かめた。

表2: 拡張方法を組み合わせた場合の精度比較

手法	val_acc	acc
Sh+Ga	0.5171	0.6882
Sh+Ga+Pe	0.5061	0.6612
Sh+Ga+Pe+Sa	0.4848	0.6666
Sh+Ga+Pe+Sa+Cn	0.4652	0.6260
Sh+Ga+Pe+Sa+Cn+CC	0.4158	0.6160

組み合わせても最も精度の良かったShの最大値とほぼ同等程度の精度しか出ず、組み合わせる拡張方法を増やすほど精度が悪くなっていくという結果になった。

5 考察

ShとCCを比較すると、accの数值は大きく変わらないのに対して、val_accには大きな差が出ている。今回使用した石器データの分類では色みや明るさよりも形状が最も重要な特徴として捉えられていると考えられる。val_accの値が全体的に高くない理由としては、本実験は最も極端な例では1枚から500枚に拡張するクラスもいくつかあり、複数回同一の拡張処理を重ねる必要があるためその過程で拡張したデータの質が低下していると考えられる。

6 まとめと今後の課題

今回の実験の条件ではShの拡張方法を単体で用いることが最も良い結果を得られた。今回使用したデータは分類の観点が用途、形状、制作過程が中心であり、形状の変形が最も拡張をする上で良いデータになりやすかったという結果であった。

今回の実験では学習モデルを統一して学習を行った。今後の課題として、機械学習では学習を行うデータによって最適なパラメータが異なる。今回のようにパラメータを固定して行った比較では十分とは言えない。それぞれの拡張方法に合わせた最適なパラメータを求めた上でそれ以外の条件を固定し精度の比較を行い、組み合わせた拡張についてもその様に学習を行って今回の実験と比較してみる必要がある。今回の実験で使用した以外にも拡張方法は存在、提案されているためその方法を試すことも改善の手段として挙げられる。今回の実験ではShが最も良い精度を出したので、他にも変形を行う拡張方法の実験を行う価値があるものと考えられる。

参考文献

- [1] 上田佳祐、菌頭元春、飯山将晃, “CNNと文字のアスペクト比を用いたくずし文字認識”, 京都大学学術情報リポジトリ (KURENAI), 2018-03, <http://hdl.handle.net/2433/230597>, (最終閲覧日: 2020年1月9日)
- [2] 福留拓也, 文字認識器の学習データ自動生成ルールの獲得, <http://www.comm.tcu.ac.jp/otani-lab/research/master/0883119.pdf>
- [3] 福島県教育委員会, ”まほろんのホームページ”, 福島県文化財センター白河館文化財データベース, <http://www.mahoron.fcp.or.jp/search.html>, 2001 (最終閲覧日: 2020年1月9日)
- [4] Karen Simonyan, Andrew Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition”, Cornell University, 2015-4-10, <https://arxiv.org/abs/1409.1556> (最終閲覧日: 2020年1月9日)