

## データ圧縮アルゴリズムを活用した画像認識 AI における効率的学習手法の開発

機械電子科 松下五樹\*

### Development of an efficient learning method for image recognition AI by using a data compression algorithm

MATSUSHITA Itsuki

In image recognition AI, a large amount of training data is required to ensure high performance, and the huge costs associated with data collection are a barrier to the introduction of AI. In this study, the authors propose a model that can efficiently learn from a small amount of data by incorporating an image data compression algorithm into image recognition AI. The results of a validation performed using CIFAR-10 indicated that the proposed method was able to reduce the training data by up to 55.6% compared to conventional methods. The results of predicting an object's hardness from microstructural images revealed that the proposed method was able to improve the prediction accuracy and reduce the training time compared to conventional methods.

Keywords : image recognition, algorithm, data compression, artificial intelligence

画像認識AIにおいて、高い性能を得るために大量の学習用データが必要であり、莫大なデータ収集コストがAI導入の障壁となっている。本研究では、画像のデータ圧縮アルゴリズムを画像認識AIに組み込むことにより、少ないデータから効率的に学習を行うことが出来るモデルの作成を試みた。CIFAR-10<sup>1)</sup>を用いた検証の結果、提案手法は通常手法と比較して、最大で55.6%の学習データ削減効果を得ることが出来た。また、金属組織写真から硬さの予想を行った結果、提案手法は通常手法よりも予測精度が向上し、学習時間を削減出来る傾向が見られた。

キーワード：画像認識、アルゴリズム、データ圧縮、人工知能

#### 1 はじめに

人工知能（以下、AI）は様々な技術分野において活用が進んでおり、その一つに画像認識分野がある。AIを活用した画像認識（画像認識AI）は、現在では人間よりも高い認識性能を得られる例も多く報告されている。画像認識AIでは、識別性能を高めるために大量の学習データ（数千～数十万枚の画像）が必要であり、少ない学習データでは十分な性能が得られず実用性が低下してしまう。この学習データの収集には莫大な時間とコスト（人件費、材料費等）が掛かるため、画像認識AIの普及を阻害している要因の一つとなっている。

そこで本研究では、画像データの圧縮アルゴリズムに着目した。データ圧縮においては、出来るだけ画像の見た目を損なわずに、情報量を削減することが求められる。このためデータ圧縮アルゴリズムでは、画像の

持つ情報を見た目への影響が大きい情報と、見た目への影響が小さい情報に分別し、後者を削ることでデータ量の圧縮を実現している。一般的に画像認識AIで使用されている畳み込み処理では、画像データの1ピクセルごとの情報に対して学習を行っている。この処理に圧縮アルゴリズムを導入することにより、学習の対象を情報量の多いデータに厳選すれば、画像の特徴を損なうことなく、より効率的に学習を進められるのではないかと考えた。

#### 2 方法

Python用機械学習ライブラリのPyTorchを使用し、画像圧縮アルゴリズムを組み込んだレイヤーをカスタムレイヤーとして作成した。作成したカスタムレイヤーを使用し、画像認識モデルを構築した。

\*現 工業技術研究所 機械電子科

## 2.1 学習データ数削減効果の検証

構築した提案手法の性能を評価するため、畳み込み層等を重ねたモデル（以下、通常手法）と、jpeg等のファイルで使用されている圧縮アルゴリズムを畳み込みに導入したモデル（以下、提案手法）を用いて性能比較を行った。モデル構造はVGG16<sup>2)</sup>を参考に、畳み込み層3層と最大値プーリングを重ねてブロック構造とし、ブロックを2~4個繋げた後に全結合層を通してから出力させた。比較を行うため、通常手法と提案手法ではブロック数やフィルターサイズ等のパラメータを同等としたモデルを使用した。学習データには、画像認識問題用データセットとして一般公開されているCIFAR-10<sup>1)</sup>を用いた。CIFAR-10は32×32ピクセルのカラー画像で、画像に写っている物体を10種類のラベルに分類する問題を取り扱うデータセットである。データセット合計60,000枚のうち、学習用データ枚数を100~50,000枚で変化させた際の、検証用データ10,000枚に対する分類正答率を6分割の交差検証により評価した。

## 2.2 実験データに対する性能評価

金属組織写真からの硬さの予測を行った。材料は

機械構造用炭素鋼S35Cを使用し、熱処理や塑性加工を行って種々の金属組織とした16種類の試料を作製し、金属組織写真の撮影およびマイクロビックカース硬さ(HV)の測定を行った（図1）。組織写真は元画像(2,592×1,944ピクセル)から240×240ピクセルのサイズで48枚(縦6枚×横8枚)を切り出した後、グレースケール化、回転や反転による画像拡張を行い、総データ数を6,144枚とした。金属組織の画像から硬さの数値をモデルに予測させる回帰分析を、4分割の交差検証により行った。なお、学習に用いた金属組織写真の枚数は変化させず一定とし、観察倍率は500倍で固定とした。

## 3 結果

### 3.1 学習データ数削減効果の検証

図2に学習データ数を変化させた場合の分類正答率の変化を示す。学習データ数が少なくなるほど、いずれのモデルにおいても正答率は低下しているが、提案手法では通常手法に比べ、学習データが少ない場合の正答率が高くなった（図2(a)）。各々の学習データ数の結果について正答率の差を取ると（図2(b)）、

	①：熱処理	②：①+圧縮	③：②+600°C炉冷	④：②+830°C炉冷
①-A 895°C 炉冷	173HV	314HV	198HV	136HV
①-B 895°C 空冷	141HV	329HV	218HV	170HV
①-C 895°C 水冷	191HV	816HV	258HV	186HV
①-D 895°C 水冷 + 600°C 水冷	180HV	337HV	232HV	164HV

図1 学習データとして使用した金属組織写真および硬さ

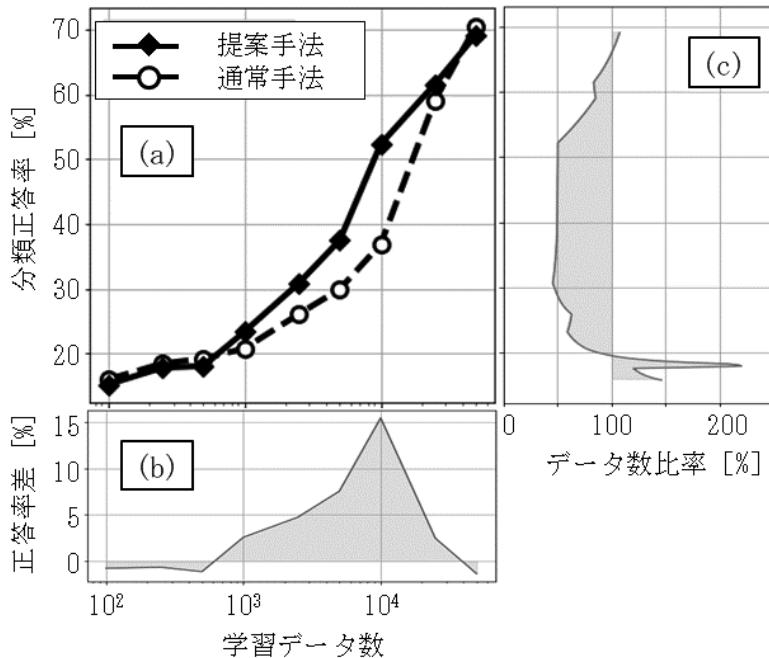


図2 CIFAR-10を用いた検証結果。(a)は2種類のモデルにおいて学習データ数を変化させた際の分類正答率の変化、(b)は2種類のモデルでの正答率の差を取ったグラフ、(c)は各正答率における2種類のモデルの学習データ数の比率を示す。

データ数が10,000枚の条件において、提案手法が通常手法に対して最大15.5%正答率が上昇していた。学習データ数が50,000枚の場合と、500枚以下の場合は通常手法の方が提案手法よりも正答率が高くなっているが、その差は1%程度であった。今回の条件下においては、提案手法はいずれのデータ数においても、通常手法と比較して正答率は向上もしくは同等となった。

図2(c)は各正答率に相当する2つのモデルのデータ数の比を示しており、一定の正答率を得るために必要なデータ数が、通常手法に対して提案手法がどの程度かを割合で表している。正答率30~50%の範囲において比率は50%程度となっていることから、この範囲ではデータ数を通常手法の半分程度に削減しても、提案手法では同等の正答率を得られることが分かった。

### 3.2 実験データに対する性能評価

提案手法と通常手法を用いて、金属組織写真からマイクロビックカース硬さ(HV)を予測した際の性能比較を行った。各モデルでブロック数や畳み込みフィルターサイズ等のパラメータを変化させた際の結果を図3に示す。縦軸に示す評価指標には、予測値と正答値の平均二乗誤差を用いており、図中の下に位置する点ほど実験値と予測値の差が小さく、性能が高いこと

を示す。全体的な傾向として、提案手法は通常手法よりも予測時の誤差が小さく、高い性能を得られたことが分かる。また、学習に要した時間(図3横軸)についても、提案手法を使用したモデルの方が非常に短時間で学習を完了している他、時間のバラつきも小さいことが分かる。この中で最も誤差の小さかったモデル(フィルターサイズ: 3×3、ブロック数: 2)の平均平

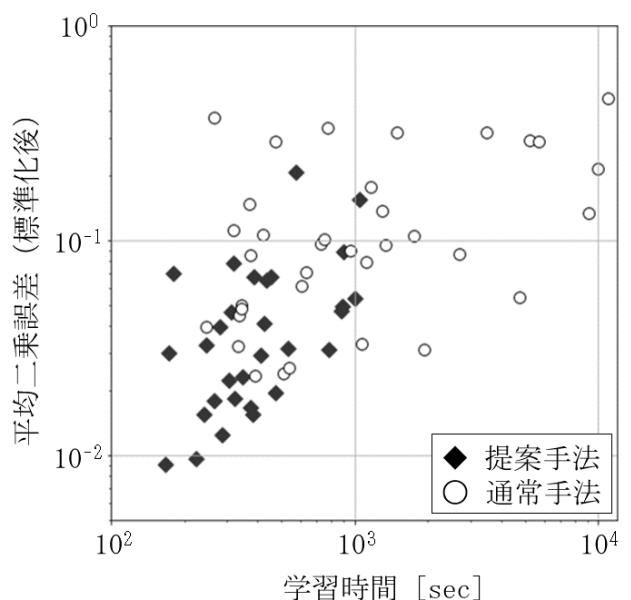


図3 金属組織写真からの硬さの予測性能と学習時間の分布

方二乗誤差 (Root Mean Square Error、以下 RMSE) は 11.8HV で、およそ 8% 程度の誤差で予測出来ていることになる。硬さ測定は、同じ試料であっても測定箇所で値の変化が生じやすい測定手法であることを考慮すると、画像からの特性予測の精度としては実用に足るものであると考えられる。

#### 4 考察

提案手法と通常手法の性能比較を行った結果、条件によるものの、学習データを削減した場合の精度向上の効果が見込めることが示唆された他、学習時間を削減する効果が確認された。

以上のような結果になった要因の一つとしては、学習で調節する必要のあるパラメータが少ないことが挙げられる。提案手法では圧縮アルゴリズムを用いて次元削減を行っているため、結果として学習パラメータも削減されており、モデル形状によっては通常手法の 1 割以下のパラメータ数となっている。このように少ない

パラメータ数で学習させることで、データ数が少ない場合にも過学習が起きにくく、計算負荷が軽くなつたことにより処理時間が短縮されたと考えられる。

#### 5 まとめ

画像圧縮アルゴリズムを画像認識 AI に適用した結果、学習用のデータ数が少ない場合にも、通常手法と比べて正答率は向上もししくは同等であった。

また、作成したモデルで実際の金属組織写真からの特性値予測を行った結果、通常手法よりも学習時間を削減した上で、予測精度を向上させることができた。

#### 参考文献

- 1) Alex K. : Learning Multiple Layers of Features from Tiny Images, (2009).
- 2) Karen S. et al. : Very Deep Convolutional Networks for Large-Scale Image Recognition, ICLR2015 (2015).