

k-means 法を用いたデータセット構築と 畳み込みニューラルネットワークによるその学習精度の評価

宮田 優一[†] 上手 洋子[†] 西尾 芳文[†]

[†] 徳島大学工学部 〒770-8506 徳島県徳島市南常三島 2-1

E-mail: †{y.miyata,uwate,nishio}@ee.tokushima-u.ac.jp

あらまし 近年、ドローンに搭載されたカメラによって空撮が以前より簡単になった。また、畳み込みニューラルネットワーク (CNN) はディープラーニングの一つで、画像認識によく使われるネットワークである。この CNN の発展により、ドローンは農業など、様々な分野への応用に向けた研究が進められている。ドローンに搭載したカメラのバッテリーやメモリは限りがあり、集められたデータを有効に使用することが重要である。本研究では、集めた画像データからより有効なデータセットを作るためにクラスタリングを用いた。今回は特に、k-means 法を用いて画像を 2 クラスタに分類した。このクラスタを用いて様々なデータセットを構築した。このデータセットを CNN に学習させ、クラスタリングする前と後のデータセットで学習精度や計算コストの違いを検討する。

キーワード ニューラルネットワーク, 画像分類, クラスタリング

Building Datasets Using k-means Clustering and its Evaluation of Training Accuracy by Convolutional Neural Networks

Yuichi MIYATA[†], Yoko UWATE[†], and Yoshifumi NISHIO[†]

[†] Electrical and Electronic Engineering, Tokushima University,
2-1 Minamijosanjima, Tokushima, 770-8506 Japan

E-mail: †{y.miyata,uwate,nishio}@ee.tokushima-u.ac.jp

Abstract In recent years, aerial photography became easier than before by using the camera loaded in the drone. Also, convolutional neural network (CNN) is one of deep learning and is a network often used for image recognition. With the development of CNN, drones are being researched for applications in various fields such as agriculture. However, the drone camera's battery and memory are limited. It is important to use the collected data effectively. In this study, we used clustering to make more efficient data sets from the collected image data. This time, we used k-means method to classify the image into two clusters and constructed data sets. This data set is trained using CNN. We examine differences in learning accuracy and computational cost on the data set before and after clustering. We compare the differences in learning accuracy and computational cost with the data set before and after clustering.

Key words Neural Networks, Image Classification, Clustering

1. ま え が き

日本では無線通信技術である 5G 環境の整備が進められており、あらゆる分野でネットワークがさらにつながり、情報やデータを入手しやすくなっている。そのため、インターネットで画像などのデータを多く集めることができるようになったが、AI に学習させる際には、これらの膨大なデータの中から有効なデータを選び、より効率的に使う技術が必要になると予想され

る。また、空撮や物資の輸送などに使われるドローンは様々な分野で私たちの目に触れる機会が多くなった。これは AI、IoT の普及やドローンが安価で手に入る事が要因のひとつであり、今後ますます身近な存在になることが予想される。農業の分野では、作物の収穫量を上げるために畑にドローンを使って、カメラからの画像を AI で分析し、生育状況の把握や適切な農薬の散布などを行っている [1]。また、林業の分野においては、農作物と同様に木々の成長具合や本数の計算などへの応

用が期待されている。さらに、野生動物の増加の影響で、農作物や植林した木の苗の被害も深刻化しているのが現状である。そのため、生態系維持の観点からドローンを使った野生動物の頭数の管理に向けた研究も行われている [2] [3]。これらの問題を解決するためには、人の手では限界があり、ドローンの使用が欠かせない。しかし、ドローン自体やドローンに搭載したカメラのバッテリーには限りがある。そのため、AI に学習させる際には、集めた画像データを効果的に使用することが重要である。そこで、学習のためのデータセットを構築する際にクラスタリングを用いる。クラスタリングとは入力データから特徴の類似性を判断し、自動的に分類する手法であり、教師なし学習に区別される。本研究では、特に、k-means 法を用いて画像データを 2 つに分類し、このクラスタを用いて様々な組み合わせのデータセットを構築する。このデータセットを CNN に学習させ、クラスタリングする前と後のデータセットを学習精度および計算コストで評価する。また、クラスタリングによって分類した画像の割合を変えてデータセットを作ることにより、より質の高いデータセットを提案する。

2. k-means 法によるクラスタリング

k-means 法は、非階層型クラスタリングの一種であり、クラスタリングの手法として最も広く使われる手法の一つである。k-means 法は (1) 式の評価関数を最小化するクラスタの中心を見つけることで、データを任意のクラスタ k 個に分割する。

$$f(\{C_k\}) = \sum_{k=1}^k \sum_{x_i \in C_k} (\bar{x}_k - x_i)^2. \quad (1)$$

k-means 法は、各データ点から最も距離が近いクラスタ中心との距離の総和が、最小となるようなクラスタ中心を求めることで、クラスタリングを行う。

この手法は一般的に 1 次元のデータを扱う。k-means 法のアルゴリズムを以下に示す [4] [5]。

- [ステップ 1] 任意の初期重心点をデータセットからランダムに集めた少量のサンプルデータから決定する。
- [ステップ 2] 各サンプルから最も近い距離の重心点を計算によって求め、クラスタを構成する。
- [ステップ 3] ステップ 2 で求めたクラスタごとに重心を求め、ステップ 2 を再度実行する。ステップ 2~3 を決められた回数繰り返し実行し、クラスタに大きな変化がなくなるまで計算する。

k-means 法は、適切なクラスタ中心を求めて分割するというアルゴリズムで、計算コストが小さいというメリットから、よく用いられる。本研究で用いる手法は画像データを 3 次元から 2 次元に変換し、クラスタリングを行っている。画像は Google の画像検索で検索ワードに一般名称に画像を入力し、画像を収集した。図 1、2 に動物(シカ)が写っている画像と背景のみの

画像を示す。このように、画像は動物が写っている画像を 200 枚、背景のみの画像を 200 枚集めた。また、画像サイズは 432×324 である。



図 1 シカが写っている画像
Fig. 1 Image with object.



図 2 背景のみの画像

Fig. 2 Background image.

3. 畳み込みニューラルネットワークの学習

CNN は主に画像認識に利用されるニューラルネットワークの一つであり、物体検出や自然言語処理などの分野にも広く利用されている。また、CNN はいくつかの層から構成されており、それぞれの層で画像の特徴を検出し、学習を行う。CNN の構成は一般的に、畳み込み層、プーリング層、全結合層から成り立っている。畳み込み層はフィルタを使って画像から特徴を抽出し、特徴マップを生成する。プーリング層は、一つの画像を各領域ごとに比較し、その中の最大値または平均値を抽出することで画像のサイズを小さくしている。全結合層は、特徴を抽出した画像データを一つのノードに結合する。出力層への計算結果を利用し、フィルタ、重み、バイアスを更新する役割がある。図 4 に本研究で用いた CNN の構成を示す。このネットワークは入力層と出力層の他に、隠れ層として 2 つの畳み込み層、2 つのプーリング層、2 つの全結合層から構成されている。このネットワーク内では、画像は 32×32 に圧縮されて処理が進む。学習回数は 200 回である。

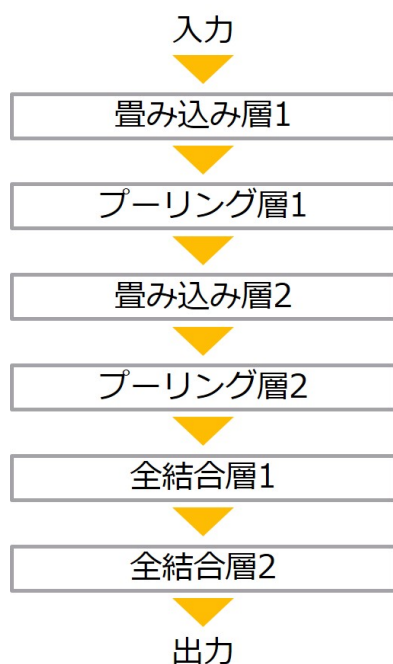


図 3 CNN の構成
Fig. 3 CNN Configuration.

4. 実験結果と検討

4.1 クラスタリングによるデータセット構築

k-means 法を用いて動物がいる画像 200 枚と背景のみの画像 200 枚をそれぞれクラスタリングした。その結果、動物がいる画像は 110 枚と 90 枚の 2 つに分けられ、110 枚をクラスタ A、90 枚をクラスタ B とする。同様に、背景のみの画像は 116 枚と 84 枚に分けられ、116 枚をクラスタ C、84 枚をクラスタ D とする。それぞれの分けられた画像の特徴を見ると、画像全体が茶色いものと緑のものに分けられていた。クラスタリングする前と後の画像を用いて以下のように様々な組み合わせでデータセットを作成した。また、テストデータは学習データ以外の画像 10 枚用意し、各データセットに共通で使用した。

- [データセット 1] : クラスタリング前 (400 枚)
- [データセット 2] : クラスタリング前をランダムで抽出 (200 枚)
- [データセット 3] : クラスタ A とクラスタ C (226 枚)
- [データセット 4] : クラスタ B とクラスタ D (174 枚)
- [データセット 5] : クラスタ A とクラスタ D (194 枚)
- [データセット 6] : クラスタ B とクラスタ C (206 枚)
- [データセット 7] : クラスタ A と B、クラスタ B と C のそれぞれ 50% の枚数をランダムで抽出 (200 枚)

4.2 学習精度と処理時間の比較

構築した各データセットを用いて CNN を学習させ、クラスタリングする前と後の学習精度を比較する。表 1 に各データ

セットにおける学習精度とテスト精度を示す。表 1 より、学習精度はどのデータセットもほぼ 100% に達している。テスト精度においては、各データセットによってばらつきが見られた。クラスタリングを行っていないデータセット 1 は学習に用いた画像が 400 枚のため、高いテスト精度が得られた。また、データセット 2 はデータセット 1 の半分の画像の枚数のため、テスト精度はデータセットよりも低い。次に、クラスタリングを行ったデータセットのテスト精度について比較する。クラスタ A 同士の画像を用いたデータセット 3 のテスト精度が最も高く、その他の組み合わせのデータセットではテスト精度が低かった。さらに、各データセットに使用した画像の枚数は違うが、データセット 4 はデータセット 5 よりも画像が少ないが、テスト精度はデータセット 4 の方が高い結果となった。そのため、画像枚数が多ければ多いほど精度が上がるという相関関係は見られず、ある程度画像に依存することが考えられる。同じ枚数の画像を使用しクラスタリングする前と後とで構築したデータセット 2 と 7 を比較すると、テスト精度はデータセット 7 が高かった。これにより、前処理としてのクラスタリングによる画像内の特徴の多様性がネットワークの学習精度に影響することが分かった。

表 1 各データセットにおける学習精度とテスト精度

Table 1 Average of learning and test accuracies of each data sets.

データセット	1	2	3	4	5	6	7
学習精度 [%]	100	99.7	99.6	99.0	100	99.6	99.6
テスト精度 [%]	94.5	93.5	97.0	66.0	55.0	62.5	96.5

次に、学習にかかった時間を比較し、データセットの学習効率の有効性を評価する。表 2 に各データセットにおける学習にかかった時間と学習効率を示す。学習効率はテスト精度を処理時間で割ることによって算出した。表 2 より、データセットに使用した画像が多ければ多いほど処理にかかる時間が長いことが分かる。学習にかかる時間は使用する画像の枚数に依存することが言える。また、学習効率による評価では、データセット 7 が一番高く、より少ない画像と短い時間で高い精度を出すデータセットであることが分かった。

表 2 各データセットにおける処理時間と学習効率

Table 2 Processing time and learning efficiency of each data sets.

データセット	1	2	3	4	5	6	7
処理時間 [秒]	30.07	16.29	17.77	14.13	15.53	16.43	16.30
学習効率	0.24	0.47	0.43	0.38	0.28	0.30	0.48

図 4 は各データセットにおける学習精度の経過である。横軸はエポック数、縦軸は学習精度である。図 5 より、学習精度はどのデータセットでもエポック数 200 までにはほぼ 100% に到達した。しかし、エポック数 0~50 では、精度の上がり方に違いが見られた。エポック数が 50 の時点でもっとも精度が高かったのはデータセット 5 であった。また、データセット 1 と 7 の比較では、ほぼ同様の学習の進み方であり、データセットの画像枚数は学習の経過にはほぼ影響しないことが言える。

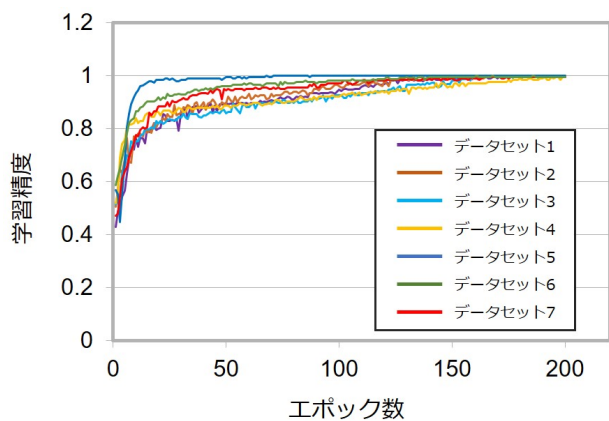


図 4 各データセットにおける学習精度の経過

Fig. 4 Line chart showing learning accuracy of each data sets.

5. ま と め

本研究では、k-means 法を用いて画像データを特徴ごとに 2 つに分類し、分けられた画像を使って様々なデータセットを構築した。各データセットを CNN に学習させ、クラスタリングをする前と後で学習精度にどのような影響があるのかを調査した。クラスタリング後のデータセットにおいて、クラスタリング前よりもテスト精度の低いものがあったが、2 つのクラスターをそれぞれ 50% ずつ使用したデータセットはクラスタリング前より高いテスト精度が得られた。このデータセットは処理時間から算出した学習効率も最も高く、より質の高いデータセットであると言える。本研究により、色や形だけでなく、人間の目では判断できない特徴をクラスタリングによって分類し、学習データとして使用することでより質の高いデータセットを得られることが示唆された。今後の課題として、クラスタリングによってさらに細かく特徴を分類し、CNN に学習させた際に画像内のどの特徴が学習に大きく寄与しているのかを調査する。

文 献

- [1] J. Rebetz, H.F. Satizabal, M. Mota, D. Noll, L. Buchi. “Augmenting a convolutional neural network with local histograms - A case study in crop classification from high-resolution UAV imagery”, ESANN 2016 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 27-29 April 2016.
- [2] Camiel R. Verschoor, Pascal Mettes, Kitso Epema, Lian Pin Koh, Serge Wich. “Nature Conservation Drones for Automatic Localization and Counting of Animals”, European Conference on Computer Vision, ECCV 2014: Computer Vision - ECCV 2014 Workshops pp 255-270 2014.
- [3] Koichi Nakajima, Katsuhiko SEO, Shinji Kitagami, Tetsuo Siotsuki, Hisao Koizumi. “Development of Unmanned Aerial Vehicle Systems for Monitoring and Prevention of Bird and Animal Damage and its evaluation”, Information Processing Society of Japan, IPSJ SIG Technical Report, Vol.2015-MBL-76 No.10 2015.
- [4] V. Birodkar, H. Mobahi, and S. Bengio. “Semantic redundancies in image-classification datasets: The 10 % you don’t need”, ArXiv, 1901.11409, 2019.

- [5] Takashi Onoda, Miho Sakai, Seiji Yamada. “Experimental Comparison of Clustering Results for k-means by using different seeding methods”, The 25th Annual Conference of the Japanese Society for Artificial Intelligence, 2011.