

平滑化画像データにおける k -means 法を用いた画像データの自動ラベル付与 Automatic image data labeling using k -means method in smoothed image data

棚本 侑宏[†] 伏見 卓恭[‡] 大久保 誠也[§] 斉藤 和巳^{† § ¶}
Yukihiko Tanamoto Takayasu Fushimi Seiya Okubo Kazumi Saito

1. はじめに

近年, IoT (Internet of Things) 技術が急速に発展している. それにともない, 農業・医療・教育などの専門知識と熟練が不可欠な分野においても, 専門家や熟練者がどのような状況でどのような活動をしているかに関する情報を, 容易に入手可能となってきた. 一方, 特に農業分野では担い手の高齢化による労働力不足が深刻化しており, 作業の合理化や技術の継承が課題となっている.

本研究の目的は, 農産物の生産性向上に向けて, 環境情報オンライン可視化システムのプロトタイプ構築であり, その一環として, 熟練農家に特有な行動の解析にも焦点を当てる. 具体的には, 農業環境データを収集・分析することにより, データの可視化並びに熟練農家特有の行動の解析を行う. 農業環境データとして, 温度・湿度等のセンサーにより値を直接得られるものと, 画像のように値を直接得られないものを使用する. このうち, 画像データは後からの分析により, さまざまな情報を得ることができる. 例えば, 高品質なバラを生産する熟練バラ農家では, 花の日焼けなど品質低下を抑制するための行動として, ビニールハウス天井に設置したカーテンを開閉する. この開閉情報は, 専用のセンサーをあらかじめつけずとも, 画像から得ることができる. しかしながら, 画像データの分類を手動で行うことには限界がある. このように, 画像などの分類が実現できれば, カーテンの開閉だけでなく, 熟練農家の多様な行動の分析への道も開かれると期待できる.

このような画像分類問題では, 一般に, 教師あり学習として定式化され, 深層学習やサポートベクターマシンにより, 高精度な分類器を得ることが期待できる. これらの手法を用いるためには, 学習データの作成が必要である. 学習データとして多くのデータが必要となる場合があり, 人手による作業では限界がある. そこで本研究では, 教師なし学習により画像の分類を行う. 具体的には, 画像などのデータをクラス数 c に分類する問題において, c より大きな k での k -means 法を適用し, 得られた k 個のクラスを c 個のクラスに併合する方法を提案する.

今回の研究での提案手法は, [1] で提案した手法を改良したものである. [1] で提案した手

法は, k -means 法で得られる各クラス重心ベクトルの集合に対し, 重心ベクトル間の距離に基づき, 最近傍 (1-nn: nearest neighbour) グラフを構築する. その結果の連結成分数 n が c のときには, 各成分に含まれるクラスを同じクラスに属すとし終了する. 一方, n が c より小さいときには, 連結成分数 n が c に等しくなるまで, 距離の大きい順にリンク削除を繰り返す. 逆に, n が c より小さいときには, 連結成分数 n が c に等しくなるまで, 第2近傍, 第3近傍と広げつつ距離の小さい順にリンク追加を繰り返す.

この手法は概ねうまくいくものの, 実験によりいくつかの問題が残っていた. 例えば, 撮影した状態の原画像を使用していたことである. この原画像集合はクラスタリングを行うと, night と判定されてしまう分析に使わずに真つ暗な画像が全体の 40 % を占めていた. そのため, 早朝等の暗めの画像はカーテンの開閉があるにも関わらず, night と判定されてしまっていた. そこで, 本研究では, 前処理として画像データを平滑化することにより改善した手法を提案する.

提案法の有効性を検証するには, どの程度のサンプル数 s が必要かとともに, どの程度の大きさにクラス数 k を設定すべきかが重要なポイントとなる. 本研究では, 実際にバラ農家のビニールハウスで撮影された画像データを利用し, 実証的に提案法の有効性を評価する. また, 実験結果から, 平滑化処理を適用しても, ある程度大くのサンプル数 s を利用し, クラス数 k も妥当な範囲の設定で, 比較的安定して高い正答率が得られること示す.

2. 関連研究

本章では, 農業画像データの分類に関する研究と, 農業画像から特徴の自動取得に関する関連研究について述べる. 農業画像データに対し機械学習を適用する研究も多様に行われている. 例えば, 畳み込みニューラルネットワークを用いた野草画像の分類 [2] では, 少ない訓練データからでも水増し方法によって 76 % 以上の精度で分類が可能であることを示した. 大豆の生育情報を自動取得する画像センシング手法の開発 [3] の研究では, 大豆の花, 子実の画像を認識するシステムを提案した. Single Shot MultiBox Detector (SSD) による高速検出を可能とし, 同一画像内に含まれる複数の花や子実を検出できることを示した.

[†]神奈川大学 理学部

[‡]東京工科大学 コンピュータサイエンス学部

[§]静岡県立大学 経営情報学部

[¶]理化学研究所 革新知能統合研究センター

3. 提案手法

データの自動分類を教師無し学習を用いて行った際、データ内に分類しやすい特徴が存在すると、意図しない分類が行われてしまう場合がある。たとえば、カーテンの開閉状態で分類をするために2つのクラスに分類することを考える。このとき、カーテンの開閉よりも明るさ等に強い特徴があった場合、そちらを重視した分類が行われてしまう場合があり、各クラスにカーテン開と閉の2つの画像が混じってしまう。しかしながら、より細かいクラスに分割していくと、あるクラス数に達した時点で、各クラスの中はカーテンの開閉の片方のみが含まれるようになることが期待できる。

そこで、本研究では、ラベルなし画像データが与えられたときに、指定したクラス数以上にクラスリングを行い、その後、重心間の距離より類似したクラスを併合し、得られたクラスにラベリングを行う手法を提案する。また、前回までの研究 [1] では、night の枚数が全体の 40% を占めていた。このため、画像に対して平滑化を行うことで、画像の輝度を調節し、真っ暗であった画像が open, close に移動することが期待でき、より精緻な分析が可能になる。

提案手法は、従来の k -means クラスターリングに k -NN 法を併せた方法である。要求するクラス数 c より多くのクラス数 k に k -means 法でクラスターリングを行い、その後、 k 個のクラス間で類似しているクラス同士を併合することにより、最終的に指定クラス数 c に分類する。具体的には、以下の手順で行う。

入力：分類する要素の集合 \mathcal{X} ならびに分割数 c

1. \mathcal{X} に対して平滑化を行う。
2. 整数 k を決定する。
3. k -means 法で c よりも大きな値である k で k 個のクラスにクラスターリングを行う。
4. k 個の各クラスの近傍クラスを求める。具体的には、クラスターリングにより得られた各クラスの重心ベクトル座標に対してそれぞれの重心間のユークリッド距離の二乗を求める。 N 次元空間ベクトル p と q 間でのその計算式は以下である。

$$d(p, q) = (p_1 - q_1)^2 + \dots + (p_N - q_N)^2$$

5. 求めた各クラス間の距離に基づき、各クラスの重心と最近傍クラスの重心にリンクを張り、最近傍グラフを構築する。
6. 併合・分割を行うことで、 c 個のクラスにする。具体的には、最近傍グラフの連結成分数が、

- c と等しければ、各クラスを含むクラスを同類なクラスであるとする。

- c より大きい時、各クラスの第2近傍、第3近傍…と近傍クラスの考慮する範囲を広げていき、重心間の距離が一番小さいかつ同じクラスに属していないという制約の下でリンクを繋ぐ。これを連結成分数が c と等しくなるまで繰り返す。
- c より小さい時、既に存在しているリンクの中でクラス間距離が大きなものから順にクラス間のリンクを削除していく。これを連結成分数が c と等しくなるまで繰り返す。

7. 作成された c 個の新たなクラスに含まれている各クラスに、新クラスのラベルを付与する。

ステップ 2~7 は [1] と同様の処理である。

4. 評価実験

評価実験により、提案手法の特徴と有効性の評価を行った。具体的には、静岡県内にある2件のバラ農家のビニールハウス（ハウス A と B）から画像データを収集し、得られた画像に対して提案手法を適用した。そして、得られたクラスターリングの精度を評価した。画像データは取得した状態の原画像データと平滑化を行った画像データの2種類のデータ集合に対して実験を行い、比較した。また、 k -means におけるサンプル数とパラメータ k が精度に与える影響も評価した。

本実験では、ビニールハウスで撮影した画像データを、“カーテン開け (open)”・“カーテン閉め (close)”・“夜 (night)” の3種類に分類する。使用する環境情報データは、ビニールハウスに環境情報収集デバイスを設置することにより収集した。赤外線カメラモジュール (Raspberry Pi PiNoir CameraModule V2 赤外線カメラモデル使用) は、5分ごとにサイズ 480x360 の画像を撮影する。撮影された画像の例を、図 1 に、平滑化を行った画像の例を図 2 に示す。本実験で用いたデータは、2019年6月21日0時0分から2019年8月14日0時00分までのデータである。1日あたり144毎の写真があり、それが108日分であるため、各農家に対して15552個の画像データとなる。

実験に用いた画像データにおける平滑化前のラベル数の統計量を表 1 に、平滑化後のラベル数の統計量を表 2 に示す。これらラベルは本研究を遂行するために人手で付与したものである。表 1 より、定時に画像を撮影しているため、各ハウスで合計は一致し、これらハウスは比較的隣りに位置するため night 数は、ほぼ等しい数となっている。一方、open 数と close 数に関しては違いが見られ、ハウス A の方が比較的 open 数が多いことが分かる。表 2 より、平滑化を行うことで、平滑化前は night ラベル

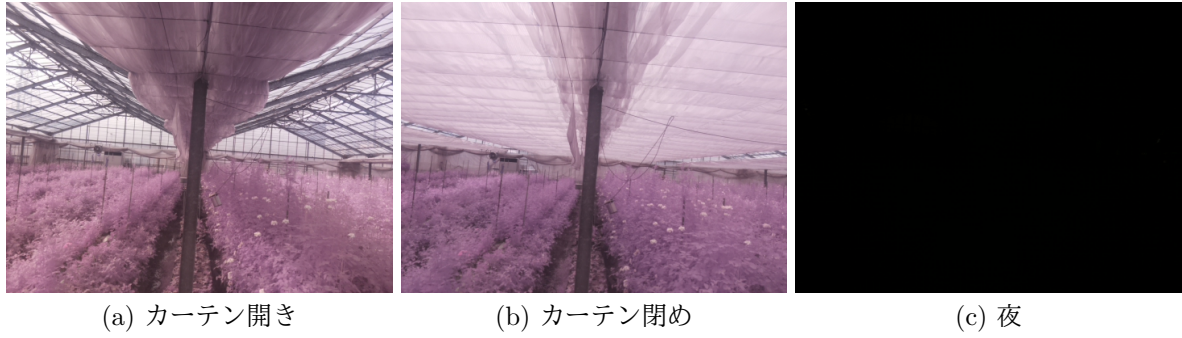


図 1: 平滑化前画像データ例

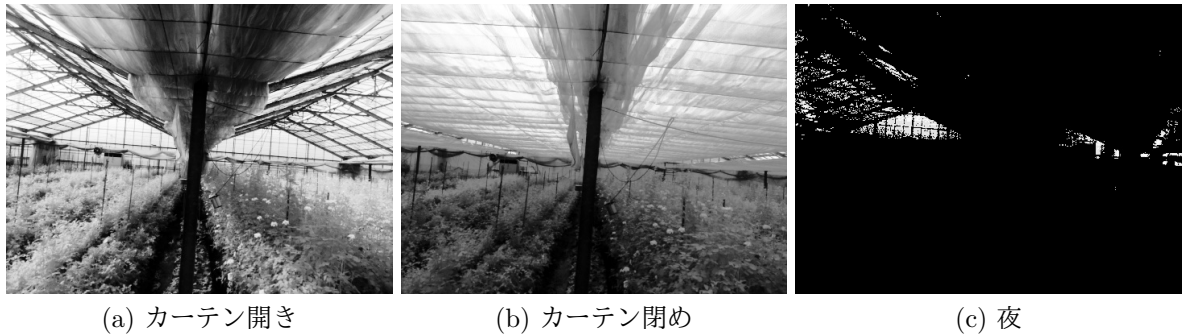


図 2: 平滑化後画像データ例

表 1: 平滑化前のラベル数の統計量

	open 数	close 数	night 数	合計
ハウス A	7,130	2,331	6,091	15,552
ハウス B	4,343	5,102	6,107	15,552

表 2: 平滑化後のラベル数の統計量

	open 数	close 数	night 数	合計
ハウス A	7,330	2,559	5,663	15,552
ハウス B	4,557	5,260	5,735	15,552

であった画像数百枚は新たに open や close のラベルに移動したことが分かる。

提案手法により c 個のクラスに分類した後 ($c = 3$), 各クラスに含まれる正解画像ラベルから精度 (*Accuracy*) を求めた。具体的には, 各クラス $\mathcal{L}(i)$ ($i \in \{1, \dots, c\}$) ごとに含まれているサンプルデータの正解ラベルから, open, close, night のどのクラスに該当するか調べた。ここで, 得られた各クラスごとの open, close, night の個数をそれぞれ O_i, C_i, N_i とし, 各クラスはそれぞれ $\arg \max\{O_i, C_i, N_i\}$ によって得られた値で該当クラスを求めた。この時, 各クラスにおいて誤りとなる要素数 D_i を以下の式により求めた。

$$D_i = |\mathcal{L}(i)| - \max\{O_i, C_i, N_i\}$$

そして, 精度 (*Accuracy*) を次式で求めた。

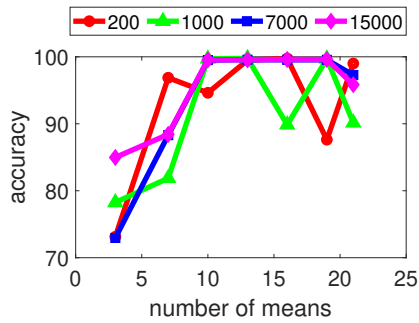
$$Accuracy = \left(1 - \frac{\sum_{i=1}^c D_i}{\sum_{i=1}^c |\mathcal{L}(i)|}\right) \times 100$$

平滑化前の結果を図 3 に, 平滑化後の結果を図 4 に示す。横軸が k , 縦軸が精度 *Accuracy*, 各色の直線が k -means を実行する際に使用したサンプルデータ数を意味する。 k の値は, $k \in \{3, 6, 9, 12, 15, 18, 21\}$ とした。 $k = 3$ のときは, k -means のみを用いた従来の手法と見なせる。

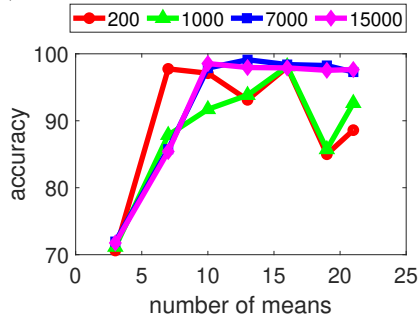
5. 考察

表 1 と 2 から, 平滑化後はハウス A では night の画像 428 枚, ハウス B では 372 枚の画像が open, close に移動した。ハウス A では, 428 枚のうち約 46% が open へ, 残り 54% が close へと移動した。ハウス B では, 372 枚のうち約 58% が open へ, 残り 42% が close へと移動した。この結果から平滑化は原画像の時のクラスタリングと比較して, より精緻な環境分析には重要であると考えられる。ハウス A の k -means のみを用いた手法では, 平滑化前は 70% ~ 85% であるのに対して, 平滑化後はサンプル数が小さい時でも 95% 以上の精度を出している。しかし, k が 7 でサンプル数が 200 枚, 1000 枚の時は, 精度が約 90% に下がっている。

平滑化後の結果は, 平滑化前に比べてサンプル数が十分な値でなく k の値を変化させた時の精度が, 比較的安定していることが分かる。



(a) 平滑化前のハウス A での正答率評価



(b) 平滑化前のハウス B での正答率評価

図 3: 提案法による自動ラベル付与結果の評価

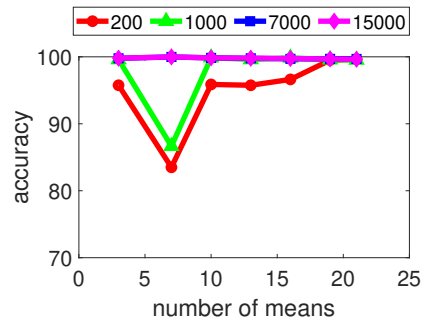
ハウス B は、平滑化前後で、 k -means のみの時の精度に大きな違いは見られなかった。そして、平滑化後のデータ集合の方が安定して高い精度でクラスタリングを行うことが可能である。しかし、平滑化後のデータ集合において、サンプル数が 200 で、 k が 13 ~ 16 の時、精度が大きく下がってしまっている。また、同データにおいては、平滑化前の方が比較的に高い精度でクラスタリングが行われている。

以上より、本研究の提案手法は、一定数のサンプルデータ (図 3, 図 4 では 7000 枚以上) に対しては、安定して高い精度でクラスタリングを行うことができる。さらに、画像データ集合に対して、平滑化処理を行うことで、 k の値を変動させても安定した精度を保つことが可能である。よって、さらなる実証実験が必要なものの、複雑な画像データが含まれているデータ集合に対しても有用であると期待できる。

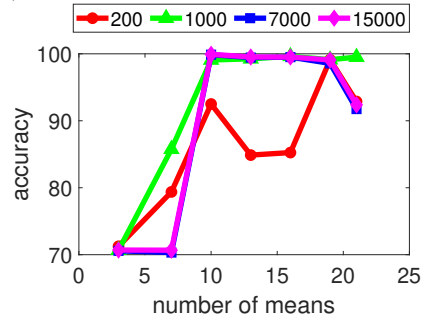
6. おわりに

本研究では、画像データ集合へのラベル付与のため、 k -means 法と k -NN 法を併せたクラスタリング手法を提案した。また、提案手法の有効性と特性を調べるために、画像の平滑化を行ったときと行わなかった時の計算機実験による比較を行った。その結果、提案法は従来の k -means 法と比較して安定して高い精度でクラスタリングを行うことができることが明らかとなった。

さらに、前回の研究 [1] と比較すると、night と判定されていたデータが数百枚 open, close に移動し、より精緻な分析が可能になった。



(a) 平滑化後のハウス A での正答率評価



(b) 平滑化後のハウス B での正答率評価

図 4: 提案法による自動ラベル付与結果の評価

今後の課題として、 k -means 法の k の今回の実験より値を大きくした時も安定して高い精度でクラスタリングをすることができるのかを検証することなどが挙げられる。

謝辞 本研究は、JSPS 科研費 (C)(No.18K11441) の助成を受けた。

参考文献

- [1] 棚本侑宏, 伏見卓恭, 岩崎清斗, 大久保誠也, 斉藤和巳. "k-means 法を用いた画像データの自動ラベル付与," 情報処理学会第 199 回知能システム研究会, 2020.
- [2] 渡邊葵, 櫻井彰人, "畳み込みニューラルネットワークを用いた花画像の分類," 情報処理学会第 78 回会全国大会, 2016.
- [3] 大村和暉, 八幡壮, 小澤誠一, 大川剛直, 村上則幸, 辻博之, "大豆の生育情報を自動取得する画像センシング手法の開発: Single Shot MultiBox Detector の導入," 人工知能学第 32 回会全国大会, 2018.