

## SVM を用いた海苔の等級判定におけるカーネル関数の性能比較

## Comparing SVM Kernel Functions for Grading Seaweed Quality

○出井俊成<sup>1</sup>, 平石秀史<sup>2</sup>\*Toshinari Idei<sup>1</sup>, Hidefumi Hiraishi<sup>2</sup>

Abstract: In this research, we evaluate the effectiveness of four kernel functions (linear, polynomials, gaussian, sigmoid) of the support vector machine (SVM) in tasks to grade seaweeds. SVM is known to achieve high precision in classification tasks with small datasets. The previous research experimentally shows that SVM with linear kernel exhibits more than 90% accuracy in grading seaweeds. This research explores the possibility of enhancing accuracy with different kernel functions.

## 1. はじめに

現在、海苔の等級判定は各地の漁業組合の有資格者が視覚・嗅覚・味覚・触覚などを用いた検査により実施している。しかし、有資格者の高齢化による後継者不足が進んでおり、海苔の等級判定を自動化することが要望されている。海苔の加工に関する状況に関しては[1]に詳しい。

海苔の等級判定の自動化へ向けた取り組みとして、深層学習と多数決方式を組み合わせた手法[2]や、海苔の画像データを水増しすることで深層学習の精度改善を試みた報告[3]がある。一方で、深層学習をはじめ、機械学習により高精度の分類を行うためには一般に多数のデータセットが必要とされる。海苔は季節性の産物なため、高品質で十分な量のデータを入手することは困難とされている。このようなデータの不足を補うための取り組みとして、[4]では少ないデータ量でも高精度の分類が可能となることが知られているサポートベクトルマシン (SVM) を用い、さらに画像の色空間を HSV に変換した画像データを用いることで高精度の等級判定が実現できる可能性が報告されている。また、グローリー株式会社は DX 化への取り組みの一環として、AI 技術を用いて海苔の等級判定を自動化する実験を開始したと発表している[5]。

本研究では[4]を受けて、SVM を用いた判定精度の改善に向けた予備実験として、様々なカーネル関数を用いた場合の判定精度の変化に関する比較実験を行った。

## 2. 実験手法

SVM は、与えられた学習データに対し、適切な超平面を定めて分離する機械学習モデルである。超平面を定める際に、各クラスのデータから距離が最大になるように定めることで、少ないデータセットでも高い汎化性能を達成できる可能性をもつモデルである。SVM は厳密な線形分離を行うハードマージンと、誤差を許容した分離を行うソフトマージンという二種類の基礎的なモデルがある。また、SVM を用いて更に柔軟な分類を可能にするための手法にカーネル法がある。カーネル法とは、学習データを高次元に写すと同時に、学習過程で現れる内積計算を高次元に写したデータを明示的に使わずに行える方法である。カーネル法を用いることで、低次元では線形分離できないデータを高次元化することで線形分離可能になる場合がある。

今回の実験では、海苔の等級 (1~5 等級) に対応する 5 クラス分類を SVM により行う。特に SVM においてよく使われる以下の 4 つのカーネルに関して実験を行う。各カーネルに関する以下の説明は[6]に準ずる。

## ① 線形カーネル

$$k(x_i, x_j) = x_i^T x_j$$

## ② 多項式カーネル

$$k(x_i, x_j) = \gamma(x_i^T x_j + c)^d$$

## ③ ガウスカーネル

$$k(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$$

## ④ シグモイドカーネル

$$k(x_i, x_j) = \tanh(\gamma x_i^T x_j + c)$$

SVMを構築するにあたり、用いる特徴量ベクトルは海苔の画像データのHSV色空間を直接用いる。海苔の画像データ(3400×3250pixel)(1~3等級は45枚, 4等級は44枚, 5等級は46枚)を縦横共に0.1倍に縮小したものを用いた。特徴ベクトルの次元は33万程度となる。画像を学習データ(1~3等級は30枚, 4等級は29枚, 5等級は31枚)テストデータ(各等級15枚ずつ)に分け、4つのカーネルで等級判別を行い、正答率を出力した。また、実験に用いたプログラムではSVM構築のためにopenCVのライブラリを用いた。

### 3. 実験結果

実験結果を表1に記載する。次数4の多項式カーネルとガウスカーネルにおいては、 $\gamma$ の値によっては判定結果が全て5等級と判定されたことから $\gamma$ の値を調節して実験を行った。実験結果(表1)から海苔の等級判別において、一番正答率が高く判別されたカーネルは線形カーネルであった。多項式カーネルについては次数2~4で行ったが正答率は92%と全部同じ正答率になった。ガウスカーネルでは正答率は9割未満となり、シグモイドカーネルでは $\gamma$ の値を0.1から $10^{-16}$ まで変えて行ったが全部5等級と判別された。

### 4. 結論と展望

今回の実験結果から、海苔の等級判別においては線形カーネルや多項式カーネルが適しており、ガウスカーネルやシグモイドカーネルは適していない可能性が示唆される。

今回の実験では[4]との比較をするため、画像の色空間をHSVで固定した上で、各カーネル関数の正答率を比較した。色空間を変えることで判定精度の改善が可能かを検討するために、RGB色空間をはじめとした各種の色空間とカーネル関数の様々な組み合わせに対しても比較実験を行いたい。

また、今回の実験では前述のとおり、画像の各ピクセルの色情報を特徴量として用いた。これにより、分類タスクの複雑さ(単純さ)に比べ、特徴量ベクトルの次元数が膨大になっている可能性が考えられる。この問題点を解決するためには、画像から適切な特徴量を抽出し、次元数の少ない適切な特徴量ベクトルを設定することが課題である。例えば、画像のピクセルを頂点とみなし、隣接ピクセル間の色の変化を基準に対応する頂点同士に辺を張ることでグラフを構築し、そのグラフの不変量の特徴ベクトルに用いることが考えられる。このようなグラフの連結度を調べて、海苔の画像の滑らかさなどを指標として用いることができないかを検討することが今後の課題である。

### 5. 参考文献

[1] 工藤盛徳, 稲野達郎, 高岡則夫, 小磯潮, 加工海苔入門 改訂版, 日本食糧新聞社, 2009.

[2] 鈴木大智, 栗野俊一, 深層学習による海苔の品質判定のデータ分割の有効性, 令和3年度日本大学理工学部学術講演会予稿集, P-1, pp. 831-832, 2021

[3] 阿部舜耶, 平石秀史, 転移学習による海苔の品質判定におけるData Augmentationの有効性, 令和4年度日本大学理工学部学術講演会予稿集, P-3, pp. 830-831, 2022

[4] 鈴木開渡, 平石秀史, サポートベクトルマシンによる海苔の品質判定におけるHSV色空間の有効性, 令和5年度日本大学理工学部学術講演会予稿集, P-11, pp. 847-848, 2023

[5] 日本経済新聞(2023年9月20日), “グローリー、海苔検査業務のDX化を目指した実証実験を開始” [https://www.nikkei.com/article/DGXZRSP662228\\_Q3A920C2000000/](https://www.nikkei.com/article/DGXZRSP662228_Q3A920C2000000/).

[6] 田村孝廣, やさしく学べるサポートベクトルマシン, オーム社, 2022

	線形	2次多項式	3次多項式	4次多項式	ガウス	シグモイド
正答率	0.93	0.92	0.92	0.92	0.89	0.2
$\gamma$ の値	1	1	1	$10^{-9}, 10^{-10}$	$2 \times 10^{-9}$	$10^{-1} \sim 10^{-16}$

表1: HSV色空間を特徴ベクトルとした場合の各カーネル関数での正答率(小数点第3位で切捨)