

卒業論文

# サポートベクタマシンを用いたラッ プフィルムの不良品検出

-畳み込みニューラルネットワークを特徴抽出器として用いた場合-

F117031 清水 竜樹

指導教員 永田寅臣 教授

2021 年 3 月

山口東京理科大学 工学部 機械工学科



## 概要

筆者らは畳み込みニューラルネットワーク (CNN) やサポートベクタマシン (SVM) を簡単に設計できるツールを開発している．本研究ではこのツールを用いて，ラップフィルムの製造工程で発生する不良品の検出を試みる．まず，ラップフィルムの全体画像に対してテンプレートマッチングを適用し，フィルム部分のみを抽出する．特徴抽出器として学習済みの CNN である AlexNet と VGG19 を用いた SVM を使用し，汎化性を高めるため多数の良品と不良品の画像を用いて学習を行う．実験では，特徴抽出器とする CNN のほか，それぞれについてカーネル関数を gaussian と polynomial とした場合について訓練及びテストを行った．次に，各カーネル関数のコストパラメータ  $C$  を変更し，より精度の高い結果を得られるパラメータを探索した．



# 目次

|       |                           |    |
|-------|---------------------------|----|
| 第 1 章 | 緒言                        | 1  |
| 第 2 章 | 人工知能                      | 3  |
| 2.1   | 機械学習 . . . . .            | 3  |
| 2.2   | 深層学習 . . . . .            | 5  |
| 第 3 章 | サポートベクタマシン                | 6  |
| 3.1   | サポートベクタマシンの仕組み . . . . .  | 6  |
| 3.2   | カーネル法とカーネル関数 . . . . .    | 7  |
| 3.3   | ソフトマージンとハードマージン . . . . . | 7  |
| 第 4 章 | 画像処理                      | 8  |
| 4.1   | テンプレートマッチング . . . . .     | 8  |
| 4.2   | 画像オーギュメンテーション . . . . .   | 9  |
| 第 5 章 | 実験                        | 10 |
| 5.1   | 使用器具 . . . . .            | 10 |
| 5.2   | 方法 . . . . .              | 10 |
| 5.3   | 結果 . . . . .              | 10 |
| 第 6 章 | 考察                        | 11 |
| 第 7 章 | 結言                        | 13 |
|       | 謝辞                        | 14 |
|       | 参考文献                      | 15 |



# 第 1 章

## 緒言

様々な工業製品の検査工程においては一部で自動化が進んでいるものの、それぞれの製品の品質管理に精通した検査員の目視検査に頼るところが大きい状況である。最近では、画像認識に特化した人工知能を製品の欠陥検出に応用しようとする試みがなされており、その中にはサポートベクタマシンも含まれる。

これまでもサポートベクターマシン (SVM) を製品の欠陥検出に応用した研究がある。例えば、西村らは HOG 特徴量を用いて、SVM を電子部品の欠陥検出に使用した場合の課題についての研究を行っている [1]。また、Yang らは SIFT 特徴量を用いたサポートベクタマシンによる輸送用パッケージの欠陥検出を試みている [2]。

本研究では、ラップフィルム品の欠陥検出を行うことができる SVM を提案する。SVM では特徴抽出器から取得した特徴ベクトルを用いて学習を行い、不良品と良品を判別する境界面の作成に必要なサポートベクトルを決定する。ベクトルデータの高次元化のためにカーネル法を用いている。実験では、特徴抽出器のほか、カーネル関数と正則化係数を変更し、より精度の高い分類が可能となる SVM を探索する。特徴抽出器には AlexNet あるいは VGG19 を用い、カーネル関数にはガウス関数と多項式関数の二通りをそれぞれ組み合わせて 2 クラス分類を行う SVM (TCSVM) を設計する。

図 1.1 のような良品と不良品からなるラップフィルム品の訓練用データセットを用意し、それぞれの画像から 4096 次元の特徴ベクトルを抽出し、SVM の学習を行う。その後、学習後の SVM が未学習のデータセットをどの程度正しく分類することができるかを評価する。

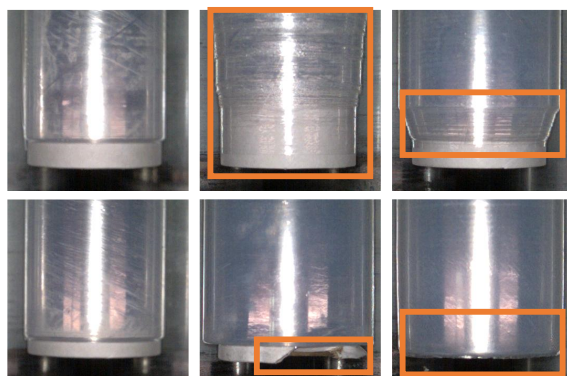


図 1.1. Samples of training images without and with defects (marked with orange rectangle).



## 第 2 章

# 人工知能

人工知能とは、人工的に作られた知能である。この時、知能という言葉が含む意味は様々である。

機械が知能を持つことがあるのかという議論は、1950 年にアラン・チューリングが発表した論文が始まりであるといわれており、チューリングの提唱によれば、人間が質問者となりコンピュータが回答者となる「模倣ゲーム」において、質問者が回答者を人間であると誤認識する回答をできるコンピュータは知能を持っているということになる。この「模倣ゲーム」は、現在では「チューリング・テスト」として人工知能の一つの指標ともされている。

現在使用されている人工知能 (AI) という言葉は、1956 年にダートマス大学で行われたダートマス会議で作られたといわれている。このダートマス会議後から 1980 年頃まで、人工知能の研究は盛んになる。

2000 年代に入ると、コンピュータの処理能力の向上や、インターネットの発達などの要因により情報収集や情報交換が簡単になった。また、画像認識や音声認識、自然言語処理などの技術へも関心が高まり、身近なものになった。これらやセンサなどの普及により、実世界のデータを大量に計測することなども可能となったため大量のデータ (ビッグデータ) に基づく機械学習が可能となった。2010 年代には画像認識や音声認識、自然言語処理の分野においても機械学習が必須の技術と考えられるようになり、現在の人工知能の研究においては、機械学習は欠かせないものとなっている。

### 2.1 機械学習

機械学習の中には教師あり学習と教師なし学習という 2 つの学習方法がある。教師あり学習とは、ある入力に対して特定の出力を予測したい場合で、入力出力のペアの例が入手できる際に有効な機械学習の手法である。教師データは、入力に対し出力が何になるかという正解となるデータを必要とする。このデータを学習した機械に対し未知のデータを入力し、学習をもとに正解を予測することが教師あり学習である。教師あり学習は訓練セットを作るために人手が必要になるものの、学習を終えることができれば人手がかかる作業を省略することなども可能になる。

教師あり学習はさらにクラス分類と回帰に大別される。この二つのどちらを適用するかは、予測結果に関わる。回帰は予測結果が連続性を持つ場合に有効な手法であり、数値の予測などに用いられる。一方クラス分類は、あらかじめ定められた選択肢の中からクラスラベルを予測することが目的であり、データがどの集合に属するかの予測に用いられる。

教師なし学習は教師あり学習と違い学習に教師情報を用いない全ての種類の機械学習が含まれる。教師なし学習では、アルゴリズムには入力データだけが与えられ、データから知識を抽出することが要求される。そのため、の手法で代表的なものが、入力データを人間や機械にとってわかりやすい新しいデータを作る教師なし変換と、似ているデータ同士を同じグループに分けるクラスタリングである。

教師なし変換の利用法として挙げられるものの一つが次元削減である。次元削減は、多数の特徴量で構成されるデータを入力として、本質的な特徴を表す特徴量でそのデータを表す要約方法を見つけることである。一方、クラスタリングは、データを似たような要素から構成されるグループに分ける手法である。クラス分類との相違点は、クラスタリングには正解となるデータが与えられないという点である。そのため、クラスタリングの結果が必ずしも予測に使えるとは限らない。

教師なし学習のアルゴリズムに与えられるデータは正解となるデータが含まれていない。そのため、アルゴリズムの出力結果が正しいのか、ユーザが求める結果を出力しているのかを確実に評価する方法は人間が確かめるしかないため、教師なし学習は教師あり学習用にデータを次元削減するなどの目的に用いられることが多くある。

教師なし学習のクラスタリングでは、正解となるデータなどは必要ないものの、クラス分類のような結果が出るとは限らなかった。しかし、クラス分類の学習には教師データが必要となるという点でコストがかかる。それを解決するための手法に半教師あり学習がある。半教師あり学習 1 つの方法としては、少量の正解ラベルのついたデータを用意する。そのデータをもとに、ラベルの付いていないデータの正解を予測し、予測されたデータの中から正解である確率が高いものを正解ラベルの付いたデータに加え再度同じ学習を繰り返すというものである。

機械学習の中にはさらに強化学習というものが存在する。強化学習は、一連の行動の最後に報酬として評価値を得て、その評価値を得るに至るまでの過程の行動それぞれについて報酬を分配して与えるという学習方法である。

今回の研究で用いているものは教師あり学習によるクラス分類である。

## 2.2 深層学習

## 第 3 章

# サポートベクタマシン

サポートベクタマシン (SVM) は 2 クラス分類問題の代表的手法であり，未知データに対して高い予測精度を持つ分類器が構築可能であることが知られている [7]．本章では，SVM の基本的な仕組みである分類境界となる超平面の決定方法，非線形分類を可能とするカーネル法，さらに分離可能でないデータを SVM で分類する際に適用されるソフトマージンとハードマージンの差異について述べる．

### 3.1 サポートベクタマシンの仕組み

一般に，訓練集合を分離できる分類境界は複数存在し得るため，様々な分類境界の作成方法が考えられるが，SVM ではそれぞれのクラスのデータと分類境界の距離（マージン）が最も離れるように分類境界を決定するマージン最大化と呼ばれる方法で境界を決定している．この時，マージン最大化は分類境界と分類境界から最も近くに存在するデータとの距離を最大化することで実現可能である．また，分類境界の決定に使用されるデータをサポートベクタと呼び，サポートベクタは少なくとも 1 つ存在する．

SVM では，特徴抽出器を使用し各入力画像から多次元特徴ベクトル  $x = [x_1, x_2, \dots, x_n]^T$  を抽出する．この特徴ベクトルが SVM の学習に使用され，学習によって  $f(x) = 0$  で与えられる超平面と呼ばれる決定境界が得られる． $f(x)$  は，これから分類される特徴ベクトル  $x$  から超平面までの符号付き距離であり，次式で与えられる．

$$f(x) = \sum_{i=1}^N \alpha_i y_i G(x_i^*, x) + b \quad (3.1)$$

ここで， $G(x_i^*, x)$  はカーネル関数， $x_i^* \in \mathbb{R}^{1 \times N}$  ( $i = 1, 2, \dots, N$ ) は学習により決定されたサポートベクトル， $N$  はサポートベクトルの数， $\alpha_i$  ( $i = 1, 2, \dots, N$ ) と  $b$  はそれぞれラグランジュ乗数とバイアスであり，これらは逐次最小最適化 (SMO) を用いた学習によって推定された SVM パラメータである．また， $y_i$  は 2 クラス学習の場合， $x_i^*$  の陽性あるいは陰性に依じて  $-1$  または  $1$  に設定されたラベルである．分類はスコア  $f(x)$  の符号によって行われ， $f(x) > 0$  の場合は陰性， $f(x) < 0$  の場合は陽性であるといったように判断される．

## 3.2 カーネル法とカーネル関数

SVM に与えられるデータは必ずしも線形分離可能とは限らず，問題によっては非線形分離を行う必要がある場合も発生する．このような問題を解決するにあたり，データを入力した際に作られた入力空間を高次元の特徴空間と呼ばれる空間に写像し，特徴空間上でマージンが最大となるように超平面を決定する方法がとられる．

式 3.1 にあるように，超平面を決定する際にカーネル関数と呼ばれる関数を使用する．このカーネル関数を使用することにより入力空間から特徴空間に写像し，非線形分離を可能とする．また，カーネル関数を使用するこれらの方法をカーネル法と呼ぶ．

## 3.3 ソフトマージンとハードマージン

## 第 4 章

# 画像処理

本研究では画像を対象とする．機械学習において，収集した画像をそのまま扱うことは稀であり，何らかの処理を施して CNN などでの処理が可能となるように変換する．本章では，ターゲットとする画像の中から必要となる一部分を抽出するテンプレートマッチングと，機械学習を行う際に過学習を防ぐことを目的としてデータを加工する画像オーギュメンテーションについて述べる．

### 4.1 テンプレートマッチング

本節では，治具を含むラップフィルム品の全体画像に対してテンプレートマッチングを適用し，ラップフィルム部分のみを抽出する画像処理を行う．様々な課題に対する画像処理で広く利用されているテンプレートマッチングは，撮影されたワークの中で欠陥が含まれやすい目標領域を抽出するためにも非常に有効である．これを使用して製品の欠陥と関係しない治具等を取り除き，欠陥があらわれる製品部分のみを抽出する [3] ．

テンプレートマッチングの処理では， $(M, N)$  のサイズのテンプレートをターゲット画像内でラスタースキャンさせて相関係数の高い位置を検出する場合，周辺領域でもマッチング評価ができるようにパディング処理を行う．テンプレートとパディングにより拡張されたターゲット画像内の同面積の領域との相関係数  $\alpha(u, v)$  は，次式から計算される．

$$\alpha(u, v) = \frac{s_{it}(u, v)}{s_i(u, v)s_t(u, v)} \quad (4.1)$$

$$s_{it}(u, v) = \sum_{y=v}^{v+N-1} \sum_{x=u}^{u+M-1} \left\{ f(x, y) - \bar{f}_{u,v} \right\} \left\{ t(x-u, y-v) - \bar{t} \right\} \quad (4.2)$$

$$s_i(u, v) = \sqrt{\sum_{y=v}^{v+N-1} \sum_{x=u}^{u+M-1} \left\{ f(x, y) - \bar{f}_{u,v} \right\}^2} \quad (4.3)$$

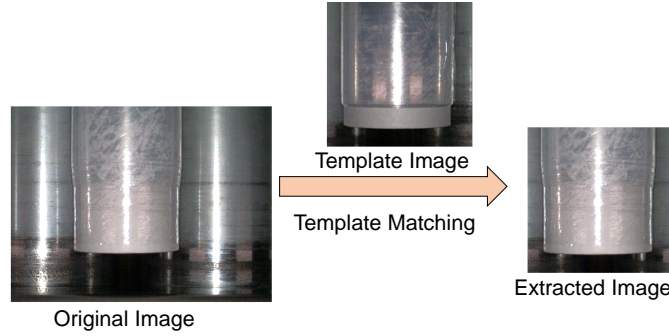


図 4.1. An example of extracted image using the template matching technique.

$$s_t(u, v) = \sqrt{\sum_{y=v}^{v+N-1} \sum_{x=u}^{u+M-1} \{t(x-u, y-v) - \bar{t}\}^2} \quad (4.4)$$

ここで、 $(u, v)$  はターゲット画像内におけるテンプレート左上コーナーの座標である。  $s_t(u, v)$  と  $s_i(u, v)$  はそれぞれ、テンプレート内とターゲット内比較領域の標準偏差であり、  $s_{it}(u, v)$  は共分散である。  $f(x, y)$  は拡張された画像内の  $(x, y)$  におけるグレースケール 256 階調値を正規化した値である。  $t(x-u, y-v)$  はテンプレート内の  $(x-u, y-v)$  における同様の値である。  $M$  と  $N$  はそれぞれ、テンプレートの幅と高さである。  $\bar{t}$  と  $\bar{f}(u, v)$  はそれぞれ、テンプレート内のグレースケール値の平均値と、ターゲット画像内のテンプレート真下の領域のそれである。式 (4.1) で与えられる相関係数  $\alpha(u, v)$  は、テンプレートをターゲット画像内の左上から右下までラスタースキャンさせることで計算される。ラスタースキャン後、テンプレートと最もマッチする領域、すなわち最も大きな値  $\alpha(u, v)$  を持つ領域が抽出される。図 4.1 には、テンプレートで抽出されたラップフィルム品の画像の例を示している。今回の実験では、テンプレートマッチングによりラップフィルム品の画像解像度を特徴抽出器として用いる CNN の入力層に適合させやすいように  $640 \times 480$  から正方形の  $347 \times 347$  にダウンサイジングして用いた [4]。

## 4.2 画像オーギュメンテーション

## 第 5 章

# 実験

### 5.1 使用器具

### 5.2 方法

### 5.3 結果

汎化性能を評価するために、良品 4035 枚と不良品 21 枚の未学習の画像を用意し、学習終了後の SVM に分類させた。表 5.1, 5.2, 5.3, および 5.4 には、AG, AP, VG, および VP それぞれの分類結果の混同行列を示す。なお、表内の Actual と Predicted はそれぞれ、実際のラベルと予測されたラベルを表す。AG, AP, VG, および VP 全てに関して、分類実験に使用した画像の枚数が 4056 であることを考えると誤認識の枚数が最大 6 というのは分類性能の高い結果であると言える。

表 5.1. Confusion matrix of AG ( $C=0.5$ ).

| Act. \ Pre. | Pre. |      |
|-------------|------|------|
|             | NG   | OK   |
| NG          | 20   | 1    |
| OK          | 2    | 4033 |

表 5.2. Confusion matrix of AP ( $C=0.5$ ).

| Act. \ Pre. | Pre. |      |
|-------------|------|------|
|             | NG   | OK   |
| NG          | 20   | 1    |
| OK          | 0    | 4035 |

表 5.3. Confusion matrix of VG ( $C=0.5$ ).

| Act. \ Pre. | Pre. |      |
|-------------|------|------|
|             | NG   | OK   |
| NG          | 20   | 1    |
| OK          | 5    | 4030 |

表 5.4. Confusion matrix of VP ( $C=1$ ).

| Act. \ Pre. | Pre. |      |
|-------------|------|------|
|             | NG   | OK   |
| NG          | 21   | 0    |
| OK          | 4    | 4031 |

## 第 6 章

## 考察

一般に，学習を行った SVM の性能を検討するには次式のような評価指標を用いる [8]．

$$Ac = \frac{TP + TN}{TP + FN + FP + TN} \quad (6.1)$$

$$Pr = \frac{TP}{TP + FP} \quad (6.2)$$

$$Re = \frac{TP}{TP + FN} \quad (6.3)$$

$$F = \frac{2 \cdot Re \cdot Pr}{Re + Pr} \quad (6.4)$$

ここでは不良品を陽性，良品を陰性とし，全体のデータの中で不良品を不良品として分類したものを  $TP$ ，良品を良品として分類したものを  $TN$ ，不良品を良品として分類したものを  $FN$ ，さらに良品を不良品として分類したものを  $FP$  として値を決定した．これらの値から式 (6.1)，(6.2)，(6.3)，および (6.4) で与えられる Accuracy( $Ac$ )，Precision( $Pr$ )，Recall( $Re$ )，および  $F$  値 ( $F$ ) をそれぞれ算出する．今回の分類結果をもとに算出した各評価指標を表 6.1 に示す．

これら 4 つの指標のうち Accuracy は，全ての予測のうち正しく分類できたものはどれだけかを表す指標である．Precision は，陽性と予測した画像が実際に陽性であった割合を示しており，Recall は，陽性の画像を陽性と予測できた割合を表す指標である．さらに  $F$  値は，Precision と Recall のバランスを示す指標である．

表 6.1. Comparison of result.

|      | AG    | AP    | VG    | VP    |
|------|-------|-------|-------|-------|
| $Ac$ | 0.999 | 0.999 | 0.999 | 0.999 |
| $Pr$ | 0.909 | 1     | 0.8   | 0.84  |
| $Re$ | 0.952 | 0.952 | 0.952 | 1     |
| $F$  | 0.930 | 0.976 | 0.870 | 0.913 |



従来、画像認識を応用したラップフィルムの欠陥検出は難しい課題とされてきたが、表 6.1 の Accuracy の値より今回設計した SVM の分類性能は優れているといえる。次に、Precision に着目すると AG, AP は 0.9 以上を記録しているのに対し、VG, VP は 0.9 未満であり、本実験の条件においては AlexNet を特徴抽出器とした SVM の方が良品を不良品と誤認識することが少なかったということになる。さらに、Recall に関しては、分類実験用のデータセットに含まれる不良品の画像が少なかったことに起因して値が変動しやすく、VP 以外の SVM でも誤認識枚数は 1 枚であるが、指標に大きな差が生じた。よって、Recall の値のみに着目すれば VP が優れているように思えるが、表 5.1, 5.2, 5.3, 5.4 から、大きな差はないと言える。

特徴抽出器ごとに指標を比較すると、Precision と F 値の 2 点で AG より AP の方が優れており、VP は VG に Precision, Recall, F 値の 3 点で優れていることがわかる。さらに AP と VP を比較すると、AP は Recall では劣るものの、Precision と F 値では VP より優れているという点から AP の方がより良い性能を有していると考えられる。よって、本実験の条件下では Precision および F 値から、特徴抽出器としては VGG19 より AlexNet の方が適しており、カーネル関数としてはガウス関数より多項式関数の方が適していたと推測できる。今後、さらに不良品の画像を収集可能になれば、その不良品画像を使った追加学習を行うことで、より分類性能の高い SVM を構築できるものと期待される。

## 第 7 章

# 結言

本研究では，ラップフィルム品の欠陥検出を行うことができる SVM について検討した．実験では，特徴抽出器として用いる CNN，カーネル関数，分離平面の-margin 内のデータに与えるペナルティの程度である正則化係数の値を変更しながら，より分類性能の高い SVM の設計を目指した．特徴抽出器には AlexNet あるいは VGG19 を，カーネル関数にはガウス関数と多項式関数をそれぞれ組み合わせて 2 クラス分類を行う SVM (TCSVM) を設計した．評価実験の結果，今回の条件下では特徴抽出器には AlexNet を，カーネル関数には多項式関数を使用し，正則化係数を 1 以下にして設計した SVM が最も性能が高かった．

実際の製造ラインにおいて不良品が良品と誤認識されることはリコールなど重大な事態につながりかねない．一方，良品が不良品と誤認識されることは生産工程における大きな損失につながる．本研究で使用したアプリケーションは，不良品検出用の SVM 構築に関して有効なツールであることが確認できたため，今後はラップフィルム以外の製品の欠陥検出用 SVM の構築にも適用できるのかの検証を行いたい．

# 謝辞

本研究は，山口東京理科大学大学院基礎工学研究科基礎工学専攻で行われたものである．

## 参考文献

- [1] 西村 晃紀, 柳部 正樹, 青戸 勇太, 長谷 智紘, 森山 健, 前田 俊二, “HOG 特徴量を電子部品検査に適用した場合の課題検討,” 2017 年度精密工学会秋季大会学術講演会講演論文集, pp. 219–220, 2017.
- [2] X. Yang, M. Han, H. Tang, Q. Li, X. Luo, “Detecting Defects With Support Vector Machine in Logistics Packaging Boxes for Edge Computing,” *IEEE Access*, Vol. 8, pp. 64002–64010, 2020.
- [3] F. Nagata, K. Tokuno, K. Mitarai, A. Otsuka, T. Ikeda, H. Ochi, K. Watanabe, M.K. Habib, “Defect detection method using deep convolutional neural network, support vector machine and template matching techniques,” *Artificial Life and Robotics*, Vol. 24, No. 4, pp. 512–519, 2019.
- [4] 中島 健斗, 永田 寅臣, 渡辺 桂吾, “畳み込みニューラルネットワークを用いたラップロール製品の不良品検出,” 第 36 回ファジィシステムシンポジウム 論文集, MC2-4, 5 pages, 2020.
- [5] 中島 健斗, 永田 寅臣, 渡辺 桂吾, “畳み込みニューラルネットワーク (CNN) とサポートベクターマシン (SVM) を用いた微小な欠陥を持つ不良品検出の基礎研究,” ロボティクス・メカトロニクス講演会 2019 講演論文集, 2A1-Q05, 4 pages, 広島国際会議場, 2019.
- [6] MATLAB, “<https://jp.mathworks.com/>”
- [7] 竹内 一郎, 鳥山 昌幸, “サポートベクトルマシン,” 講談社, 2015.
- [8] 中川 裕志, “東京大学工学教程 情報工学 機械学習,” 丸善出版, 2015