

## スリーブ型はんだ装置の圧力を用いた不良検出におけるアルゴリズムとカーネルの精度比較

## Comparison of Algorithm and Kernel Accuracy in Defect Detection Using Pressure in Sleeve Soldering Equipment

渡邊 幹<sup>1)</sup>上野 秀剛<sup>2)</sup>

Kan Watanabe Hidetake Uwano

## 1 はじめに

電子基板と素子を接合するはんだ付けの良否は接合部の強度に影響するため、電気自動車など振動のある環境における信頼性に重要な影響を及ぼす。そのため、はんだ付けの不良を検出する検査は重要な工程である。主要な検査方法として作業者が目視でスルーホールに十分なはんだが充填されているか確認する方法があるが、作業者の経験や能力に依存してしまう。また、素子の小型化によって素子の集積度が増大し、目視による検査が困難になっている。そのため、はんだ接合部の検査を自動で行う手法が研究されている。

はんだ付けを行う装置の1つとしてスリーブはんだ付け装置がある。スリーブはんだ付け装置はスルーホールに熱したセラミック製の円筒(スリーブ)をかぶせ、スリーブ内にはんだ片を落とすことによってはんだを融解し、はんだ付けを行う。スリーブ内ではんだを溶融させるためボールはんだの飛散が無く、一定量のはんだを接合部に供給できるため、はんだの品質が安定する利点がある。また、スリーブ内に窒素ガスを充填させることではんだの酸化を防止し、溶けたはんだがスルーホールに流れやすくなり、イモはんだを防止することができる。このとき、スリーブと基板の接触やはんだ片の融解によってスリーブ内の圧力が変化する。

著者らは先行研究においてはんだ付けの良否を自動的に判定することを目的に、スリーブ型はんだ付け装置によるはんだ付けの様子を記録した圧力変化に対して Random Forest と SVM で推定し、SVM で 95.5% の精度ではんだの良否を分類できた [1]。本研究ではさらに分類精度を向上させることを目的として、SVM とその発展手法である RVM を対象に精度比較を行う。また、それぞれのアルゴリズムにおけるカーネル関数を適宜変更することによってその精度からデータ分布の特徴に関する知見を得ると同時に、より分類精度の高い機械学習の設定を明らかにする。

以下、2 章で関連研究について説明し、3 章で準備として装置の説明や機械学習について示す。4 章で用いるデータセットや実験について紹介し、5 章で結果と考察を述べる。

## 2 関連研究

はんだ付けの良否を識別する既存研究として、はんだ付けが完了した基板の画像から特徴量を抽出する手法が

一般的である。Wu[2] は、基板画像からはんだ接合部を定位した後、画像特徴を抽出し機械学習を行った。5 つの機械学習の手法(決定木,k 近傍法,SVM,NN,Random Forest) について精度を比較した結果、Random Forest で 5 種類の不良はんだを 100% の精度で予測できた。Wenting ら [3] は物体検出手法の 1 つである YOLO を用いて基板画像中から定位したはんだ接合部の特徴を用いて正常はんだと不良はんだの分類をした結果、基板画像から微少かつ間隔の狭いはんだ接合部の検出に成功し、すべてのデータセットにおいてアノテーションの誤り率は 1.5% 未満であった。基板画像を用いた従来研究では基板 1 枚に存在するすべてのはんだ付けが完了してから検査を行うのに対して、スリーブはんだ付け装置の圧力変化を用いる提案手法では個々のスルーホールに対するはんだ付けを完了した直後から、はんだ付けを中断することなく検査できるため、リアルタイムな検出が可能となる。また、スリーブはんだ付け装置の圧力波形は、はんだ付けを構成する各工程における装置の動作の違いを反映する。そのため、各工程において不良はんだと良はんだの間で生じる差を圧力波形から識別することで異なる原因に基づく不良を検出できる可能性がある。

機械学習におけるカーネル法は、SVM の成功等により多様な分野で利用されている [4]。カーネル法を用いる場合の判別性能はデータ分布に適したカーネル関数を選択できたかに影響されるため、判別対象となるデータの特徴がわかっていない場合は異なるカーネル関数を比較する必要がある。石垣ら [4] は紙幣の弾き音に注目し、そのスペクトル構造を特徴量として機械学習を用いた疲労紙幣の検出を行った。機械学習のアルゴリズムとして SVM と RVM を採用し、13 種類のカーネル判別器の検出精度を比較した結果、RVM より SVM のほうが 3.3~4.0% 高く、両方のアルゴリズムにおいて Hellinger, Bhattacharyya, KL カーネルの平均精度が 99% 以上だった。近藤ら [5] は機械学習を用いて脳 MRI 画像からの年齢推定を行った。アルゴリズムに SVM, RVM, LDA を選択、特徴量として T1 強調画像なるものから画像処理を行ったものを用いる。実験の結果、RVM の MAE が SVM よりも 0.496 小さく、LDA よりも 2.951 小さかった。脳 MRI 画像から得た特徴量は異なる年齢間で分布に重なりが多く、疎にサポートベクトルを選択する RVM の特徴が精度向上に寄与したと考えられる。

本研究では、はんだの良否判別において SVM と RVM を用いた際のカーネル関数の違いによる精度を比較する。本研究が対象とするスリーブ型はんだ付け装置の圧力変化は良はんだ・不良はんだ間のデータの重なりが発生する可能性が高い。そのため、先行研究 [1] でもっとも精度の高かった SVM に加え、SVM と比較してデータ

1) 奈良工業高等専門学校 システム創成学専攻 情報システムコース. Department of Advanced Information Engineering, National Institute of Technology (KOSEN), Nara College.

2) 奈良工業高等専門学校 情報工学科. Department of Information Engineering, National Institute of Technology (KOSEN), Nara College.

の重なりや外れ値に強い RVM を対象として複数のカーネル関数を用いた際の精度を比較し、もっとも精度の高い組み合わせを明らかにする。

### 3 準備

#### 3.1 スリーブ型はんだ装置の自動検査法

現在のポイントはんだ付けにはコテはんだ付け、レーザーはんだ付けがある。コテはんだ付けは、加熱したコテ先ではんだを溶かし、ピンとプリント基板の接合部に供給する方法である。レーザーはんだ付けは、レーザーがピンとプリント基板に照射し、発熱することで接合部の温度をはんだの融点まで上昇させ、装置からはんだを供給する方法である。

スリーブはんだ付けは、ほかのポイントはんだ付けと比較して3つの利点があり、優れた手法である。コテはんだ付けに比べはんだの広がりや濡れ性がよく、コテ先の摩耗及び交換の必要がない。レーザーはんだ付けに比べ、はんだの飛散がなく、一定量のはんだ片を供給できるため、品質が安定している。また、窒素ガスが酸素を遮断することにより、はんだ表面の酸化を防ぐ特徴がある。

図1は、はんだ付けを行っている際のスリーブ内の様子と圧力変化を示す。(a)-(c)は、それぞれ(a)スリーブと基板の接触、(b)はんだ片落下、(c)融解途中のはんだ片を示す。(d)は良はんだの際のスリーブ上昇を、(d')は不良はんだの際のスリーブ上昇の状態を示している。はんだ片がスリーブ上部で切断された時刻を  $t=0s$  とする。(a) $t = -0.57s$  にスリーブと基板が接触することで窒素ガスの流路が狭くなり、圧力が上昇する。(b) $t = 0s$  にはんだ片がスリーブ内に投入されることでさらに流路が狭くなり、圧力が上昇する。(c) $t = 1.2s$  にはんだ片が溶融するとスルーホールを塞ぐことで最高圧力まで上昇する。(d),(d') $t = 2.5s$  にスリーブの上昇が始まり圧力が低下する。(d)良はんだの場合、はんだはスルーホール付近で山形に凝固するためスリーブが上昇するとすぐに窒素ガスが流出し圧力が低下する。一方で(d')不良はんだの場合はイモはんだになることが多く、良はんだと比較してより基板から離れた位置で球状に凝固するため、スリーブが上昇してもすぐにはガスの流路が出来ず、圧力が低下するまでに時間がかかる。

著者らはこれまでの研究においてはんだ付けの良否による圧力変化のタイミングが異なる点に着目した判別手法を提案した [1]。圧力が最高値から降下を始める時間やその降下速度を特徴量とした 11 個の指標を元に機械学習によってはんだの状態を分類した。機械学習のアルゴリズムとして Random Forest と SVM を用いた結果、「部品ランド側にはんだあり」と「スルーホール内にはんだあり」と「スルーホール内にはんだなし」の 3 パターン分類における精度は Random Forest で 82.7%、SVM で 87.1% だった。また、「スルーホール内にはんだあり」と「スルーホール内にはんだなし」の 2 パターン分類における精度は Random Forest で 94.5%、SVM で 95.5%、「部品ランド側にはんだあり」と「スルーホール内にはんだあり+なし」の 2 パターン分類では Random Forest で 89.9%、SVM で 93.3% だった。

本研究では、「スルーホール内にはんだあり」と「スルーホール内にはんだなし」の 2 パターン分類を機械学習を用いて行う。先行研究 [1] で最も優れた判別制度を

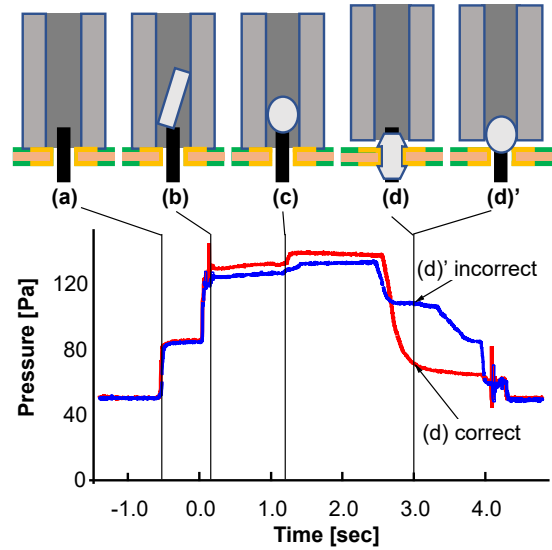


図 1: はんだ付け中のスリーブ内の圧力変化

示した SVM に加えて、SVM の持つ特性を引き継ぎながら問題点を解消した RVM を用いる。また、複数のカーネル関数を用いた際の精度を比較し、最も精度の高い組み合わせを明らかにする。

### 3.2 機械学習

#### 3.2.1 Support Vector Machine

Support Vector Machine (SVM) は機械学習モデルの一種である。SVM は 2 値クラス判別のための線形学習機である [4]。SVM はサポートベクトルによって境界を生成し、分類を行う。そのため、最大マージンの分類境界はサポートベクトルの位置によってのみ決まり、それ以外のデータ点はどこに動かしても分類境界に影響を与えない [6]。このアルゴリズムは教師あり学習で分類と回帰に関して扱うことができるが、本研究では分類で用いる。SVM には以下の 3 つの特徴がある。

1. データの次元が大きくなっても分類精度が良い。
2. 最適化すべきハイパーパラメータが少ない。
3. 学習データが増えると計算量が膨大になる。

SVM は分布間のマージンが最大となるようなサポートベクトルに基づいて境界面を引く。そのため、データ分布に重なりや外れ値の少ないデータの場合に高精度が期待できる。また、カーネル関数の選択によって非線形な境界面を引くことができる。

#### 3.2.2 Relevance Vector Machine

Relevance Vector Machine (RVM) は回帰及び分類問題を解くために提案された疎なカーネルベースのベイズ流学習手法を用いた機械学習である [6]。SVM の持つ特性の多くを引き継ぎながら数々の問題点を克服し、確率モデルとして用いることができる点が最大の特徴である。また、RVM には SVM よりもさらに疎なモデルが得られることが多く、SVM と同等な汎化能力を持ちながら、より高速な予測が可能となる [6]。RVM には関連ベクトルというサポートベクトルと似た役割のもの存在している。RVM には SVM と比較して以下の 4 つの特徴がある。

1. 確率を出力とする。

2. 超平面を構成するために必要となるパラメータが少ない。
3. 外れ値に影響されにくい。
4. データの重なりに強い。

SVM がはっきりと 2 値分類を行うのに対して、RVM は分類結果を確率的に出力する。そのため、RVM はどのはんだが何点のはんだなのか点数付けを行うことができる。よって、良否の基準となる値を決めて実験を行う。

また、SVM が特徴分布によっては多くのサポートベクトルが選択され複雑な境界面になってしまう一方で、RVM はより疎にサポートベクトルを選択し、境界面を引くため、特徴分布に対して SVM より単純な境界面を引く。そのため、データの分布が複雑であった場合 SVM よりも高精度が期待できる。本研究においては、良はんだと不良はんだのデータ分布に重なりや外れ値が多い場合において RVM のほうがより精度が高くなると考えられる。

### 3.2.3 カーネル関数

カーネル法は、主にパターン認識の分野で広く利用されている非線形処理を実現するための手法である [7]。また、SVM や RVM、主成分分析等の線形モデルで表現される手法に広く組み込まれている。カーネル法では、データを入力空間から高次元の特徴空間内に写像を行い写像先で内積計算を行う。よって、高次元空間における線形モデルの適用は、もとの空間での非線形処理と等価になる。つまり、カーネル法を線形モデルに組み合わせることで非線形モデルへと拡張することが可能となる。カーネル関数は、カーネル法で特徴空間内での内積計算で用いる関数のことである。そのため、カーネル関数とはアルゴリズムにおける主要となる重要な関数である。

カーネル関数ではそれぞれハイパーパラメータを選択する必要がある。ハイパーパラメータはカーネル関数がどのような動作をするのかを決定するパラメータである。例えば、ハイパーパラメータの 1 つであるコストパラメータ (C) は値が低いと誤分類を許容し、値が高くなるにつれて誤分類を許容しなくなる。C は SVM を構成するときの 2 次計画式にて現れる。

本研究ではモデル学習と予測、及びパラメータの最適化を行う際に実装が容易な以下の 3 種類のカーネル関数を用いる。

- RBF カーネル
- Polynomial カーネル
- Linear カーネル

## 4 実験

### 4.1 データセット

株式会社アンド製のスリーブはんだ付け装置であるスマートショットを用いてはんだ付けを行う際の圧力変化を計測する。はんだ付けを行っている際のスリーブ内圧力を 500Hz で計測する。はんだ対象として 20x16 の計 320 個のスルーホールがあるユニバーサル基板を用いる。

スリーブはんだ付け装置の設定を実務で使う場合と同じくした場合、不良はんだが発生しにくく、学習に十分な数のデータを得ることが難しい [1]。不良はんだを発

表 1: データ件数

ラベル	良否	点数	データ件数	抽出件数
$TH_{NG}$	×	0	3	3
		1	390	390
		2	47	47
$TH_{OK}$	△	4	989	75
		5	546	75
		6	482	75
		7	172	75
		8	163	75
	○	9	334	75
合計			3,126	890

生させる方法としてスリーブ温度を低下させる方法やはんだ量を減らす方法等が考えられる。しかし、スリーブの温度や投入するはんだ量を変化させた場合に発生する不良は、実務で使用する際に発生する不良とは原因や特徴が異なる恐れがある。そこで本研究では実環境においても不良の原因となっている予熱不足に着目する。予熱とはスリーブを下降させて基板と接触させることで予め基板を熱することである。はんだ付けに適した予熱時間や温度は対象の基板の大きさや種類に依存するため、新しい設計の基板を対象とした際に予熱不足による不良が発生する。本研究ではエアポンプによる送風で基板を冷却するとともに、はんだ付け間の時間間隔を通常の 1 秒未満から 30 秒にすることで熱が不十分になりやすい環境を発生させる。また、直前のはんだ付けの熱の影響を受けまいよう、同じ行にあるスルーホールを連続ではんだ付けせず、1 行ずつ空けてはんだ付ける。

個々のはんだ付け結果に対して、はんだ付け装置の開発者の 1 人が目視で 0~9 点の点数付けを行う。0~2 点はスルーホールにはんだが無い不良はんだ、4~9 点はスルーホールにはんだがある良はんだとした。

表 1 に取得したはんだの良否数を示す。 $TH_{NG}$  が不良はんだ、 $TH_{OK}$  が良はんだで、点数が高いほど良いはんだづけであることを表す。良否の × はバックフィレットなしの不良はんだ、△はバックフィレットなしの良はんだ、○をバックフィレットありの良はんだとする。ただし、学習にはバックフィレットの有無を区別せず、ラベルを学習させる。各点数のデータ件数に着目すると、不良はんだの合計が 440 件であるのに対して、良はんだの合計が 2,686 件と数に偏りがあり、データのすべてを用いると学習が偏り、予測精度に影響を与える可能性がある。本研究では、ランダム抽出を行い学習に用いる良はんだと不良はんだのデータ数をそろえる。また、抽出するデータの差が結果に影響しないようにするため、ランダム抽出を行ったデータセットを複数作成し、同じデータセット間でアルゴリズムやカーネルを比較する。実験で取得した計 3,126 件の全データに対して、不良はんだから全件、良はんだから各点数 75 件をランダムに抽出し、計 890 件を 1 つのデータセットとする。同様の操作を 100 回繰り返し、100 個のデータセットを作成する。RVM を用いた予測では点数を判別結果として用いる。SVM を用いた予測では良否ラベルを判別結果として用いる。

## 4.2 特徴量抽出

はんだの良否の特徴が圧力に表れると考えられる工程として節 3.1 にて (a)~(d) の工程を挙げた。はんだが不良はんだになる事例は、はんだ片を溶かす基板を予熱不十分な環境において発生しやすい。基板が予熱不十分な状態になるとはんだ片に十分な熱が伝わらないため溶融が始まる時間が遅くなる。はんだ付け中にスリーブが上昇するタイミングは毎回一定であるため、溶けたはんだがスルーホールに流れ込むよりも先にスリーブの上昇が始まり、ピンの途中で冷めて凝固することで不良はんだになりやすい。このような不良はんだが発生するとき、良はんだと比較してスルーホールが塞がる時間が遅くなるため圧力がピークに達する時間も遅くなる。また、スリーブが上昇した後もピンの途中にあるはんだがガスの流路を塞いでいるため、良はんだより圧力低下が遅くなる。そのため最高圧力を保持する時間や圧力の低下速度を特徴量とすることで、良否を分類できると考えられる。

本研究でははんだの良否を識別する特徴量として、圧力波形の尖度、ピーク長、圧力降下にかかる時間の 3 種類を用いる。尖度は分布の尖り度合いを表す値であり、尖った波形ほど大きい値となる。そのため、はんだ片の溶融速度の差を表す特徴量として有用と考えられる。ピーク長ははんだが溶けて最高圧力に達した後、最大圧力値を持続した時間長である。溶けたはんだがスルーホールを塞いでいる状態を最高圧力とし、はんだ付け開始から最初に「最高圧力\*0.9」となる点を始点、「最大圧力\*0.9」未満となる点を終点とする。ピーク長は基板のスルーホールが溶けたはんだで埋まっている時間を表し、はんだがどのように溶けているかを反映すると考えられる。圧力降下にかかる時間は、最高圧力から指定の割合まで圧力が降下するのに要する時間と定義する。1 回のはんだづけに必要な工程が始まってから完了するまでに測定される圧力のうち、圧力計測を開始して 0.02 秒から 0.04 秒間の平均値を最低圧力と定義する。最低圧力はスリーブが基板に接触する前の状態を意味し、最高圧力はスリーブが基板に密着して、かつ、溶融したはんだがスルーホールを塞いだ（スルーホールを埋めていない場合は部分的に塞いだ）状態を意味する。また、はんだ付けが終了する際は、スリーブが基板から離れ接触する前の状態になるため最低圧力まで低下する。スリーブが基板から離れ、最高圧力から最低圧力まで低下する際、スルーホールを埋めていない場合では傾きが緩やかになるか、低下する時間が遅くなると考えられる。本研究では最高圧力を 100%、最低圧力を 0%とし、最高圧力から降下し、指定の 10 段階 (90%, 80%, ..., 10%, 5%) に到達する時間 ( $T_{90\%}$ ,  $T_{80\%}$ , ...,  $T_{10\%}$ ,  $T_{5\%}$ ) を圧力降下の様子を表す特徴量として用いる。

## 4.3 分類

100 個のデータセットに対して節 4.2 で説明した特徴量を用いて機械学習による分類を行う。学習アルゴリズムには SVM と RVM を用いる。カーネル関数には RBF カーネルと Polynomial カーネル、Linear カーネルを用いる。100 個のデータセットそれぞれに対して各ハイパーパラメータの組み合わせを変更しつつ 10 分割交差検証を 10 回行い、ハイパーパラメータのチューニングを行う。チューニングには R 言語のパッケージである

表 2: SVM の予測精度

カーネル	accuracy			FN
	最高	最低	平均	平均
RBF	100.0%	98.2%	99.7%	0.3%
Polynomial	100.0%	99.9%	100.0%	0.0%
Linear	100.0%	99.1%	100.0%	0.0%

caret の train<sup>1)</sup>を用いる。

SVM の予測結果は良はんだ/不良はんだのいずれかである。RVM の予測結果は良はんだを 1、不良はんだを 0 とした点数で出力するため 0.5 を閾値として設定し良はんだと不良はんだを分類し精度を求める。2 アルゴリズムと 3 カーネルの組み合わせである計 6 組に対してそれぞれハイパーパラメータのチューニングを行い、精度を比較すると最もよいアルゴリズム、カーネル関数の組み合わせを見つける。SVM の RBF カーネルと RVM の 3 カーネルにおいてハイパーパラメータのチューニングを TuneLength=100 で実行した。SVM の Polynomial カーネルは TuneLength を用いずに実行した。SVM の Linear カーネルは C の値を 15 の範囲を 0.001 刻みで探索して実行した。

## 5 結果と考察

### 5.1 SVM と RVM の比較

表 2 に SVM を用いた際のカーネルごとの分類精度を示す。100 個のデータセットに対して予測を行った時の最高 accuracy と平均 accuracy, 最低 accuracy を示す。また、不良はんだを良はんだとして誤って予測した割合である FalseNegative の平均を示す。太字の値は各列の最高精度を表す。すべてのカーネルで最高 accuracy が 100.0%, Polynomial と Linear カーネルで平均 accuracy が 100.0%, 平均 FN が 0.0%だった。最低 accuracy は Polynomial カーネルが 99.9%と最も高く、全ての指標を通じて Polynomial カーネルが最も精度が高かった。一方で、最低、平均 accuracy と平均 FN で RBF カーネルが最も精度が高く、最低 accuracy は Polynomial カーネルと比較して 1.7%低かった。

図 2 に SVM で RBF, Polynomial, Linear の各カーネルを用いた際の 100 個目のデータセットに対する予測結果を示す。横軸は各はんだの ID を示し、縦軸は各はんだの予測で  $TH_{OK}$  が良はんだ,  $TH_{NG}$  が不良はんだを示す。点数 0 点~2 点の不良はんだを ×, 4 点~7 点の良はんだ (バックフィレット無し) を △, 8 点~9 点の良はんだ (バックフィレットあり) を ○ で表す。いずれのカーネルにおいても共通して、高精度で正しく分類している。RBF カーネルにおいて 4 点のはんだを 1 つだけ誤分類しており、良否はんだの境界は誤分類が起きやすい可能性がある。

表 3 に RVM を用いた際のカーネルごとの分類精度を示す。全ての指標において RBF カーネルがもっとも精度が高く、平均 accuracy が 98.6%の精度ではんだの良否を識別できた。一方で Linear カーネルが全ての指標においてもっとも低く、平均 accuracy が 93.1%と RBF カーネルと比べて 5.5%の差が見られた。

図 3 に RVM で RBF, Polynomial, Linear の各カーネルを用いた際の 100 個目のデータセットに対する予測

1) <https://www.rdocumentation.org/packages/caret/versions/6.0-92/topics/train>

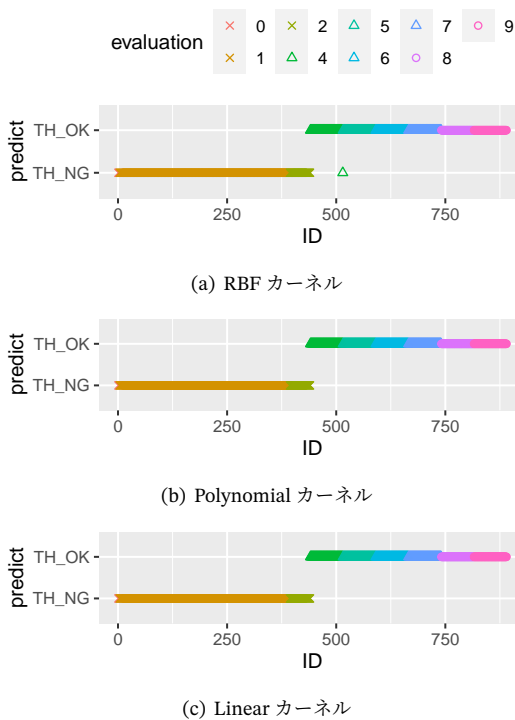


図 2: SVM の予測結果

表 3: RVM の予測精度

カーネル	accuracy			FN
	最高	最低	平均	平均
RBF	99.6%	97.5%	98.6%	0.8%
Polynomial	94.3%	92.5%	93.5%	5.9%
Linear	93.7%	92.2%	93.1%	6.3%

結果を示す。水平の黒線は閾値の 0.5 を示す。縦軸は各はんだに対する予測点数を示す。縦軸の値が高いほど良はんだ、低いほど不良はんだと分類している。いずれのカーネルにおいても共通して、採点が高いはんだほど不良はんだとして分類される数が減少しており、8 点~9 点の良はんだ（バックフィレットあり）はすべてが良はんだとして分類された。一方で、良はんだと不良はんだの境界に相当する 4 点（良はんだ）と 2 点（不良はんだ）は予測結果が縦に広く分布している。手動採点が 0 点と 1 点のはんだ（いずれも不良はんだ）においては、一部のデータについて予測点数が高くなっており良はんだと誤認識された。

カーネル間の比較では RBF カーネルがもっとも精度が高く、平均 accuracy で 5%他のカーネルより高かった。不良はんだの検出を行う本研究において、5%の精度の差は大きく、カーネルの選択が重要であることが分かる。今後、より適したカーネルの調査を継続する必要がある。図 3 に示した結果は RBF カーネルと比べて Polynomial と Linear カーネルで採点が 1,2 点の不良はんだを良はんだと分類している数が多く、誤分類が発生したはんだ付けの圧力波形を分析することで誤分類の原因を特定することは重要である。

アルゴリズムとカーネルの組み合わせの比較では、SVM の Polynomial カーネルでもっとも精度が高い。ま

表 4: 実験のハイパーパラメータ

ハイパーパラメータ	値	回数
degree	1	0
	2	72
	3	28
scale	0.001	0
	0.01	0
	0.1	100
C	0.25	25
	0.50	35
	1.00	40

た、SVM はどのカーネルを用いても RVM よりも精度が高い。RVM で精度が低かった原因は今後調査する必要がある。

## 5.2 ハイパーパラメータと過学習

本節では、本研究の結果について過学習の恐れがないか考察する。過学習とは構築した予測モデルが、学習に用いたデータに対して過剰に最適化された状態であり、過学習が発生しているモデルは他のデータに対する予測精度が低下する。本研究では過学習を防止するために 10 分割交差検証とハイパーパラメータチューニングを用いているが、以下の 2 点については今後の検証が必要である。

### ・チューニング不足

表 2 にもっとも精度の高い組み合わせであった SVM の Polynomial に対する 3 種類のハイパーパラメータのチューニング結果を示す。各行はチューニングによって各ハイパーパラメータに対して選択された値と回数を示す。本稿でチューニングに用いた caret パッケージの train 関数では 3 パラメータそれぞれに 3 通りの値しか試しておらずチューニング不足の可能性がある。

### ・不良はんだのデータ数不足

本研究では実験で取得した良はんだと不良はんだの数を揃えるため、良はんだについてランダム抽出を行った 100 個のデータセットを用いて評価を行った。しかし、不良はんだは数が少ないため、すべてのデータセットで不良はんだのデータが同一になっており、過学習が発生している可能性がある。各データセットに対して交差検証を用いることで対策を行っているが、今後不良はんだに対してもランダム抽出を行い、評価する必要がある。

## 6 おわりに

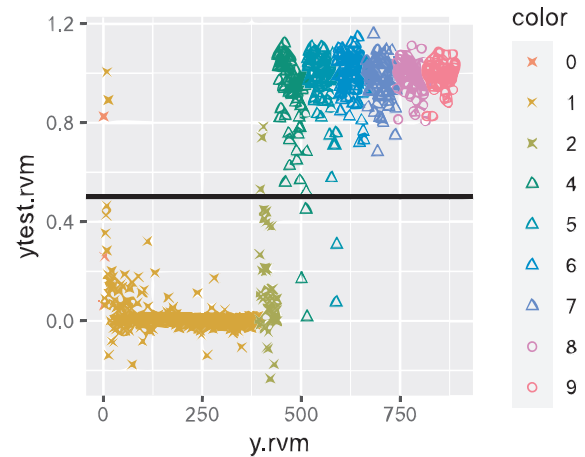
本研究ではスリーブ型はんだ装置を用いたはんだ付けを対象に、スリーブ内の圧力変化から特徴量を抽出し、機械学習による良はんだ、不良はんだの分類を行った。機械学習ではアルゴリズムとして SVM と RVM を用い、RBF, Polynomial, Linear の 3 種類のカーネルを組み合わせた 6 パターンの分類精度を比較した。実験の結果、SVM の Polynomial カーネルにおいて最高 accuracy が 100.0%、最低 accuracy が 99.9%、平均 accuracy が 100.0%、平均 FN が 0.0%とっとも優れた分類精度だった。

今後の課題として、SVM と RVM の双方で良はんだと不良はんだの境界に相当する、手動採点 4 点と 2 点の

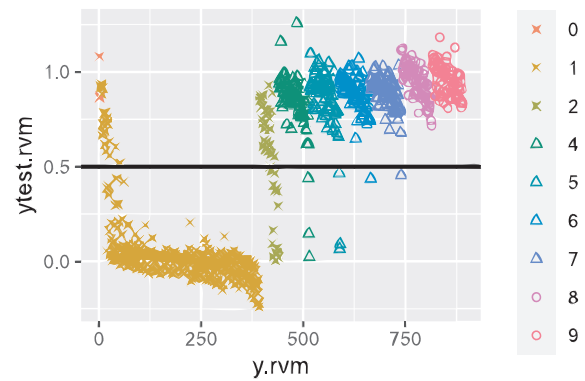
はんだに精度の向上がある。RVM においては手動採点 4 点と 2 点、および 1 点の一部で予測結果が良否の間で広く分布しており、今回抽出した特徴量以外の要素を学習する必要があると考えられる。また、不良はんだのランダム抽出による予測モデルの変化を分析することで、過学習の有無の検証と影響を明らかにすることも重要である。本稿の実験でははんだ対象としてユニバーサル基板を用いたが、はんだ対象の基板の大きさや素子の数は予熱に必要な時間に影響するため、異なる特性を持つ基板を対象に評価を行うことも提案手法の汎化性能を明らかにすることも興味深い発展である。

#### 参考文献

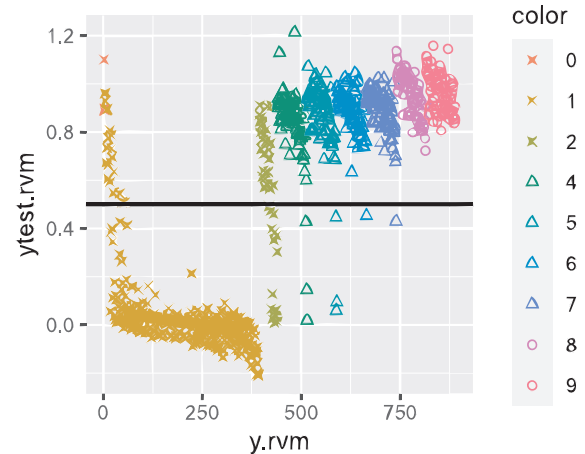
- [1] Y. Fukumitsu, K. Nakamichi, H. Uwano, and H. Fukuoka, "Defect Classification Using Pressure Change of Sleeve Soldering Machine," In The 5th International Conference on Robotics, Control and Automation (ICRCA), 2021.
- [2] H. Wu, "Solder joint defect classification based on ensemble learning", *Soldering & Surface Mount Technology*, Vol.29, 2017.
- [3] W. Dai, A. Mujeeb, M. Erdt, and A. Sourin, "Souldering defect detection in automatic optical inspection", *Advanced Engineering Informatics*, Vol.43, p.101004, 2020.
- [4] 石垣司, 樋口知之" 音響信号を利用したカーネル判別器による疲労紙幣検出", 計測自動制御学会論文集, Vol.44, No.5, pp.444-449, 2008.
- [5] 近藤千裕, 伊藤康一, 呉凱, " 脳局所特徴量に基づく年齢推定手法と脳 MRI 画像データベースを用いた性能評価", 映像情報メディア学会技術報告, Vol.38, No.32, pp.15-18, 2014.
- [6] C.M.Bishop, 元田浩, 栗田多喜夫, 樋口知之, 松本裕治, 村田昇, " パターン認識と機械学習 下 ベイズ理論による統計的予測", 丸善, 2012.
- [7] 渡邊真也, " カーネル法", 情報と知能, Vol.21, No.4, p.537, 2009.



(a) RBF カーネル



(b) Polynomial カーネル



(c) Linear カーネル

図 3: RVM の予測結果