



卒業研究報告書

令和4年度

研究題目

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

指導教員 上野秀剛 准教授

氏名 渡邊 幹

令和5年01月26日 提出

奈良工業高等専門学校 情報工学科

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

上野研究室 渡邊幹

電子基板と素子を接合するはんだ付けの良否は機器の信頼性に影響するため、はんだ付けには高い品質と信頼性が求められる。そのため、はんだ接合部の不良を検出する検査は重要な工程である。現在主流となっている検査は作業者の目視によるもので、電子基板上の素子集積度が増加していることに伴い検査が困難になっており、機械による検査の必要性が向上している。先行研究ではスリーブ型はんだ付け装置による余熱不十分な環境で計測したデータセットを用いた機械学習によってはんだの良否を分類した。スリーブ型はんだ付け装置ははんだ表面の酸化防止のために、スリーブ上部よりはんだに対して窒素ガスの噴射を行うため、スリーブ内圧力が変化する。そのため、スリーブ内に設置している圧力センサによってスリーブ内圧力の測定が可能である。先行研究は機械学習のアルゴリズムである Random Forest と SVM を比較し、SVM のほうが精度が高く、95.5% の確率で良否を分類できた。本研究では、実務での使用に耐えうる様にさらなる精度の向上を目的とする。具体的な方法としては先行研究で最も精度の高かった SVM と、その発展手法である RVM とのアルゴリズム比較を行う。RVM は 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たアルゴリズムである。SVM が 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。また、それぞれのアルゴリズムのカーネルを変更することによる精度の比較を行う。用いるカーネルとしては RBF カーネル、Polynomial カーネル、Linear カーネルを用いる。実験でははんだ装置開発者が 3326 件のはんだ付けを行った際のスリーブ内圧力波形を用いて機械学習による良否の分類を行った。はんだ対象物はユニバーサル基板にてスリーブ型はんだ装置を用いてはんだ付けを行ったものである。また、実務で使う際と同様の設定ではんだ付けを行うと不良はんだが発生しにくく、学習に十分な数を得ることが難しい。そのため、実務の環境でも不良はんだ発生原因となっている余熱不十分な状況に着目し、不良はんだの発生を意図的に増加させた。計測したはんだ付け全てに対してはんだ装置の開発者が 0~9 点で採点する。個々のはんだ付けにおけるスリーブ内の圧力変化から得られる特徴量と、採点結果をデータセットとする。データセットに対して機械学習の手法を用いた予測を行い、その精度を計測する。学習データを抽出する際にははんだの良否数をそろえて学習させるため、ランダム抽出を用いる。なお、先行研究においても学習に用いるデータをランダム抽出しているが、著者が同分析を再現した結果、ランダム抽出による精度への影響が認められた。そのため、本研究ではデータ抽出による結果の差が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 100 個作成し、実験を行う。本研究で用いる RVM は予測結果が採点結果に相当する値で出力されるため、良否の基準を設定する必要があり、基準として 0.4~0.7 の 4 パターンを用意して比較する。また、先行研究と異なり、ハイパーパラメータの探索範囲を指定せずにチューニングを行う。実験の結

果, SVMのPolynomialカーネルを選択したときに100個のデータセットに対して最高accuracyが100.0%, 最低accuracyが99.9%, 平均accuracyが100.0%, 平均FNが100.0%と最も優れた分類精度を示した. 考察では, 過学習について考察するために予測対象を学習に用いたデータセットではなく, 全データに対して予測を行った. 結果は最高accuracyが100.0%, 最低accuracyが98.5%, 平均accuracyが99.6%, 平均FNが100.0%と精度の低下は小さく, 過学習の影響は見られなかった. また, 学習時のハイパーパラメータについても100個のデータセットに対して6通りの組み合わせしか存在しておらず, SVMのPolynomialでは過学習の可能性は少ないと考えられる.

目次

1	はじめに	1
2	関連研究	3
3	準備	4
3.1	スリーブはんだ付け [1]	4
3.2	スリーブ型はんだ装置の自動検査法	5
3.3	機械学習	6
3.3.1	Support Vector Machine	6
3.3.2	Relevance Vector Machine	6
3.3.3	カーネル関数	7
4	実験	9
4.1	データセット	9
4.2	特徴量抽出	10
4.3	分類	12
5	結果と考察	14
5.1	SVMとRVMの比較	14
5.2	学習に使った以外のデータを対象にした推定	19
5.3	ハイパーパラメータと過学習	21
6	おわりに	23
	謝辞	24
	参考文献	25

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機器の丈夫さや信頼性に影響する。はんだ付けの良否は基板と素子の接合がどれだけ強固に行われるかに影響し、基板に空いたスルーホールにどのくらいはんだが充填されているかで判断できる。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。現在主流となっている検査手法は作業者の目視に依存する方法であるため、作業者の経験や能力、そのときの体調によって精度が左右される問題がある。また、素子の小型化が進むにつれて基板上の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ため、集積度が高い場合にも検査可能となる検査の必要性が高まっている。そのため、はんだ接合部の検査を自動で行い、高品質な検査を行うことができる自動検査に関する研究が行われている [1]。

上野研究室における先行研究 [1] では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データセットと個々のはんだ付けの良否を目視によって確認したデータセットを用いた機械学習によってはんだの良否を分類した。スリーブ型はんだ付け装置とは、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、はんだ付けを行う装置である。また、常にスリーブ上部よりはんだに対して窒素ガスの噴射を行うため、スリーブ内圧力が変化する。先行研究はスリーブ内の圧力変化から抽出した特徴量を用いて機械学習のアルゴリズムである Random Forest と SVM を用いて推定を行った結果、SVM のほうが精度が高く、95.5% の確率ではんだの良否を分類できた。よって、この検査手法を用いることではんだ付けの直後に検査を行うことができる。はんだの不良を発見した際には警報を発することによってどのはんだが不良はんだであるかを表示することが可能となった。

本研究は先行研究の分類精度をさらに向上させ、より実務での使用に適した誤判定の少ない分類手法の構築を目的とする。誤判定を減らすことによって、警報による装置の一時停止を減らすことができ、また、製品の信頼性を守ることもつながる。

本研究では先行研究で最も精度の高かったアルゴリズムである SVM とその発展手法である RVM を対象に精度比較を行う。RVM は 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たアルゴリズムである。また、SVM と RVM が用いるカーネル関数についてガウスカーネル、多項式カーネル、線形カーネルの 3 つの精度を比較する。カーネル関数とは、アルゴリズムにおける最も主要な関数である。カーネル関数は識別対象となる現象のデータ分布によって最適な関数に変化する。先行研究が対象としたデータセットは特徴量が 11 個あるため、分布を図示することが困難であり、詳細なデータ分布が明らかになっていない。本研究で

は異なるカーネル関数ごとに分類を行い，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

カーネル学習判別機械は，SVMの成功等により多様な分野にて利用されるようになってきている [2]. その判別性能は，与えられたデータに対して適切なカーネル関数を使用するカーネル法によってより優秀な判別性能を持つことができた. 適切なカーネル関数とはデータ分布に適したカーネル関数のことで，データ分布がわかっている場合は，それらに合わせたカーネル関数を用いるとよい判別性能を得ることができる. しかし，データ分布がわかっていない場合は，適切なカーネル関数を選ぶことが難しい. そのため，異なるカーネル関数を比較し，適切な関数を発見することでデータ分布を理解することができる.

石垣ら [2] 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た機械学習を用いたアプローチによって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 [3] は機械学習を用いて脳MRI画像からの年齢推定を行った. 機械学習のアルゴリズムとしてSVM, RVM, LDAを用いた. 特徴量はT1強調画像なるものから画像処理を行い，270個の特徴を抽出したものをを用いる. 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. この結果は異なる年齢間で特徴の分布に重なりが多く，RVMの特徴の1つである，疎にサポートベクトルを選択することが影響していると考えられる. また，従来手法と提案手法で特徴量も数千から270個に減少し提案手法の有効性が確認できた.

本研究では，はんだの良否判別においてSVMとRVMを用いてカーネル関数の変更による精度比較を行う. アルゴリズムでは，先行研究 [1] で最も精度の高かったSVMに加え，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たRVMを実装することによって精度の向上を狙う. また，カーネル関数の変更によってより適したカーネル関数の設定を行い，精度の向上を狙う.

3 準備

3.1 スリーブはんだ付け [1]

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、はんだの悲惨がなく、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)は良好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) $t=-0.57s$ 、(b) $0.03s$ および(c) $1.2s$ において圧力上昇が確認できる。また、 $t=2.5s$ において圧力降下が確認できる。(a) $t=-0.57s$ における圧力上昇は、スリーブと基板の接触により窒素ガスの流路が狭くなるためである。(c) $t=1.2s$ における圧力上昇は、融解したはんだがピンと基板の隙間を埋めるためである。(d),(d') $t=2.5s$ における圧力降下は、はんだ付け完了後にスリーブ上昇により、窒素ガスが流れ出したためである。不良はんだの場合はイモはんだになることが多く、圧力が下がりやすい。

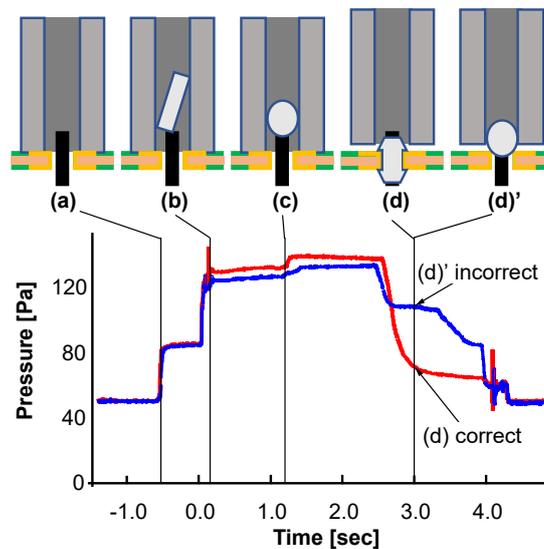


図1 はんだ付け中のスリーブ内の圧力変化

3.2 スリーブ型はんだ装置の自動検査法

はんだ付けを行う装置の1つとしてスリーブ型はんだ付け装置がある。この装置は熱したセラミック製の円筒（スリーブ）を基板のスルーホール上にかぶせ、スリーブ内にはんだ片を落とし融解させる。その後、熱したスリーブを基板のスルーホールにかぶせ、スリーブ内に落下させたはんだ片を融解し、はんだがスルーホールを埋めることによってはんだ付けを行う。その後、一定時間でスリーブが上昇し、次のはんだ対象のスルーホールにスリーブを移動する。スリーブ内に一定量の窒素ガスを投入することによりはんだの酸化を防止することで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。そのため、スルーホールに接する側であるスリーブ下部から常にガスが漏出する。この時、スリーブの基板への密着、はんだ片の落下、はんだ融解等の工程でスリーブ内の穴が狭まったり塞がれることで、圧力の変化につながる。はんだ付け中のスリーブ内の圧力値は、圧力孔に取り付けられたセンサにより測定している。

福光 [1] ははんだ付けの良否によって圧力の変化に違いがあると考えた。例えば、ピン頭頂部でイモはんだになった場合、スリーブが上昇し基板から離れ始めても、イモはんだがスリーブ内の穴を塞ぎ続けることになり、圧力が降下しないことがある。しかし、良はんだの場合はスリーブが上昇するとスリーブと基板の隙間から窒素ガスが抜けるため圧力が降下するという違いがある。そのため、良はんだと不良はんだで圧力変化のタイミングが異なる点に着目した。圧力波形から良はんだと不良はんだで特徴が異なる箇所を特定し、特徴量として抽出した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際の精度を評価した。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RandomForest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RandomForest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RandomForest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] では、SVMが最も優れた判別結果を示したため、本研究でもSVMを用いて実験を行う。また、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たRVMを用いてSVMとの精度比較を行う。同時に、アルゴリズム比較だけでなく、カーネル関数の変更によってどの程度精度が変化するのか調べる。アルゴリズムとカーネルの組み合わせをそれぞれ評価し、最も精度の高い組み合わせを明らかにする。

3.3 機械学習

3.3.1 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[2]。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[4]。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4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。

3.3.2 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[4]。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[4]。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が存在している。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。

3.3.3 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[5]。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によって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

RBFカーネルはガウス関数のことで式(3.3.1)を用いるものである。ハイパーパラメータ σ はRBFカーネル上での境界の引き方に影響する。 σ の値が小さいとより単純な境界(直線)を引き、値が大きいと複雑な境界を引く。ハイパーパラメータとしてSVMでは σ とCを与え、RVMでは σ のみを与える。

$$k(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (3.3.1)$$

Polynomialカーネルは多項式展開のことで式(3.3.2)を用いるものである。ハイパーパラメータdegreeが何次の多項式で用いるかを決め、式(3.3.2)ではdとして表

す. scaleは定数であり, 式(3.3.2)では1として表す. ハイパーパラメータとしてSVMに degree, scale, Cを与え, RVMでは degree, scaleのみを与える.

$$k(x_i, x_j) = (x_i \cdot x_j + l)^d \quad (3.3.2)$$

Linearカーネルは線形回帰のことで式(3.3.3)を用いるものである. ハイパーパラメータとしてSVMにCのみを与える. RVMではハイパーパラメータを用いない.

$$k(x_i, x_j) = x_i \cdot x_j \quad (3.3.3)$$

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。実務で使う際と同様の設定ではんだ付けを行うと不良はんだが発生しにくく、学習に十分な数を得ることが難しい[1]。不良はんだを発生され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ると、実務で使用する際に発生しない不良の原因を学習する恐れがある。そのため先行研究[1]では、実務の環境でも不良はんだ発生原因となっている余熱不十分な状況に着目した。余熱不十分な状況は、実際の製造環境における不良の種類のうち、実験で再現できる状況である。はんだ付けの時間間隔を実用では1秒未満であったところ30秒以上あけることで直前のはんだ付けによる余熱を減らすとともに、エアポンプによる送風で冷却することで余熱不十分な状況を再現することで不良はんだを発生させた。はんだ対象として20*16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[1]。スルーホールは前回はんだ付けの熱の影響を受けないよう、同じ行を連続で行わず、一行ずつ開けて行う[1]。

はんだ付けを完了した基板の各はんだ結果に対して、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。表1より、0~2点のはんだを不良はんだ「スルーホールにはんだなし」とし、4~9点のはんだを良はんだ「スルーホールにはんだあり」とすると開発者が決定した。

表1 各点数におけるはんだ良否

点数	はんだ良否
0	不良はんだ
1	
2	
4	良はんだ
5	
6	
7	
8	
9	

表2に取得した計3126件のはんだ付けの点数分布を示す。表の左から2列目は各はんだの点数を示し、値が高いほど良いはんだであることを表す。良はんだと不良はんだの件数に着目すると、不良はんだが440件であるのに対して、良はんだ

が2686件と数に偏りがあるため、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。なお、先行研究[1]においても学習に用いるデータをランダム抽出しているが、著者が同分析を再現した結果、ランダム抽出による精度への影響が認められた。本研究ではアルゴリズム間、カーネル間の比較を行う際にデータ抽出による結果の差が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3126件のデータに対して、各点数から表3のデータ件数の示す件数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そのまま用いる。SVMを用いた予測では0~2を不良はんだ TH_{NG} 、4~9を良はんだ TH_{OK} としたラベルを判別結果として用いる。

表2 取得したデータ

点数		データ数
不良はんだ	0	3
	1	390
	2	47
良はんだ	4	989
	5	546
	6	482
	7	172
	8	163
	9	334
合計		3126

4.2 特徴量抽出

特徴量抽出は先行研究とほぼ同じ条件で行う。先行研究[1]との違いは、はんだの良否の判別基準がスルーホールにはんだがあるかないかになったことである。先行研究[1]では、スルーホールへのはんだ充填である良否はんだの区別が可能な特徴が圧力に表れると考えられる工程として節3.1や図1にて(a)~(d')の工程を挙げた。はんだ付けの結果、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余熱不十分な環境にすると発生しやすい。基板を余熱不十分な環境にするとはんだ片の溶融が遅くなり、溶けたはんだがスルーホールにはんだが流れ込むことが遅くなるため不良はんだになりやすい。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、不良はんだの場合は良はんだの場合

表3 抽出件数

予測分類			データ件数
良否	良否ラベル	点数	
不良はんだ	TH_{NG}	0	3
		1	390
		2	47
良はんだ	TH_{OK}	4	75
		5	75
		6	75
		7	75
		8	75
		9	75
合計			890

合と比べて最高圧力を維持する時間長が短く、波形が尖った形状になると考えられる。これは節3.2にもあるように、不良はんだの場合ははんだがスリーブと基板の隙間を埋めることで圧力が降下せず、良はんだの場合は隙間からガスが抜けることによって圧力が降下する違いがあるため圧力変化が異なることが原因である。そのため最高圧力を保持する時間や波形の形状を特徴量とすることで、良否はんだの判断基準であるスルーホールにはんだあり、なしを分類できると考えられる。

これらの工程におけるはんだの良否の特徴を抽出できる特徴量として、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、はんだ片の溶融が遅いはんだを検出する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。はんだが溶け切って基板の穴を塞いでいる状態を最高圧力としたとき、最初に「最高圧力*0.9」となる点を始点とし、再び「最大圧力*0.9」となる点を終点とする。始点と終点における時間の差分をピーク長として算出する。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。そのため、はんだ片の溶融が遅いはんだを検出する特徴量として有用と考えられる。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。図2[1]に圧力変化の例を示す。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。また、外れ値を考慮して、上位20番目から40番目の平均値を最高圧力と定義する。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($T_{90\%}, T_{80\%}, \dots, T_{10\%}, T_{5\%}$)を圧力降下の様子を表す特徴量として用いる。

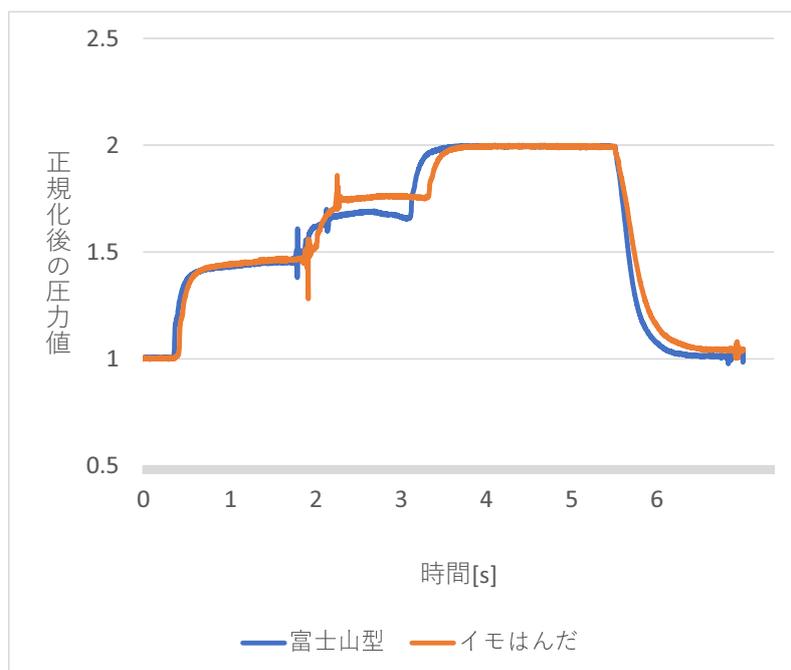


図2 良はんだ(富士山型)と不良はんだ(イモ)の圧力変化

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。ハイパーパラメータは節3.3.3で説明したものを用いる。各ハイパーパラメータの組み合わせを変更しつつ, 10点交差検証を10回行うことによって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caretのtrain¹を用いる。ハイパーパラメータのチューニングについて, 本研究は先行研究[1]と比べて異なる点が2つある。

1. ハイパーパラメータの範囲を指定しない

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

先行研究 [1] では福光らがハイパーパラメータの範囲を指定してチューニングしていたのに対して，本研究ではR言語の caret パッケージの train に一任して行う．メリットとしてはハイパーパラメータの最適な値を指定した範囲より細かい値まで探索してくれる点にある．

2. チューニング済の学習モデルで予測を行う

先行研究 [1] ではチューニング時の平均精度で議論していたが，本研究ではチューニングの結果得られた最適なパラメータを用いてデータを予測する．予測対象は100個あるデータセットのうちの学習に用いたデータセットである．メリットとしては多くの偏ったデータを対象として精度比較を行える点にある．

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の予測結果ははんだの良否を点数として出力するため良否の基準を設定する必要がある．本研究では学習済モデルが出力する値域の最大値が1.2であるため，中央の0.6と手動採点の中央である0.5，およびその前後である0.4～0.7の0.1刻みで良はんだと不良はんだに分類し精度を求める．SVMのRBFカーネル，Polynomialカーネル，LinearカーネルとRVMのRBFカーネル，Polynomialカーネル，Linearカーネル，計6つの組み合わせで比較し，精度を出すことで最もよいアルゴリズム，カーネル関数を見つける．

5 結果と考察

5.1 SVMとRVMの比較

図3, 図4, 図5にRVMでRBF, Polynomial, Linearの各カーネルを用いた際の予測結果を示す。水平の点線はそれぞれ0.4~0.7の基準を示す。横軸は学習に用いたはんだのIDを示し、縦軸は各はんだの点数を示す。また、各点の色がはんだ付け装置開発者による点数付けの結果を示す。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、良はんだの手動採点が高いはんだほど完成度の高いはんだである。手動採点が高いはんだになるにつれて機械学習でも不良はんだであると分類されるはんだが減少している。そのため、より完成度の高いはんだになるにつれて予測結果が縦に狭く分布している。そのため、機械学習による予測もより確実になってることがわかる。一方で、良はんだの手動採点が低いほど完成度の低いはんだであり、予測結果が縦に広く分布している。よって機械学習による予測も不確かなものであることがわかる。これらのことから、手動採点の点数が応じて、学習に用いる圧力波形の特徴もより良はんだであると分かりやすいものになっていると考えられる。良はんだ(4点)と不良はんだ(2点)の境界においては予測結果が縦に広く分布している。これは、良はんだと不良はんだの境界では圧力変化による違いが少ないことを意味している。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、少数ではあるが図の左側に予測結果の点数が高く、良はんだと予測されたケースが見られる。これは、予測結果の縦への広がり最大となっているため、現在の特徴では学べていない不良はんだの特徴があると考えられる。

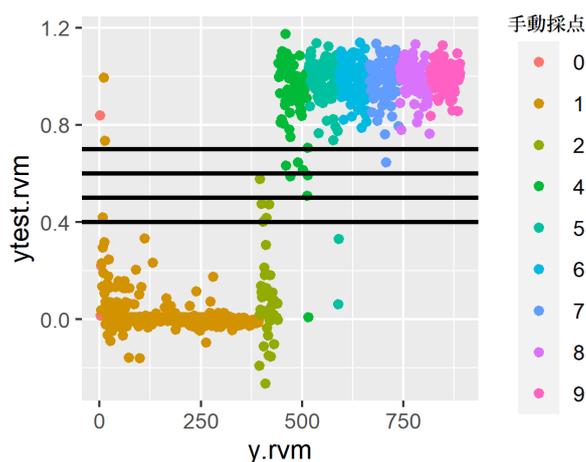


図3 RVM, RBFカーネルの予測結果

表4にそれぞれのカーネルを用いた際の分類精度を示す。RVMの列で用いたカーネルと基準の値を表し、太字の値はその列の精度における最高精度である。

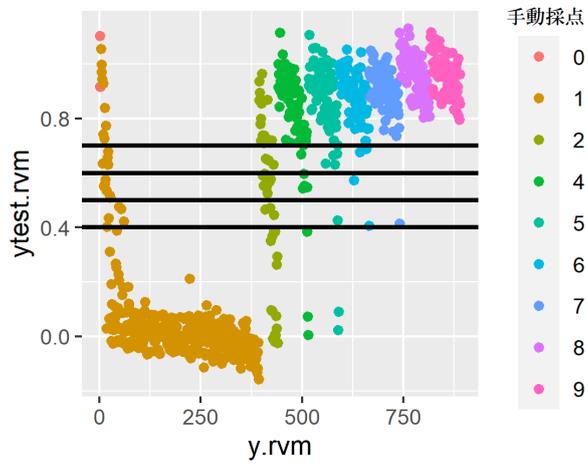


図 4 RVM, Polynomial カーネルの予測結果

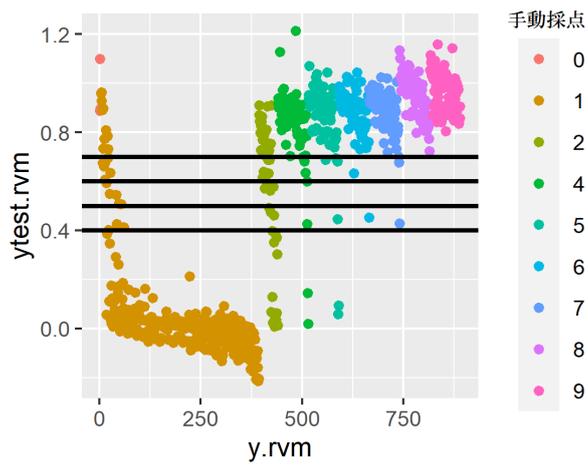


図 5 RVM, Linear カーネルの予測結果

RBFカーネルにおいて、最高accuracy, 最低accuracy, 平均accuracy, 平均FNがもっとも高かった。また、RBFカーネルの基準0.5を良否の境界とした時に最高accuracy, 最低accuracy, 平均accuracyの精度がもっとも高く、不良はんだの検出漏れである平均FNは境界値が0.7のときにもっとも高い。そのため、最高accuracy, 最低accuracy, 平均accuracyの精度を優先して運用するか、不良はんだの検出漏れである平均FNを優先するかによって基準となる値を変更することができると考えられる。結論として、RVMでは最も精度の高い組み合わせはRBFカーネルである。また、基準は0.5か0.7のときとなる。

RVMでは、予測結果の縦方向への分布の仕方で予測の確かさを視覚的に理解することが可能である。これは2値判別を行うSVMとの明確な差である。今回の結果を得たことによって現在の特徴による学習では良否を誤分類しやすい2か所を発見することができた。その2か所は0点と1点のはんだと、2点と4点のはんだの良否の境界のはんだである。現在特徴抽出されている尖度、ピーク長、圧力降下だけでは学習できていないのか、またはこれらのはんだの特徴のうちどこかに外れ値があることによって誤分類している可能性が発見できた。そのため、0点～4点における誤分類されたはんだの特徴と正しく分類されたはんだの特徴の違いを発見することによってより精度を高められると考えている。

表 4 RVM 結果

RVM		最高 accuracy	最低 accuracy	平均 accuracy	平均 FN
カーネル	基準				
RBF	0.4	99.3%	96.7%	98.4%	98.9%
	0.5	99.6%	97.5%	98.6%	99.2%
	0.6	99.4%	97.0%	98.5%	99.5%
	0.7	99.4%	96.3%	98.1%	99.6%
Polynomial	0.4	93.6%	91.5%	92.7%	93.1%
	0.5	94.3%	92.5%	93.5%	94.1%
	0.6	95.4%	92.7%	94.4%	95.5%
	0.7	95.6%	92.9%	94.5%	96.9%
Linear	0.4	93.3%	91.6%	92.4%	92.9%
	0.5	93.7%	92.2%	93.1%	93.7%
	0.6	94.4%	92.6%	93.6%	94.5%
	0.7	95.3%	92.7%	93.9%	96.2%

図6, 図7, 図8に100個目のデータセットで学習し、良否はんだを予測した際の結果を示す。横軸は学習に用いたはんだのIDを示し、縦軸である TH_{OK} は良はんだを示し、 TH_{NG} は不良はんだを示す。また、各点の色がはんだ付け装置開発者による点数付けの結果を示す。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できていることがわかる。RBFカーネルの時のみ4点のはんだが1つ

だけ誤分類されていることから、RVMの時と同じく良否はんだの境界は誤分類が起きやすい可能性がある。

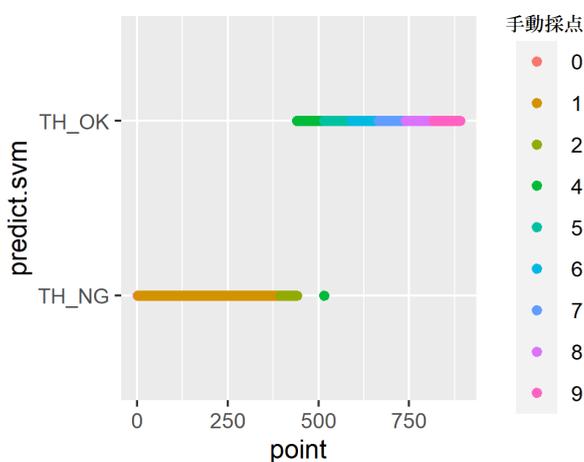


図 6 SVM の rbf カーネル出力結果

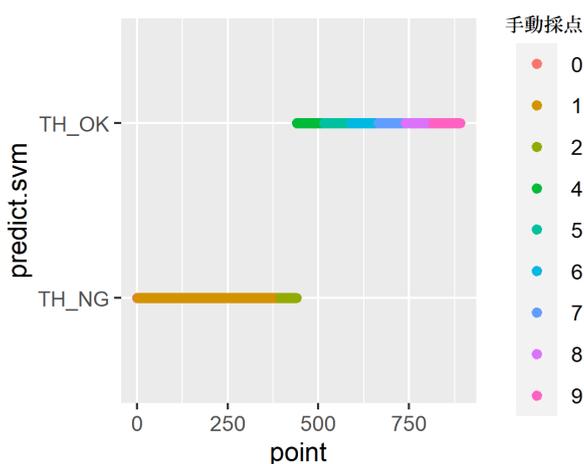


図 7 SVM の polynomial カーネル出力結果

表5では、良否はんだの分類においてSVMを用いた時の精度をまとめている。SVMの列では用いたカーネル関数を示し、太字の値はその列の精度における最高精度である。SVMではすべてのカーネルで最高accuracyとして100%が求められた。また、平均accuracy、平均FNではPolynomialカーネルとLinearカーネルが100%を示した。結果、Polynomialカーネルを選択したときに最低accuracyも含めてすべての精度において最も高くなることわかる。よってSVMではこの3つのカーネルのうち最適カーネルがPolynomialだと言える。

表6にSVMとRVMのそれぞれの最適カーネル精度を示す。組み合わせのうちアルゴリズムでは用いたアルゴリズムであるRVMとSVMを示している。カーネル

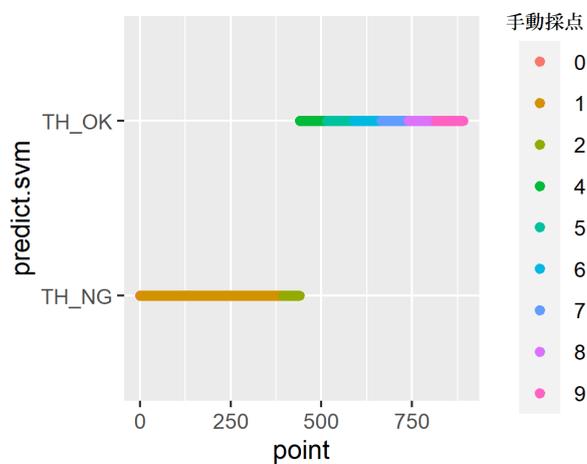


図 8 SVM の Linear カーネル出力結果

表 5 SVM 結果

SVM	最高 accuracy	最低 accuracy	平均 accuracy	平均 FN
RBF	100%	98.2%	99.7%	99.7%
Polynomial	100%	99.9%	100%	100%
Linear	100%	99.1%	100%	100%

表 6 SVM と RVM のそれぞれの最適カーネル精度

組み合わせ			最高 acc	最低 acc	平均 acc	平均 FN
アルゴリズム	カーネル関数	基準				
RVM	RBF	0.5	99.6%	97.5%	98.6%	99.2%
		0.7	99.4%	96.3%	98.1%	99.6%
SVM	Polynomial		100%	99.9%	100%	100%

関数ではそれぞれのアルゴリズムでの最適カーネルを示している．基準はRVMでは最高 accuracy, 最低 accuracy, 平均 accuracy の精度を優先して運用するか, 不良はんだの検出漏れである平均 FN を優先するかによって基準となる値が変化するため 0.5 と 0.7 の時の 2 つを示している．SVM では基準の値を用いないため空欄である．フォーマットの都合上, accuracy を acc と略して記載している．太字の値はその列の精度における最高精度である．このことからわかるように最高 accuracy, 最低 accuracy, 平均 accuracy, 平均 FN 精度のすべての項目で SVM の組み合わせが RVM の組み合わせを上回っている．よって, SVM と RVM のそれぞれの最適カーネルで比較すると最も良い組み合わせは SVM の Polynomial カーネルとなる．

5.2 学習に使った以外のデータを対象にした推定

5.1 の結果より, SVM の Polynomial カーネルが最も良い精度を出していることがわかった．本研究では 100 個あるデータセットのうち学習に用いたデータセットを予測を予測対象としている．節 5.2 では, 学習データは章 4 と同様に 100 件のデータセットを用いるが, 予測対象として同データセットを使用せずに表 2 の取得した全データを用いる．

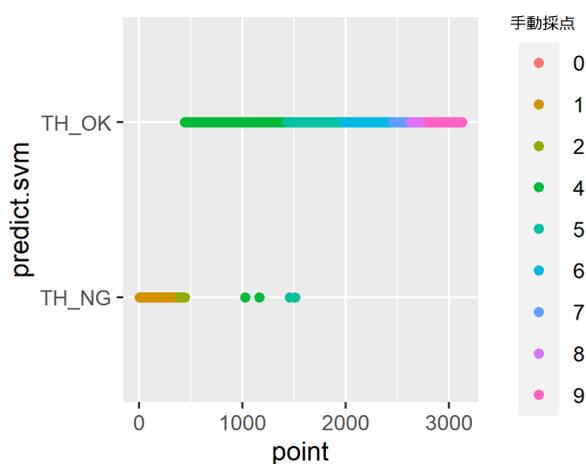


図 9 SVM の Polynomial カーネル全データ予測 1

図 9 と図 10 に 100 個目のデータセットで学習し, 表 2 を対象に良否はんだを予測した際の結果を示す．横軸は学習に用いたはんだの ID を示し, 縦軸である TH_{OK} は良はんだを示し, TH_{NG} は不良はんだを示す．また, 各点の色がはんだ付け装置開発者による点数付けの結果を示す．図 9 と図 10 より, 分類精度は図 7 である学習データを対象に予測した場合に比べて劣っていることが視覚的にわかる．また, 良はんだでは誤分類が生じているが, 不良はんだでは誤分類は生じていない．

表 7 に, SVM の Polynomial カーネルによる取得データの予測精度結果を示す．最高 accuracy, 平均 FN では 100% と節 5.1 と同じ結果を残しているが, 最低 accuracy と平

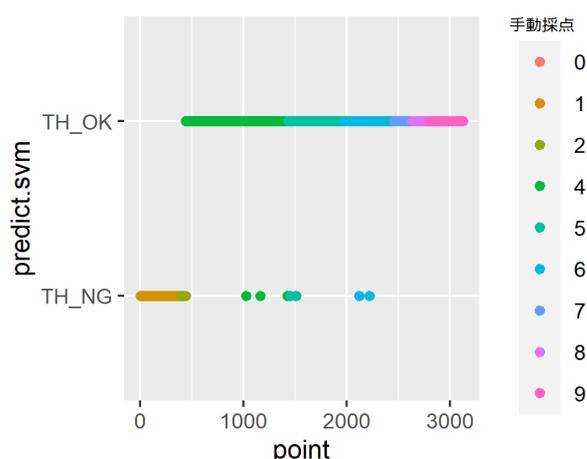


図 10 SVM の Polynomial カーネル全データ予測 2

表 7 SVMpolynomial 全データ予測

SVM	最高 accuracy	最低 accuracy	平均 accuracy	平均 FN
Polynomial	100%	98.5%	99.6%	100%

均 accuracy においては低下がみられる。

精度変化の原因として考えられるのはランダム抽出である。学習データは良はんだのみランダム抽出を行って作られている。そのため、良はんだとなるための特徴に多様な種類があったと考えられ、その場合は、抽出されていない良はんだの特徴は学べないということになる。そのように考えると、図9と図10において予測が失敗している点が異なっている理由も理解できる。また、4点の良はんだは表2からもわかる通りに絶対数が多いため、より強くランダム抽出による影響を受けていると考えられる。したがって、誤分類が生じる回数も多くなっている。このランダム抽出の影響が良はんだのみに反映されるとは考えられないため、不良はんだにおいてもランダム抽出を行った際には誤分類が生じるものであると考えている。

今後の課題として表2で良否はんだの数に偏りがあることがわかる。そのため、ランダム抽出を行わずに学習させることはできないので、より良否の均等がとれたデータを学習に用いたいと考えている。その学習データをSVMのPolynomialカーネルで学習させたとすると、節5.1と同程度の予測精度を得られると考えている。

5.3 ハイパーパラメータと過学習

5.1の結果より、SVMのPolyが最も良い精度を出しており、5.2からはランダム抽出によって予測精度に影響している可能性を考えた。しかし、本研究では先行研究[1]との変更点としてハイパーパラメータをコンピュータ側に一任している。そのため、過学習による精度の向上が考えられるので、その影響に対して考察する。

表8実験のハイパーパラメータ

ハイパーパラメータ	値	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

表7からハイパーパラメータの過学習による精度の影響を考察する。ハイパーパラメータの列にPolynomialのハイパーパラメータを示す。値はそれぞれのチューニングによってハイパーパラメータで選択された値である。数の列ではハイパーパラメータの値が選択された回数を示している。SVMのPolynomialカーネルにおいてハイパーパラメータは3種類あり、それらのハイパーパラメータも3通りの値を持っているため、チューニングは27通りとなる。しかし、今回のチューニングにおいてはdegree = 1の時と、scale = 0.001, scale = 0.01が選択されることがなかったため、実際には6通りとなる。

過学習とは、ハイパーパラメータがその学習データでのみ最適化されている状態である。そのため、他のデータに対しての予測精度が低下する。結論としては、SVMのPolynomialカーネルではハイパーパラメータをコンピュータ側に一任しても過学習による影響は少ないと考えている。理由は2つある。1つはハイパーパラメータの組み合わせが27通りあって6通りしか選択されていない点である。学習データは100個用意しているが、その学習データに最適化されているとしたら100通り近く組み合わせがないと最適化されているとは考えられない。最大値の27でもなく、6通りであるため過学習しているとは考えていない。2つはこの表7は節5.2にて学習した際のハイパーパラメータをまとめたものだからである。過学習が生じていた場合その学習データに最適化されていて高精度が求められるが、他のデータセットでは予測精度が低下する。確かに、表5のSVMのPolynomialと表7での精度の低下は起きている。しかし最低accuracyで1.4%、平均accuracyで0.4%

と小さい違いしかない。また、ランダム抽出による影響もあると考える。これらの2つの理由から過学習による影響は少ないとみている。

ランダム抽出による影響のない良否の偏りのないデータで学習した際には過学習がどの程度影響しているかの定量的な値を出せると考えているが、現段階では示せない。よって今後の課題として検討していきたいと考えている。

6 おわりに

本研究ではスリーブ型はんだ装置を用いてはんだ付けしたはんだを対象にスリーブ内の圧力変化から特徴量の抽出を行い、機械学習で良はんだ・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を用いた。カーネル関数としてRBFカーネル、Polynomialカーネル、Linearカーネルの3種類を用いた。アルゴリズムとカーネル関数の組み合わせである計6パターンでの分類精度を比較した。結果として、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100.0%と最も優れた分類精度だった。

今後の課題としては、良否はんだの境界である2点～4点のはんだにおいて両方のアルゴリズムで誤分類が目立ったため、特徴を抽出する箇所について議論する必要があると考えている。今回使用したスリーブにおいては良否はんだを最高精度100%と高精度で予測することができたため、他のスリーブにおいても同様の実験を行いたい。また、良否の均一なデータを取得で来た際にはランダム抽出を行わなくてよいため、チューニングによる過学習の影響についてより深く検討することが可能と考えている。

謝辞

本研究を進めるにあたり、多くの方々のご助力をいただきました。この場を借りてお礼を申し上げます。

株式会社アンド様には、スリーブ型はんだ付け装置であるスマートショットによるはんだデータの提供、はんだ付け装置の知識など多岐にわたって本研究へのご協力誠に感謝いたします。指導教員の上野秀剛准教授には、研究を始めるにあたって多くの知識やアドバイス、発表資料作成など様々な面でご指導をいただきました。査読教員である山口賢一教授はじめ、松村寿枝教授、岡村真吾准教授、岩田大志准教授には中間発表で大変貴重な意見を賜りましたこと深く御礼申し上げます。

参考文献

- [1] 福光嘉伸, "スリーブはんだ付け装置の圧力変化を用いた自動検査法", 令和3年度奈良工業高等専門学校専攻科特別研究論文.
- [2] 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).
- [3] 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).
- [4] 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
- [5] 渡邊真也, "カーネル法", 情報と知能, Vol.21, No.4, p.537