

システム創成工学専攻
情報システムコース

Department of Systems Innovation
Advanced Information System Course

令和6年度 専攻科特別研究論文

スリーブはんだ付け装置の圧力特徴量の選択と学習量削減
による不良検出精度への影響

Effect of Pressure-Features Selection and Learning
Reduction on Defect Detection Accuracy of Sleeve
Soldering Equipment

指導教員名 上野 秀剛 准教授

論文提出者名 渡邊 幹

独立行政法人 国立高等専門学校機構

奈良工業高等専門学校 専攻科

National Institute of Technology, Nara College
Faculty of Advanced Engineering

スリーブはんだ付け装置の圧力特徴量の選択と学習量削減による不良検出精度への影響

Effect of Pressure-Features Selection and Learning Reduction on Defect Detection Accuracy of Sleeve Soldering Equipment

渡邊 幹

Watanabe Kan

独立行政法人 国立高等専門学校機構

奈良工業高等専門学校 専攻科 システム創成工学専攻 情報システムコース

大和郡山市矢田町 22 番地 (〒 639-1080)

National Institute of Technology, Nara College, Faculty of Advanced Engineering
22 Yata-cho, Yamatokoriyama, Nara 639-1080, Japan

Abstract: An increase in element density of the printed wired board requires an automatic inspection method for incorrect soldering. In previous work, the authors developed the classification method for correct/incorrect soldering based on the pressure change of the sleeve soldering machine. The device melts the solder piece placed in the heated sleeve that covers the through-hole and pin. The pressure in the sleeve changes during each process, such as solder melting or through-hole closing. In this paper, the author evaluates the classification accuracy of the proposed method with feature selection and a reduced number of training data. The classification model with reduced features can run on a cheaper microcomputer or process quickly, improving the method's applicability. Our experiment uses 2,686 correct and 440 incorrect solders labeled by the device developer, and four different dataset sizes are created by random sampling. The result of the evaluation shows the feature selection and data reduction improve the classification accuracy.

Keywords: Sleeve-Soldering, Machine Learning, Feature Selection;

関連業績リスト

1. 渡邊幹, 上野秀剛, “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,” 情報処理学会 第 22 回情報科学技術フォーラム, 2023.
2. Kan Watanabe, Hidetake Uwano, Hiroshi Fukuoka, “Comparison of Algorithm and Kernel Accuracy in Defect Detection Using Pressure in Sleeve Soldering Equipment,” In The 8th International Conference on Robotics, Control and Automation, 2024.

目次

1.	はじめに	1
2.	関連研究	3
3.	準備	5
3.1	スリーブはんだ付けの自動検査法	5
3.2	機械学習	7
3.3	分類境界の可視化	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	11
5.	結果と考察	13
5.1	分類結果	13
5.2	分類境界の可視化結果	14
6.	おわりに	20
	参考文献	22

目次

3.1	はんだ付け中のスリーブ内圧力変化	6
4.1	先行研究 [1] の特徴量	11
5.1	特徴量選択の有無によるガウスカーネルの分類境界の変化（データセット 400 件#100）	16
5.2	特徴量選択の有無による多項式カーネルの分類境界の変化（データセット 400 件#100）	17
5.3	特徴量選択の有無による線形カーネルの分類境界の変化（データセット 400 件#100）	18

表目次

4.1	実験データ	10
5.1	全特徴量を利用した分類精度	13
5.2	特徴量を選択した分類精度	14
5.3	特徴量選択確率 (データセット 400 件)	15
5.4	選択された特徴量 (データセット 400 件#100)	19

1. はじめに

電子基板と素子を接合するはんだ付けの不良は基板の動作不良を引き起こすため、はんだ付けの不良を検出する検査は重要な工程である。また、はんだ付けによって素子と基板が電氣的に接続されている場合でも、基盤に空いたスルーホールに十分なはんだが充填されていないと接合部の強度が下がり、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。加えて、はんだ付けを行う面と反対の面にはんだの山であるフィレットおよびバックフィレットが形成されることで基板を表面と裏面からはんだで挟み込む形となり、より接合部の強度の高いはんだ付けとなる。はんだ付けの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、フィレット／バックフィレットが形成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。また近年、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。

現在主流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を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ポイントはんだ付けの1つであり、他の手法と比較して3つの利点がある。コテはんだ付けに比べはんだの広がりを表す濡れ性がよく、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、スリーブ上部からスリーブ内に流入する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。このとき、スリーブ下部が基板と接触したりはんだ片が融解することで窒素ガスの流出経路が変化し、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、スリーブはんだ付け装置によるはんだ付けの際のスリーブ内圧力変化を特徴量として Random

Forest と SVM(Support Vector Machine) を用いた推定手法を提案し, SVM で 95.5% の精度ではんだの良否を分類した [1]. 現状は実用的な分類精度ではないが, 今後の研究によって目標分類精度である 100% に近づいた場合, 最低でも人力検査と同等の精度が求められた場合に製造現場のラインで用いられるスリーブはんだ付け装置へ不良はんだ検出用のコンピュータが搭載される. 十分な精度を持つモデル生成に必要なデータ数, 特徴量が少ないほど搭載するコンピュータがより安価で済むため製造現場への適用が容易となる. 本研究では小さい学習モデルで十分な精度を維持するために使用する学習データの削減と使用する特徴量の選定を行う.

2. 関連研究

はんだ付けの良否を識別する方法として、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提案されている。Wu [2] は、基板画像のはんだ接合部の画像特徴を抽出し機械学習によって不良はんだを識別する手法を提案した。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た。

本研究の提案手法で用いるカーネル法は、SVM の成功等により多様な分野で利用されている。Karal [4] は SVM におけるカーネル関数と k-fold cross-validation の有効性を示す実験を行った結果、分類精度がカーネル関数の変更で最大 17.4%、cross-validation の交差回数変更で最大 16.7% 変化した。Wijayanti ら [5] は SVM を用いて twitter 上のいじめを分析するために 4 つのカーネルを使用し、分類精度を比較した。その結果、シグモイドカーネル選択時に 83.6% と最も高い精度を示した。このように検知、予測などにおける分類において SVM はよく用いられ [6] [7] [8] [9] [10] [11]、カーネル関数の変更による精度比較の研究も盛んにおこなわれている [12]。

また、すべての特徴量を用いずにモデルを作成する特徴量選択という手法が存在する。本研究では提案手法を製造現場で利用しやすくすることを目的に、学習に必要なデータ数の削減と特徴量選択を行う。KR ら [13] は Leukemia, Colon データセットにおける膨大な特徴量を選択することによる SVM-RFE を用いた分類精度を改善した。それぞれ特徴量を選択することによって Leukemia では平均 9.2%、Colon では 8.2% の向上が確認できた。このように特徴量選択による精度向上を目指す研究もあれば [14] [15]、選択手法の提案を行う研究も存在する [16]。

本研究で著者らは提案手法を製造現場で利用しやすくすることを目的に、学習に必要なデータ数の削減と特徴量選択を行う。一般に不良はんだのデータを十分な数用意するのは時間のかかる作業であり、十分な分類精度を持つモデル生成に必要なデータ数が少ないほ

ど製造現場への適用が容易になる。モデルによる判別を行う機能をはんだ付け装置に組み込む場合、装置に内蔵されたコンピュータで特徴量の算出を行うため、より少ない特徴量から高い予測精度を出せることが望ましい。また、少ない特徴量で構成されたモデルはメモリ容量が小さいため用いるコンピュータの金額を抑えることができ、実装のハードルを下げることができる。そこで、モデル生成に用いるデータ数を削減すると共に、特徴量選択をしたときの分類精度を実験で比較する。実験では SVM を用いて 3 種類のカーネル (ガウス, 多項式, 線形) を比較する。先行研究が用いていた 22 種類の特徴量に対する特徴量選択と、学習データ数を段階的に減らした際の分類精度を比較し、より製造現場で利用しやすく、精度が維持される条件を明らかにする。

3. 準備

3.1 スリーブはんだ付けの自動検査法

図 3.1 にスリーブはんだ付け装置によってはんだ付けを行う際のスリーブ内の様子と圧力変化を示す。図下部のグラフは横軸が時間、縦軸がスリーブ内の圧力を示す。図上部に各時間における基板とはんだ付け対象のピン、スリーブの断面、およびはんだを表している。(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)' は不良はんだの一種であるイモ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降下する。(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しており、スリーブが上昇するとすぐに接合部とスリーブ間に隙間が生じるため窒素ガスが流出し圧力が低下する。一方で (d)' イモはんだの場合は溶解したはんだがスルーホールに流入せず、基板上部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できず、圧力が低下するまでに時間がかかる。

福光 [1] はスリーブはんだ付け装置において、はんだ付けの良否によってスリーブ内の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

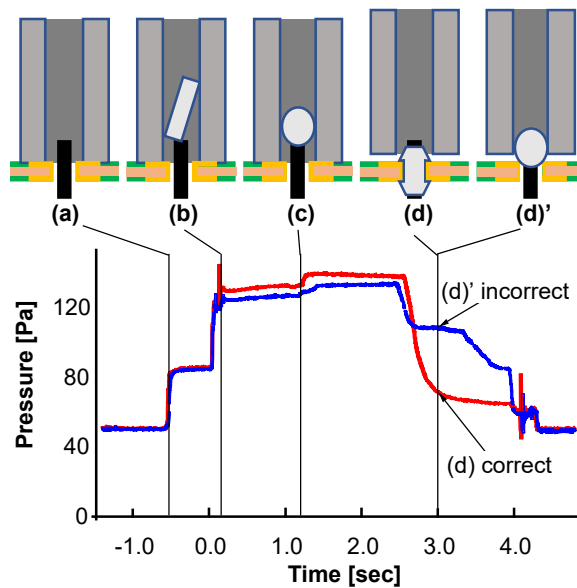


図 3.1 はんだ付け中のスリーブ内圧力変化

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。

本研究で著者らは提案手法を製造現場で利用しやすくすることを目的に、学習に必要なデータ数の削減と特徴量選択を行う。一般に不良はんだのデータを十分な数用意するのは時間のかかる作業であり、十分な分類精度を持つモデル生成に必要なデータ数が少ないほど製造現場への適用が容易になる。また、モデルによる判別を行う機能をはんだ付け装置に組み込む場合、装置に内蔵されたコンピュータで特徴量の算出を行うため、より少ない特徴量から高い予測精度を出せることが望ましい。そこで、モデル生成に用いるデータ数を削減すると共に、特徴量選択をしたときの分類精度を実験で比較する。実験では SVM を用いて 3 種類のカーネル (ガウス, 多項式, 線形) を比較する。先行研究が用いていた 22 種類の特徴量に対する特徴量選択と、学習データ数を段階的に減らした際の分類精度を比

較し、より製造現場で利用しやすく、精度が維持される条件を明らかにする。

3.2 機械学習

本研究では機械学習に Support Vector Machine (SVM) を用いる。SVM は 2 値クラス判別のための線形識別器である。SVM はサポートベクトルによって 2 クラス間の境界を生成し、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。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。

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。本手法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。そのため、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。本研究ではカーネル関数はガウス、多項式、線形の 3 種類を用いる。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。チューニングには R 言語のパッケージである `caret` の `train`^{*1}を用いる。

特徴量選択はすべての特徴量を使用せずに、アルゴリズムの選択によって用いる特徴量を特徴量選択の手法は Backward Selection を用いる。この手法は全特徴量で学習したモデルと選択した特徴量で学習したモデルを比較して必要ない特徴量を除去していく手法である。本研究ではモデル容量の削減を目的として特徴量の選択を行うため、精度が減少しなくなるまで特徴量の除去を繰り返す。

3.3 分類境界の可視化

SVM による分類結果として、分類境界を示す。特徴量に対して主成分分析を用いることによって 2 つの主成分を作成する。縦横軸を 2 つの主成分とした 2 次元の図を用意する。その特徴量をもつテスト対象となるデータをプロットすることによって 2 つの主成分

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

からなる分布図が完成する。2つの特徴量からなる学習データを作成し、SVMによる学習済みモデルを作成する。上記分布図にモデルに含まれる境界をプロットすることによってテストデータに対する分類境界を確認することができる。主成分分析を用いて特徴量を2次元に落とし込んでいるため、可視化した2次元の境界領域はもとの境界領域と同値ではないことに注意したい。

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。圧力はスリーブ内に圧力センサーを配置し 500Hz で計測する。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。

実際の利用環境において発生しうる不良はんだの原因の 1 つである予熱不足の状況を再現して不良はんだのデータ件数を確保する。そのため、予熱不足以外で発生する不良はんだが含まれるワークに適用すると、分類精度が低下する恐れがある。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。遅くなることによってスルーホールに正しくはんだが流れ込みづらくなり、不良はんだが発生しやすくなる。実験では基板に対してエアポンプで常温空気を送風し、はんだ付け間の時間間隔を 30 秒と長くとり、隣接したスルーホールを避けることで予熱の効果を小さくする。投入するはんだ片はスルーホールの径を考慮した標準的な量を用いる。

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 人が目視で良否の分類を行う。3,126 件のはんだ付けの内、2,686 件が良はんだ、440 件が不良はんだだった。不良はんだの内、1 件は圧力データの形状が他と大きく異なるため除外した。本研究ではデータ件数の少ない不良はんだについて、200 件をパラメーターチューニングと特徴量選択 (Tr) に使い、200 件を性能評価 ($Test$) として用いる。収集したデータに対してランダム抽出を行い良否はんだ各 50, 100, 150, 200 件とデータ数と揃えたデータセット Tr を 4 パターン (Tr_{50} , Tr_{100} , Tr_{150} , Tr_{200}) 作成する。テストデータ $Test$ は学習に使用していないはんだデータを良否各 200 件用いて計 400 件とする。本研究では収集したデータ数に偏りがあるため、ランダム抽出を実施しデータ数をそろえる必要がある。良否はんだ各 200 件 (Tr) をパラメーターチューニングと特徴量選択に使い、良否はんだ各 200 件 ($Test$) を

表 4.1 実験データ

	Recorded	Tr_{50}	Tr_{100}	Tr_{150}	Tr_{200}	Test
Failure	440	50	100	150	200	200
Success	2686	50	100	150	200	200
Total	3,126	100	200	300	400	400

性能評価として用いる。データセット (Tr) として良否はんだ各 50, 100, 150, 200 件からなる計 100, 200, 300, 400 件のデータセットを 4 パターン (Tr_{50} , Tr_{100} , Tr_{150} , Tr_{200}) 作成する。テストデータ (Test) は学習に使用していないはんだデータを良否各 200 件用いて計 400 件とする。同様の操作を各 100 回繰り返し、4 パターンのデータセット (Tr) を 100 組を作成する。表 4.1 に実験で用いるデータセットの概要を示す。はんだ付けごとに計測した圧力変化から、先行研究で用いた圧力波形の尖度、ピーク長、圧力降下にかかる時間、圧力降下の傾きの 4 種類 22 個を特徴量として抽出する。

4.2 特徴量抽出

本研究では先行研究 [1] が利用した圧力波形の尖度、最大値、ピーク長、圧力降下に掛かる時間の 4 種類 (図 4.1) を特徴量とした機械学習による分類を行う。**尖度**は分布の尖り度合いを表し、グラフの波形が尖っているほど大きい値となる。熱量不足によってはんだ片の溶融が遅れてイモはんだとなった場合、図 3.1(d)' のように圧力降下に時間がかかり、良はんだと比べてグラフの変化が緩やかになる。本手法でははんだ付け開始から終了までの圧力変化データを用いて正規分布の尖度が 0 になるように調整された尖度であるフィッシャー尖度を求める。**ピーク長**は最高圧力に達して持続した時間長であり、はんだ融解した後に基板と接しているスリーブが上昇するまでを表していると考えられる。ピーク長は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。**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を表す。1 回のはんだ付けにおいて、スリーブが基板に接触していない状態で圧力が最低値となり、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状態で最高値となる。このとき、図 3.1 に示したように良はんだでは圧力が素早く降下するのに対して、不良はんだでは圧力が低下するまでに時間がかかる。

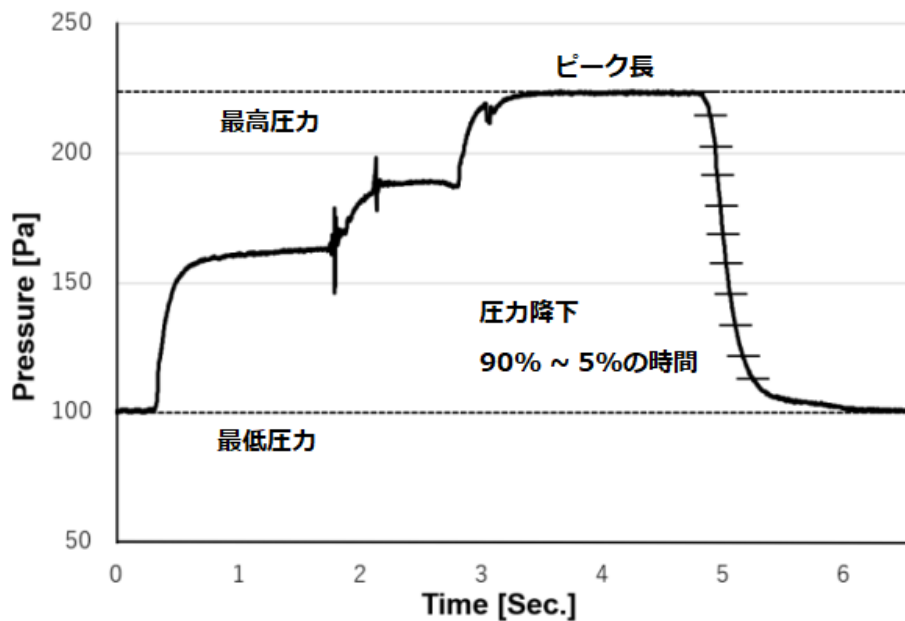


図 4.1 先行研究 [1] の特徴量

最高圧力を 100%, 最低圧力を 0% とし, 最高圧力に到達後, 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. 圧力降下時の傾きは指定された 10 間隔ごとの圧力降下率を示す. 不良はんだでは, 良はんだに比べて圧力降下の開始が遅れるため圧力が素早く降下し, 傾きの特徴に反映される. 圧力降下時の傾き $S_{X\%-Y\%}$ は, 圧力が X% から Y% に降下する差異を経過時間で割った値として定義される. 指定の 10 段階 (90 %, 80 %, ..., 10 %, 5 %) に隣接する 2 つの %, すなわち $S_{90\%-80\%}$, $S_{80\%-70\%}$, ..., $S_{10\%-5\%}$ を選択し, 全体の圧力降下の傾き $S_{90\%-5\%}$ を求める.

4.3 分類

それぞれの Tr に対して SVM を用いた良はんだ/不良はんだを分類するモデルを作成する. 1 つのデータセットに対して 22 件の特徴量を用いて学習モデルを作成し, 10 回の 10 分割交差検証でハイパーパラメータをチューニングする. チューニング後のモデルに対して精度に基づいた特徴量を選択し, 再度 Tr を用いた学習とハイパーパラメータのチューニングを行う. 作成されたチューニングと特徴量選択後のモデル 400 個を

用いて *Test* に対する分類精度を求める。チューニングには R 言語のパッケージである *caret* の *train*^{*1}を用いる。学習アルゴリズムは SVM, 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。

また、本研究では 4 種類の特徴を 22 件の特徴量としているが、より重要な箇所の特徴量を選択して学習することも可能である。この特徴量選択によって特徴量を減少させるとモデルのもつ容量が減少し、ノイズとなる特徴を排除できた場合には分類精度が向上することもある。モデルを構成するデータが少ないほど実機搭載時に必要なコンピュータのメモリ容量が下がるため、特徴量を減少させつつ分類精度を維持する必要がある。そのため、最大 22 件の特徴量を持つ学習モデルで求めた分類精度と特徴量を減少させたモデルで求めた分類精度を比較するために *Backward Selection* を用いる。本研究ではすべての特徴量を用いて学習したモデルによる分類精度と、精度が減少しなくなるまで特徴量の除去を繰り返し最終的に残った特徴量で学習したモデルによる分類精度を比較する。

同時に、分類境界を可視化することによって特徴量選択によって不必要な特徴を除去した際に分類境界の変化を確認する。また、分類境界を確認することができた場合正しく分類できているはんだ、正しく分類できていないはんだを確認することが可能となる。正しく分類できていないはんだを特定することができた場合、現在の 22 件の特徴で捉えることのできなかつた特徴を持つ波形を確認することができる。そのため、分類境界を確認することは今後の精度向上のために重要なタスクといえる。

しかし、本研究での分類境界は 22 件の特徴量からなる 22 次元の超平面として生成されている。本節では分類境界を可視化するために主成分分析を用いて 22 次元の特徴量を 2 次元にまで圧縮する。同時に、特徴量選択を行ったあとの選ばれた特徴量に対しても主成分分析を用いて 2 次元にまで圧縮する。ここで注意したいのが、本手法で作成するモデルは主成分分析により圧縮された 2 次元の特徴量から学習された点である。本来の特徴量で学習したモデルより分類精度が低下しているはずであるため、本来の分類境界でないと留意したい。

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

5. 結果と考察

5.1 分類結果

表 5.1 にすべての特徴量を用いたときの分類精度を示す。各列はモデル構築に用いたデータセットを示す。Test の 3 行は各モデルによってモデル構築に用いていない 400 件のデータ (データセット Test) を分類したときの精度を表す。太字の値は各行の最良を表す。平均して精度が最も高いのはガウスクーネルで、最も精度が高いのは Tr_{150} を用いたときの 93.0% である。精度に学習件数ごとの影響が少ないのが多項式カーネルを用いた場合で、すべての学習件数で最も精度が低いのが線形カーネルを選択したときである。いずれのカーネルでもデータ数が多いと分類精度が高く、データ数が少ないとわずかに精度が低下する。

表 5.2 に特徴量選択を行ったときの分類精度を示す。特徴量選択によってガウスクーネルは最大 0.8% と比較的小さい分類精度の低下を発生させた。多項式カーネルは最大 0.4% の精度低下、線形カーネルに至っては分類精度がすべての学習件数で上昇している。しかし、最も精度が高いカーネル関数はガウスクーネルとなっており、特徴量選択の有無で変化はない。特徴量選択によりのこった特徴量の個数は学習のたびに变化するため全体の結果としては記載しない。

表 5.1, 5.2 より、カーネル関数によっては精度が上昇することもあることがわかる。ま

表 5.1 全特徴量を利用した分類精度

カーネル関数	Tr_{50}	Tr_{100}	Tr_{150}	Tr_{200}
Test ガウス	92.4%	92.9%	93.0%	92.9%
Test 多項式	92.5%	92.6%	92.6%	92.6%
Test 線形	91.7%	92.2%	92.5%	92.5%

表 5.2 特徴量を選択した分類精度

カーネル関数		Tr_{50}	Tr_{100}	Tr_{150}	Tr_{200}
Test	ガウス	91.6%	92.6%	92.6%	92.8%
	多項式	92.1%	92.4%	92.5%	92.4%
	線形	92.0%	92.5%	92.7%	92.6%

た、分類精度の低下も大きくなく、最大 0.8% 最小 0.1% であるため影響は少ない。学習件数も Tr_{50} の時点で他の Tr より最大 1.0% 最小 0.1% と変化が少なく、さらに学習件数を減少させることも可能であるといえる。

表 5.3 に Tr_{200} に対して特徴量選択を行ったときに選択された変数を示す。縦軸はそれぞれの特徴量 22 件である。横軸は学習に用いたカーネル関数である。それぞれの列には学習全体を通して変数が選択された確率を示しており、50% 以上の選択率を示した場合太字としている。圧力降下の終わり際を示す変数 ($T_{10\%}$, $T_{5\%}$, $S_{10\%-5\%}$) はいずれのカーネルでも用いられる可能性が低い変数である。これらの必要のない特徴量を除去し、のこりの特徴量のみで学習することでモデル容量を削減することが可能である。また、尖度 (f_kurtosis) や圧力降下の中間を示す変数 ($T_{50\%}$, $T_{40\%}$, $T_{30\%}$, $S_{70\%-60\%}$) や一部の圧力降下の終わり際の傾きを示す変数 ($S_{20\%-10\%}$) は選択される可能性の高い変数である。これらの重要度の高い特徴量のみで同等の分類精度が求められた場合、多くの特徴量は必要ないものであり、よりモデル容量の削減が可能である。

5.2 分類境界の可視化結果

図 5.1, 図 5.2, 図 5.3 に Tr_{200} (#100) に対して各カーネルが決定した境界を示す。それぞれの図にある (a) は特徴量選択なし, (b) は特徴量選択ありの場合を示す。表 5.4 にはそれぞれの特徴量選択によって選択された特徴量に \circ を示している。可視化のために、各グラフは分類に用いられた特徴量 (選択無しの場合は 22 件の特徴) に対して主成分分析を行い、2つの主成分分析を作成している。2つの主成分を構成する特徴やその重みはグラフごとに異なる。また各グラフにおける2つの主成分の寄与率は特徴量選択なし (a) の時 77.1%, 図 5.1(b) が 81.7%(特徴量 13 個), 図 5.2(b) が 80.9%(特徴量 14 個), 図 5.3(b) が 72.5%(特徴量 11 個) であり、表 5.1, 5.2 の結果を完全には反映していない事に注意され

表 5.3 特徴量選択確率 (データセット 400 件)

	Linear	Polynomial	RBF
f_kurtosis	84%	97%	93%
f_maxwave	50%	67%	42%
T _{90%}	46%	65%	70%
T _{80%}	38%	49%	66%
T _{70%}	32%	54%	80%
T _{60%}	46%	75%	89%
T _{50%}	65%	83%	93%
T _{40%}	70%	85%	87%
T _{30%}	50%	70%	86%
T _{20%}	43%	63%	77%
T _{10%}	43%	45%	39%
T _{5%}	12%	20%	22%
S _{90%-80%}	22%	45%	75%
S _{80%-70%}	29%	82%	78%
S _{70%-60%}	56%	85%	65%
S _{60%-50%}	48%	77%	78%
S _{50%-40%}	39%	69%	65%
S _{40%-30%}	36%	67%	64%
S _{30%-20%}	31%	64%	58%
S _{20%-10%}	53%	77%	75%
S _{10%-5%}	20%	16%	25%
S _{90%-5%}	47%	55%	44%

たい。図 5.1 にはガウスクーネルの特徴が強く出ている。特徴量選択によって少し良はんだの領域を狭めるような変化が確認できる。現状では、良否の境界あたりを分類するための新たな特徴を発見できるとガウスクーネルがきれいな分類境界を引く可能性も高い。しかし、学習データへの過学習が心配される。図 5.2, 5.3 は多項式、線形カーネルであるが、変化はないように見える。しかし、分類境界の引き方が似ていることから表 5.3 での

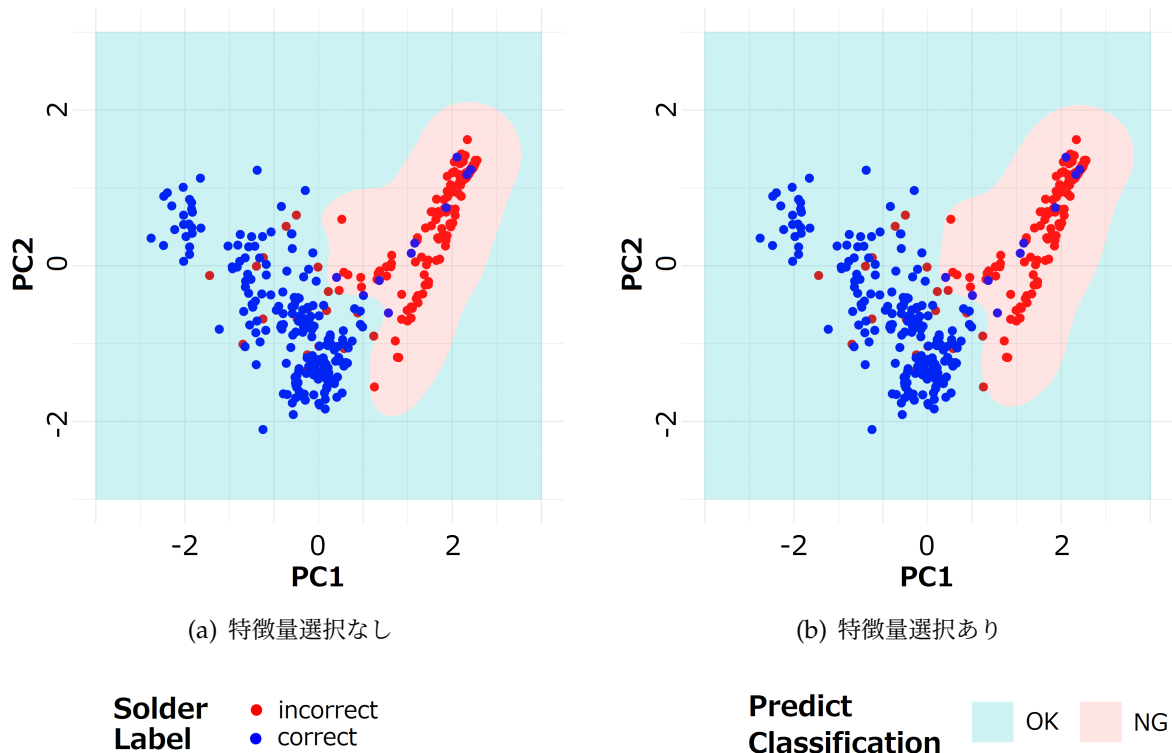


図 5.1 特徴量選択の有無によるガウスカーネルの分類境界の変化 (データセット 400 件#100)

精度に違いが小さい原因を確認できた。境界付近のはんだについて分類するのは困難だが、これらの特徴を新たにとらえることができた場合、過学習の心配なく良否を分類することが可能かもしれない。本データセットでは左側に良はんだ、右側に不良はんだが集中する領域が存在しており、いずれのグラフにおいても左右を分割する位置に境界が引かれている。また、境界付近の良はんだと不良はんだが混在している領域や、良が多い領域に少数の不良が存在し、不良が多い領域に少数の良が存在する。これらの誤分類されたはんだを特定し、波形の特徴を確認することで新たな特徴となる箇所を発見できる可能性が高い。また、表 5.4 には 5.3 で示した必要のない特徴も選択されており、一部の重要である変数が選択されていない。しかし、分類境界には大きな変化がなく特徴量選択を行わなかった (a) にそれぞれ似ている。そのため、距離の近い特徴量 ($T_{50\%}$, $T_{40\%}$, $T_{30\%}$) が似た意味をもつ特徴である可能性がある。尖度 ($f_kurtosis$) や圧力降下の中間を示す変数 ($T_{40\%}$, $S_{70\%-60\%}$) や一部の圧力降下の終わり際の傾きを示す変数 ($S_{20\%-10\%}$) 4 つでも十分な精度が求められる可能性が高い。

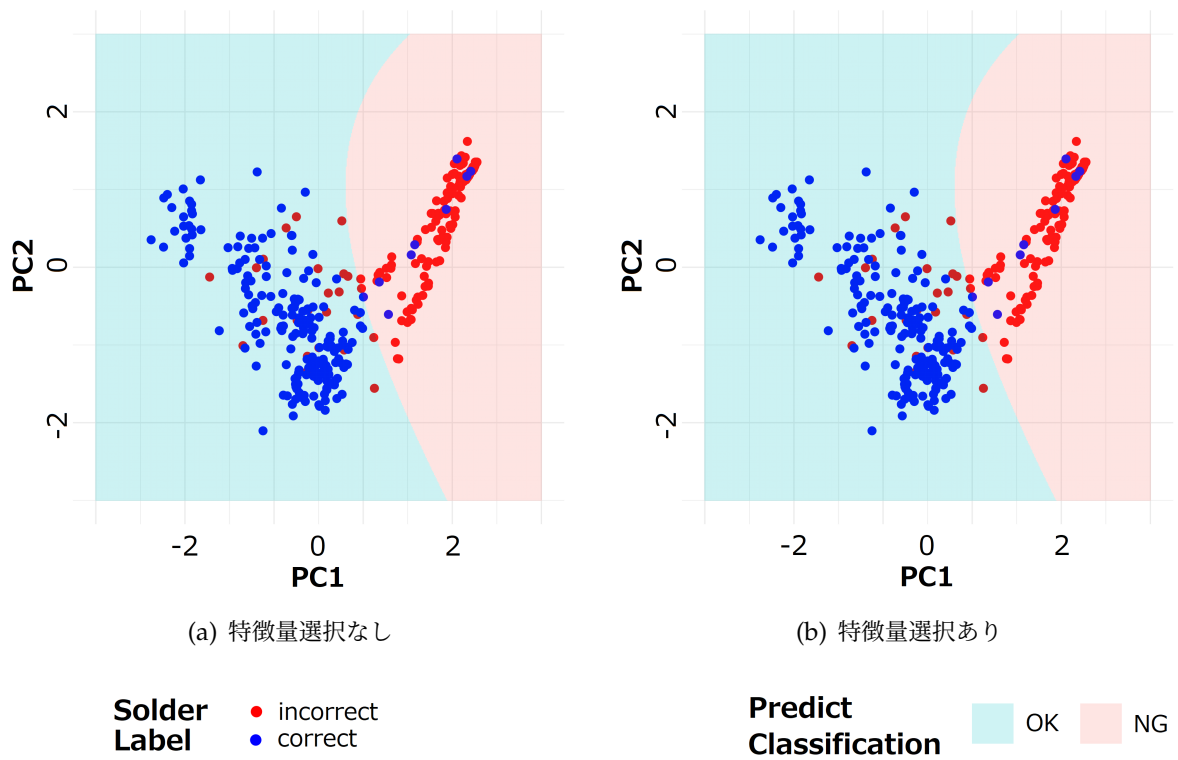


図 5.2 特徴量選択の有無による多項式カーネルの分類境界の変化 (データセット 400 件#100)

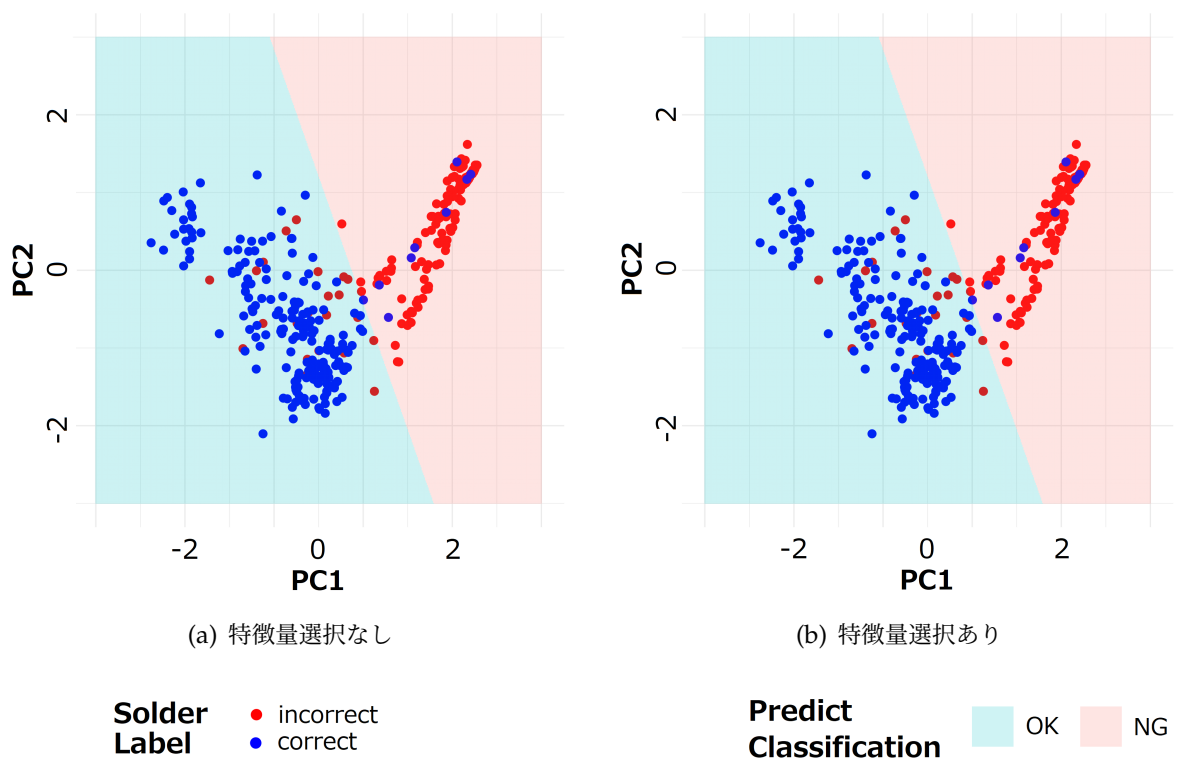


図 5.3 特徴量選択の有無による線形カーネルの分類境界の変化 (データセット 400 件#100)

表 5.4 選択された特徴量 (データセット 400 件#100)

	Linear	Polynomial	RBF
f_kurtosis	○	○	○
f_maxwave_		○	
T _{90%}	○		○
T _{80%}			
T _{70%}			
T _{60%}	○	○	○
T _{50%}		○	○
T _{40%}	○	○	○
T _{30%}			○
T _{20%}	○	○	○
T _{10%}	○	○	
T _{5%}			
S _{90%-80%}			
S _{80%-70%}	○	○	○
S _{70%-60%}		○	○
S _{60%-50%}	○	○	○
S _{50%-40%}	○		○
S _{40%-30%}		○	
S _{30%-20%}			
S _{20%-10%}		○	○
S _{10%-5%}	○	○	○
S _{90%-5%}	○	○	

6. おわりに

本論文ではスリーブはんだ付け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、SVMによる良はんだ、不良はんだの分類を行った。RBF, Polynomial, Linear の3種類のカーネルを用いて、学習に用いるデータ数と特徴量選択の有無を変化させたときの分類精度を比較した。実験の結果、学習に用いていないデータセットを対象とした分類精度は学習件数を100,200,300,400件と増減させてもほとんど変化せず、特徴量を削減しても分類精度は最大0.8%と大きく低下しなかった。また、特徴量選択によって必要のないであろう特徴3つを発見できたため、これらの特徴は必要ない可能性が高い。

今後はデータ数が100件400件でも精度差が小さかったため、よりデータ数を減少させることによって学習コストの削減が可能である可能性が高い。また、今回重要であると示された特徴6個による学習によって分類精度が低下しなかった場合、必要な特徴量は6つ、もしくは4つといえるためよりモデル容量の削減が可能である。今後、分類精度を向上させるためには特徴量を追加する必要がある。現在利用している特徴量では良の領域に不良が含まれたり、その反対の分布が存在している。そのため、データの分布上の良はんだの特徴を持つ不良はんだや、不良はんだの特徴を持つ良はんだの圧力波形から新たな特徴量を調査する必要がある。

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、3年間研究を行う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。査読教員である市川嘉裕助教をはじめ、松尾賢一教授、内田真司教授、山口賢一教授には特別研究論文発表において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] Y. Fukumitsu, K. Nakamichi, H. Uwano, and H. Fukuoka, "Defect classification using pressure change of sleeve soldering machine", *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, "Soldering defect detection in automatic optical inspection", *Advanced Engineering Informatics*, Vol.43, p.101004, 2020.
- [4] Ö. Karal, "Performance comparison of different kernel functions in SVM for different k value in k-fold cross-validation", *Innovations in Intelligent Systems and Applications Conference (ASYU)*, 2020.
- [5] S. S. Wijayanti, U. Ema, and Y. Ainul, "Comparison of kernels on support vector machine (SVM) methods for analysis of cyberbullying", *International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, 2022.
- [6] M. W. P. Indi, A. Novianty, and A. L. Prasasti, "Automatic first arrival picking on P-wave seismic signal using support vector machine method", *International Conference on Information and Communication Technology (ICoICT)*, 2020.
- [7] S. Shobitha, P. M. Amita, B. Niranjana Krupa, and G. K. Beng, "Cuffless blood pressure prediction from PPG using relevance vector machine", *International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT)*, 2017.
- [8] H. Shi, X. Kang, and T. Meng, "Modeling of pressure-flow in the flow facility

- based on support vector machine", Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC), 2020.
- [9] Y. Yamasari, A. Qoiriah, N. Rochmawati, I. M. Suartana, O. V. Putra, and A. I. Nurhidayat, "Exploring the kernel on SVM to enhance the classification performance of students' academic performance", Fifth International Conference on Vocational Education and Electrical Engineering (ICVEE), 2022.
- [10] L. Ilagan, "Comparative analysis of KNN, Random Forrest, decision tree and SVM in application to millimeter wave chemical detection", International Conference on Engineering and Emerging Technologies (ICEET), 2022.
- [11] H. Cho, J. Chun, S. Song, and S. Jung, "Radar target classification using the relevance vector machine", IEEE Radar Conference, 2014.
- [12] A. Goel and S. K. Srivastava, "Role of kernel parameters in performance valuation of SVM", Second International Conference on Computational Intelligence & Communication Technology (CICT), 2016.
- [13] K. Kr, A. R. Kv, and A. Pillai, "An improved feature selection and classification of gene expression profile using SVM", International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICT), 2019.
- [14] Md. S. Islam, Md. A. Mohammad M. Rahman, "Feature Selection and Classification of Spinal Abnormalities to Detect Low Back Pain Disorder using Machine Learning Approaches", International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), 2019.
- [15] A. U. Haq, J. Li, M. H. Memon, "Heart Disease Prediction System Using Model Of Machine Learning and Sequential Backward Selection Algorithm for Features Selection", International Conference for Convergence of Technology (I2CT), 2019.
- [16] X. Liu, J. Tang, "Mass Classification in Mammograms Using Selected Geometry and Texture Features, and a New SVM-Based Feature Selection Method", IEEE Systems Journal, 2013.