

博士論文

外観検査における見逃しを抑制する機能を持つ AI の開発と
企業への新技術導入の実践

光産業創成大学院大学

岡崎元樹

博士論文要旨
外観検査における見逃しを抑制する機能を持つ AI の開発と
企業への新技術導入の実践

本論文は、筆者が所属する自動車用クラッチ部品メーカーが、自動車業界を取り巻く情勢の激しい変化に対応し、競争力を維持・強化し、将来的な成長を継続することを目的として取り組んだ、光技術を応用した新しい外観検査方法の開発と企業への新技術導入の実践について論じている。

自動車業界は大きな変革の時代を迎えており、自動車部品メーカーは情勢の激しい変化に対応し、将来的な成長を継続するため、DX（デジタルトランスフォーメーション）を始めとした新たなデジタル技術の活用を推し進めている。その取り組みの一つに人間の感覚的な作業の自動化があり、AI（人工知能）の活用が期待されている。本研究では光技術と AI を用いて、従来技術では自動化が難しかった外観検査の自動化システムの構築と新たなデジタル技術の活用を促進するための組織体質造りを行った。

製品の重要な品質保証項目に完成品の外観があるが、現状は人の感覚に頼った目視検査が欠かさないため、目視検査と同等以下の見逃し率を実現する自動外観検査装置の開発が望まれている。一般に外観に現れる欠陥は多様であるうえ、その判定基準を明確にすることが難しい場合が多い。したがって、従来のルールベースの画像検査装置では判定基準を定めることが容易ではなく、完全な自動化は困難である。これに対し、AI は正常品と欠陥品のデータを与えることで、機械学習により複雑なルールを定めることが可能である。本研究では AI による画像認識の主要な技術である CNN（畳み込みニューラルネットワーク）を用いて、正常品と欠陥品を分類可能な外観検査装置の開発に取り組んだ。CNN を単純に適用した場合であっても外観欠陥分類の正解率 90% を達成し、外観検査自動化の可能性を確認できた。しかしながら、1) 学習中の正解率が安定しない、2) 見逃しと誤警告を同時に減少させることは困難である、3) 入力データの品質が分類性能に影響する、等の課題が確認された。学習が安定しない要因は、正常品と複数の欠陥種を同時に判別しようとする、判定基準があまりに複雑になってしまうことであると考えられた。そこで、多クラス分類のタスクを個々の欠陥種ごとに分類する 2 値問題に置き換え、単純化することで解決を試みた（開発手法 1：2 クラスアンサンブル手法）。開発した手法は正常品と欠陥品の 2 クラス分類で、通常が多クラス分類法と比較して正解率 0.25 ポイント向上の成果を得た。量産環境での実用性を確認するために、実際の量産を想定したデータセットを用いた多クラス分類実験を行った。その結果、正解率 95.69%、見逃し率 1.62%、誤警告率 1.87% の分類性能を得ることに成功し、量産環境においても、自動化の可能性を確認できた。次に、見逃し抑制に取り組んだ。外観検査の最も重要な機能は「見逃し」の発生を最小限に抑えることである。この解決にあたり、人の目視検査での正常品と欠陥品の発見プロセスの違いに着目した。正常品はいずれの欠陥にも該当しない場合のみ検出する消極的なプロセスであるのに対して、欠陥品はいずれかの欠陥に該当する可能性がある場合に検出される積極的なプロセスである。このプロセスの違いを正常品と欠陥品の取り扱い非対称性と定義し、CNN モデルに取り入れることで、欠陥の見逃し低減と誤警告低減を両立させる手法を開発した（開発手法 2：非対称

ラベルスムージング手法)。実験では従来 CNN 手法で 1.62% (82 件) あった見逃しを 0.08% (4 件) まで低減することができた。この手法は学習の過程で与える正解の提示方法の工夫により、CNN モデルの性格を誘導することに成功したとすることができる。最後に、入力データの品質向上に取り組んだ。学習における入力画像の撮影時には光技術を応用し、欠陥部を強調したり、ノイズを低減したりすることが可能である。照明の当て方やフィルタの利用など、いくつかの手法を試したところ、本研究で対象とするクラッチディスクの検査においては偏光フィルタの利用が非常に有効であることを見出した(開発手法3: 画像データの質向上による分類性能向上手法)。これによりノイズの少ない画像を取得することが可能になり、性能向上が確認できた。これらの技術は実際の量産工程における欠陥検出の性能向上と見逃し低減に有効な手法であり、その活用が期待できる。

これらの技術開発に続き、開発した手法を実際の生産現場に導入するまでの課題整理を行った。一般に、企業への AI 技術の導入では、「AI に対する理解不足」と「AI 導入効果への不安」が大きな課題となっている。本研究では AI 導入の実践活動として、1) 生産技術エンジニアを対象とした講習会の開催と、2) AI システムの内製化を行い、AI 導入の課題解消を試みた。1) 講習会の開催においては、自らの手を動かしながら実際のシステムを構築するアクティブ・ラーニングを取り入れ、AI に対する理解度向上を図り、成果を得ることができた。2) システム内製化では、これを実現したことで、実証実験を行える環境が整備された。これにより、実際に動作する AI を目の当たりにし、導入の効果の見通しを得られるようになることで、導入効果への不安を取り除く体制が構築できた。これらの活動により、複数の AI 導入案件を獲得するなど、企業への AI 導入の端緒につくことができた。本研究での組織体質構築の活動は、新しい技術の導入を促進する組織作りの実践事例となった。

本研究では外観検査自動化のための画像検査システム、特にクラッチディスク完成品における外観欠陥検査の自動化技術開発とその技術導入を促進するために組織構築の実践を行った。本研究で開発した技術は人の視覚を代替し、人を超える性能を実現する「光×AI」テクノロジーとして、製造業に限らずさまざまな分野での活用が期待され、新しい光産業の創成に寄与することが期待できる。

Abstract

Development of AI Function to Reduce Missed in Appearance Inspection and Practical implementation of New Technology in Companies

This study discusses the development of a new appearance inspection method using AI and optical technology and the implementation of the new technology in the organization, which was undertaken by a manufacturer of automotive clutch components to enhance competitiveness and continue future growth in response to the rapidly changing situation in the automotive industry.

The automotive industry is undergoing a period of significant change, and automotive component manufacturers are pushing to utilize new digital technologies to respond to the rapidly changing and continue future growth. One of the efforts to utilize digital technology is the automation of human sensory tasks, and the use of AI (artificial intelligence) is expected as new technology. In this study, an automated system for appearance inspection, which has been difficult to automate with conventional technology, is realized using AI and optical technology. Moreover, the creation of an organizational structure to promote the use of new digital technology is also discussed.

The appearance inspection of finished products is important for quality assurance. However, visual inspection, which relies on the human senses, has been indispensable. Therefore, the development of an autonomous appearance inspection system is desired. It is difficult to define rules with rule-based image inspection because appearance defects are diverse and it is difficult to define the criteria for their determination. In contrast, AI can define complex rules through machine learning by providing data on normal and defective products. At first, the Convolutional Neural Networks (CNN), an AI-based image recognition technology, is adopted to automate appearance inspections. In an experiment on appearance defect classification, a 90% correct answer rate was achieved. The possibility of automating appearance inspections was confirmed. However, the following issues were identified: 1) the correct answer rate is not stable during training, 2) it is difficult to reduce both missed and over-detection at the same time, and 3) the low quality of the input data affects classification performance. The reason for the lack of stability was thought to be that the criteria became too complex when trying to discriminate between normal products and multiple defect types at the same time. Therefore, we replace the task of multiclass classification with a binary problem of classifying each defect species to simplify it. (Two-class Ensemble Method). The developed method achieved a 0.25% improvement in the rate of correct answers in the two-class classification of normal and defective products compared to the multi-class classification method. To confirm its practicality in a mass production environment, we conducted a multi-class classification experiment using a data set that simulates actual mass production. As a result, we succeeded in obtaining a classification performance of 95.69% correct, 1.62% missing rate, and 1.87% over-detection rate, and confirm the feasibility of automation in a mass production environment. Next, we worked on missing suppression. The most important function of appearance inspection is to minimize the occurrence of missed items. In this study, we focused on the difference in the detection process between normal and defective products in human visual inspections to suppress missed products. The finding of normal products is a passive process that detects only when none of the defects are applicable, whereas that of defective products is an

active process that detects when any of the defects may be applicable. This difference in process is defined as asymmetry in appearance inspection. By learning the asymmetry in the CNN model, we have developed to both reduce the number of defects missed and over-detections (Asymmetric Label Smoothing Method). In our experiments, we were able to reduce the missed rate from 1.62% with the conventional CNN method to 0.08%. The proposed method successfully induces the character of the CNN model by devising a method of providing the correct answer given to the AI. Finally, we worked to improve the quality of input data. When capturing input images for learning, photonic technology can be applied to enhance defect areas and reduce noise. After trying several techniques, including illumination and the use of filters, we found that the use of polarizing filters was very effective in the appearance inspection of clutch disks. (Improving Data Quality with Photonic Technology) As a result, it became possible to obtain images with less noise, confirming improved performance. These developed techniques are expected to be effective in improving the performance of defect detection and reducing the number of missed defects in actual mass production processes.

Following the development of these technologies, we summarized the issues to introduce the developed methods to the actual production sites. In general, the issues in implementing AI technology are "lack of understanding of AI" and "uncertainty about the effectiveness of AI implementation". In this study, as practical activities to introduce AI, 1) training sessions were held for production technology engineers, and 2) AI systems were produced in-house. The training sessions incorporated active learning, in which participants built actual systems while working with their own hands, to improve their understanding of AI, and the results were successful. The in-house production of the system has created an environment in which demonstration experiments can be conducted. This has enabled us to see AI in action, gain a perspective on the effects of implementation, and establish a system to eliminate anxiety about the effects of implementation. These activities led to the introduction of AI to the company, including the acquisition of several AI implementation projects. The organizational structure-building activities in this study served as a practical example of how to create an organization that facilitates the introduction of new technologies.

In this study, we developed an image inspection system for the appearance inspection of finished clutch disks and implemented an organizational structure to promote the introduction of this technology. The technology developed in this study is expected to be utilized not only in the manufacturing industry but also in various other fields as a "photonics×AI" technology that replaces human vision and achieves performance beyond that of humans. This technology is expected to contribute to the creation of a new photonics industry.

目次

第 1 章 序論	1
1.1 本研究の背景	1
1.1.1 新技術開発の必要性	1
1.1.2 人作業自動化の新しい技術	5
1.1.3 研究の対象「AI を用いた完成品外観検査の自動化」	6
1.1.4 新技術導入体質の改善必要性	10
1.2 本研究の目的	12
1.3 本論文の構成	12
第 2 章 AI を用いた外観欠陥の分類手法	14
2.1 画像外観検査装置システムの課題	16
2.2 AI の概要と画像認識分野の進歩	17
2.2.1 AI の概要(1)	17
2.2.2 AI の概要(2)	20
2.2.3 画像認識における AI 技術の進歩	22
2.3 欠陥分類実験 1：従来手法の適用（実験環境）	27
2.3.1 画像取得方法	28
2.3.2 データセット	29
2.3.4 実験結果	34
2.3.5 実用的な CNN 完成品外観検査装置の開発に向けた課題	35
2.4 欠陥分類実験 2：開発手法 1 2 クラスアンサンブル手法	37
2.4.1 2 クラス分類アンサンブル手法	38
2.4.2 データセット	39
2.4.3 実験結果	40
2.4.4 2 クラスアンサンブル手法のまとめ	44
2.5 欠陥分類実験 3：量産環境での試験設備の導入	45
2.5.1 画像取得方法	45
2.5.2 データセット	46
2.5.3 学習モデル	49
2.5.4 実験結果	52
2.6 本章のまとめ	55

第 3 章 見逃し抑制に焦点を当てた新手法の開発	57
3.1 見逃し抑制に焦点を当てた新手法の開発	58
3.1.1 新手法開発の仮説	58
3.1.2 正解ラベルの付与方法による CNN モデルの性格誘導	59
3.1.3 検査工程における非対称性の導入	61
3.1.4 非対称ラベルスムージング手法	61
3.1.5 非対称ラベルスムージング手法の解釈	63
3.2 欠陥分類実験 4：開発手法 2 非対称ラベルスムージング手法	70
3.2.1 データセット	70
3.2.2 学習モデル	70
3.2.3 実験結果	70
3.2.4 考察	77
3.3 本章のまとめ	78
第 4 章 撮像方法変更による CNN 分類性能向上の取り組み	80
4.1 欠陥分類実験 5：開発手法 3 画像データの質向上による分類性能向上手法	81
4.1.1 CNN におけるデータの重要性	81
4.1.2 撮像方法	82
4.1.3 データセット	86
4.1.4 学習モデル	86
4.1.5 実験結果	86
4.2 偏光撮像による画像の特徴	86
4.3 開発手法との組み合わせによる精度向上の効果	89
4.3.1 データセットと学習モデル	89
4.3.2 実験結果	89
4.4 本章のまとめ	91
第 5 章 組織内での新技術導入活動	92
5.1 新技術導入基盤の構築	94
5.1.1 研究方法	94
5.1.2 講習の概要	95
5.1.3 アンケート調査	96
5.1.4 調査結果と考察	97
5.1.5 新技術導入基盤構築活動のまとめ	100

5.2 新技術導入に関する導入障壁へのアプローチ	100
5.2.1 システム内製化の必要性	101
5.2.2 システム内製化の手段	103
5.2.3 システム内製化のまとめ	106
5.3 新技術導入活動実践の成果と実用化に向けた取り組み	107
5.3.1 新技術導入活動実践の成果	107
5.3.2 開発した技術の実用化に向けたロードマップ	108
5.4 本章のまとめ	109
第6章 結論	110
参考文献	113
謝辞	117
業績目録	118

第1章 序論

1.1 本研究の背景

1.1.1 新技術開発の必要性

世界第3位となっている日本のGDPのうち、約2割は製造業が占めている（図1.1）。製造業は日本の経済を支える基幹産業であり、その成長が期待されている。製造業においては、世界レベルでの生産性、品質、性能、コスト及び供給力をめぐる競争は益々激しくなっており、特にドイツのIndustrie4.0、中国の中国製造2025、日本のConnected Industriesに代表される産業革新への取り組みが進んでいる。一方で、貿易摩擦や度重なる災害、天候不順などによって、世界経済での不確実性が高まっており[2]、製造業にとっては今後の見通しが立てづらい状況となっている。そのような中、企業はさらなる競争優位性を見出し、製品を安定的に提供するとともに、環境の変化に柔軟に対応できる体質であることが求められている。

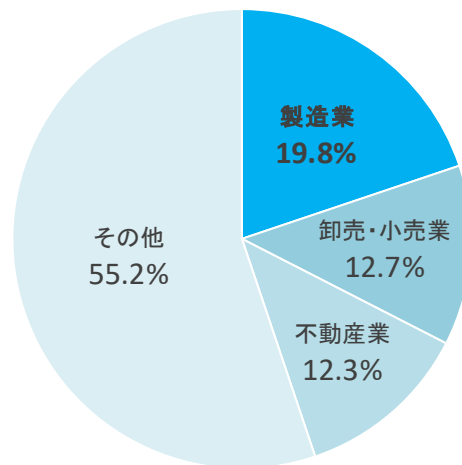


図 1.1 日本の業種別 GDP 構成比(2020 年)
(出典：内閣府 2020 年度国民経済計算(3) 経済活動別国内総生産 [1] をもとに作成)

特に自動車業界においては MaaS (Mobility as a Service)、CASE (Connected : コネクテッド、Autonomous : 自動運転、Shared & Service : シェアリング/サービス、Electric : 電動化) といったキーワードが示すように自動車そのものや社会的役割が様変わりしてきている。自動車業界は急速に進む技術革新にも後押しされ、大きな変革の時代となっており、生き残りのための競争力と変革力が求められている。

筆者が所属する企業は自動車用クラッチを主力製品とした自動車部品メーカーであり、製造業および自動車業界全体の変化の影響を大きく受ける。自動車業界では電気自動車や燃料電池車といった次世代自動車へのシフトが進んでいる（図 1.2）が、これらの次世代自動車はトランスミッション

部品を持たないことが多いため、自動車用クラッチを含むトランスミッション分野は、将来的な市場の縮小が予測されている。自動車部品メーカーは取り巻く情勢の変化に対応し、競争力を維持・強化し、事業の多角化を図り、将来的な成長を継続するための取り組みを早急に進めていく必要がある。



図 1.2 次世代自動車の割合
(出典：日本の自動車工業 2020 [3])

このような情勢に対応し、成長を継続するため、デジタルトランスフォーメーション (Digital Transformation 以下、DX) への取り組みの重要性が高まってきており[4]、製造業をはじめ、多くの企業が取り組みを始めている (図 1.3)。DX とは、経済産業省によって発表された「DX レポート」などを通して広く認知されるようになった概念で、経済産業省の発表した「DX 推進指標」[69] (2019) では「企業がビジネス環境の激しい変化に対応し、データとデジタル技術を活用して、顧客や社会のニーズを基に、製品やサービス、ビジネスモデルを変革するとともに、業務そのものや、組織、プロセス、企業文化・風土を変革し、競争上の優位性を確立すること」とされている。経済産業省が 2018 年に発表した DX レポート[4]では、今後 DX の推進が進まない場合、2025 年には年間最大 12 兆円の経済損失が見込まれる「2025 年の崖」が訪れるとされており、製造業に限らない多くの国内企業は DX の推進に追われている。筆者が所属している企業の生産技術部においても、競争力の維持・強化のために高効率生産ライン、製品価値の向上、プロダクト変革を生み出すことが求められており、DX への取り組みを進めている。

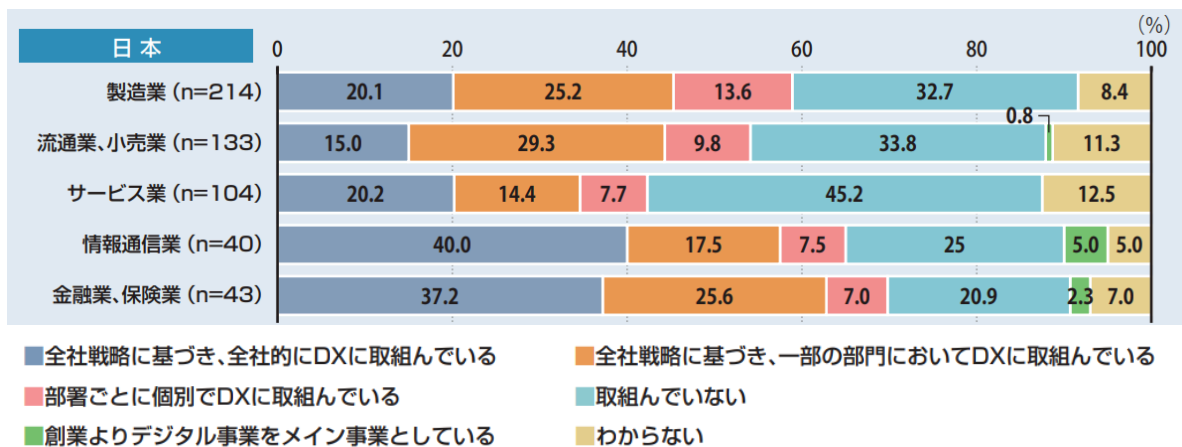


図 1.3 業種ごとの DX への取組状況（業種別）
 （出典：DX 白書 2021 [5] より作成）

DX は企業が具体的なアクションを設計しやすいようにデジタイゼーション、デジタルライゼーション、デジタルトランスフォーメーションの 3 つの異なる段階に分解されている（図 1.4）。データのデジタル化を行うデジタイゼーション、業務・製造プロセスのデジタル化を行うデジタルライゼーションを工場で実現した形としてスマートファクトリーがある。スマートファクトリーとは製造現場の様々な装置に取り付けたセンサなどからデータを収集・分析することで、製造現場を最適にコントロールし、製品価値、生産性、品質の向上を継続的に発展させる工場のことである。スマートファクトリーを構築するためには、これまでの改善活動に加えて、デジタル技術を活用したプロダクト変革、省人化、予測分析などを推し進める必要がある（図 1.5）。自動車部品メーカーをはじめとしたすべての製造業にとって、スマートファクトリーを構築することは、情勢の変化に対応し、競争力を維持・強化し、事業の多角化を図り、将来的な成長を継続するための取り組みとして急務となっており、その取り組みが進められている。

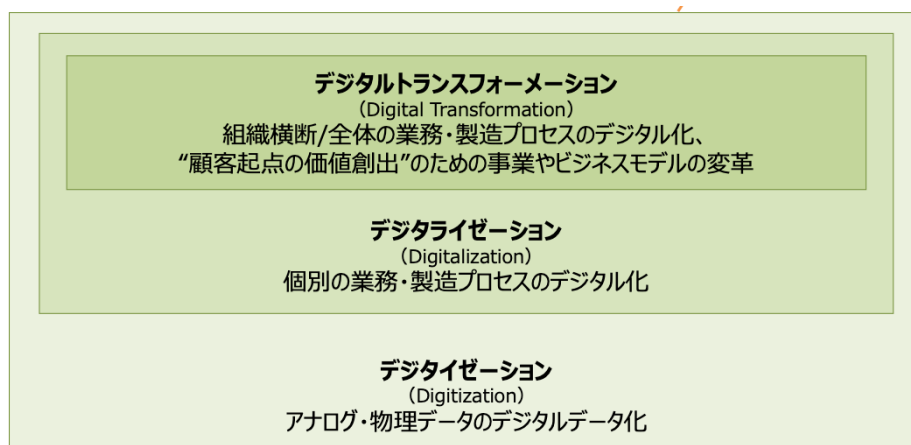


図 1.4 DX における 3 つの段階
 （出典：DX レポート 2 中間とりまとめ [6]）

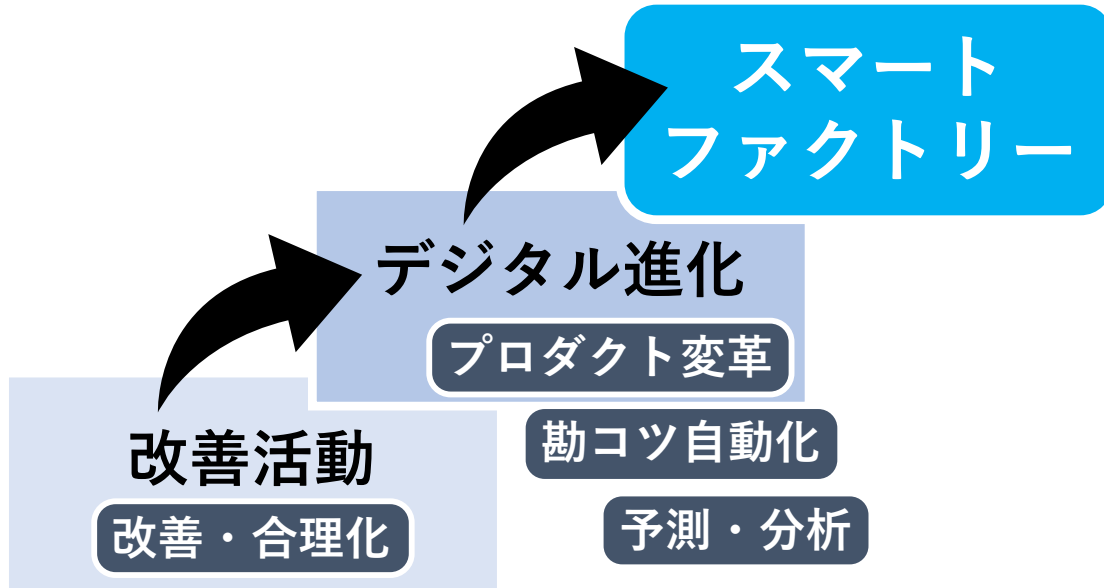


図 1.5 スマートファクトリーの概要

スマートファクトリー構築を進めていくための課題のひとつに人作業の自動化がある。人に頼った作業からはデータをリアルタイムで正確に収集することが難しい。人作業を自動化し、作業に関連するデータが収集可能になれば、データを分析し、製造現場を最適にコントロールし、新たな付加価値創造につなげることができる。また、人作業は人件費によるコスト圧迫が避けられない上、能力の個人差が大きいこと、教育・訓練が難しいこと、作業者の確保・増員が困難であることが課題となっており、多くの企業が抱える共通の問題となっている（図 1.6）。

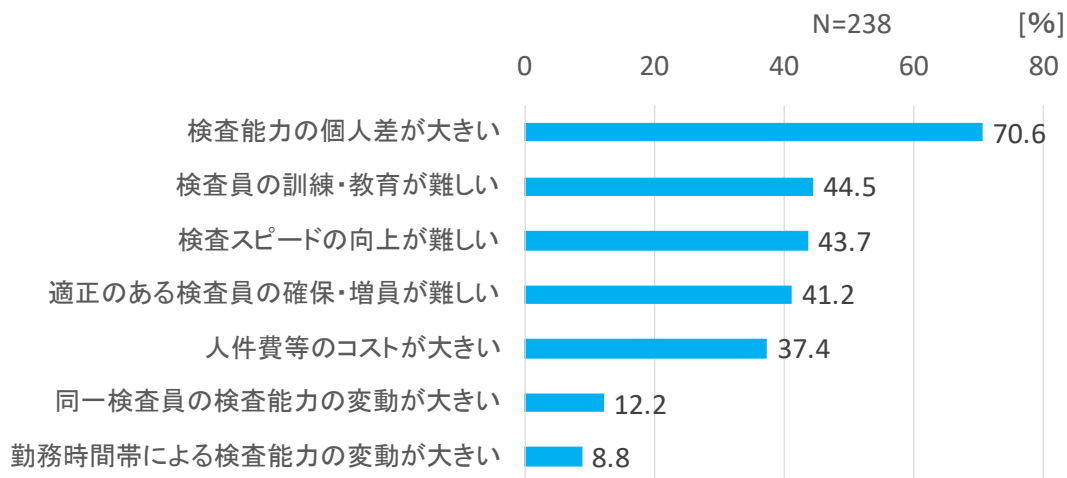


図 1.6 人作業の困難点（検査の例）
 （出典：周辺視目視検査法の理解と導入のためのヒント [7] をもとに作成）

製造業の生産現場ではセンサやロボットなどを活用することにより作業を自動化するファクトリーオートメーション（Factory Automation 以下 FA）化が進められ、生産効率アップやコストダウンを実現してきた。しかし、未だに生産ラインには人間の介入が必要な作業が存在している。これらの作業は、従来技術での自動化が難しく、人に頼らざるを得ない。人に頼った作業の自動化はスマートファクトリー構築、即効性のある生産効率改善、品質安定化、コスト削減のためにも、急務となっており、自動化を実現する新しい技術の探索や開発が必要である。

1.1.2 人作業自動化の新しい技術

従来の技術では自動化が難しい作業の代表的なものに、人の視覚や嗅覚、聴覚などの感覚から得られた情報を処理し、判断する作業がある。これらの作業は情報の入力を人の感覚に頼っている上、その判断においても人の勘やコツに頼っていることから「勘コツ作業」と呼ばれ、自動化困難な作業とされている。中でも、視覚を用いた作業の自動化は、例えば検査工程において、強く望まれている（図 1.7）。

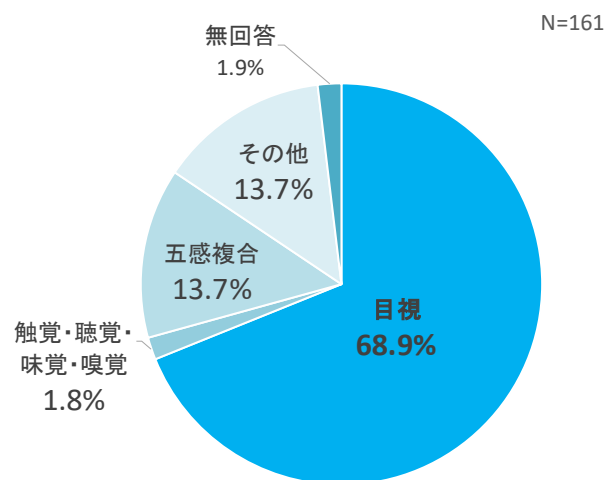


図 1.7 検査工程における自動化したい方法
(出典：ものづくり企業の生産現場における検査の自動化促進可能性調査 [8] をもとに作成)

人の視覚に頼った作業を代替し、自動化するための主要な方法として、光技術である画像認識システムがある。製造業において光技術は、加工、計測、画像診断をはじめとして、広く利用されており、自動化のための重要な技術となっている。画像認識システムは照明、カメラ、画像処理装置などからなり、対象を画像として撮像し、処理などを行うことで、人の視覚に変わる作業を行うものである。画像認識システムにおける撮像技術、画像処理技術、判定アルゴリズムの研究開発は盛んに行われ、画像認識システムは多くの分野、工程で導入されてきた[9]。画像を取得する技術や機器の性能は年々向上し、新しい色分離方法の開発、光量範囲の拡大、空間情報の高精細分離化、高速化、小型化、低価格化が進んでいる[10][11]。更には赤外光、紫外光領域に感度を持ったカメラや専用の照明も充実してきており、これらの波長域を利用した、人間の視覚を拡張し、超越する技術

も身近なものになりつつある。画像認識システムにおける光技術の向上によって画像の高品質化や特徴の顕在化が可能となったことで、製造現場における画像認識システムの精度向上につながっている。しかしながら、2000年代後半からの画像処理システムに関する特許件数は伸び悩んでおり（図 1.8）、技術進歩が減速していることも見て取れる。画像認識システムで人間の視覚を代替し、勘コツに頼った作業を自動化するためには、新しい技術の開発が必要となっている。



図 1.8 画像処理技術の日本特許件数（公開、登録）
（出典：パテント・インテグレーション レポート [12]）

近年、そのような勘コツ作業の自動化を実現しうる新しい技術として人工知能（Artificial Intelligence 以下、AI）が注目されている。AI のなかでも、機械学習、深層学習といった技術を用いることによって、様々な分野で従来の手法では実現し得なかった成果が上げられている。特に、画像認識の分野では、深層学習によって大きな成果が得られている。AI により画像認識の精度は大きく向上し、従来手法の性能を大きく凌駕し、人間の認識率を上回る性能も報告されており[34][54]、製造業においてもその活用が期待されている。

「光技術」である光計測・センシング技術や画像取得技術と、大きく技術進歩している「AI」を用いた画像認識技術を掛け合わせた新たな画像検査技術を開発することにより、スマートファクトリー構築の障害となっている人の視覚に頼った勘コツ作業の自動化にむけた大きなステップとなることが期待できる。

1.1.3 研究の対象「AI を用いた完成品外観検査の自動化」

人の視覚に頼った勘コツ作業の例として完成品の外観検査が挙げられる。現在、完成品外観検査の多くは人による目視検査が行われている。人による目視検査での外観検査は、導入に関するイニシャルコストを低く抑えることができる、設備の保守が必要ない、少量多品種への適応が容易であるなどの利点がある。一方で、データ活用、品質保証、労働環境の面で課題があり、自動化が急務となっている。

人作業であることによって特に問題視されるのは、作業の情報が取得されず、データの活用ができないため、スマートファクトリー化が進まないことである。人作業では外観に現れる欠陥などから、その場で設備状態を把握し、対処するなど、高度で柔軟な判断を行うことができるなどの利点がある一方、その情報は作業者の頭の中で処理され、他者と共有されることはない。また情報が客観的なデータとして保存されていないため、後に作業時の状況と判断を正確に紐付けて分析する事ができず、分析による効率化につなげるなどの情報活用ができない。その上、判断プロセスを共有できていないため、作業が属人化してしまうという問題もある。完成品はその製造ラインの状態を写す鏡であり、完成品検査から得られるデータは生産効率最適化、予防保全への活用が期待できることから、正確なデータをリアルタイムで収集・分析することが強く求められている工程でもある（図 1.9）。完成品の外観検査が人作業化していることは良質なデータを活用することができず、データの重要性が増す現代においては、自社の宝をみすみす捨てているようなものである。

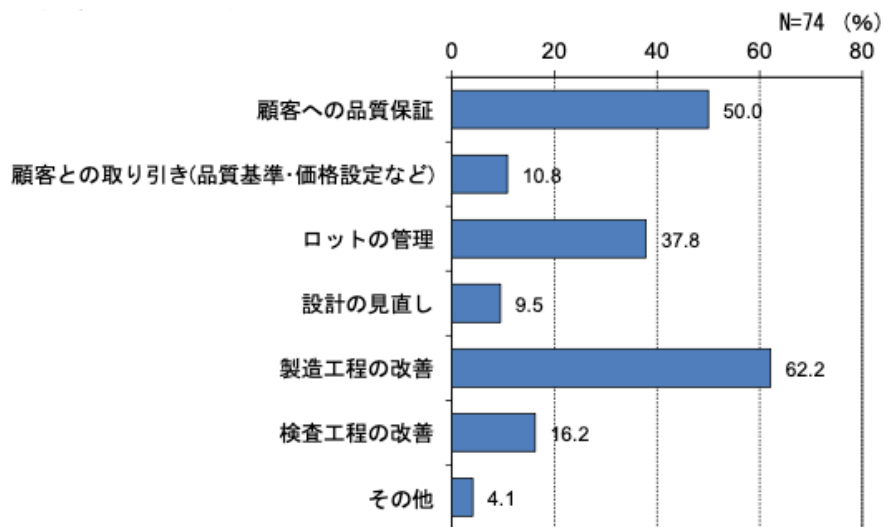


図 1.9 検査工程の数値データ活用方法
 (出典：ものづくり企業の生産現場における検査の自動化促進可能性調査 [8])

また、人による外観検査は検査の重要な機能である品質保証能力にも問題を抱えており、改善が望まれている。製造業の現場では、品質保証のために完成品を検査し、欠陥のある製品を取り除く必要がある。特に外観に関わる欠陥は、見た目でも認識できるために、機能性能への影響に関わらず、見た目品質として製品品質保証上、大きな意味を持つ。しかし、人による検査は、人間の感覚で判定を行うため、ヒューマンエラーが避けられず、検査員による判定の基準にばらつきが生じてしまう。さらに、これらはその日の検査員の体調、精神的疲れ、検査台・検査室の照明といった環境によっても影響を受ける。これに対し、実際の現場では環境の安定化、ツールの充実、教育の徹底などを行っているが（図 1.10）、これらヒューマンエラーや判定基準のばらつきを完全になくすことは難しく、品質面での問題が避けられない。

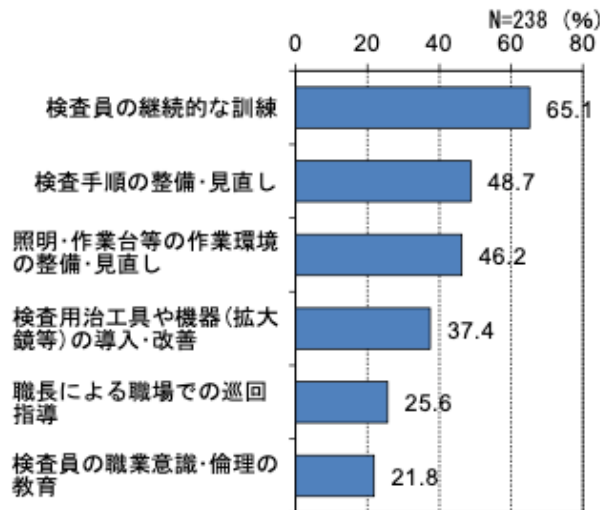


図 1.10 人による検査の改善への取り組み
 (出典：ものづくり企業の生産現場における検査の自動化促進可能性調査 [8])

さらに、完成品の外観目視検査はその労働環境にも問題を抱えている。製造業の生産ラインでの外観不良発生率は、生産技術部の技術投入、現場エンジニアのメンテナンス、現場作業員の適切なオペレーションにより、極限まで低くなっている。高効率を極めた製造ラインでの日々の生産数は何万にも達するが、その内、外観欠陥のある製品は1/100,000を下回ることもある。人による検査ではコスト効率が非常に悪いうえ、検査員には長時間にわたって欠陥を見逃さないための高い集中力と、まれに発生する欠陥を発見しなければならないという精神的な緊張を強いることとなる。

目視検査は、自動化に比べて利点もあるが、上述のデータ活用、品質保証能力、労働環境の課題に加えて、人材の確保、人的コストの負担を強いられるため、その自動化が急務となっている。しかしながら、製造の現場では外観検査の自動化は容易ではなく、検査員による目視検査に頼っている、もしくは自動検査と目視検査を併用している工程が少なくない。

筆者が所属する企業で行われているクラッチディスクの外観検査はこのような外観検査の自動化が難しく、人に頼らざるを得ない検査の一例である。クラッチディスクとは自動車用クラッチを構成する部品の一つで、鉄製の芯板にペーパー摩擦材と呼ばれる複合材料が貼り付けられたものである(図 1.11)。クラッチはその用途によって多様に分類され(図 1.12)、それぞれ形状や要求性能が異なる。クラッチディスクはその主要部品であり、特にペーパー摩擦材はクラッチディスクの性能を大きく左右する重要部品となっている。そのため、ペーパー摩擦材に欠陥が発生した場合は、それらを取り除く必要がある。ペーパー摩擦材に発生する欠陥は、傷による毛羽立ち、打痕による凹みや変形、めくれによる盛り上がり、浮きなど種類が多様であるうえ、欠陥の発生場所、特徴、大きさは様々である。外観に現れる欠陥には機能性に影響するもの、外観上の美観を損なうものなどが含まれる。また、欠陥種によっては、欠陥の有無だけでは判断されず、程度によってその扱いが異なる場合もある。一般的な画像検査装置を用いて、これらの欠陥を個別に検出する場合の難易度はそれぞれ異なるが、すべての欠陥について個別の設定が必要であるのに加えて、一つの欠陥をと

っても、多様な欠陥に対して人手でルールを定めることには多大な労力を必要とする。そのため、自動化が現実的でなく、目視検査に頼っている状況となっており、検査の精度向上が望まれている部品のひとつとなっている。

クラッチディスクをはじめとした製造部品の多くは外観検査を必要とし、その自動化が求められている。外観検査が自動化できていないことによる課題は以下のようにまとめることができる。

- データ収集：正確なデータをリアルタイムで収集・分析することができないため、生産効率最適化、予防保全へのデータ活用が進まない。
- 品質保証能力：ヒューマンエラーが避けられない、作業員や環境により判定の基準にばらつきが生じてしまう。
- 労働環境：長時間にわたって欠陥を見逃さないための高い集中力と、まれに発生する欠陥を発見しなければならないという精神的な緊張を強いることとなる。

外観検査の自動化は、中期的に進めていく必要のあるスマートファクトリー構築のためのみならず、短期的な品質保証能力確保、人的コスト削減、労働環境改善という側面からも強く望まれている。外観検査は従来技術では自動化が難しかったが、技術進歩の目覚ましい光技術と AI を用いることによって自動化実現の可能性がある。外観検査の自動化技術は、クラッチディスクに限らない多くの製造部品における課題を解決し、競争優位性を見出し、製品の安定的提供を実現し、環境変化に対応できる体質を構築する糸口となることが期待できる。



図 1.11 自動車用クラッチディスクの外観
(出典：株式会社 エフ・シー・シー HP)
<https://www.fcc-net.co.jp/tech/segment-disk/>

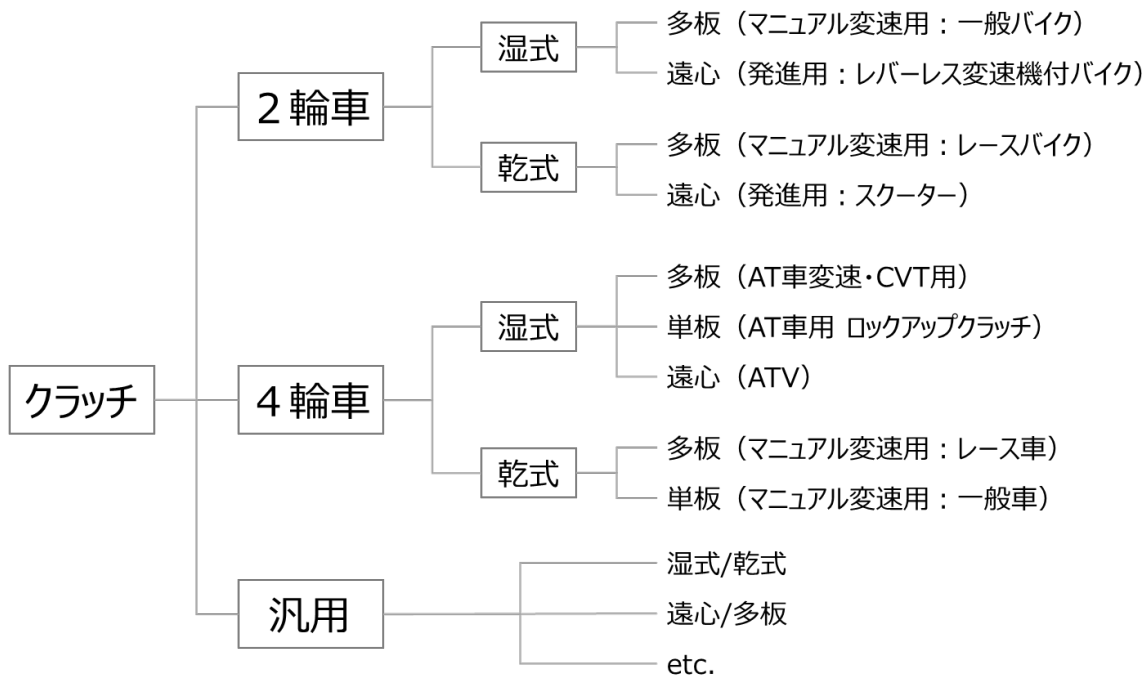


図 1.12 クラッチの種類
 (出典：株式会社 エフ・シー・シー HP を参考に作成)
<https://www.fcc-net.co.jp/clutch/allkind.html>

1.1.4 新技術導入体質の改善必要性

DX、スマートファクトリーのキーテクノロジーであり、その活用への期待が高まる AI は急速に導入が進んでいるが、未だに導入に至っていない企業は少なくない。AI 市場は高い成長率を維持し、大きく発展すると予測されており、企業にとっての重要なツールとして引き続き位置付けられ、高い関心が持たれていることは疑う余地がない。しかし、AI 導入に関心を持つ企業は多い一方で、導入や実証実験は進んでいない。導入検討にあたっての課題は、情報処理推進機構 (Information-technology Promotion Agency 以下、IPA) によるアンケート結果によれば、自社内に AI についての知識、人材が不足していることや導入効果への不安、高額な導入費用などとなっている (図 1.13)。2020 年調査と 2021 年調査を比較してみると、「自社内に AI についての理解が不足している」の割合が減少したのに対し、「AI 人材が不足している」の割合が増加している。また、「導入効果が得られるか不安である」、「導入費用が高い」の割合が減少しているのに対し、「AI の導入事例が不足している」、「利用できそうな業務がない」の割合は増加している。これらの結果は、2020 年に AI 導入検討初期の漠然とした検討段階にあった企業が、web の情報、セミナー、書籍、各種サービスの普及などを通して、AI の実態を把握できたことにより、より具体的な導入検討のフェーズに入った結果と考えることができる。しかしながら、AI についての理解不足、導入効果への不安が完全に払拭されているわけではなく、AI についての理解促進および導入効果の不安解消が大きな取り組み課題であることに変わりはない。AI を導入して DX の推進を加速させるためには、その技術開発のみならず、企業内で AI 導入に向けた課題への取り組みが必要となっている。

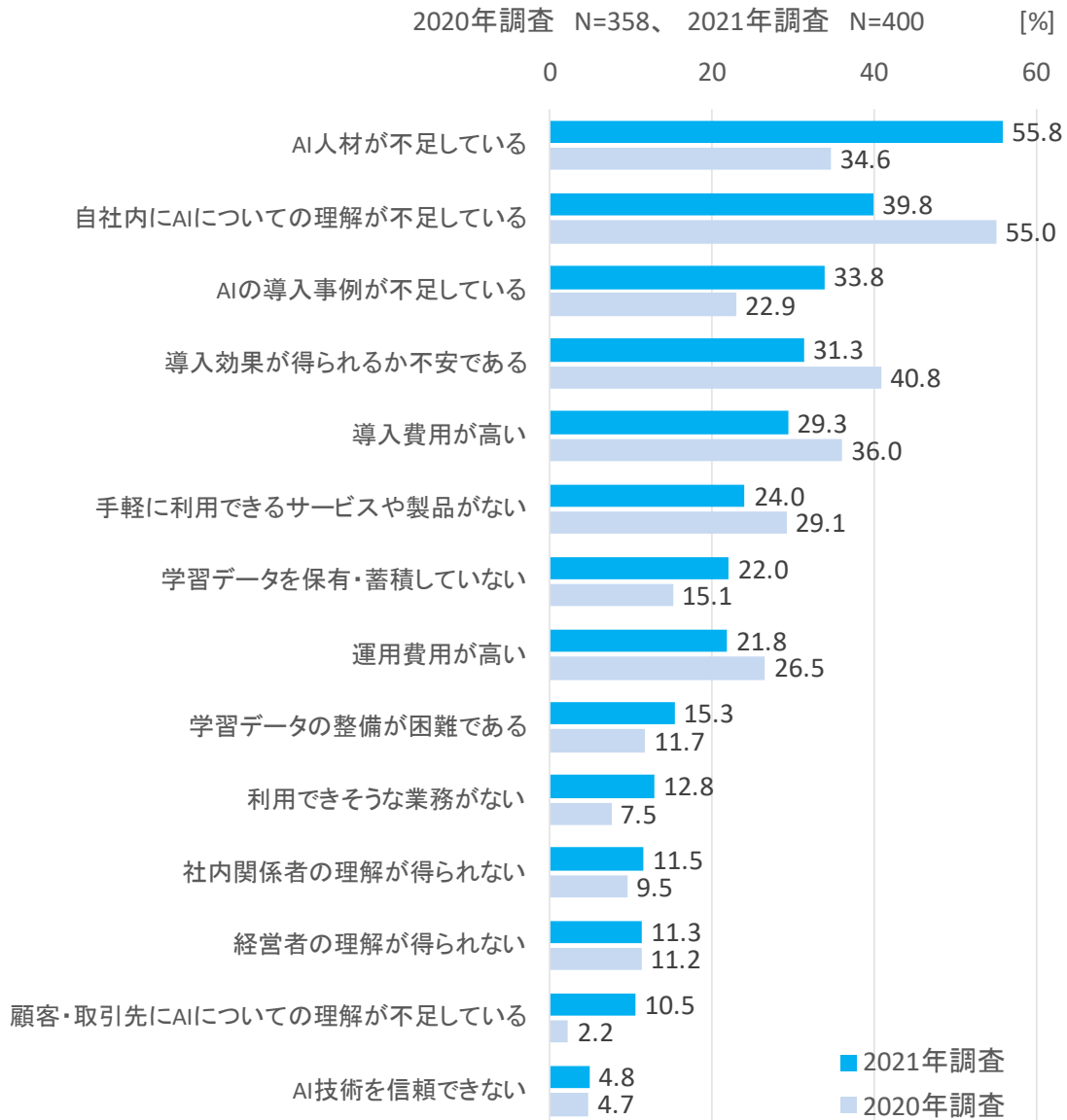


図 1.13 AI 導入検討するにあたっての課題
 (引用：AI 白書 2020 [13] p306 図 3-4-21 および DX 白書 2021 [5]p283 図 42-53 をもとに作成)

1.2 本研究の目的

本研究は以下の2つを研究目的とする。

- ① 従来技術では自動化が難しかった人に頼った作業の自動化を実現するため、光技術と AI を用いた新しい技術の開発に取り組む。その一例として、クラッチディスクの完成品外観検査に取り組み、勘コツ作業である外観検査の自動化システムを構築する。
- ② DX、スマートファクトリーの構築を加速するために、新しい技術を受け入れる組織体質造りを行う。AIにおける導入の課題を解消する活動を行い、開発したAI技術の導入事例を創出する。

生産技術部として DX を推し進めスマートファクトリー構築を実現するため、光技術と AI をかけた新しい技術を開発するとともに、新しい技術の導入基盤を構築することで、人に頼っていた作業を自動化することを目指す。企業はさらなる競争優位性を見出し、製品を安定的に提供するとともに、情勢の変化に柔軟に対応可能な体質となることで、継続的に成長することができる。

1.3 本論文の構成

第 1 章では、本研究の背景として製造業および自動車業界が置かれている状況とスマートファクトリー構築の必要性を述べ、キーテクノロジーとして「光技術×AI」を提案し、研究の対象を完成品の外観検査に定め、目的をクラッチディスクの完成品外観検査の自動化と新技術を導入するための組織作りと定めた。

第 2 章では研究の対象である完成品外観検査に関する概要と筆者のケースでの課題を整理し、課題解決のツールである AI に関する概要と AI による画像検査の可能性について述べたあと、実際に AI を用いた外観検査実験の結果をまとめる。

第 3 章では AI 外観検査実験で明らかになった課題を解決するために開発した新しい AI アーキテクチャである非対称ラベルスムージング手法について、着想と解釈、アーキテクチャの詳細、実験概要と結果、その有用性について述べる。

第 4 章では光技術側からのアプローチとして各種照明方法による AI 特性の選択性について、各種欠陥に対して光技術を用いた実験結果から評価する。

第 5 章では変わりゆく状況の中で求められる、新技術の導入を受け入れる組織作り活動の実践内容について述べ、第 6 章で本論をまとめる (図 1.14)。



図 1.14 本論文の構成

第2章 AIを用いた外観欠陥の分類手法

製造業における製品品質とは機能性能、耐久性、安全性、環境配慮など様々な品質を含んでいる。製造業における競争激化の中にあつては、機能性能、コストはもとより、品質をめぐる競争も激しくなっており（図 2.1）、品質は重要な経営課題と認識されている。これらの製品品質保証のためには、よりよい製品設計を行い、適切な生産状態を維持管理するとともに、製品を検査し、欠陥のある製品を取り除く必要がある。検査は目的や場所、性質などで多種多様に区分され（表 2.1）、その検査項目には数量検査、寸法検査、重量検査などがあるが、重要な検査項目の一つに外観検査がある。外観に現れる欠陥は機能性能、耐久性、安全性に影響する場合がある。また、影響がない程度であっても、使用者が見た目で欠陥の存在を認識できるため、製品の良し悪しの判断に直結する。そのため、外観検査は企業の品質管理への信頼を確保するためにも非常に重要な検査項目となっている。製品品質をめぐる競争激化の中にあつて、製造ラインでの外観検査は、ますますその重要性を増している。

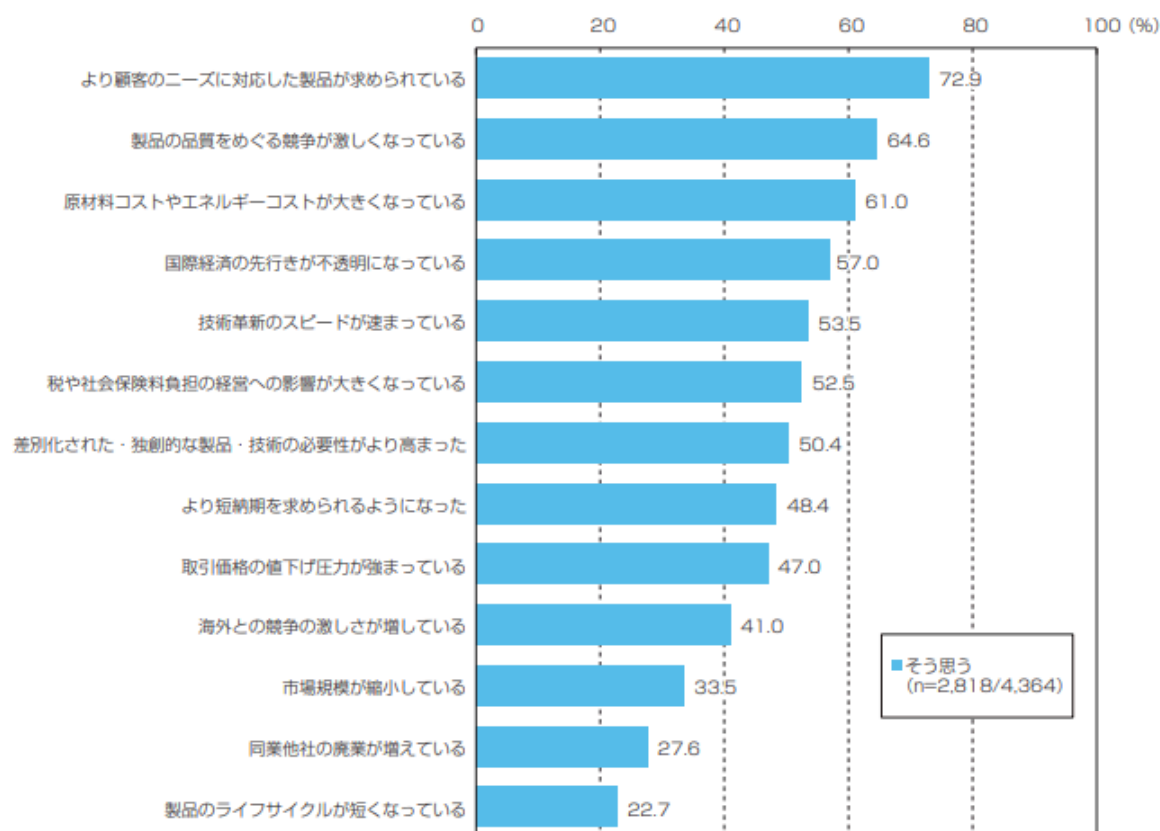


図 2.1 事業環境・市場環境の状況認識
(引用：2020年版ものづくり白書 [2] p139 図 221-3)

表 2.1 検査の種類

製造プロセスによる分類	検査の性質による分類
受入検査	破壊検査
工程内検査	非破壊検査
完成品検査	耐久検査
出荷検査	官能検査
検査頻度による分類	検査項目による分類
全数検査	数量検査
抜き取り検査	寸法検査
間接検査	重量検査
無試験検査	性能検査
検査場所による分類	外観検査
定位置検査	
巡回検査	
持ち込み検査	

本研究の対象となっているクラッチディスクの完成品外観検査はクラッチディスク生産ラインの最終工程で集積されたすべての完成品に対して行われている。実際のクラッチディスク生産ラインでは、最終工程前にひずみ検査機によるクラッチディスクの厚みや反りの検査、摩擦材位置検査装置（図 2.2）による摩擦材の位置や寸法の検査が自動で行われている。しかし、外観に関わる欠陥はこれらの検査装置では検出できないため、検査員による目視検査が行われている。



図 2.2 摩擦材位置検査機の外観

2.1 画像外観検査装置システムの課題

筆者の所属企業でも FA 化を進めるうえで、ルールベース画像検査システムを用いた目視検査の自動化への取り組みを続けてきた。しかしながら、必ずしも目視検査の自動化は容易ではなく、現在まで、外観に関わる勘コツ項目を自動で検査することは完全には達成されていない。

ルールベース手法では、人（生産技術エンジニア）が画像データを基に前処理と特徴量抽出を行い、判定アルゴリズムを作成することで、対象画像に対する判定結果を得ている（図 2.3）。画像検査システムで言えば、画素数、輝度、位置、色調などの数値を特徴量として判定アルゴリズムを作成する。検査対象の範囲や特徴量を数値として、「右上の 20 画素×20 画素の範囲で輝度 100 を超える画素が 10 画素以上あった場合は欠陥と判定する」のような定義ができる。そのため、人が設定したルールに該当すれば、100%正しい判定を行うことができるうえ、判定根拠が明確であるという利点がある。また、個別のルールのみについて、判定の閾値を調整することができるため、実際の導入環境に合わせ、その場で閾値を調整するなどの対応が可能である。

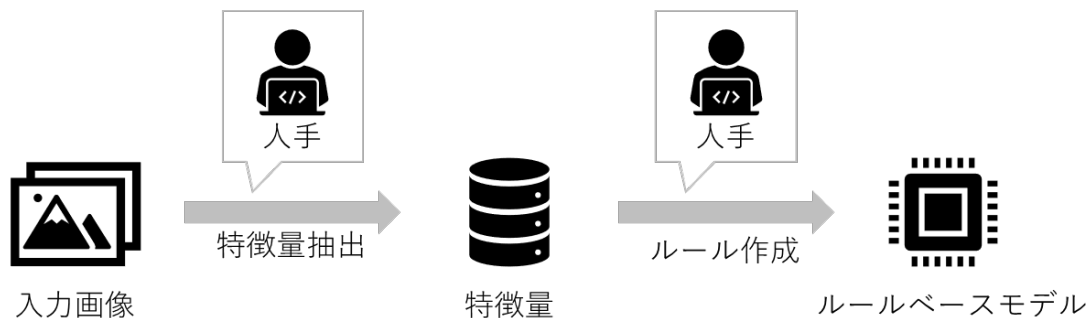


図 2.3 ルールベース手法の画像判定モデル構成手順

一方、欠陥種類が多い場合や、欠陥そのものが多様性を有するため、判断基準が複雑になる場合は、非常に多くのルールを設定する必要があるため、ルールの設定や修正が難しいという課題がある。また、検査装置を現場で運用するにあたっては、製品の出来栄バラツキや照明状態の変化による特徴量の変動が発生する可能性がある。外観検査の項目によっては、現場での特徴量変動が誤判定を引き起こすことも考えられる。欠陥が明確で、ルールが単純な場合は、これらの変動に対してルールの再設定や閾値調整が可能な場合もあるが、変動に対して柔軟に設定を与えることは容易ではなく、変動があった度に作業員による調整が発生してしまうようでは実用性に欠ける。特に、クラッチディスクの外観検査においては、ペーパー摩擦材に発生する外観に関わる欠陥種類が多様であるうえ、欠陥の発生場所、特徴、大きさは様々で、程度によって扱いが異なる場合もある。一般的なルールベース画像検査装置では、すべての欠陥について個別の設定が必要であるのに加えて、一つの欠陥をとっても、多様な欠陥に対して人手でルールを定めることには多大な労力を必要とするため、すべての項目について自動化することが現実的でなく、目視検査に頼っている項目も存在している。

2.2 AI の概要と画像認識分野の進歩

これに対し、新たな外観検査の手法として AI が注目されている。ルールベース手法では人手でルールを設定するのに対して、AI を用いた手法では、AI 自身が画像データを基に、自動的に特徴量抽出と判定アルゴリズムとなる学習モデルを生成し、対象画像に対し判定結果を出力する点で大きく異なる。本節では AI の概要と画像認識分野における AI の技術進化について述べる。

2.2.1 AI の概要(1)

AI の定義は研究者によって異なり、明確な定義は存在しないが、「知的な機械、特に知的なコンピュータプログラムを作る科学と技術」[14][15]、「一般に人間の知性と結び付けて考えられる、推論、学習などの機能を遂行するモデル及びシステムを取り扱う学際分野」[16]などと定義されており、一般的に、さまざまな知覚や知性を人工的に再現するものという意味合いで理解されている。AI という言葉は 1956 年にアメリカのダートマス大学で開催されたダートマス会議で、計算機科学者・認知科学者のジョン・マッカーシー教授によって提案されたのが初めてであると言われている。現在、AI は教育、金融、医療、エンターテインメントなど様々な分野で研究・実用化が進められている。内閣府が提唱する超スマート社会実現に向けた取り組みである Society5.0[17]においても、AI は競争力向上と基盤技術の戦略的強化における基盤技術に位置づけられており、技術の強化が進められている。

近年の AI 技術進歩は目覚ましいが、AI はその言葉が生まれた 1950 年代後半から 2022 年現在まで、ブームと冬の時代（停滞期）を繰り返して、徐々に進化してきた（図 2.4）。1950 年代後半から 1960 年代にかけて、第一次 AI ブームが始まった。第一次ブームでは「推論」と「探索」が中心的に研究され、明確なルールや定義付けがある問題を解くことができるようになった。しかし、現実社会で起こっているさまざまな要因が絡む複雑な問題を解くことはできなかった。そのため、第一次 AI ブームは徐々に下火となり、1970 年代には、冬の時代（停滞期）を迎えた。第二次 AI ブームが到来したのは、1980 年代～90 年代である。「エキスパートシステム」により第一次 AI ブームでは解くことのできなかった現実的な問題の解決も可能になると期待された。エキスパートシステムとは、「人間の専門的知識で構築された知識ベースからの推論によって、特定の領域又は応用分野の問題を解決するように設計されている知識ベースシステム」[16]とされ、あらかじめ専門家が考え得る限りの状況を予測して対処方法や判断を用意し、ルールとして取り込み、推論を行うことで、専門家と同等の問題解決・判断が可能となるシステムである。エキスパートシステムには大きな期待が寄せられたが、すべての問題を正確に解くには課題があった。エキスパートシステムは取り込むルールが多ければ多いほど正確性は向上するが、必要となる情報をすべて人の手でコンピュータに取り込まなければならぬため、多大な労力を必要とする。また、ルールが膨大になるとルール同士での矛盾や一貫性の無さが発生してしまうという課題もあった。こうしたエキスパートシステムの限界から、AI への期待は再度低下し、1995 年頃から AI は再び冬の時代を迎えた。第三次 AI ブームは 2000 年代から始まった。ブームの要因を作ったのは、AI 自身が大量のデータから知識を獲得する「機械学習（マシンラーニング）」の実用化が進んだことである。さらに、2006 年には多層のニ

ニューラルネットワークを用いることによって、従来よりも高精度な特徴抽出が可能になる「深層学習（ディープラーニング）」が提唱され[18]、ブームに拍車をかけた。第3次ブームは2022年現在も続いている。

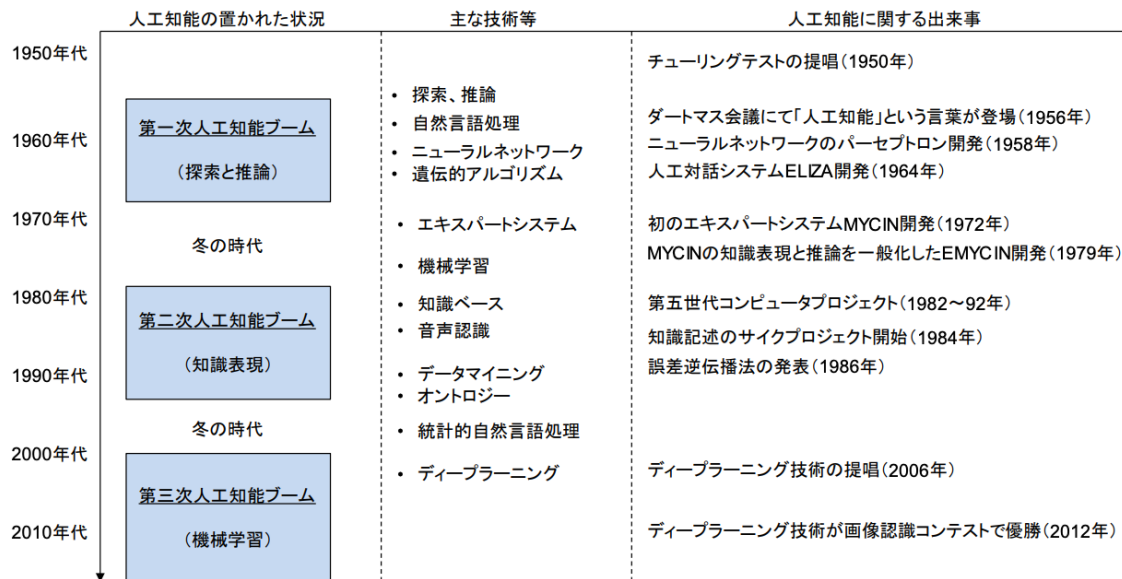
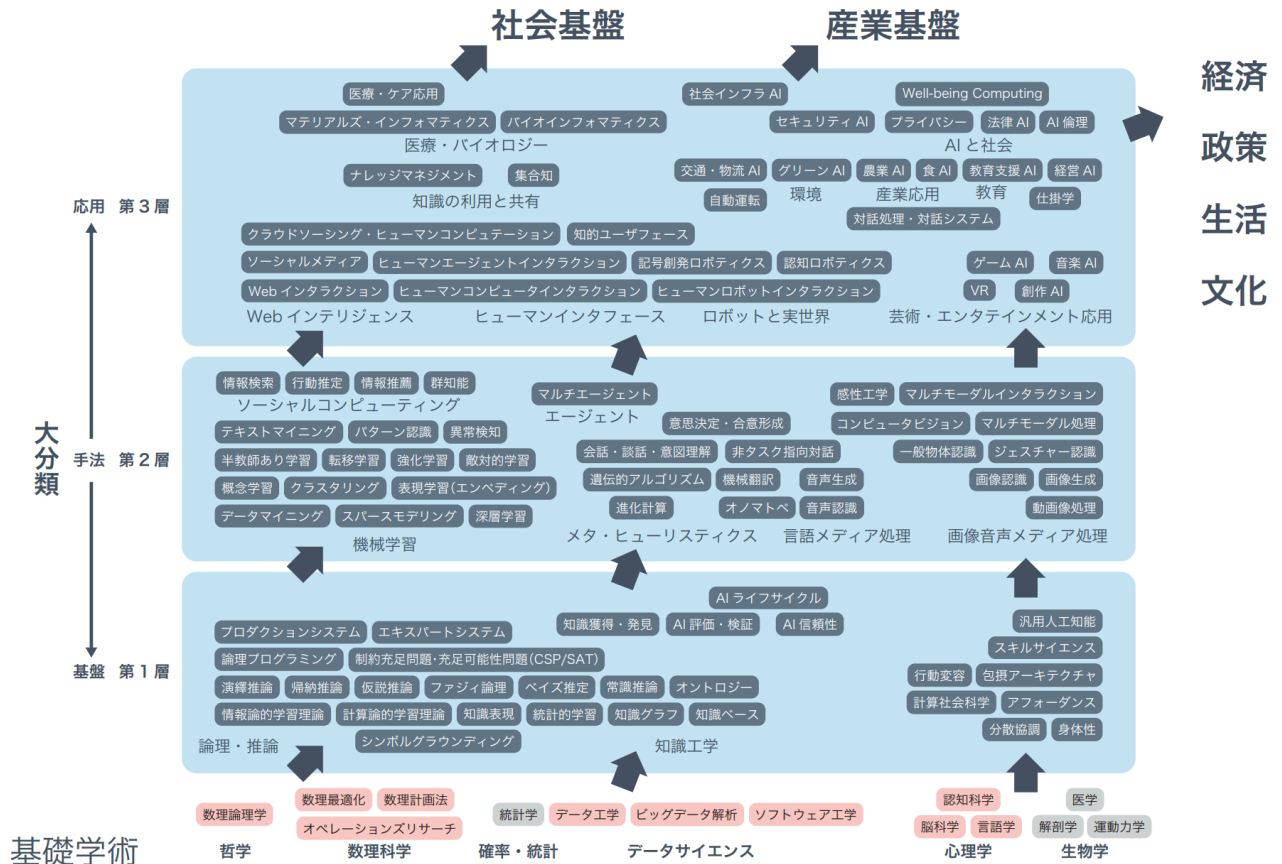


図 2.4 AI の歴史

(引用：ICTの進化が雇用と働き方に及ぼす影響に関する調査研究[19] p15 図 2-1-1-5)

理論自体は1980年代に提唱されていたAIが現代になって注目されている背景には、機械学習や深層学習などのアルゴリズム進化があるが、その他にビッグデータの利用が可能となったこと、計算機の処理能力向上が挙げられる。機械学習の性能を高めるためには大量のデータが必要となるが、1990年代の初めごろから普及してきたインターネット[20]により、世界中から大量のデータを集めることが可能になった。また、計算機の処理能力が飛躍的に向上したことで、大量のデータを高速で処理することが可能になった。特にGPU(Graphic Processing Unit)の性能向上は、機械学習の実用化に大きく貢献している[21]。GPUによって大規模な並列処理(同時に複数の計算処理を行うこと)が可能となり、機械学習の実用化を可能にするレベルに達している。

このような背景から、AI・機械学習・深層学習の応用は、教育、金融、医療、環境、エンターテインメントなど様々な分野で進んでいる(図 2.5)。AIの具体的な機能には、識別、予測、実行があり(図 2.6)、音声、言語、画像、数値、デザインなど様々な要素を扱うことができる。今後、技術の進歩による性能向上や機能の組み合わせによって実用化が加速していくと見込まれている。



※ 解釈の一例：確率・統計、論理学、数理科学に基づき帰納推論が生まれ、機械学習技術が発展、予測の技術により、社会インフラ AI が発展し、インフラ維持や環境保全、経済活動の発展に貢献

図 2.5 基盤領域から手法・応用領域への展開
(出典：AI マップ β 2.0[22])

識 別	音声認識	予 測	数値予測	実 行	表現生成
	画像認識		マッチング		デザイン
	動画認識		意図予測		行動最適化
	言語解析		ニーズ予測		作業の自動化

図 2.6 人工知能 (AI) の実用化における機能領域
(引用：ICT の進化が雇用と働き方に及ぼす影響に関する調査研究[19] p19 図 2-2-2-2)

AI、機械学習、深層学習は新しい技術としてこれからますます社会での活用が進み、生活を豊かにすることが期待されるが、社会への影響力が大きいことから、適切な開発と活用が求められる。AI の活用に関して、総務省の AI ネットワーク社会推進会議（報告書 2018 別紙 2-1）[23]では、以下のように課題がまとめられている。

- 生命・身体・財産の安全に関する課題（事故の防止など）
- AI による判断の正当性や公平性に関する課題（差別、生命倫理との関係など）
- プライバシーに関する課題（プライバシーの尊重、プロファイリングなど）
- 人間と AI との役割分担等に関する課題（人間の判断の介在、関係者感の協力など）
- AI に対する受容性に関する課題（利用者に対する説明責任など）
- AI の判断のブラックボックス化に関する課題（事故が発生したときの原因究明など）
- セキュリティに関する課題（ハッキング対策など）
- AI 間の連携に関する課題（AI 間の交渉・調整など）
- AI が学習するデータに関する課題（データの正確性など）

AI を有効に活用して社会を豊かにするためには、このような課題を十分に理解し、技術開発を進める必要がある。また、人、社会システム、産業構造、イノベーションシステム、ガバナンス等、AI を有効かつ安全に利用できる社会を構築する「AI-Ready な社会」への変革を推進する必要がある、このことは内閣府が公表している人間中心の AI 社会原則[24]にまとめられている。

2.2.2 AI の概要(2)

AI、機械学習、深層学習の関係は図 2.7 のとおりである。AI は人間の知能を模倣したシステムや機械を指す、広い意味の言葉である。AI と機械学習は同じ意味で使われることもあるが、厳密には言い分ける必要がある。機械学習とは、入力されたデータからコンピュータがルールを発見し、そのルールを新たなデータに当てはめることで、新たなデータに対する予測を行う手法である。従来のルールベース手法では、どのような特徴に注目するのか、どのような判定ルールとするのかを、人間が一つ一つ決定、調整していく必要がある（図 2.3）。一方で機械学習の場合、どのような特徴に注目するのかを与えることで、与えられたデータをもとに自身でパターンを学習してルールを定めることができる点で従来のルールベース手法と大きく異なる（図 2.8）。また、深層学習とは、機械学習手法の一つであり、多数の層で構成されるニューラルネットワークを用いて行う機械学習のことである。深層学習は、特徴を人間が指示することなく、自らが与えられたデータから注目する特徴を学習する点で大きなブレークスルーとなったとされている（図 2.9）。機械学習、深層学習ではデータの量や質が非常に重要となる。特に深層学習モデルでは特徴抽出、ルール作成までを与えられたデータをもとに行うため、データの「質」が精度に大きく影響する。データの質とは、データ数やデータに付与されている情報の正解性などが含まれる。質の良いデータが十分にあれば、これまでのルールベース手法では複雑すぎて解けなかった問題が解けるようになる可能性がある。

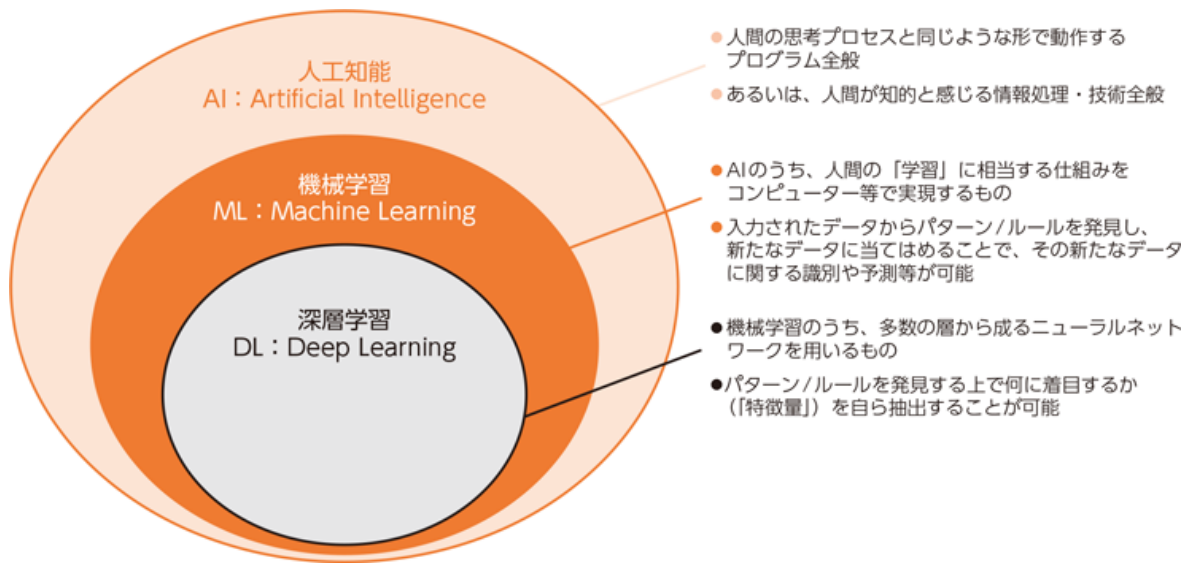


図 2.7 AI・機械学習・深層学習の関係
(引用：令和元年版情報通信白書 [20] p83 図 1-3-2-1)

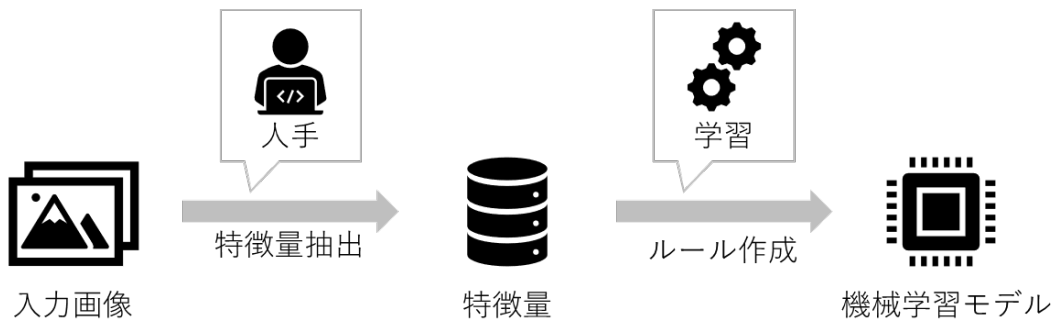


図 2.8 ルールベース・機械学習・深層学習手法の画像判定モデル構成手順



図 2.9 ルールベース・機械学習・深層学習手法の画像判定モデル構成手順

機械学習は与えられるデータや目的によって分類される（図 2.10）。機械学習は、真実のデータや人間による判別から得られた正解に相当する「教師データ」の与えられ方によって教師あり学習、教師なし学習、強化学習の 3 つに大別される。教師あり学習とは、入力データと人が与えた正解を対にして、教師データを既知の情報として学習に利用し、未知の入力データに対して正しい予測が可能となるようなルールを定めていく手法である。教師あり学習はさらに回帰と分類に分けられる。分類の主な目的は、Yes か No かのようにデータが属するクラスを予測することである。予測するクラス数が 2 クラスの場合は 2 値分類もしくは 2 クラス分類と呼ばれ、3 クラス以上の場合が多クラス分類と呼ばれる。回帰の主な目的は、連続値の予測である。例えば、過去の気象情報をもとに明日の天気が晴れか曇りか雨かを予測する場合は分類問題であり、明日の最高気温を予測する場合は回帰問題となる。教師なし学習は、正解を与えず学習させる方法である。入力データを学習させることで入力データの特徴やパターンなどを覚え、新しく入力されたデータが学習した特徴やパターンに当てはまるかどうかを判断する手法である。教師なし学習では、クラスタリングや次元削減を行うことが多い。強化学習では、試行錯誤を通じて、報酬（評価）が得られる行動や選択を学習する。例えば、迷路（環境）を進むロボットが歩行距離（報酬）を最大化するために、迷路の進み方を学習していくようなケースが考えられる。

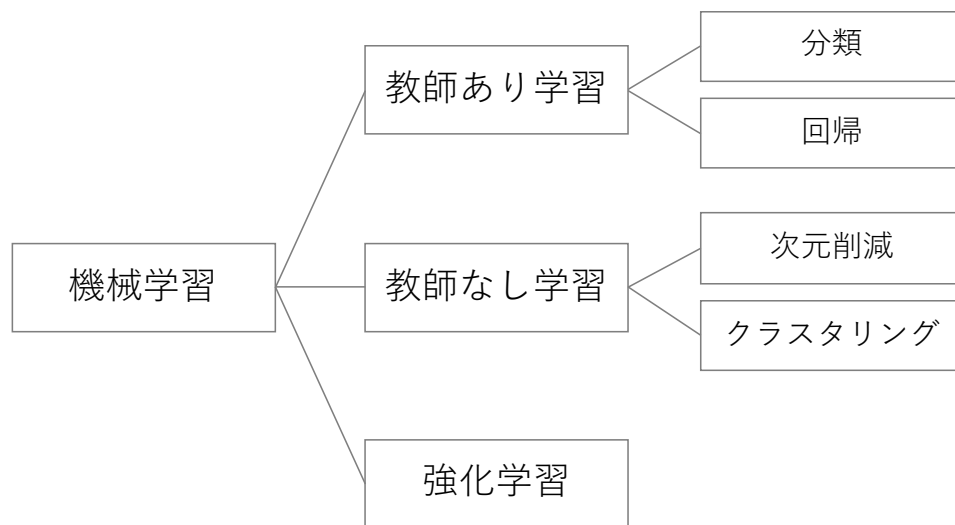


図 2.10 機械学習の分類

2.2.3 画像認識における AI 技術の進歩

AI・機械学習・深層学習の応用は、教育、金融、医療、エンターテインメントなど様々な分野で進んでおり（図 2.5）、もちろん製造業も例外ではない。具体的な技術では、音声認識、機械翻訳、テキストマイニングなど多岐にわたり、我々の日常にもこれらの技術を応用したシステムが溢れている。様々な技術のなかでも、大きく進歩した技術に画像認識がある。画像認識の中心的な技術に深層学習を用いた手法である畳み込みニューラルネットワーク (Convolutional Neural Networks : 以下 CNN) があり、非常に大きな成果を上げている。

CNN は福島らが提案したネオコグロトロン[25]を発展させた多層のニューラルネットワークである。一般的な CNN は主に畳み込み層 (Convolution Layer) とプーリング層 (Pooling Layer)、全結合層 (Dense Layer) という 3つの個性的な機能を備えた層を積み上げることで構成されており、入力画像を畳み込み層とプーリング層で何回か処理した後に、全結合層によって出力を得る構成となっている (図 2.11)。

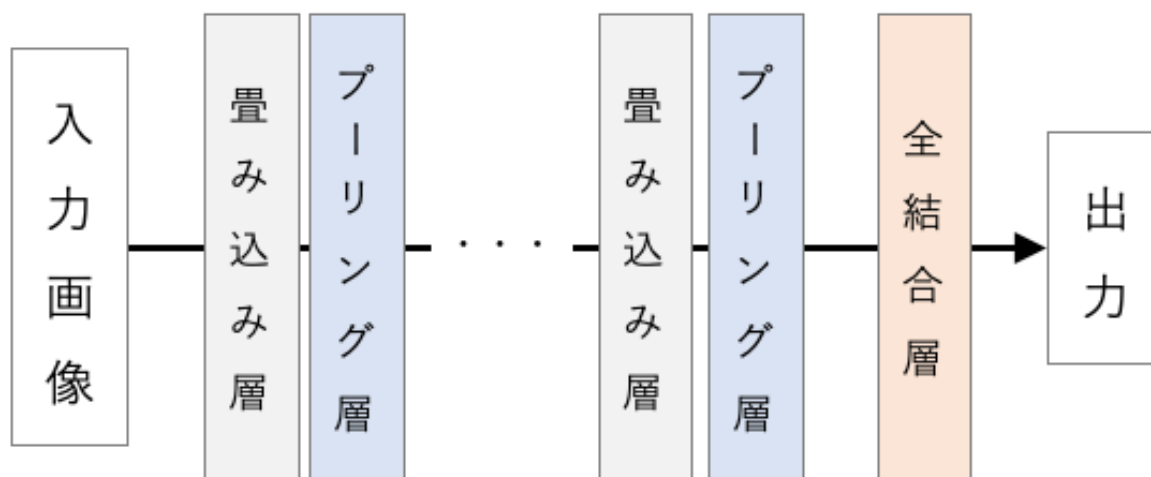


図 2.11 畳み込みニューラルネットワークの構造

CNN の核となるのが畳み込み層である。畳み込み層はフィルタ (またはカーネル) と呼ばれるパラメータを用いて入力画像に対する畳み込み演算を行うことによって、入力画像の特徴を捉えていく。畳み込み演算は一般的な画像処理でのぼかし処理やエッジ抽出で行われるもの基本的には同じ処理となる。畳み込み層におけるフィルタは通常、ランダムに初期化されており、それ自体が学習によって更新されるパラメータとなる。フィルタのサイズ、演算ごとのフィルタの移動量を示すストライドの値、入力データの周りがある数値で埋めるパディングの有無などがパラメータとして人の手によって与えられる。畳み込みの演算の結果に対してバイアス (一定の値) を加え、活性化関数を作用させたものが出力となる。畳み込み層によって得られた出力を特徴マップと呼ぶ。活性化関数 f にはシグモイド関数やハイパボリックタンジェント関数などの種類があるが、近年は、ReLU (Rectified Linear Unit : 正規化線形関数) [26]がよく用いられる。ReLU を改良した活性化関数もいくつか提案されている[27][28]が、標準的な ReLU が幅広く利用されている。入力 x に活性化関数 ReLU を作用させた出力は次の式 2.1 で表される。

$$f(x) = \begin{cases} 0 & (x \leq 0) \\ x & (x > 0) \end{cases} \quad (2.1)$$

畳み込み層の基本的な構造を図 2.12 に示す。畳み込み層への入力画像 x は 3次元配列 $W \times W \times N$ として与えられる。 N をチャンネル呼び、最初の入力層では画像がグレースケールの場合で $N = 1$ 、カ

ラーの場合で $N = 3$ 、それ以降の中間層では入力となる特徴マップの数となる。フィルタを h 、フィルタサイズを $m \times m \times N$ 、ストライド量を s 、バイアスを b とすると畳み込み演算によって得られる出力 x' はサイズ $W' \times W'$ ($W' = (\frac{W-m}{s} + 1)$)の2次元配列として求められる。 x' を x'_{ij} ($i \in [1, W'], j \in [1, W']$)と表現すると、以下の式 2.2 で表すことができる。

$$x'_{ij} = \sum_{k=1}^N \left[\sum_{p=1}^m \sum_{q=1}^m x_{s(i-1)+p, s(j-1)+q, k} h_{pqk} \right] + b \quad (2.2)$$

畳み込み演算によって得られた x_{ij} に活性化関数 f を作用させた出力である特徴マップ y_{ij} は以下の式 2.3 で表される。

$$y_{ij} = f(x'_{ij}) \quad (2.3)$$

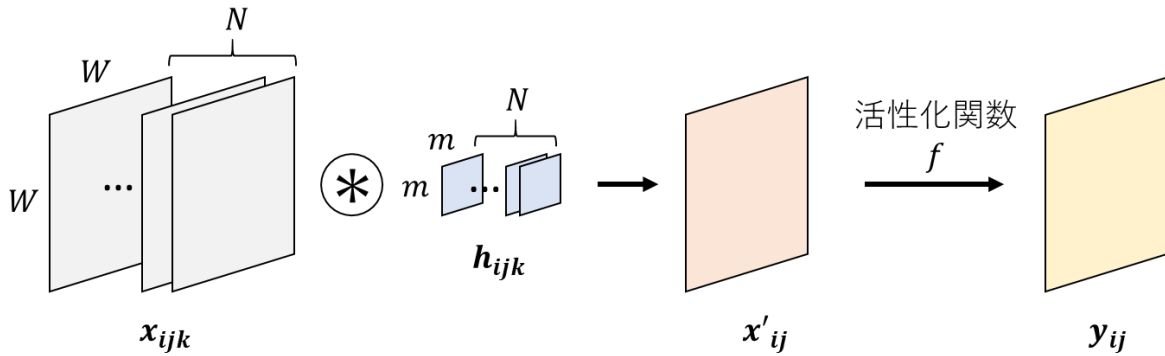


図 2.12 畳み込み層の基本構造

実際のモデルでは出力される特徴マップが複数となるように、CNN モデルが構成される。出力として得たい特徴マップの数を M とすると、 N 枚のフィルタが M セット準備され、式 2.2 の演算が M 回行われることによって所望の出力である特徴マップを得る(図 2.13)。

畳み込み層の出力として得られた特徴マップは、プーリング層によって処理が行われる。プーリング処理にはいくつかの種類があり、CNN の研究初期には平均プーリングと呼ばれるものがよく用いられていた[29][30]が、近年、よく用いられるものに最大プーリング (Max Pooling) がある[31]。最大プーリング処理は事前に決めたサイズ内での最大値を取り、それを 1 つの出力とする演算を行う。畳み込みではフィルタをパラメータとして学習し更新していったが、プーリングは学習するパラメータを持たない。プーリング層の目的は特徴マップサイズを小さくすること以外に、画像のどの位置でフィルタの応答が強かったのかという特徴の位置感度を落とすことで、平行移動などが起きても影響を受けることがないようにするという重要な役割を担っている。プーリング層への入力を y_{ij} 、プーリングサイズを $l \times l$ としたときの出力 z_{ij} は以下の式 2.4 で表される。

$$z_{ij} = \max(y_{l(i-1)+s, l(j-1)+t}) \quad \text{where } s \in [1, l], t \in [1, l] \quad (2.4)$$

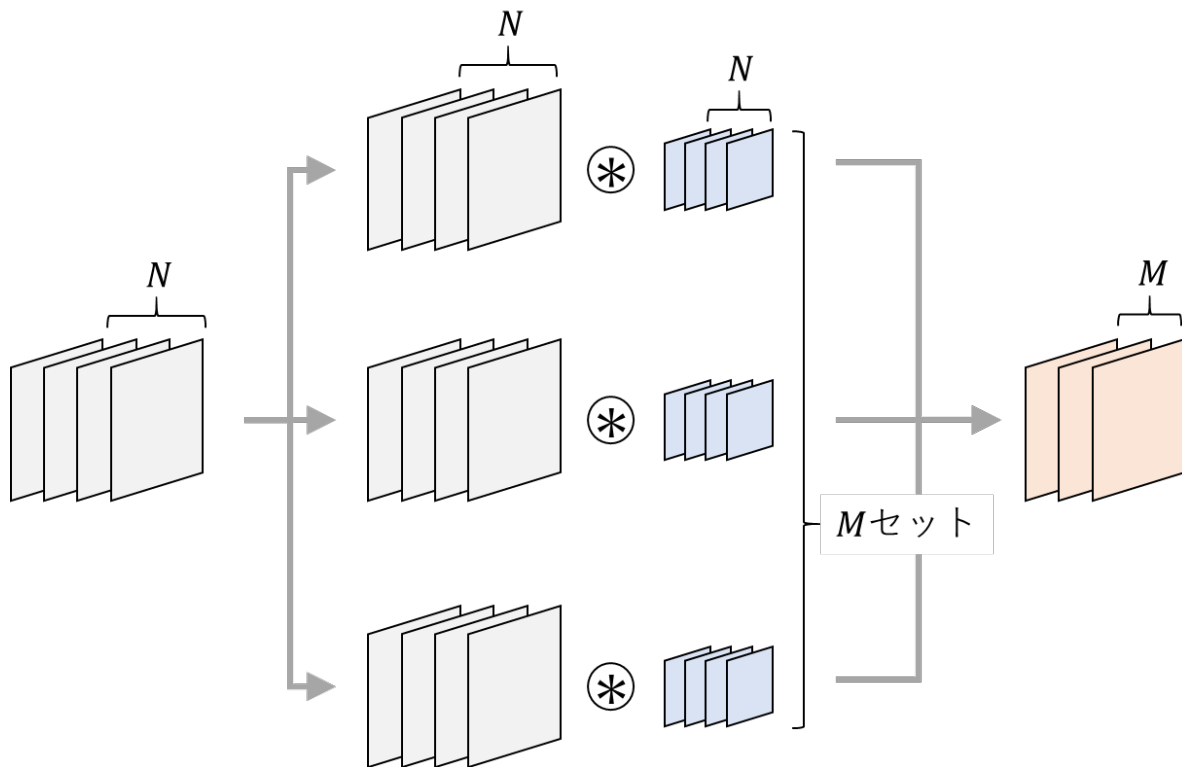


図 2.13 出力が複数チャンネルの畳み込み模式図

畳み込み層とプーリング層による演算を複数回行った後に出力された特徴マップは 1 次元配列に変換され、全結合層の入力として用いられる。全結合層は一般的なニューラルネットワーク (Neural Network) と同様の演算が行われる。全結合のある層の j 番目のユニット v_j を得るには、まず、その直前の層の値 z_i ($i = 1, 2, 3, \dots, m$) を入力とした重み付き和にバイアスが加算された z'_j が演算される。このときの z'_j は以下の式 2.5 で表される。

$$z'_j = b_j + \sum_{i=1}^m z_i w_{ij} \quad (2.5)$$

得られた z'_j に活性化関数 f を作用させた v_j が全結合層の出力となる。 v_j は以下の式 2.6 で表される。

$$v_j = f(z'_j) \quad (2.6)$$

クラス分類を行う場合、最終層では分類クラス数と同じ数のユニット n 個を配置し、活性化関数にソフトマックス関数 (Softmax Function) を用いて出力 p_j を得る。直前の全結合層の出力を n 個のユニットを持つ v_j とすると、ソフトマックス関数を用いた出力 p_j ($j \in [1, n]$) は次式 2.7 で表される。

$$p_j = \frac{e^{v_j}}{\sum_{k=1}^n e^{v_k}} \quad (2.7)$$

このとき、

$$\sum_{j=1}^n p_j = 1 \quad (2.8)$$

となり、 p_j は各クラスにおける確率と解釈することができる。最終的な予測クラスは p_j の最大値を持つインデックスを取る。予測クラス、 p は以下の式 2.9 で表される。

$$p = \operatorname{argmax}(p_j) \quad (2.9)$$

畳み込み層が持つ特徴に局所的受容野と重み共有の 2 つがある。これらの特徴によってパラメータ数が大幅に削減されており、全結合層のみでは困難であったサイズの大きい画像での学習を可能にしている。受容野とは、ある演算に影響のある入力側の領域のことである。受容野が広ければ広いほど多くの演算を必要とするため、多くのパラメータが必要となる。重み共有とは、画像を畳み込む際に畳み込みを行う位置によって別々のフィルタを用いるのではなく、画像全体で 1 つのフィルタを共有することである。全結合層では、ある一つのユニットは入力となる直前の層の全てのユニットに影響を受けるため、受容野は直前の層の全てのユニットとなる。また、ユニット同士の結合すべてに個別の重みパラメータを有するため、入力画像のサイズが大きい場合にはパラメータ数が非常に大きくなってしまい、学習が困難になる。これに対し、CNN では、畳み込み層の出力はフィルタ演算によって得られており、画像の一部のみを受容野としているため、パラメータを削減することができる。また、画像のサイズに関わらず、フィルタ（パラメータ）を共有するために、パラメータの増大を避けることができる。これらの事によって全結合層のみで演算を行うときよりもパラメータ数を大幅に削減することが可能になる。

CNN が大きな成果を上げ、注目されるようになったきっかけに 2012 年に開催された画像認識の競技会 ILSVRC (ImageNet Large Scale Visual Recognition Challenge) がある。ILSVRC とは ImageNet と呼ばれる画像数約 1,400 万枚、22,000 カテゴリからなる大規模な画像データセットに対して、画像の分類性能を競うコンペティションであり、2010 年から 2017 年まで毎年開催されていた。2012 年、トロント大学のヒントン教授のチームが開発した AlexNet[31]と呼ばれる CNN モデルが当時のスタンダードとなっていたサポートベクトルマシン (SVM) と比較して誤差率 10%以上の改善を達成した (図 2.14)。それ以降、CNN が画像認識の主流となり、毎年優秀なアルゴリズムが開発され[32][33]、2015 年に発表された ResNet[34]では、ヒトの認識の限界と言われる誤差率 5%[54]よりも低い値を達成した。

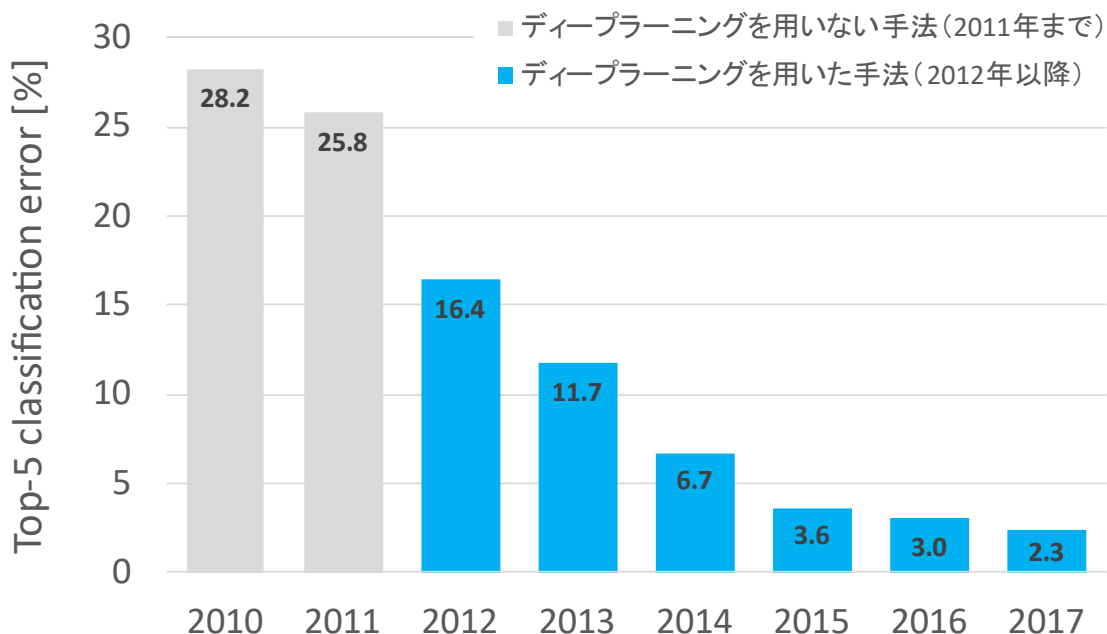


図 2.14 ILSVRC におけるトップモデルの誤差率推移
(2012 年以降は CNN を用いたモデル)

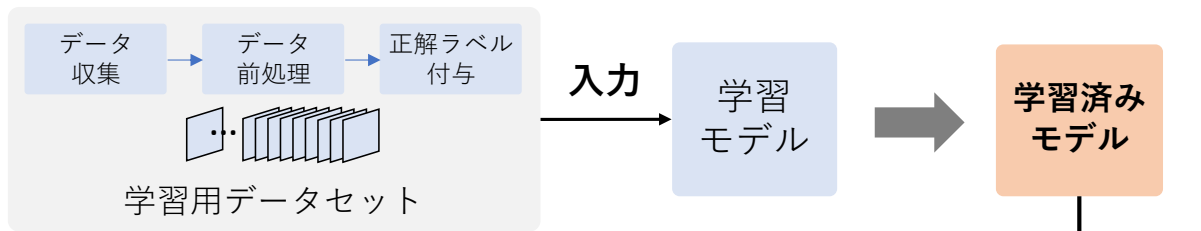
研究対象であるクラッチディスクの外観検査において、一般的なルールベース画像検査装置では、すべての多様な欠陥に対して人手でルールを定めることには多大な労力を必要とするため、自動化は完全には達成されていないが、CNN は自動化を実現するための手法として有効であると考えられる。そこで、本研究では CNN を用いて傷、変形、浮き上がりなどの欠陥種類毎に分類を行う多クラス分類法を適用し、自動化技術の構築を試みる。

2.3 欠陥分類実験 1：従来手法の適用（実験環境）

CNN を用いた画像認識のクラッチディスク外観検査への適用可能性を確認するため、テストサンプルを用い欠陥分類実験を行った。実験は CNN 教師あり多クラス分類モデルを用いて実際のクラッチディスクを撮像した画像を分類するタスクとした。

CNN 教師あり多クラス分類モデルで分類予測を得るまでの手順を図 2.15 に示す。手順は大きく学習フェーズと判定フェーズに分けられる。学習フェーズではまずモデルの学習に用いるデータセットを生成する。このとき、データの形や形式、ノイズ除去などの前処理と、人による正解ラベルの付与が行われる。生成したデータセットを入力としてモデルを学習し、成果物として学習済みモデルを得る。判定フェーズでは分類したい未知の画像を学習済みモデルに入力することで、CNN から出力される予測から、分類の判定結果を得る。

<学習フェーズ>



<判定フェーズ>

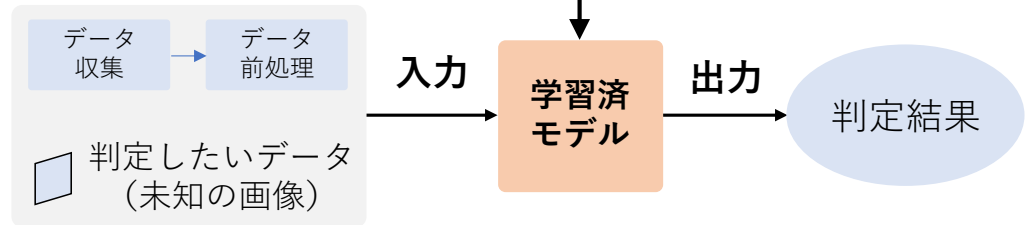


図 2.15 モデル構築手順の模式図

2.3.1 画像取得方法

サンプル画像は実際のクラッチディスク完成品を実験環境で撮像することで作成された。実験環境では撮像機器に 3,296×2,472 画素モノクロカメラ、照明には外径 300mm、内径 252mm の白色 LED リング照明（出力 27.6W）を用いた。撮像対象のクラッチディスクはリング形状であるため、対象をムラの少ない状態に照射することが可能となるよう、リング形状かつ拡散光照射タイプの照明を選定した。また、人の目視検査を代替するという点で、一般的で人間の感覚（可視光領域）と画像が比較しやすいように白色 LED を採用している。カメラおよび照明は、検査対象となる直径 180mm のクラッチディスクからの距離 471.5mm、25mm の位置に設置した（図 2.16）。撮像機器及び検査対象は外乱光の影響を抑えるために簡易的な暗箱の中に設置した。撮像は検査対象を治具上で静止させ、ペーパー摩擦材が同じ位相となるように位相合わせした状態で行われ、クラッチディスク全体を含む撮像視野 342×256.5mm の 256 階調グレースケール画像を取得した。取得した画像から、ペーパー摩擦材を貼り付けた部分の近傍を切り出した 400×100 画素の画像を 1 枚のサンプル画像とした。クラッチディスクを回転させながら、同様の撮像を行うことで 1 枚のクラッチディスクから 8 枚のサンプル画像を取得した。サンプル画像におけるクラッチディスク以外の背景部分には設備部品などが写り込まないよう治具形状が決定されている。

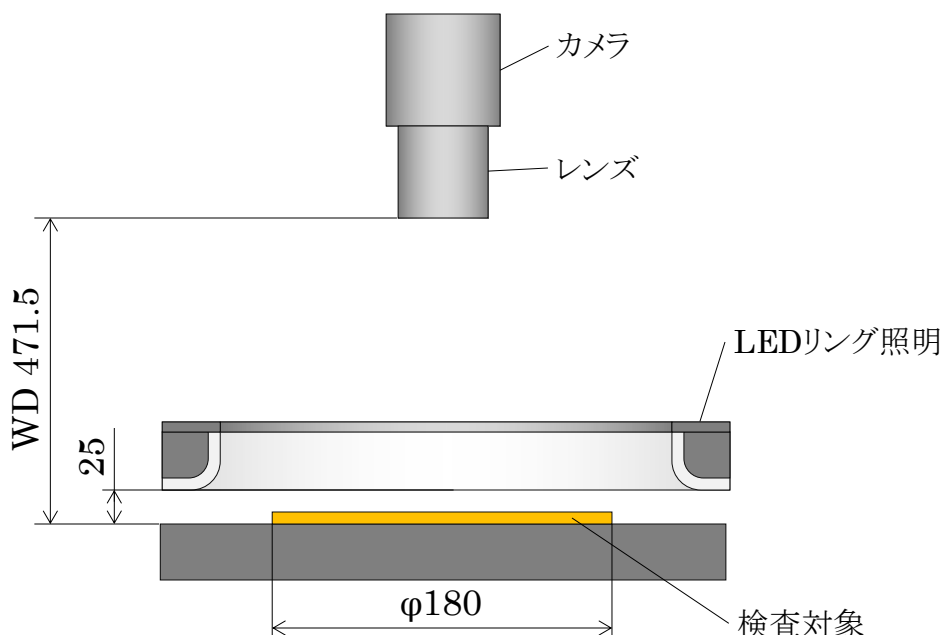


図 2.16 分類実験 1 の撮像環境

2.3.2 データセット

データセットとして正常品クラスと 4 種の欠陥品クラスを含む 5 クラスのサンプル画像が合計 850 枚収集された。各クラスから 20 枚ずつ取り分けた 100 枚の画像を検証データセットとし、残りの 750 枚を学習用の教師データセットとして用いた。すべての欠陥品サンプルは正常なクラッチディスクに人為的なダメージを与えることで作成された。クラッチディスクのペーパー摩擦材に発生する外観に関わる欠陥は種類が多様であるが、分類実験 1 では代表的な欠陥である異物、キズ、打痕、めくれの 4 種をサンプルとして用意した。データセットに含まれるクラスと各クラスのデータ数を表 2.2 に、各クラスのサンプル画像の例を図 2.17 に示し、欠陥箇所を赤枠で囲っている。すべてのサンプル画像はエンジニアによって、不適切な画像が含まれていないか、属するクラスが正しいかの確認を行った。クラス分けは目視検査において検出が望まれるレベルを基準に行われた。

表 2.2 分類実験 1 用データセットのクラスとサンプル数

Index	クラス名	総サンプル数	教師データ数	検証データ数
0	正常 (欠陥なし)	169 枚	149 枚	20 枚
1	異物	170 枚	150 枚	20 枚
2	キズ	172 枚	152 枚	20 枚
3	打痕	168 枚	148 枚	20 枚
4	めくれ	171 枚	151 枚	20 枚
合計		850 枚	750 枚	100 枚

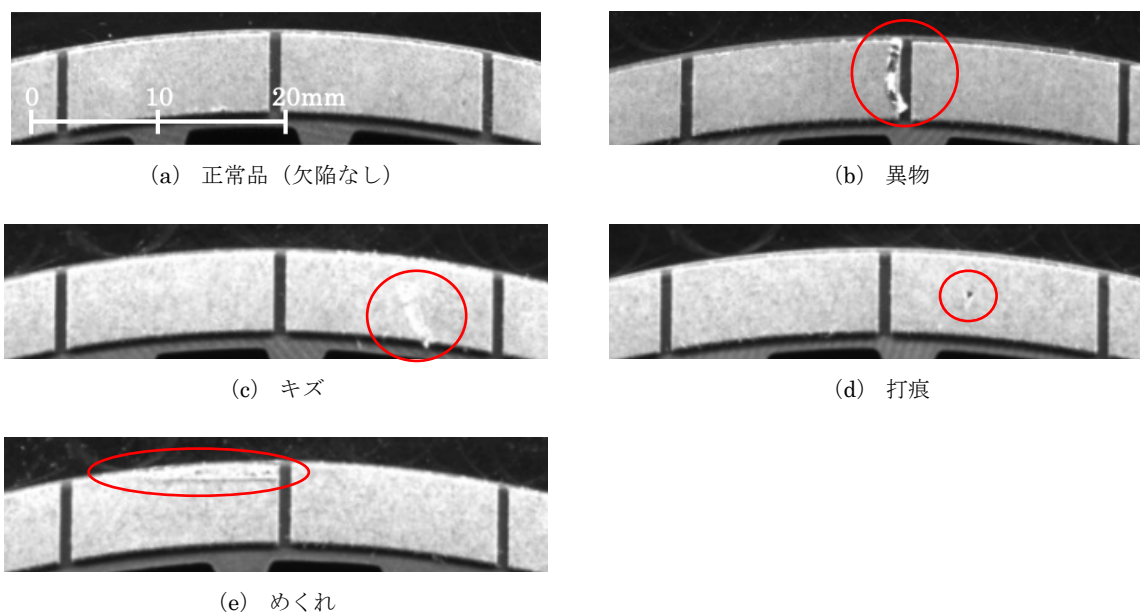


図 2.17 分類実験 1 に用いたサンプルの例

2.3.3 学習モデル

画像分類モデルの構成及びハイパーパラメータは先行研究[35]で提案されている VGGNet を参考とし、クラッチディスク外観検査の予備実験での調整を経て、本実験用に再構成した。モデルは入力にサイズが 400×400 の画像を受け取り、7 層の畳み込み層と 2 層の全結合層を経て、入力画像が属するクラスを予測する構成とした。畳み込み層は全層でフィルタサイズ 3×3 、ストライド 1 とし、活性化関数に ReLU を用いた。出力する特徴マップは畳み込みを行うごとに 32 枚、64 枚、128 枚と増やし、畳み込みの最終層では 256 枚の特徴マップを出力させた。畳み込み層の途中には 2×2 のマックスプーリング層が特徴マップを小さくするために用いられた。最初の畳み込み層を除き、畳み込み層と対にしてプーリング処理を行った。畳み込み最終層の特徴マップを全結合層に入力するために 1 次元に変換した後、2 層の全結合層を経て、最終層でソフトマックス関数を用いて 5 次元に変換し、その最大値をもつクラスを予測クラスとして用いた。全結合層では正則化効果を得るためにドロップアウト (Dropout) [36]を用いた。モデルは損失関数¹にカテゴリカルクロスエントロピー²、

¹ 機械学習における損失関数 (Loss function) とは、データとして与えられた正解値と、モデルによって出力された予測値とのズレの大きさである損失値 (Loss Value) を計算するための関数である。機械学習モデルはこの損失値を最小化するように学習を進める。損失関数には様々な種類があるが、代表的なものに、平均二乗誤差 (Mean Squared Error)、平均絶対誤差 (Mean Absolute Error) などがある。深層学習ではクロスエントロピー誤差 (Cross-entropy Error) が用いられることが多い。

² カテゴリカルクロスエントロピー (Categorical cross entropy) とは、2 つの確率分布がどれくらい離れているかを表す指標である。2 つの確率分布が似ているほど小さな値をとり、まったく同じ場合には 0 となる。

最適化アルゴリズムに **RMSProp**[37][38]を用い、1回の学習でモデルのパラメータを更新する量である学習率は 5×10^{-4} とした。モデルの総学習パラメータ数は **6,965,637**となった。分類実験1の学習モデル構成を図 2.18 に示す。汎化性能を向上させるため、教師データセットの画像はモデルに入力される前に、元画像に対し回転、左右反転、左右移動、上下移動、剪断処理、輝度シフトの加工をランダムに行う画像増強がわれた[31][39] (図 2.19)。評価には教師データセットの画像 **750** 枚を **1,000** 回繰り返し学習することで得られたモデルを用いた。画像分類モデル学習の実装にはプログラミング言語に **python**、深層学習用のライブラリに **TensorFlow** を用いた。**python** は比較的少ないコードで実装が可能な言語であり、データ分析、機械学習、深層学習の分野で広く使われている。**TensorFlow** は **Google** 社が開発している機械学習ライブラリで、特に深層学習の機能が充実している。どちらもオープンソースであり、利用者が多いことから、情報が充実しているというメリットが有る。

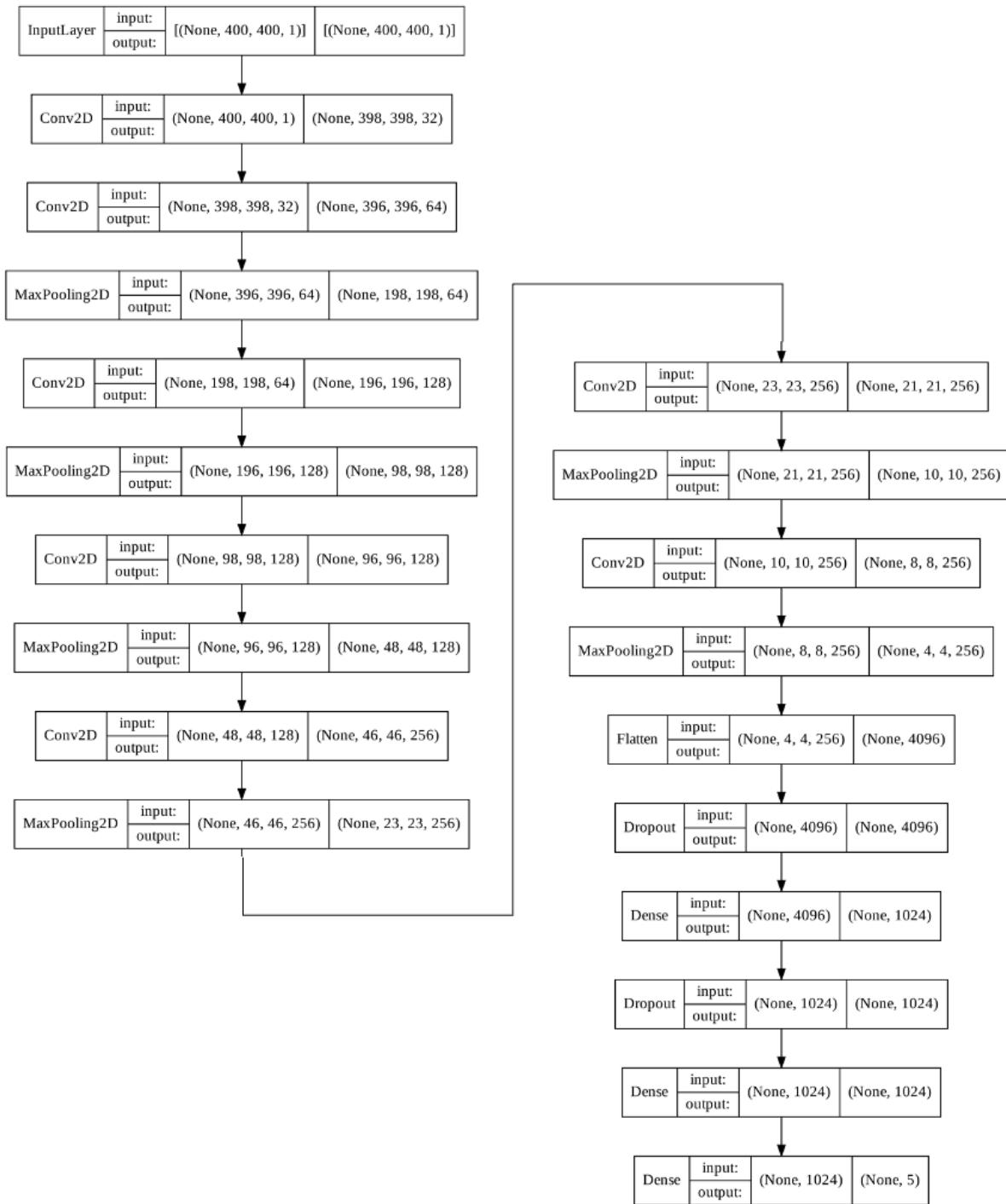


図 2.18 分類実験 1 の学習モデル構成

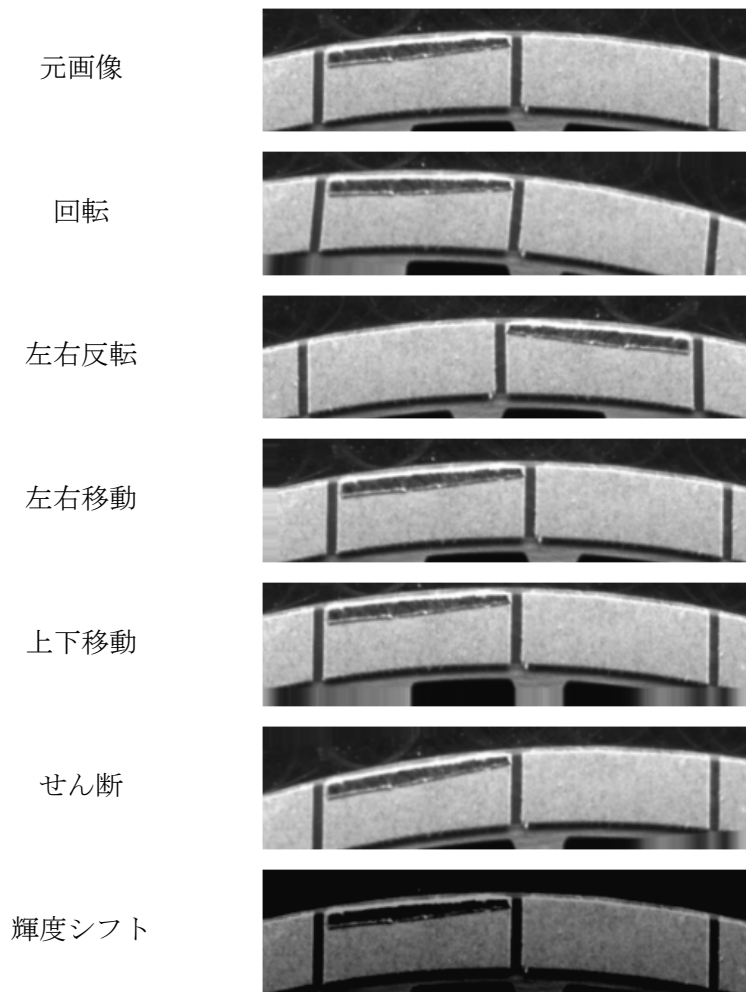


図 2.19 画像増強による画像加工の例
 (各処理の効果をわかりやすくするため、実際の処理よりも強度を高くした処理が含まれる)

2.3.4 実験結果

実験の結果、100件の検証データのうち、90件が正しいクラスに分類され、正解率は90%となった。ここでの正解率は、予測クラスが正解クラスと一致している確率のことを言う。検証データにおける混同行列³を図2.20に示す。実際には正常品であるものを欠陥品と誤認することはなく、実際には欠陥品であるものを正常品と誤認したのは1件のみとなった。正常品と欠陥品という大きな分類で考えれば、99%の正解率であると解釈することもできる。実際の量産と比較して、撮像環境が理想的であり、欠陥種が少ないという限定的な状況ではあるが、CNNを用いたクラッチディスク完成品外観検査の可能性を確認できた。

正解クラス	0_正常	20	0	0	0	0
	1_異物	1	18	1	0	0
	2_キズ	0	0	18	0	2
	3_打痕	0	1	2	16	1
	4_めくれ	0	0	2	0	18
		0_正常	1_異物	2_キズ	3_打痕	4_めくれ
		予測クラス				

図 2.20 混同行列 (分類実験 1)

³ 混同行列 (Confusion Matrix) とは、機械学習の教師あり学習モデルの性能評価、可視化によく用いられるマトリクスである。行列の行方向に正解クラスを表し、列方向に予測クラスを表す (逆の場合もある)。正しい予測ができた数が行列の対角成分に位置しており、対角成分から外れた部分は誤った予測の数を示している。混同行列によって、予測が正しかった割合 (正解率) だけではなく、例えば、キズをめくれに予測する事が多いなど、どのクラスをどのクラスに分類したかの詳細情報が得られるという利点がある。

2.3.5 実用的な CNN 完成品外観検査装置の開発に向けた課題

分類実験 1 によって、一定の性能が確認されたと同時に、実用的な CNN 完成品外観検査装置を開発する上での課題が整理された。主要な課題は以下の 4 つである。

① 学習の安定化

学習時の正解率推移を図 2.21 に、損失値の推移を図 2.22 に示す。図中の破線は学習時の教師データセットにおける結果、実線は検証用データセットにおける結果を示しており、縦軸に正解率および損失値、横軸に学習回数を示している。損失値とは正解と予測の誤差を示す値で、小さい方が良いとされる数値である。損失値の詳細については第 3 章で触れる。教師データセットでの正解率について見てみると、学習全体を通して安定しており、特に学習の後半で大きな変動はない。しかし、実際の運用では教師データセットに含まれない未知のデータに対しての性能が求められる。学習に用いられていない未知のデータである検証データセットにおける正解率は、学習全体の 3/4 が過ぎた学習回数 750 回以降においても 16% ほどの変動が出ており、学習終了のタイミングによってはモデルの精度が極端に悪化する可能性がある。損失値も正解率と同様、教師データセットでは安定しているが、検証データセットにおいては、後半でも収束せず、変動が大きくなっており、モデルの性能が不安定となっている。このような学習中の正解率、損失値の変動を少なくし、安定した学習が行えるように対策する必要がある。

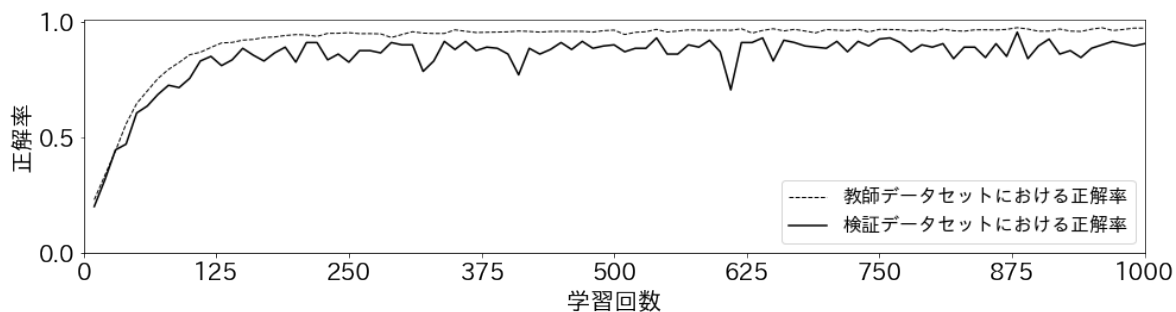


図 2.21 学習時の正解率推移

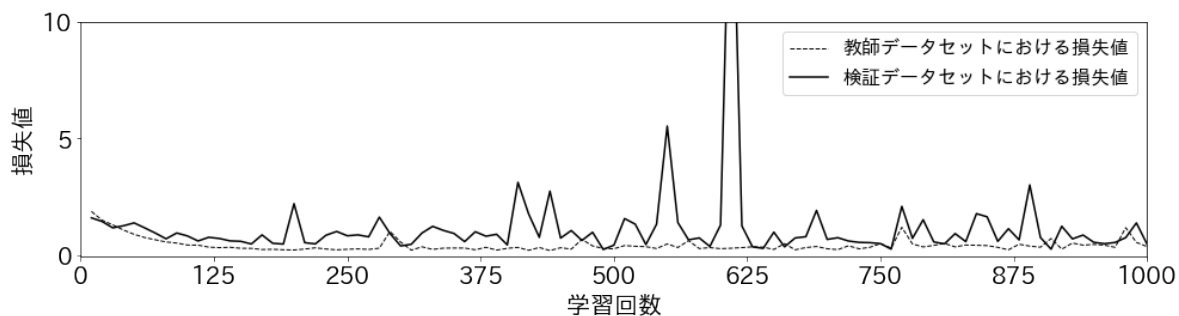


図 2.22 学習時の損失値推移

② 実際の量産環境での性能確保

分類実験 1 は、実証実験であったため、撮像環境や欠陥の種類が実際の量産工程とは異なる。実際の撮像環境では、生産機能確保のための制限があり、取得画像の背景には治具や設備部品などの写り込みや検査対象の位相のランダム性が避けられない。また、実際に発生する可能性のある欠陥は 4 種類よりも多く、その状態も多様である。そのため、分類実験 1 よりも分類が難しくなることが想定される。量産における検査工程で検出が望まれている欠陥種を想定したサンプル品を量産同等の環境で撮像した画像から作成したデータセットを用いて分類実験を行い、実際の量産環境での性能を確認する必要がある。

③ 見逃し発生の抑制

外観検査装置に求められる最も重要な機能は、欠陥品を正常と誤認する「見逃し」の最小化である。分類実験 1 の検証では、80 件の欠陥を含む画像の内、異物のある画像 1 件を正常品と判定し、見逃しとなっている（図 2.23）。欠陥品の内、見逃しとなった割合である見逃し率は 1.25%となる。実用化に向けては、この見逃しを抑制するための手法を開発する必要がある。

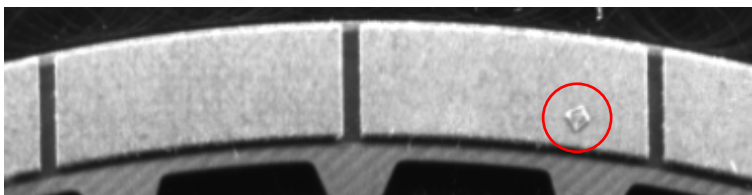


図 2.23 分類実験 1 で欠陥を見逃した画像（正解：異物）

④ 入力データの質向上

CNN を用いた画像分類学習では、CNN モデル自らが与えられた入力画像から特徴量を抽出し、分類のルールを作成する（図 2.10）ため、入力データの「質」が精度に大きく影響する。分類実験 1 で見逃しとなった欠陥画像は異物を含む画像であったが、同じ欠陥種で正確に分類できた画像は見逃した画像と比べて、異物部分とペーパー摩擦材部分との輝度差が大きく、欠陥部分を目視で認識しやすい画像が多い（図 2.24）。また、キズの欠陥は全体的に画像の欠陥部分を目視で認識することが難しい（図 2.25）。このような欠陥を顕在化させ、入力データの質を改善することで、分類性能を向上させる検討を行う必要がある。

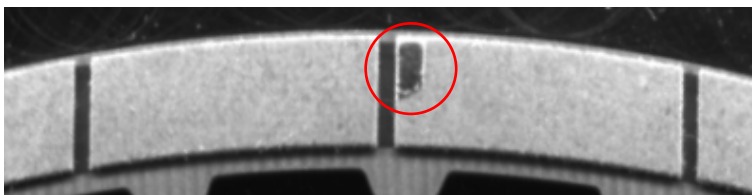


図 2.24 分類実験 1 で欠陥を正しく分類できた画像（正解：異物）

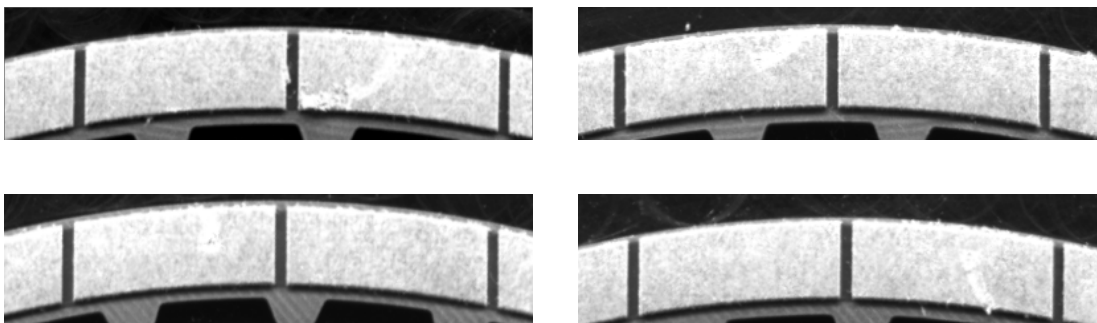


図 2.25 キズの欠陥を含む画像の例

分類実験 1 で得られた成果と課題を整理すると、以下のようにまとめられる。

<分類実験 1 で得られた成果>

- 100 件のうち、90 件を正しいクラスに分類でき、正解率 90%を達成した。
- 正常品と欠陥品の 2 クラス分類ととらえると、正解率 99%を達成した。
- 限定的な状況ではあるが、CNN を用いたクラッチディスク完成品外観検査の可能性を確認できた。

<CNN を用いた完成品外観検査装置の実用化に向けた課題>

- 学習中の正解率、損失値の変動を少なくし、安定した学習が行えるように対策する必要がある。
- 実際の量産環境を再現し、欠陥種を拡大した上で、量産想定 of 性能を確認する必要がある。
- 検査装置の重要な機能である見逃し発生の抑制を実現する技術を開発する必要がある。
- 欠陥を顕在化させるなど、入力データの質向上を行う必要がある。

2.4 欠陥分類実験 2：開発手法 1 2 クラスアンサンブル手法

分類実験 1 では正解率 90%と、正常品と欠陥品に注目した分類では 99%の成果率を達成することができた。しかし、学習時の正解率や損失値の推移をみると、変動が大きく不安定であり、改善が必要になっている。そこで、より安定して学習が行える学習手法の検討を行った。

分類実験 1 では多クラス分類法を使用して、サンプルが属するクラスのカテゴリ分けを行った。ただし、異なる外観を持つサンプルを分類するため、単一の特徴量抽出モデルではなく、複数の特徴抽出モデルを持っていると考えることができる。分類のタスクはメインのモデルがサンプルを複数モデルのどれに準拠させるかの決定タスクを負担することで、より複雑になっていると想定される。そこで、タスクを欠陥のない正常品クラスと個別の欠陥品クラスの 2 クラス分類問題に分解し、欠陥品を検出する方法を検討する。多クラス分類は、特徴空間を複数に分類することと同等であり、分類の境界はより複雑になると予想される (図 2.26)。これに対して、タスクを 2 クラス分類に分解する

ことによって、個々の問題は空間を 2 つに分割するという単純なタスク (図 2.27) となり、分類能力の向上と学習の安定性向上が期待できる。欠陥ごとの 2 クラス分類では高精度で欠陥分類が可能になる反面、学習していない欠陥は正確に分類できなくなると想定される。しかし、いずれかのモデルで欠陥と分類された場合には欠陥と判定することで、それぞれのモデルが学習した欠陥さえ確実に分類できれば、欠陥を取り除くことが可能となる (図 2.28)。

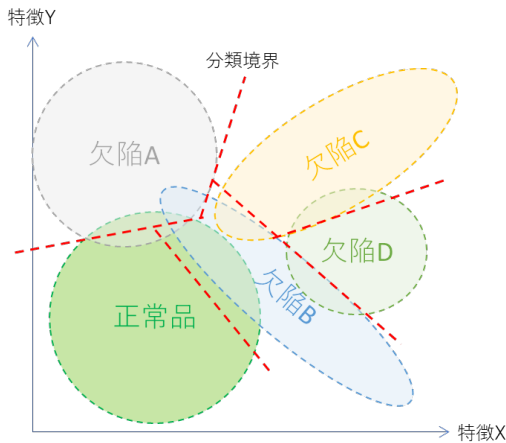


図 2.26 多クラス分類の境界線イメージ

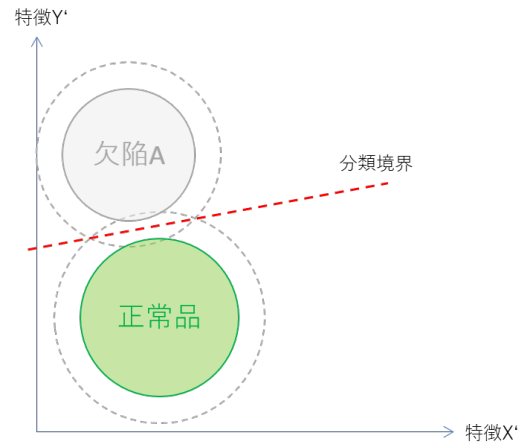


図 2.27 2 クラス分類の境界線イメージ

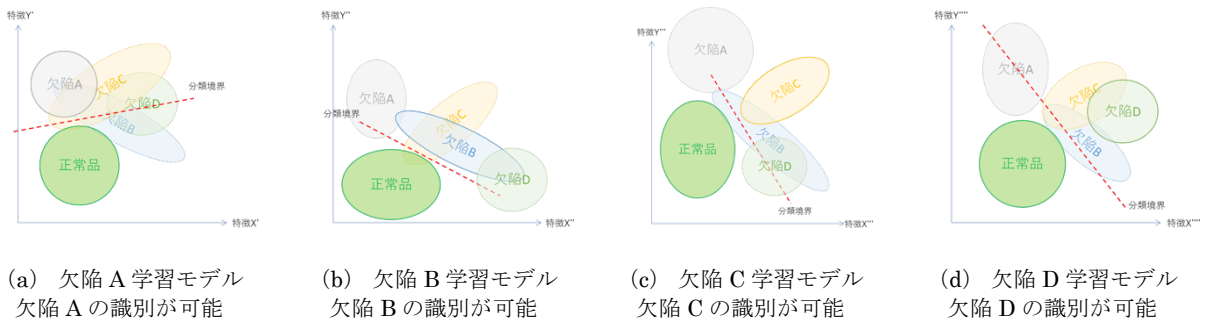


図 2.28 2 クラス分類アンサンブルのイメージ

2.4.1 2 クラス分類アンサンブル手法

実験では、通常が多クラス分類法を用いて、複数のクラスに分類した結果を正常品と欠陥品群に注目し、正常以外に分類されたサンプルすべてを欠陥として集計する多クラス分類結果集約手法と、欠陥種類ごとに 2 クラス分類を行い、いずれかのモデルで欠陥と分類されたサンプルを欠陥として集計する 2 クラス分類アンサンブル手法とで比較を行う。各分類手順を図 2.29 と図 2.30 に示す。

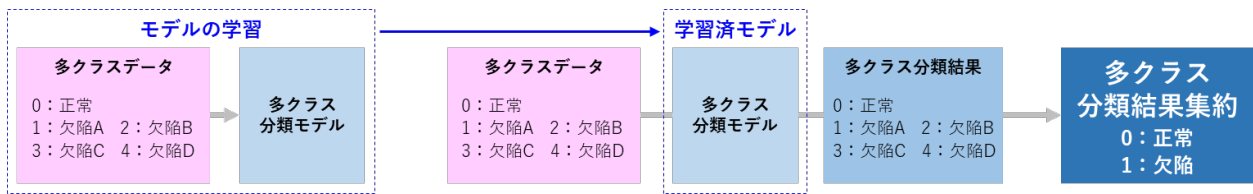


図 2.29 多クラス分類結果集約手法の手順
多クラス分類を行い、正常以外に分類されたサンプルすべてを欠陥として集計する

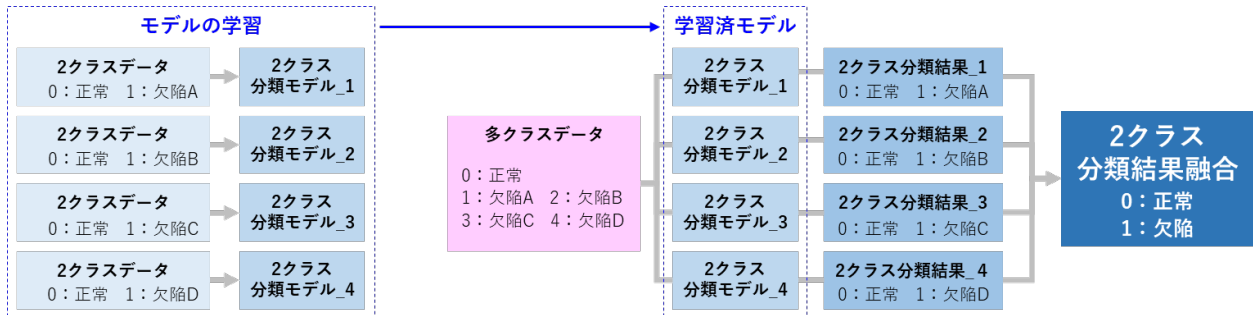


図 2.30 2クラス分類アンサンブル手法の手順
いずれかのモデルで欠陥と分類されたサンプルを欠陥として集計する

関連する研究には Hafiz らの研究[40]がある。Hafiz らは多クラス分類タスクを 2 クラス分類タスクに分解することで正解率を向上させることができ、学習の収束が早いことを利用して学習率を低く設定することでさらに精度を向上させることができると報告している。Hafiz らの研究では手書き数字の分類をターゲットとしており、どの 2 クラス分類モデルの結果が信用できるかを判断するために、ソフトマックスの値を比較し、値が最大となるモデルの結果をサンプルの分類結果として用いている。一方で、本研究は、多クラス分類問題における正常品と欠陥品群の判別をターゲットとしており、2 クラス分類モデルのいずれかで欠陥と判断されたサンプルを欠陥として集計する点で先行研究と異なる。

2.4.2 データセット

2 クラス分類アンサンブル実験は分類実験 1 で使用したデータセットを用いて行った。分類実験 1 とは異なり、データセットのすべての画像から各クラス 20 枚、合計 100 枚からなるサブセットを 8 セット作成し、1つのサブセットを検証データセット、残りを学習データセットとする試験を組み合わせを変えて 8 回行い、8つの試験結果を平均して評価する検証を行った (図 2.31)。多クラス分類モデルを学習する際には、データセットをそのまま使い、2 クラス分類モデルを学習する際には、学習データセットを正常品といずれかの欠陥のみに限定し、検証データには全ての欠陥種を含むデータを用いた。検証データセットには重複がないようにサブセットを作成した。



図 2.31 交差検証の方法

2.4.3 実験結果

2クラス分類モデルにおける欠陥ごとの正解率を表 2.3 に示す。各モデルで正常クラスと担当する欠陥クラスでの正解率は 99.38%から 100.0%であり、多クラス分類のタスクを複数の 2クラス分類に分解することによって、タスクが単純化され、個々のモデルでは非常に高い分類性能となることが確認できた。次に、多クラス分類結果集約手法および 2クラス分類アンサンブル手法の正解率を表 2.4 に示す。正解率はそれぞれ 99.38%、99.63%であった。2クラス分類アンサンブル手法は、多クラス分類結果法と比較して 0.25%の正解率向上が見られた。

2クラス分類アンサンブル手法における学習時の正解率および損失値の推移を確認すると、多クラス分類手法と比較して変動が少なくなっており、安定性が向上していることがみてとれる（図 2.32、2.33）。また多クラス分類におけるいくつかのセットでは、学習の後半で、教師データにおける損失値と検証データにおける損失値が乖離していく過学習が発生している（図 2.34）。過学習とは教師データに過度に適合してしまうことで、本来の目的である未知のデータに対しての汎化性能が低下してしまうことであり、過学習を抑制することは機械学習タスクにおける重要な観点である。一方、2クラス分類の個々のモデルでは、学習時の損失値は、多クラス分類の場合よりも早く収束し、学習の後半に過学習が発生することもなく、損失値の変動も小さくなっていることが確認できる。

表 2.3 欠陥種ごとの 2 クラス正解率

	正解率
2 クラス分類結果_1 : 正常品と欠陥 A (異物)	100.00% (320 件中 320 件)
2 クラス分類結果_2 : 正常品と欠陥 B (キズ)	100.00% (320 件中 320 件)
2 クラス分類結果_3 : 正常品と欠陥 C (打痕)	99.38% (320 件中 318 件)
2 クラス分類結果_4 : 正常品と欠陥 D (めくれ)	99.69% (320 件中 319 件)

表 2.4 2 クラス分類アンサンブル手法の正解率

	正解率
多クラス分類結果集約手法	99.38% (800 件中 795 件)
2 クラス分類アンサンブル手法	99.63% (800 件中 797 件)

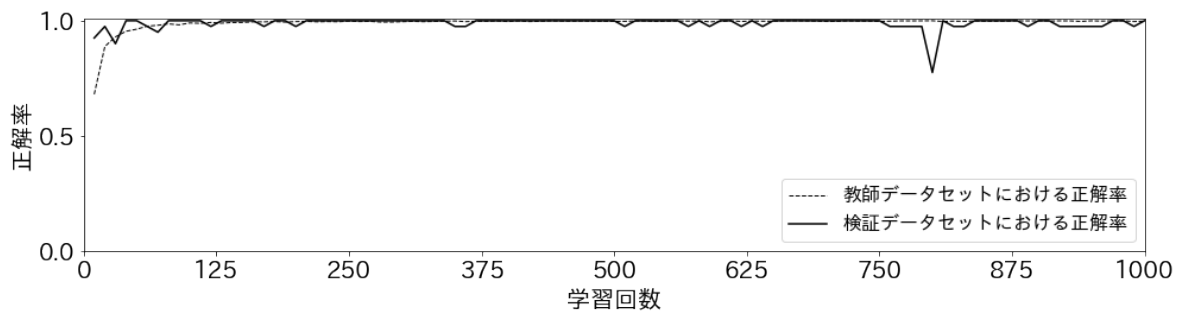


図 2.33 2 クラス分類モデル学習時の正解率推移

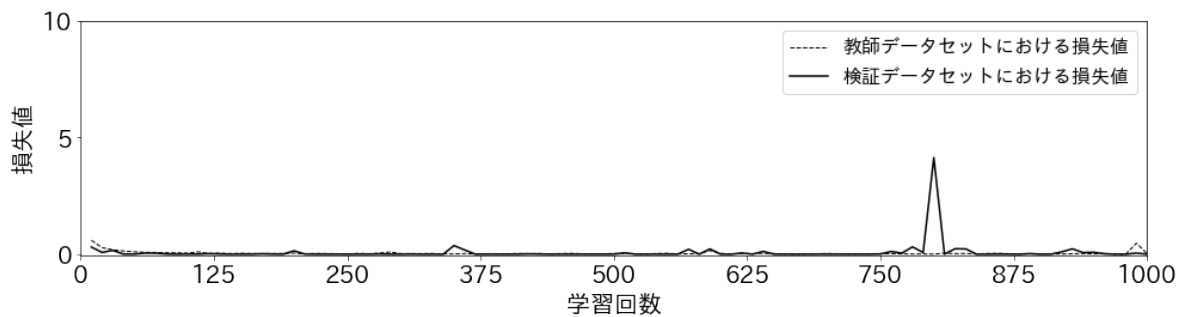


図 2.33 2 クラス分類モデル学習時の損失値推移

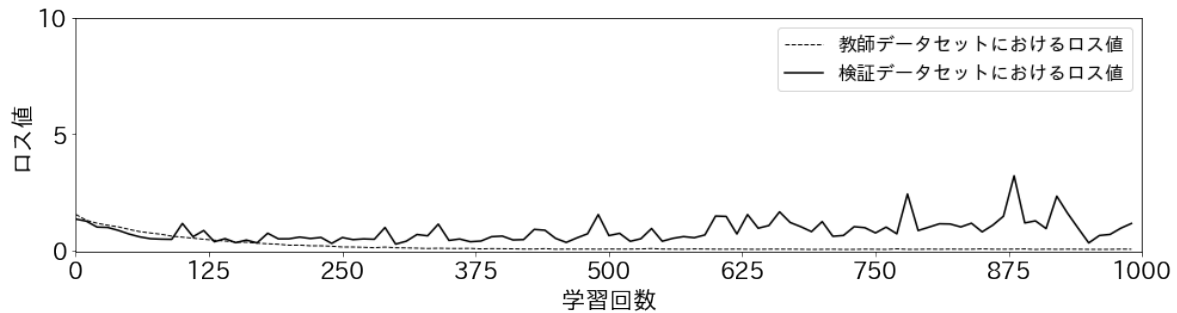


図 2.34 過学習が発生した多クラス分類モデルの学習時損失値推移の例

学習モデルがどのように特徴を抽出しているかを考察するために、t-SNE (t-distributed Stochastic Neighbor Embedding : t 分布型確率的近傍埋め込み)手法[41]を用いて、可視化を行った。t-SNE は Hinton らによって提案された次元削減手法の一つで、次元削減する際に、t 分布を用いることで、高次元空間で近くにあったものは低次元空間ではより近くに、高次元空間で遠くにあったものは低次元空間ではより遠くに表現することができる特徴を持っている。特に高次元データの可視化に適する次元削減手法として、ニューラルネットワークを用いた分類モデルの性能向上に関する研究[42]や CNN を用いた分類性能向上の研究でも活用されている[43]。可視化は、モデルの畳み込み最終層でグローバルアベレージプーリング⁴によって 256 次元に変換された特徴マップを t-SNE 手法によって 2 次元に次元削減したデータをマップすることで行った。畳み込み最終層ではデータセットの画像が持つ特徴を抽出しており、2 次元マップ化することによって各クラスの類似度などの可視化が期待できる。本実験における特徴の 2 次元マップ化は、学習済みのモデルを用いて検証データセットのすべてのデータについて行い、正解クラスごとに色分けした散布図として描画した。多クラス分類モデルの特徴量マップを図 2.35 に示す。図中で、正常クラスは青色破線で囲んだ領域に収まっており、欠陥クラス群が明確に分離できている。正常クラスと欠陥クラス群の 2 値分類として捉えた場合の分類性能が高くなることを視覚的に確認できる。しかし、欠陥クラス群内では、赤色破線で囲んだ領域で各クラスを明確に分離できておらず、このことが学習を不安定にしている原因になっていると考えられる。2 クラス分類の各モデルでの特徴量マップを図 2.36 に示す。2 クラス分類の各モデルでは、正常クラスと欠陥クラスを明確に分離できており、問題が単純化していることが見て取れる。このことによって、各モデルの学習は安定し、高い分類性能を発揮できることから、アンサンブル手法での分類性能向上が期待できる。この可視化結果は、図 2.16、図 2.17、図 2.18 で示した問題単純化のイメージとほぼ同じであり、意図した性能を発揮できていることが確認できた。

⁴ グローバルアベレージプーリング (Global Average Pooling) とは、サイズ $(w \times h \times n)$ として与えられた特徴マップに対して、各チャンネルごとに $(w \times h)$ の平均値を取り、 (1×1) にプーリングすることで、 $(1 \times 1 \times n)$ の出力を得る処理のことを言う。

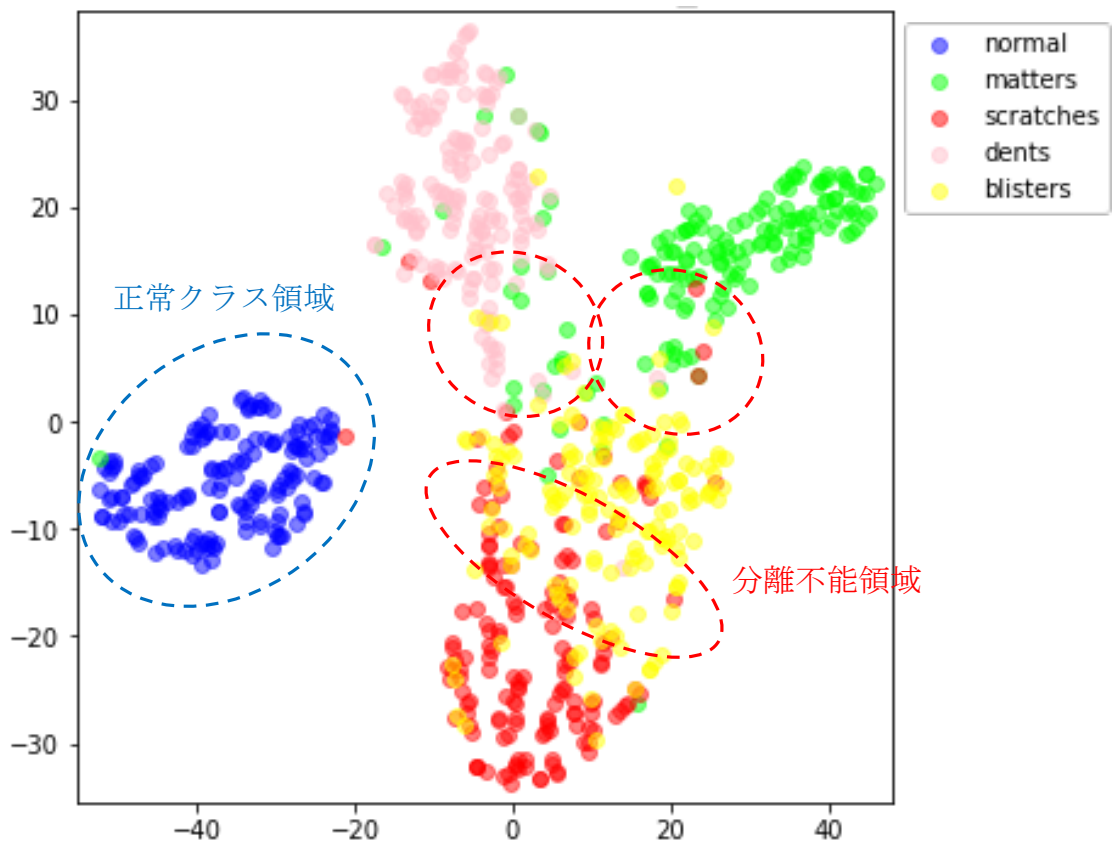


図 2.35 多クラス分類モデルの t-SNE による特徴量 2 次元マップ

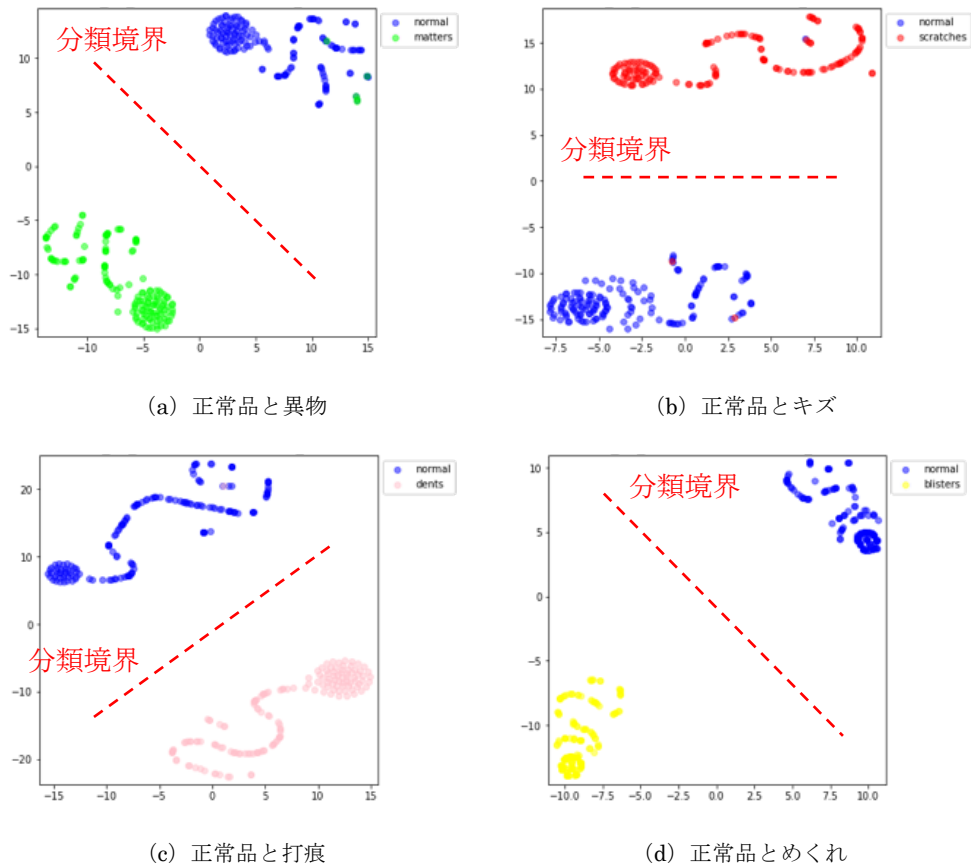


図 2.36 2 クラス分類モデルの t-SNE による特徴量 2 次元マップ

2.4.4 2 クラスアンサンブル手法のまとめ

CNN を用いた外観欠陥検出実験において、欠陥品を取り除くことに注目した 2 クラス分類アンサンブル手法と多クラス分類結果集約手法での正解率と損失値を比較した。2 クラス分類アンサンブル手法は、分類のタスクを分解し、複雑さを回避することで、多クラス分類結果法と比較して正解率を 0.25% 向上でき、学習における正解率安定化の効果が確認できた。また、個々のモデルは過学習が発生しにくく、汎化性能の向上が期待できる。また、実際の運用場面では、生産現場から検査項目の追加、削除を要望されることがある。多クラス分類モデルの場合、検査項目を追加・削除したデータセットで再度学習を行う必要があるが、そのことによって、すべてのクラスに関して分類基準・性能が再学習前と異なってしまう。対して、2 クラスアンサンブル手法は、各モデルが独立しているため、検査項目の追加・削除が容易である上、他のモデルの分類基準・性能に影響しないというメリットがある。

しかしながら、実際のクラッチディスク完成品検査における欠陥クラスは 4 クラスではなく、より多種多様な欠陥を検出する必要があり、分類する欠陥種と同じ数のモデルを必要とする構成では実用面で問題となる。具体的には、運用するモデルが多くなることによるモデル学習時間の増大、

推論時間の増大、モデル管理の複雑化である。また、検査項目の数が多くなった場合に、欠陥種類ごとに 2 クラス分類を行い、いずれかのモデルで欠陥と分類されたサンプルを欠陥として集計する手法では、見逃しを抑制することができる反面、誤警告が増加することが懸念される。そのため、クラッチディスク完成品の外観検査のように多様な欠陥は分類するための手法として適していると言えない。そのため、クラッチディスクの完成品検査における分類性能の向上、学習の安定化のためには、別の手法を検討する必要がある。分類実験 2 (2 クラスアンサンブル手法) で得られた成果と課題は以下のようにまとめられる。

<分類実験 2 (2 クラスアンサンブル手法) で得られた成果>

- 個別のモデルで最適化が可能であるため、分類性能が高い。
- 個別のモデルで学習が安定しており、過学習が発生しにくい
- 検査項目の追加・削除が容易で、他のモデル (検査項目) の分類性能に影響しない。

<分類実験 2 (2 クラスアンサンブル手法) の課題>

- 検査項目が多くなると過検出の発生確率が高くなる可能性がある。
- 検査項目が多くなると、モデル学習のコストが高い。
- 検査項目が多くなると、推論時間が増大する。
- 多くのモデルを管理する必要がある。

2.5 欠陥分類実験 3 : 量産環境での試験設備の導入

分類実験 1 及び分類実験 2 の結果から、限定的な状況ではあるが CNN によってクラッチディスク外観検査を自動化できる可能性を確認できた。しかし、実際の量産工程は、検出が望まれている欠陥は 4 種類よりも多い、取得画像の背景に治具や設備部品などが写り込むという点で、分類実験 1、分類実験 2 とは異なり、より分類が難しくなることが想定される。より量産環境に近い状況での実用性を確認するため、量産における検査工程で検出が望まれている欠陥種を想定したサンプル品を量産同等の環境で撮像した画像から作成したデータセットを用いて、分類実験 3 を行った。検査項目となる欠陥種類が多いため、分類実験 2 で開発した 2 クラスアンサンブル手法は用いず、分類実験 1 での問題点となっていた学習時の正解率の変動を抑えるための対策として、パラメータの削減、学習率の見直しを観点にモデルの再構成を行った。

2.5.1 画像取得方法

サンプル画像は製造ラインで用いられている摩擦材位置検査装置を用いて実際のクラッチディスク完成品を撮像することで作成された。撮像機器には 4,000×3,000 画素モノクロカメラ、照明には分類実験 1 で用いたものと同じ外径 300mm、内径 252mm の白色 LED リング照明を用いた。カメラおよび照明は、検査対象となる直径 160mm のクラッチディスクからの距離 360mm、25mm の位置に設置した (図 2.37)。撮像機器及び検査対象は外乱光の影響を抑えるために簡易的な暗箱の中に設置した。撮像は検査対象が治具上で静止している状態で行われ、クラッチディスク全体と背景に

治具や設備部品を含んだ撮像視野 $333 \times 250 \text{mm}$ の 256 階調グレースケール画像を取得した。取得した画像から、ペーパー摩擦材を貼り付けた部分の近傍を切り出した 640×300 画素の画像を 1 枚のサンプル画像とし、クラッチディスク中心を軸に回転させながら、同様の処理を行うことで 1 枚の全体画像から 10 枚のサンプル画像を切り出した。なお、本実験で用いたクラッチディスクは分類実験 1 で用いたものとは異なる形状をしている。

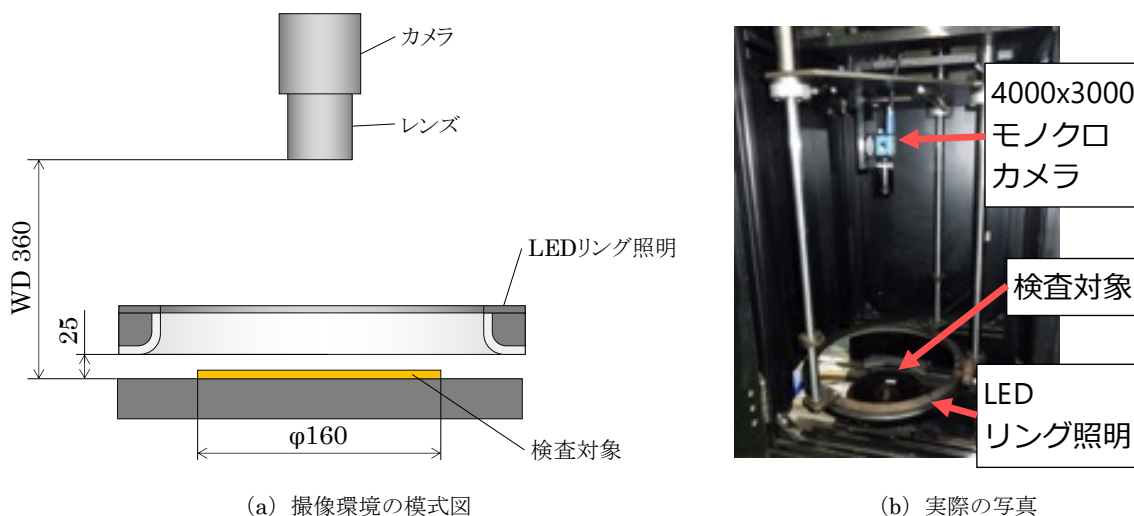
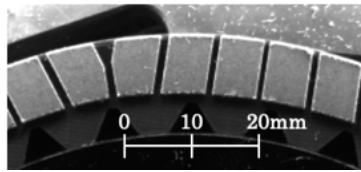


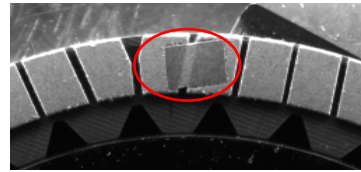
図 2.37 分類実験 3 の撮像環境

2.5.2 データセット

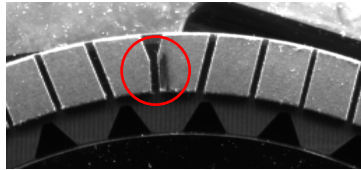
データセットとして正常品と 15 種の欠陥品を含む 16 クラスのサンプル画像が合計 7,102 枚収集された。データセットのうち 2,027 枚は正常品サンプル画像、5,075 枚は欠陥品サンプル画像である。分類実験 3 における欠陥品サンプルも、分類実験 1 と同様に正常なクラッチディスクに人為的なダメージを与えることで作成された。クラッチディスクのペーパー摩擦材に発生する外観に関わる欠陥は種類が多様であるが、実際の量産で検出が望まれている欠陥 15 種をサンプルとして用意した。各クラスのサンプル画像例を図 2.38 に示し、欠陥箇所を赤枠で囲っている。データセットに含まれる画像はエンジニアによって、不適切な画像が含まれていないか、属するクラスが正しいかの確認が事前に行われている。クラス分けは目視検査において検出が望まれるレベルを基準に行われた。これらの画像を 5 つのサブセットに分割し、そのうちの 4 つのサブセットを教師データセットとして学習に用い、残りの 1 つのサブセットを検証データセットとして検証を行う試験を、組み合わせを変えて 5 回行った後、5 回の試験結果を集計し、評価する交差検証を行った (図 2.39)。各データセット群には各クラスのサンプルが同じ割合で含まれるように調整した。データセットに含まれるクラスと各クラスのデータ数およびサブセット 1 における教師データと検証データの数を表 2.5 に示す。



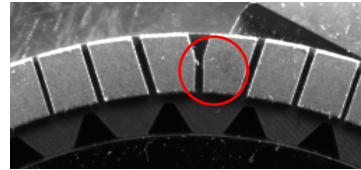
(a) 正常



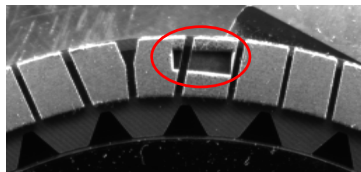
(b) コマ過多



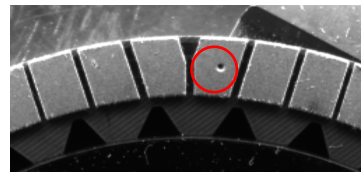
(c) 汚れ



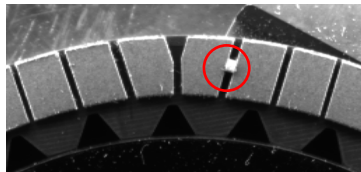
(d) オイル滲み



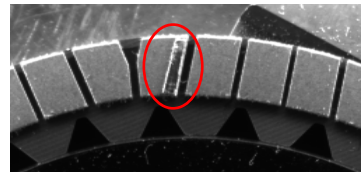
(e) 打痕:大



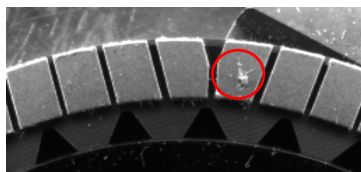
(f) 打痕:小



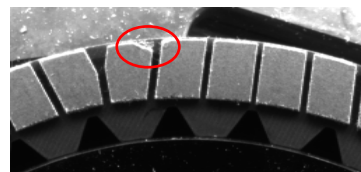
(g) 異物



(h) 異物

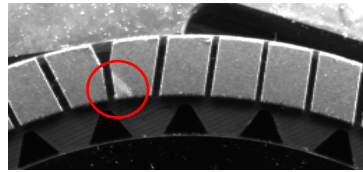


(i) キズ

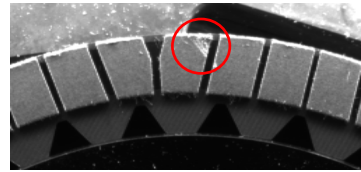


(j) 欠け

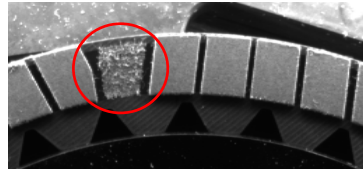
図 2.38-1 サンプル画像の例(1)



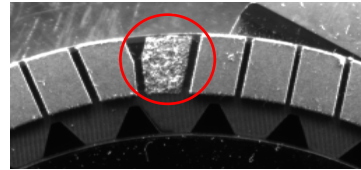
(k) めくれ



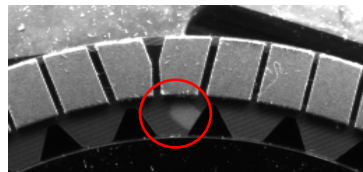
(l) 削れ



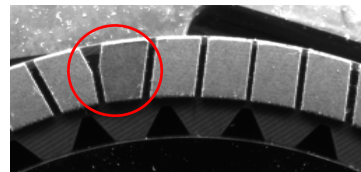
(m) 剥がれ



(n) 剥がれ



(o) 接着液カスレ



(p) 2枚重ね

図 2.38-2 サンプル画像の例(2)

	サブセット 1	サブセット 2	サブセット 3	サブセット 4	サブセット 5	
試験1回目	学習	学習	学習	学習	検証	⇒ 試験結果 1
試験2回目	学習	学習	学習	検証	学習	⇒ 試験結果 2
試験3回目	学習	学習	検証	学習	学習	⇒ 試験結果 3
試験4回目	学習	検証	学習	学習	学習	⇒ 試験結果 4
試験5回目	検証	学習	学習	学習	学習	⇒ 試験結果 5

図 2.39 交差検証

表 2.5 分類実験 3 用データセットのクラスとサンプル数

Index	クラス名	総サンプル数	サブセット 1 の教師データ数	サブセット 1 の検証データ数
0	正常 (欠陥なし)	2,027 枚	1,621 枚	406 枚
1	コマ過多	401 枚	320 枚	81 枚
2	汚れ	424 枚	339 枚	85 枚
3	オイル滲み	418 枚	334 枚	84 枚
4	打痕 (大)	421 枚	336 枚	85 枚
5	打痕 (小)	415 枚	332 枚	83 枚
6	異物 1	379 枚	303 枚	76 枚
7	異物 2	422 枚	337 枚	85 枚
8	キズ	431 枚	344 枚	87 枚
9	欠け	402 枚	321 枚	81 枚
10	めくれ	419 枚	335 枚	84 枚
11	削れ	422 枚	337 枚	85 枚
12	剥がれ 1	48 枚	38 枚	10 枚
13	剥がれ 2	53 枚	42 枚	11 枚
14	接着液カスレ	373 枚	298 枚	75 枚
15	2 枚重ね	47 枚	37 枚	10 枚
合計		7,102 枚	5,674 枚	1,428 枚

2.5.3 学習モデル

分類実験 1 と同様、学習モデルには CNN アーキテクチャを用い、入力画像を正常もしくは何れかの欠陥種に分類する多クラス分類モデルを構成した。モデル構成及びハイパーパラメータは分類実験 1 と同様、先行研究[35]で提案されている VGGNet を参考とし、分類実験 3 用データを用いた予備実験での調整を経て、本実験用に再構成した。分類実験 1 で確認された学習中の正解率の変動が大きいことについての対策として、モデルのパラメータを削減するための入力画像のサイズダウンと全結合層への入力次元数の削減及び学習率を小さくすることが反映されている。モデルは入力にサイズが 256×256 の画像を受け取り、10 層の畳み込み層と 2 層の全結合層を経て、入力画像が属するクラスの確率を予測ラベルとして返す構成とした。分類実験 1 と同様、畳み込み層は全層でカーネルサイズ 3×3 、ストライド 1 とし、活性化関数に ReLU を用いた。出力する特徴マップは畳み込みを行うごとに 32 枚、64 枚、128 枚と増やし、畳み込みの最終層では 256 枚の特徴マップを出力させた。畳み込み層の途中には 2×2 のマックスプーリング層が特徴マップを小さくするために用いられた。畳み込みの最終層ではグローバルアベレージプーリング[44]を用いて 256 次元の特徴量として抽出した。畳み込み最終層の特徴マップを全結合層に入力するために 1 次元に変換した後、2 層の全結合層を経て、最終層でソフトマックス関数を用いて 16 次元に変換し、その最大値をもつクラス

を予測クラスとして用いた。全結合層では正則化効果を得るためにドロップアウトを用いた。モデルは損失関数にカテゴリカルクロスエントロピー、最適化アルゴリズムに **RMSProp** を用い、学習率は学習中の変動が大きくなりすぎないように 1×10^{-4} に調整した。モデルの総学習パラメータ数は 2,403,312 となり、分類実験 1 で使用したモデルの 1/3 程度となった。分類実験 3 の学習モデル構成を図 2.40 に示す。分類実験 1 と同様に、教師データセットの画像はモデルに入力される前に、元画像に対し回転、左右反転、左右移動、上下移動、剪断処理、輝度シフトの加工をランダムに行う画像増強がわれた。学習は教師データセットの画像約 5,600 枚を 300 回繰り返し学習した。

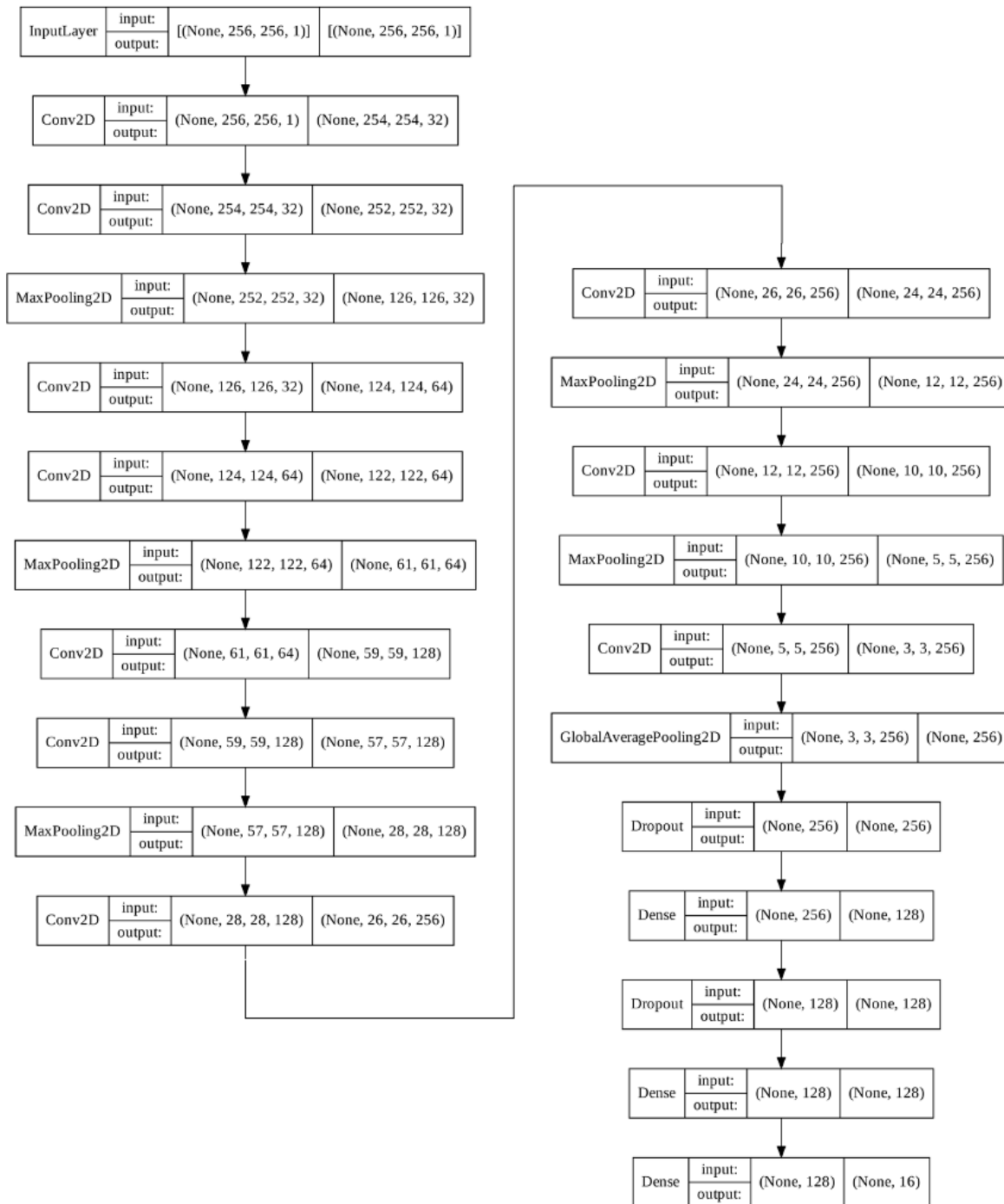


図 2.40 分類実験 3 の学習モデル構成

2.5.4 実験結果

分類実験 3 の結果を表 2.6 に示す。5 つのサブセットの平均で 95.69% の正解率となった。具体的には、5 つのサブセットの合計 7,102 件の検証データのうち、6,796 件を正しく分類できた。分類実験 1 と比較して、分類すべきクラス数は 5 クラスから 16 クラスと約 3 倍となっているうえ、画像は検査対象以外にも背景に治具や設備が写り込んでいるため、分類難易度は向上していると思われるが、正解率は 90% (100 件中 90 件) よりも良好な結果となっており、量産工程での実用性を確認できる結果となった。

分類実験 1 での課題であった正解率の変動に関して、対策によって安定性が改善されていることが確認できた。対策として、モデル構成時に学習パラメータの削減、学習率の見直しを行った。実験結果の平均値に一番近い正解率であったサブセット 4 における学習時の正解率および損失値の推移を図 2.41、図 2.42 に示す。分類実験 1 では学習全体の 3/4 を過ぎても正解率の変動が約 16% あったのに対して、対策を実施した分類実験 3 では同じく学習の 3/4 を過ぎた学習回数 225 回以降においても正解率の変動が 5.5% に収まっている。同様に損失値についても変動が抑えられており、学習終了のタイミングによってモデルの精度が極端に悪化することはなくなった。

表 2.6 分類実験 3 での正解率

	正解率
試験結果 1	93.42% (1,428 件中 1,334 件)
試験結果 2	97.33% (1,425 件中 1,387 件)
試験結果 3	94.65% (1,420 件中 1,344 件)
試験結果 4	96.19% (1,416 件中 1,362 件)
試験結果 5	96.89% (1,416 件中 1,372 件)
集計	95.69% (7,102 件中 6,796 件)

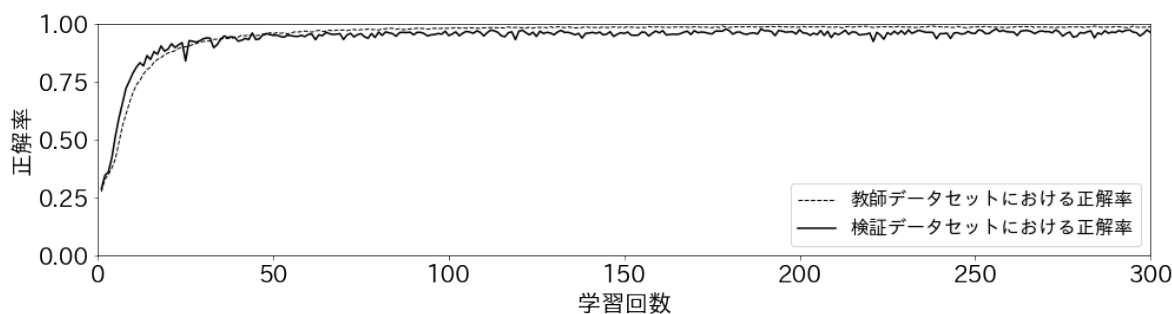


図 2.41 分類実験 3 での正解率推移 (サブセット 4)

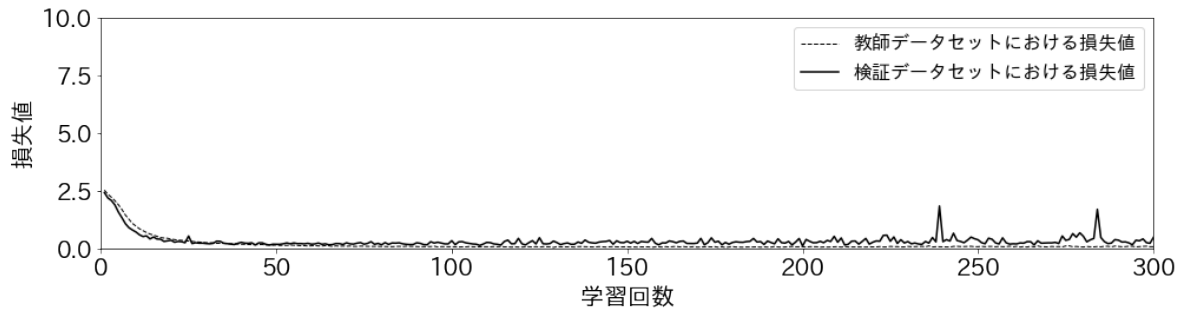


図 2.42 分類実験 3 での損失値推移 (サブセット 4)

分類の詳細について確認するため、サブセット 4 における検証データの正解クラスと予測クラスを混同行列にしたものを図 2.43 に示す。混同行列からは、以下の 3 つについて確認できる。1)めくれを削れに誤認するケースが目立つ、2)正常品とオイル滲みは相互に誤分類しやすい、3)欠けを正常と誤認しやすい。特にオイル滲みについて、欠陥は目視でも発見が難しく、正常品に現れる色ムラと正確に区別することが難しいため、目視検査においても検出が難しい項目となっている。

正解クラス	0_正常	1_コマ過多	2_汚れ	3_オイル滲み	4_打痕：大	5_打痕：小	6_異物1	7_異物2	8_キズ	9_欠け	10_めくれ	11_削れ	12_剥がれ1	13_剥がれ2	14_接着液カスレ	15_2枚重ね
0_正常	396	0	0	3	0	0	2	0	1	0	2	1	0	0	0	0
1_コマ過多	0	78	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2_汚れ	0	0	85	0	0	0	0	0	0	0	0	0	0	0	0	0
3_オイル滲み	4	0	0	79	0	0	0	0	0	0	0	0	0	0	0	0
4_打痕：大	0	0	0	0	84	0	0	0	0	0	0	0	0	0	0	0
5_打痕：小	0	0	0	0	0	83	0	0	0	0	0	0	0	0	0	0
6_異物1	0	0	0	0	0	0	76	0	0	0	0	0	0	0	0	0
7_異物2	0	0	0	0	0	0	0	82	0	0	1	1	0	0	0	0
8_キズ	0	0	0	0	0	0	0	0	85	0	0	1	0	0	0	0
9_欠け	4	0	0	0	0	0	1	0	0	75	0	0	0	0	0	0
10_めくれ	2	0	0	0	0	0	0	0	2	0	58	22	0	0	0	0
11_削れ	0	0	0	0	0	0	0	0	0	0	1	83	0	0	0	0
12_剥がれ1	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0
13_剥がれ2	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0
14_接着液カスレ	2	0	0	0	0	0	0	0	0	0	0	0	0	0	72	0
15_2枚重ね	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	7

図 2.43 混同行列 (分類実験 3 サブセット 4)

混同行列で確認できた、3つの事柄について、特徴量 2 次元マップ上でも同様に確認することができる (図 2.44)。1)めくれと削れはマップ上でのクラスの分布が重なっており、特徴が類似しているために、相互に誤分類される可能性が高いことが見て取れる。同様に、2)正常品とオイル滲みは相互に誤分類しやすいこと、3)欠けを正常と誤認しやすいことについても、各クラスの分布が近く、特徴が似通っていることが確認できる。一方で、コマ過多や打痕 (大)、剥がれは、正常品クラスと

は分布が離れているが、これらは人間の目で見ても比較的区別が容易な欠陥クラスとなっている。このことは、CNN によって抽出された特徴は人間の感覚と似通っていることを示していると考えられ、目視による外観検査の自動化技術として CNN の有効性を確認できた。

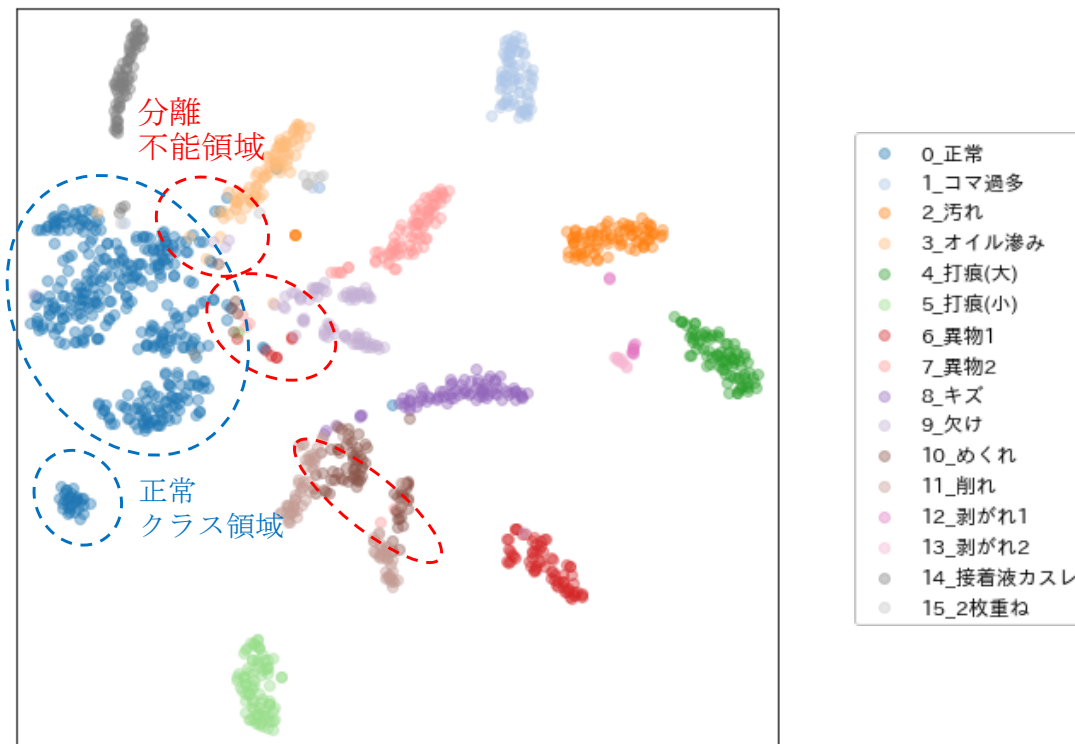


図 2.44 t-SNE での特徴量 2 次元マップ (サブセット 4)

しかしながら、外観検査装置を実際の検査工程へ導入するにあたってはさらなる精度向上が求められる。特に外観検査では、品質保証能力の確保のため、実際には欠陥がある製品を正常品と判定してしまう「見逃し」を極力減らすことが重要である。同時に、生産効率の観点では正常品を欠陥品と判定してしまう「誤警告」の最小化も求められる。分類実験 3 における見逃しの発生件数は欠陥クラス 5,075 件中 82 件で 1.62%、誤警告の発生件数は正常クラス 2,027 件中 38 件で 1.87%となっている (表 2.7)。また、t-SNE を用いた 2 次元特徴量マップを見ても、現状は図 2.44 で示したように正常品と欠陥品群の分布を明確に区分することが難しく、実用の場面においても見逃し及び誤警告の発生が避けられないと想定される。実用化に向けて分類性能を向上させるには、検証データにおける見逃しや誤警告の件数だけでなく、特徴マップにおける分布の様子にも注目し、特徴マップ上の正常品と欠陥品群の分布に明確な区分を持たせなければならない。

表 2.7 分類実験 3 での分類性能

	正解率	見逃し率	誤警告率
実験結果 1	93.42%	2.64% (1,022 件中 27 件)	0.74% (406 件中 3 件)
実験結果 2	97.33%	1.18% (1,019 件中 12 件)	0.99% (406 件中 4 件)
実験結果 3	94.65%	2.27% (1,015 件中 23 件)	3.70% (405 件中 15 件)
実験結果 4	96.19%	1.19% (1,011 件中 12 件)	2.22% (405 件中 9 件)
実験結果 5	96.89%	0.79% (1,008 件中 8 件)	1.73% (405 件中 7 件)
集計	95.69%	1.62% (5,075 件中 82 件)	1.87% (2,027 件中 38 件)

分類実験 3 で得られた成果と課題は以下のようにまとめられる。

<分類実験 3 で得られた成果>

- 平均で **95.69%** の正解率を達成した。
- 実験 1 よりも検出する欠陥項目を増やしても分類性能を維持できた。
- 撮像画像の背景に治具などの映り込みがあっても分類性能を維持できた。
- 量産想定環境においても、実用性が確認できた。

<分類実験 3 の課題>

- 検査項目のいくつかで正常品との明確な区別が難しい。
- 品質保証能力を確保し、実用化するためには、見逃しの抑制が必要となっている。
- 生産効率低下を避けるために、誤警告の抑制が必要となっている。

2.6 本章のまとめ

CNN を用いたクラッチディスクの外観検査を模擬した分類実験 1 を経て量産環境を想定した分類実験 3 を行い、一定の検査性能を達成することができた。一般的なルールベース画像検査装置では、すべての多様な欠陥に対して人手でルールを定めることに多大な労力を必要とするため、完成品の外観検査を自動化することが現実的でなかったが、CNN を用いた学習を行うことで、量産想定環境においても、多様な欠陥に対して高い精度を達成することができた。クラス分類の正解率は **95%** を超えており、完成品の外観検査が人作業となっていることで収集できていなかった生産状況の情報を収集・分析し、製品価値、生産性、品質の向上のために活用することが期待できる。同時に、人作業の課題である品質保証能力の安定化、過酷な作業からの解放へ繋がることも期待できる。本実験で得られた結果は、人作業となっている完成品の外観検査工程の自動化を実現する技術として有望なものとなった。

しかしながら、実用化のためには、検査工程の重要な機能である「見逃しの抑制」に焦点を当て、見逃しを極力少なくするための新しい技術の開発が必要となっている。

第3章 見逃し抑制に焦点を当てた新手法の開発

CNN を用いたクラッチディスクの外観検査を模擬した実験では、一定の検査性能を達成することができた。検査工程を自動化するにあたって要求される機能には検査スピードの向上、実働時間の拡大、人的コストの削減など生産の高効率化につながる機能や、設計・製造工程への情報フィードバックといった継続的な改善につなげるための機能など様々である（図 3.1）。その他にも検査員の雇用確保、人材育成といった管理的課題を解決することも自動検査機導入の目的となっている。その中でも、自動検査機に求められる最も重要な機能は、欠陥品を正常と誤認する「見逃し」の最小化であり、導入を判断するにあたって非常に重要な指標となる。検査工程そのものは顧客に対する品質保証を強化する目的で設置されており、品質保証能力をおろそかにしたままでは、生産効率を向上させる機能を持っていても検査装置としての意味を持たない。スマートファクトリー構築のための人に頼った作業の自動化を推し進めるためには、CNN を用いた外観検査装置における見逃しの最小化は至上命題である。

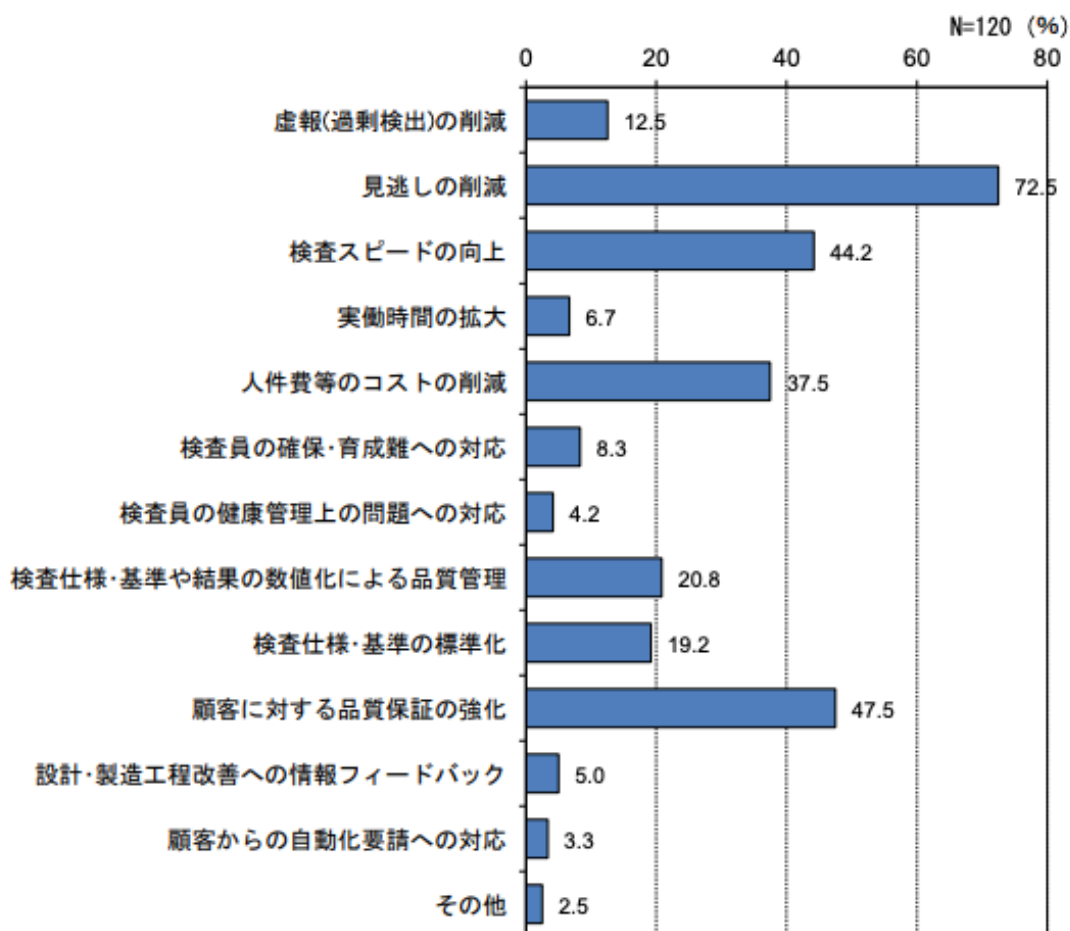


図 3.1 検査工程自動化の目的
(引用：ものづくり企業の生産現場における 検査の自動化促進可能性調査 p40 図 2.35)

ルールベースの画像検査システムで見逃しを抑制するための手段には、画像前処理方法の変更や判定に用いる特徴量の変更が考えられるが、最も一般的に用いられているのは閾値の調整である。閾値は数値であり、その大小関係が明確で理解しやすい上、設定変更が容易であることから、検査性能向上の手段としてよく用いられ、エンジニアによって緻密なチューニングが行われる。しかし、見逃しを抑制しようとするあまり、欠陥の可能性があるものを積極的に検出できるように閾値を操作した結果、正常品を欠陥品と誤認する「誤警告」が多発してしまうことがよくある。誤警告の発生は生産ラインの効率を悪化させ、人による再確認の工数を発生させるなどの悪影響をもたらす。欠陥が明確である場合や、種類が少ない場合は望んでいる検査性能を得られることもあるが、欠陥種類が多く、欠陥そのものが多様性を有するような場合は、判断基準が複雑になり、非常に多くのルールを設定する必要があるため、閾値の調整によって望んでいる検査性能を得ることは容易でない。検査工程の重要な機能である見逃しの抑制を行うためには、トレード・オフの関係にある誤警告の発生にも注意を向けなければならないが、その両立は容易ではない。このことはルールベース画像検査機における長年の課題であり、新しい技術によって解決が望まれている。

CNN を用いた外観検査の分類性能向上を目的とした取り組みは多数報告されており、それらは大きく、アルゴリズムの改善によるもの、学習データの改善によるもの、取得画像の改善によるものがある。外観検査の分類性能向上のためのアルゴリズムとして、Staar らによる深層距離学習ネットワーク[45]を用い、テクスチャ表面異常の識別能力を向上させる手法[46]、Tao らによるオートエンコーダーネットワーク[18]を連結することで、欠陥の位置特定とクラス分類の両方のタスクで精度を向上させる手法[47]、Hafiz らによる多クラス分類タスクに 2 クラス分類器のアンサンブルを用いることによって、学習の収束を早くし、学習率を下げることで分類性能を向上させる手法[40]、半澤らによるモデルの最終出力を **Feature Map** とすることで、学習画像に存在しない未知のパターンに高い対応能力を持った手法[48]などが提案されている。データの正確性に着目した手法として、Prez らは高次元のデータを **t-SNE** を用いて 2 次元空間に埋め込み、四分位範囲 (**IQR**) を用いて学習データセット内の外れ値を取り除くことによって、欠陥の検出性能を向上させる手法[43]を提案している。また、Aghaei らはオフラインの複数照明によって撮像された画像を学習することで、インラインの単一照明でも表面欠陥の検出性能向上を実現する手法を提案している[49]。CNN を用いた外観検査技術は日々進化しており、検出精度向上など一定の成果が達成されているが、欠陥品の見逃し抑制に着目し、ルールベース手法では困難だった見逃し抑制と誤警告抑制の両立を実現するための手法は報告がない。

3.1 見逃し抑制に焦点を当てた新手法の開発

3.1.1 新手法開発の仮説

CNN 多クラス分類モデルの予測結果詳細は予測ラベルとして確認することができる。予測ラベルとは多クラス分類モデルにおいて、入力データがどのクラスに属するのかを確率分布として表現したもので、一般的に、最大の確率を持つクラスが予測クラスとして出力される。分類実験 3 で得ら

れた予測ラベルの詳細を確認すると、見逃しとなったサンプルの多くで欠陥の可能性を示す確率を持っていることが確認できた (図 3.2)。しかしながら、これらは正常の可能性を示す確率が最大となっているため、正常品と判定されている。この欠陥可能性を積極的に検出することができれば、見逃しの発生を抑制できる可能性がある。

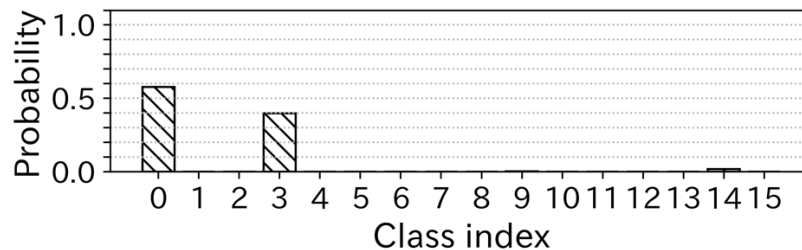


図 3.2 欠陥分類実験 3 から得られた見逃しサンプルの予測ラベルの例

3.1.2 正解ラベルの付与方法による CNN モデルの性格誘導

本研究では CNN を用いて多様な欠陥を検出するために、傷、変形、浮き上がりなどの欠陥種類毎に分類を行う多クラス分類法を適用している。CNN 多クラス分類法の学習過程で用いられる教師データの各サンプルには、分類されるべきクラスを示す正解ラベルが事前に付与されている。学習過程ではモデルに教師データを入力することで、その教師データが各クラスに分類される確率が予測ラベルとして計算される (図 3.3 (a))。次に、正解ラベルと予測ラベルから予測の誤差の大きさを示す損失値が計算される (図 3.3 (b))。そして、教師データの全サンプルに対する損失値が最小となるようにモデルが更新されることで学習が進んでいく (図 3.3 (c))。

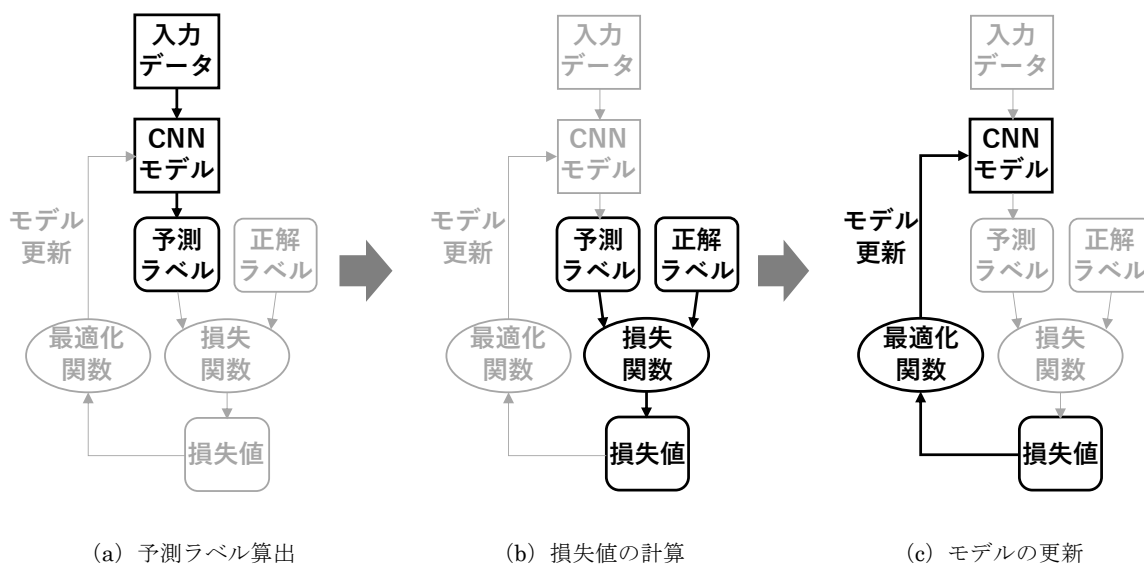
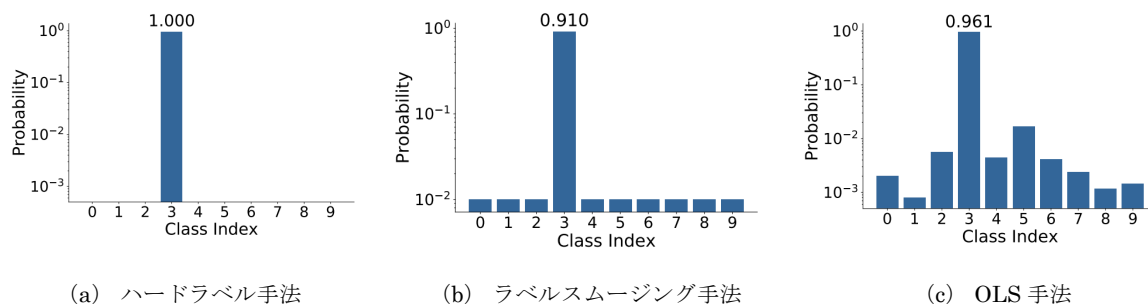


図 3.3 CNN の学習過程模式図

CNNを用いた多クラス分類モデルが学習する上での正解ラベルは、モデル更新の方向性を決定づける重要な要素となる。一般的な正解ラベルの付与方法はハードラベルと呼ばれ、そのサンプルが分類されるべき正解クラスに対応する要素のみが値を持つ確率分布として与えられる（図 3.4 (a)）。この正解ラベルの付与方法を変更することでモデルの分類性能を向上させる手法として、ラベルスムージング手法[50]が開発されている。ラベルスムージング手法では、すべてのクラスに一定の確率を与え、正解クラスに残りの確率を与えることで最も高い確率を持たせる（図 3.4 (b)）。ラベルスムージング手法を用いることで、モデルが過度に学習データに適合することを抑制でき、ハードラベル手法と比較して分類性能を向上させることができると報告されている[51][52][53]。また、ラベルスムージング手法を応用し、学習過程で正解ラベルを教師データセット内のクラス類似度に合わせて動的にスムージングすることによって画像分類の性能を向上させる OLS (Online Label Smoothing) 手法[53]が提案されている。OLS 手法は、学習の過程でモデルの予測ラベルに基づいて正解ラベルの確率分布を再構築する。OLS 手法によって学習されたモデルの教師データの正解ラベルは、例えば「猫」が正解クラスであった場合には「自動車」クラスよりも「犬」クラスにより多くの確率を持つようになる（図 3.4 (c)）。モデルの予測ラベルは、異なるクラス間の関係性を示していることがわかっており[55]、OLS 手法で学習したモデルの予測分布は、人間の予測分布に近く、より合理的で正しいクラス間の関係性を示していると主張されている。つまり、CNN モデルが分類性能を向上させるために正解ラベルを再構築した結果、人間の感覚と同様の分布となったということである。このことは逆にとれば、正解ラベルの付与方法によって、CNN モデルの性格を誘導することができるのではないかと考えられる。なお、本研究では CNN モデルに関して「特性」ではなく、「性格」と言う表現を用いている。これは、「特性」というほど厳密には表現できない、人間の持つ曖昧さを獲得していることを表現するためである。



0 - airplane	1 - automobile	2 - bird	3 - cat	4 - deer	5 - dog	6 - frog	7 - horse	8 - ship	9 - truck
--------------	----------------	----------	---------	----------	---------	----------	-----------	----------	-----------

図 3.4 正解ラベルの付与方法（正解クラスが 3 のケース）
 (引用：Delving Deep into Label Smoothing [53] Fig. 1 をもとに作成)

3.1.3 検査工程における非対称性の導入

検査員による目視の外観検査では、欠陥品の見逃しは極力避けたい。そのため、検査員は目視検査の際に、検査対象に「欠陥がないか」という観点を持ち、いずれの欠陥種にも属さず、確実に正常品と判断できるサンプルである場合のみ正常品として分類する一方、欠陥の存在が疑われるサンプルは積極的に欠陥品として分類することがよく行われている。このことは、正常品発見プロセスと欠陥品発見プロセスの違いとして考えることができる。正常品発見プロセスはいずれの検査項目にも該当しない場合のみ正常品とされる、受動的・消極的な発見プロセスであるのに対し、欠陥品発見プロセスはいずれかの検査項目に該当する可能性がある場合に欠陥品とされる、いわば能動的・積極的な発見プロセスである。本論文ではこのプロセスの違いを正常品と欠陥品の取り扱い非対称性と呼ぶこととする。本研究では、検査員による目視検査で見逃し発生を避けるため経験的に取り入れられている正常品と欠陥品の取り扱い非対称性に着目し、CNN を用いた画像検査装置に非対称性を取り込むことで、従来の閾値調整ではトレード・オフの関係にある見逃し抑制と誤警告抑制を両立した実用的な検査システムの開発を目指す。具体的には、CNN の学習過程で用いられる正解ラベルを操作することで、CNN に欠陥品を積極的に検出する機能、つまり検査員が経験的にしている外観欠陥検出における正常品と欠陥品の取り扱い非対称性を付与する手法の開発を試みる。

3.1.4 非対称ラベルスムージング手法

今、 K 個のクラスが存在する問題に対し、教師データである入力画像 x_i と対応する正解ラベル y_i からなるデータセット (x_i, y_i) が与えられたとき、各サンプル (x_i, y_i) に対して、モデルはクラス k に分類される確率 $p(k|x_i)$ を計算する。 y_i の分布を $q(k|x_i)$ と表すと損失値として用いるクロスエントロピー L を次のように書くことができる。

$$L = - \sum_{k=1}^K q(k|x_i) \log p(k|x_i) \quad (3.1)$$

正解ラベル y_i は通常、そのサンプルが分類されるべき正解クラスに対応する要素が 1、他のクラスに対応する要素は 0 となる数列として与えられる。この手法はハードラベルと呼ばれる。これは正解クラスに分類されるべき確率が 1 で、他のクラスに分類されるべき確率が 0 となる確率分布と考えることができる (図 3.5)。正解ラベル y_i として与えられる分布 $q_{hard}(k|x_i)$ は次式で表される。

$$q_{hard}(k|x_i) = \begin{cases} 1, & k = \text{ground truth class}(y_i) \\ 0, & k \neq \text{ground truth class}(y_i) \end{cases} \quad (3.2)$$

ハードラベルによる学習では教師データのサンプルが正解クラス以外に分類されるべき確率が 0 とされており、極力、正解クラスに分類されるように学習される。しかしながら、そのことにより、モデルが教師データに過度に適合した、過学習と呼ばれる状態が生じてしまうことがある。すなわち、教師データにノイズが含まれていたり、正解ラベルの設定に誤りが含まれていたりするような場合でも、教師データに極力、適合するように学習が行われた結果としてテストデータでの分類性

能低下が引き起こされる。過学習を抑制し分類性能を向上させるため、データ増強[13]、ドロップアウト[36]、カットアウト[56]など、さまざまな手法が提案されている。なかでも正解ラベルを用いたシンプルな手法として、ラベルスムージング手法[50]がある。ラベルスムージング手法では、すべてのクラスに一定の確率を与え、教師データが分類されるべき正解クラスに残りの確率を与えることで最も高い確率を持たせる（図 3.6）。本論文では提案手法との区別のため、この手法を一様ラベルスムージング手法と呼ぶことにする。一様ラベルスムージング手法は正解クラス以外に確率を持つことを許容した学習が行われるため、データノイズやラベルミスによる過学習を抑制する効果が期待できる。一様ラベルスムージング手法における正解ラベル y_i の分布 $q_{uniform}(k|x_i)$ は次式で表される。

$$q_{uniform}(k|x_i) = (1 - \epsilon)q_{hard}(k|x_i) + \frac{\epsilon}{K} \quad (3.3)$$

ここで、 ϵ は平滑化パラメータと呼ばれ、0 から 1 までの値を取り、確率としてすべてのクラスに等しく分配される。実際には 0.1 などの値が用いられることが多い[50]。

どちらの手法もすべてのクラスは対等に扱われるが、欠陥検査の場面では正常品クラスとそれ以外の複数の欠陥品クラスが存在し、これらを非対称に扱う必要がある。そこで、欠陥検査工程の持つ非対称性の概念を取り入れることを目的とし、一様ラベルスムージング手法を改良した新たな手法を開発した。開発した手法は、正常品の教師データに対する正解ラベルと、何れかの欠陥を持つ欠陥品の教師データに対する正解ラベルとで付与方針を別にしてしている。正常品教師データにはハードラベルと同様に、分類されるべき正常クラスの確率のみを持ち、それ以外の欠陥クラスには確率を持たない正解ラベルを付与する（図 3.7 (a)）。一方で、欠陥品教師データの正解ラベルには正常品の可能性を示す正常クラスの確率を持たず、分類されるべき欠陥クラスに最も高い確率を持ちつつ、それ以外の各欠陥クラスにも一定の確率を持つ正解ラベルを与える（図 3.6 (b)）。正解ラベル y_i の分布 $q_{asymmetric}(k|x_i)$ は教師データが正常品のときと欠陥品のときで別々の式で表される。正常品教師データに与えられる正解ラベルは次式で表される。

$$q_{asymmetric}(k|x_i) = q_{hard}(k|x_i) \quad (3.4)$$

一方で、欠陥品教師データに与えられる正解ラベルは次式で表される。

$$q_{asymmetric}(k|x_i) = \begin{cases} 0, & k = \text{normal class} \\ (1 - \epsilon)q_{hard}(k|x_i) + \frac{\epsilon}{K-1}, & k = \text{defect class} \end{cases} \quad (3.5)$$

提案手法は正常品と欠陥品を非対称に扱い、かつ、欠陥品の欠陥クラス間にはラベルスムージング手法を適用することから、非対称ラベルスムージング手法と呼ぶこととした。ラベルスムージングを用いた先行研究として、学習過程で正解ラベルを教師データセット内のクラス類似度に合わせて動的にスムージングすることによって画像分類の精度を向上させる手法[53]が提案されているが、本論文で提案している手法は検査工程における見逃しの抑制に焦点を当て、学習前に教師データの

正解ラベルを静的にスムージングしている点で先行研究とは異なる。

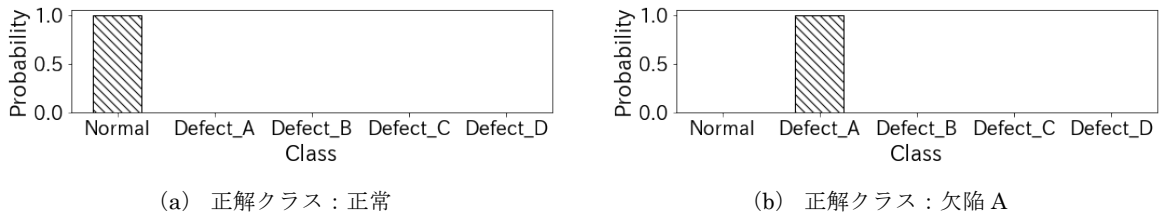


図 3.5 ハードラベル手法で与えられる正解ラベルの例

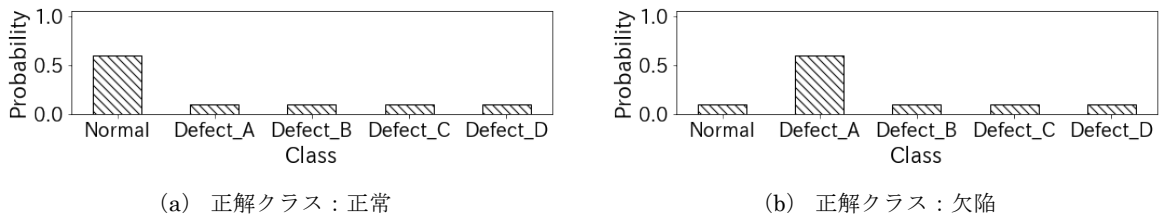


図 3.6 一様ラベルスムージング手法で与えられる正解ラベルの例

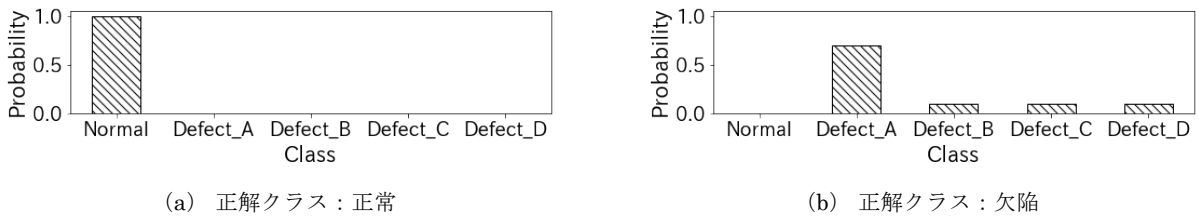


図 3.7 非対称ラベルスムージング手法で与えられる正解ラベルの例

3.1.5 非対称ラベルスムージング手法の解釈

正常、欠陥 A、欠陥 B の 3 つのクラスからなる簡易モデル ($K=3$) におけるハードラベル手法と一様ラベルスムージング手法、非対称ラベルスムージング手法での正解ラベルと損失値の分布を図 3.7、図 3.8、図 3.9 に示す。各図の (a1)、(a2)、(a3) は正解クラスが正常、欠陥 A、欠陥 B である時の正解ラベルを示す。各図の (b1)、(b2)、(b3) は (a1)、(a2)、(a3) に対応した正解ラベルが与えられているときの予測ラベルによる損失値を示し、予測ラベルの各クラスに対する確率をそれぞれの軸にとり、損失値の大きさをカラーマップで示している。図中では損失値の最小値が青、最大値が赤になるように示している。

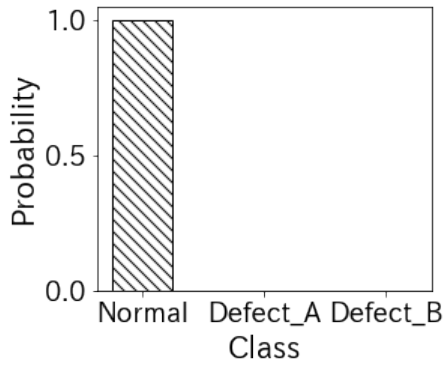
ハードラベル手法の場合、正解クラスが正常、欠陥 A、欠陥 B の何れであろうと与えられる正解ラベル $q_{hard}(k|x_i)$ は式 (3.2) より正解クラスで 1、正解クラス以外で 0 である (図 3.8 (a1)、(a2)、(a3))。式 (3.1) より求められる損失値は予測ラベルの正解クラス確率のみに依存し、正解クラス以外の確率には依存しない。つまり、損失値は正解クラスが正常、欠陥 A、欠陥 B のすべてで、予測ラベルの正解クラス確率が 1 に近づく方向に向けて、一様に減少する (図 3.8 (b1)、(b2)、(b3))。

損失値を最小にするように学習が進むと、予測ラベルは損失値マップの最小値である青い点へ向かっていく。一様ラベルスムージング手法の場合、正解クラスが正常、欠陥 A、欠陥 B の何れであろうと式 (3.3) より与えられる正解ラベル $q_{uniform}(k|x_i)$ は、正解クラスに最も高い確率 $(1 - \epsilon) + \epsilon/K$ を持ちつつ、正解クラス以外にも一定の確率 ϵ/K を持つ (図 3.9 (a1)、(a2)、(a3))。そのため、式 (1) より求められる損失値は予測ラベルの持つ正解クラス以外の確率にも影響を受ける。損失値はハードラベル手法のように予測ラベルの正解クラス確率が 1 に近づく方向に向けて一様に減少するのではなく、予測ラベルの正解クラス確率が $(1 - \epsilon) + \epsilon/K$ 、正解クラス以外の確率が ϵ/K となる点を最小値とするくぼみを持った分布となる (図 3.9 (b1)、(b2)、(b3))。一方で、非対称ラベルスムージング手法の場合、正解クラスが正常であるときと、何れかの欠陥であるときとで与えられる正解ラベル $q_{asymmetric}(k|x_i)$ が異なる (図 3.10 (a1)、(a2)、(a3))。正解クラスが正常であるときの正解ラベルは式 (3.4) より与えられ、ハードラベルと同一である。損失値は予測ラベルの正解クラス確率が 1 に近づく方向に向けて、一様に減少する (図 3.10 (b1))。正解クラスが何れかの欠陥であるときの正解ラベルは式 (3.5) より与えられる。正常クラスは確率を持たず、欠陥クラス群は正解クラスに最も高い確率 $(1 - \epsilon) + \epsilon/(K - 1)$ を持ちつつ、正解クラス以外の欠陥種にも一定の確率 $\epsilon/(K - 1)$ を持っている。損失値は正解である欠陥クラスの確率が $(1 - \epsilon) + \epsilon/(K - 1)$ 、正解ではない欠陥クラスの確率が $\epsilon/(K - 1)$ 、正常クラスの確率が 0 の点を最小値とするくぼみを持った分布となる (図 3.10 (b2)、(b3))。

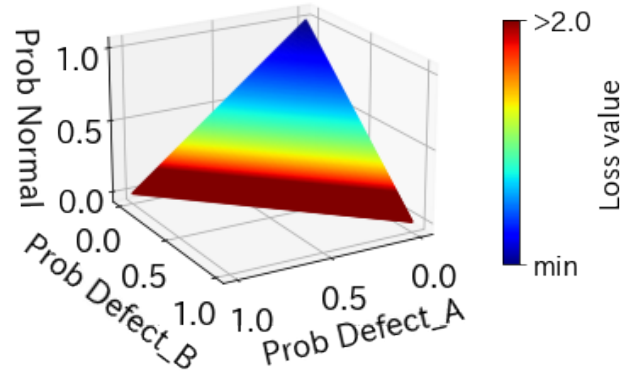
ハードラベル手法では学習がうまく進むと損失値が 0 となる点に収束し、教師データセットにおける見逃し、誤警告を最小化できる。しかし、モデルは教師データに過度に適合しやすく、汎化性能が低下する場合がある。一方で一様ラベルスムージング手法では正解クラス以外にも確率を与えることにより、過学習抑制の効果が期待できる。このとき、学習過程で目指す損失値の最小点はマップの隅から少しずれた位置となる。この点では、正解クラス以外で確率を持つことになるが、ソフトマックス関数を用いた CNN 多クラス分類法では、正解クラスが最も高い確率を持てば正しい分類が可能である。ハードラベル手法と一様ラベルスムージング手法は全てのクラスが対等に扱われ、欠陥検出工程で重視されるべき機能である欠陥品の見逃し抑制に繋がる働きかけは行われていない。

非対称ラベルスムージング手法はハードラベル手法と一様ラベルスムージング手法の特徴を併せ持つことで、誤警告を増加させずに見逃しを抑制することを実現する。まず、欠陥品クラス群にはラベルスムージングを取り入れることでモデルの過学習を抑制し、汎化性能を向上させる効果が期待できる。次に、正常品にはハードラベルを与えることで極力、正解クラスに分類されるように学習し、誤警告の発生を抑える。そして、欠陥品教師データの正常クラスと欠陥クラス間で正解ラベルの付与方針を別にするすることで教師データが欠陥品であるときの損失値に正常品の確率を 0 にする方向の勾配を持たせる。これにより、教師データが欠陥品のときには正常品と判断することがないように学習が進む。このようにして学習されたモデルは、テストサンプルにおいて欠陥品教師データが持っていたような欠陥らしさを認識した場合、積極的に欠陥品と検出する。つまり、いずれの検査項目にも該当しない場合のみ正常品、いずれかの検査項目に該当する可能性がある場合に欠陥品とされる取り扱いの非対称性の概念が導入され、見逃しの抑制効果が期待できる。従来手法、一様ラベルスムージング手法、非対称ラベルスムージング手法における正解クラスが欠陥 A であると

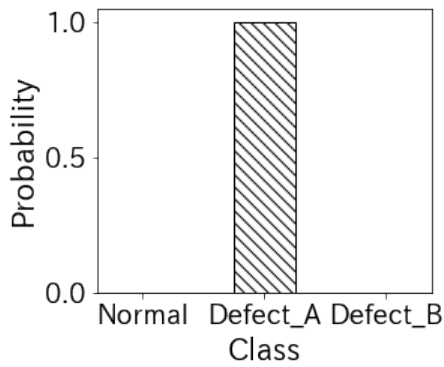
きの損失値勾配をまとめて比較し、代表の 3 点における勾配方向を矢印で注釈したものを図 3.11 として示す。非対称ラベルスムージング手法は正解の方向に勾配を持ちながらも、常に見逃し抑制の方向の勾配も持つことが見て取れる。なお、正常品は正解ラベルにはハードラベルが与えられているため、一般的には過学習によるモデルの汎化性能低下が懸念される[53]。しかし、外観検査工程における正常品は欠陥という特徴を持たず、位置的な変化のみを持ち、多様性が低いことから、教師データとテストデータが類似している。そのため、モデルの汎化性能低下が発生しにくいと考えられる。



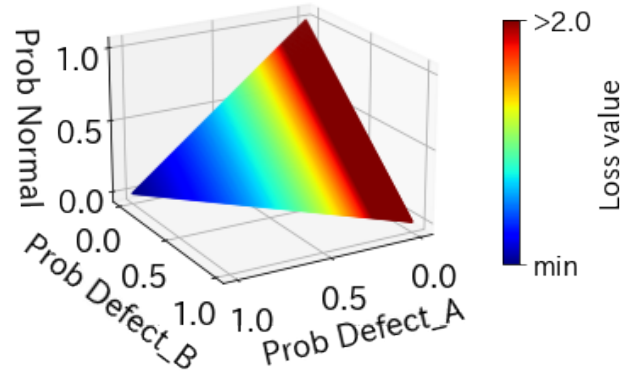
(a1) 正解ラベル
正解クラス：正常



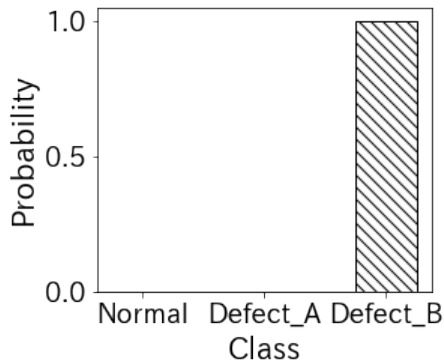
(b1) 損失値の分布
正解クラス：正常



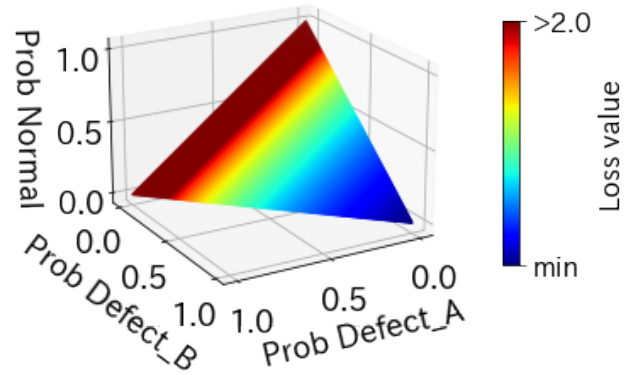
(a2) 正解ラベル
正解クラス：欠陥 A



(b2) 損失値の分布
正解クラス：欠陥 A



(a3) 正解ラベル
正解クラス：欠陥 B

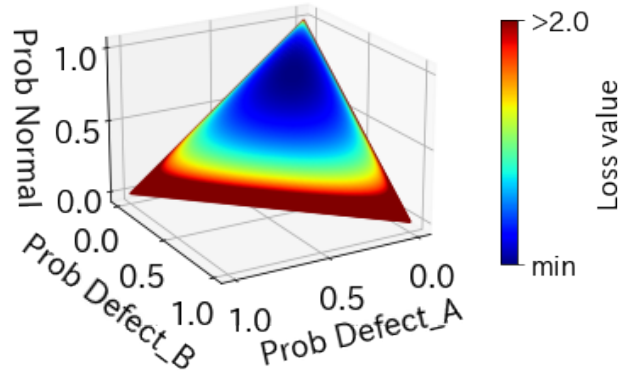


(b3) 損失値の分布
正解クラス：欠陥 B

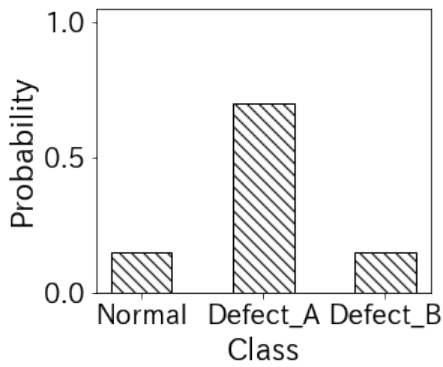
図 3.8 ハードラベル手法における正解ラベルと損失値の分布



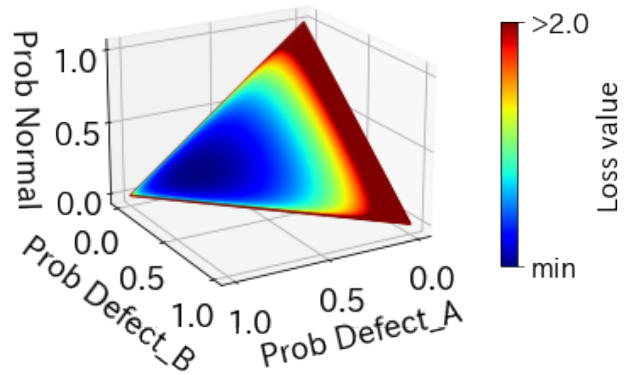
(a1) 正解ラベル
正解クラス：正常



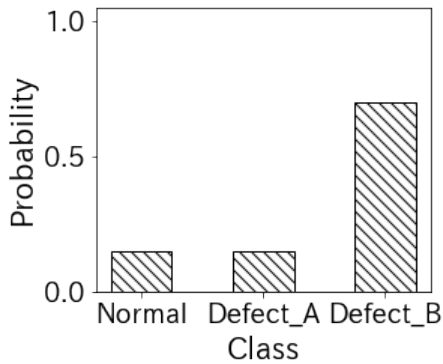
(b1) 損失値の分布
正解クラス：正常



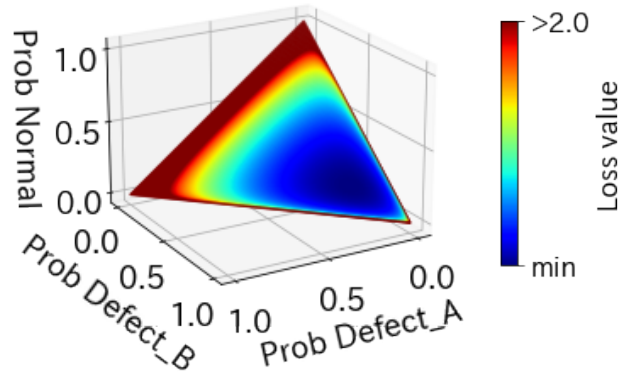
(a2) 正解ラベル
正解クラス：欠陥 A



(b2) 損失値の分布
正解クラス：欠陥 A

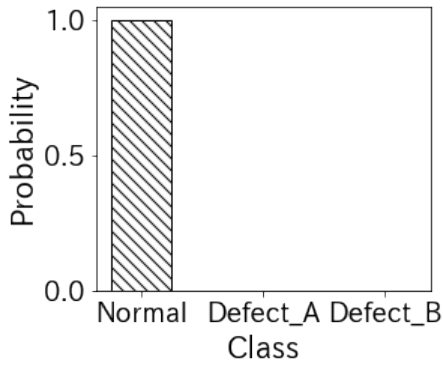


(a3) 正解ラベル
正解クラス：欠陥 B

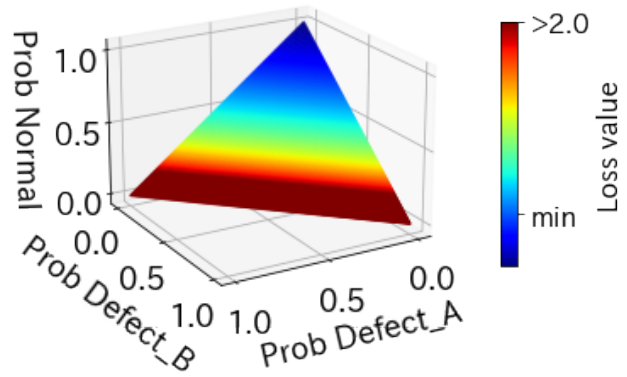


(b3) 損失値の分布
正解クラス：欠陥 B

図 3.9 一様ラベルスムージング手法における正解ラベルと損失値の分布



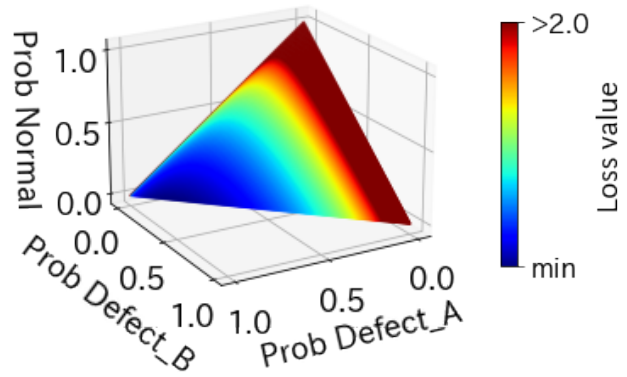
(a1) 正解ラベル
正解クラス：正常



(b1) 損失値の分布
正解クラス：正常



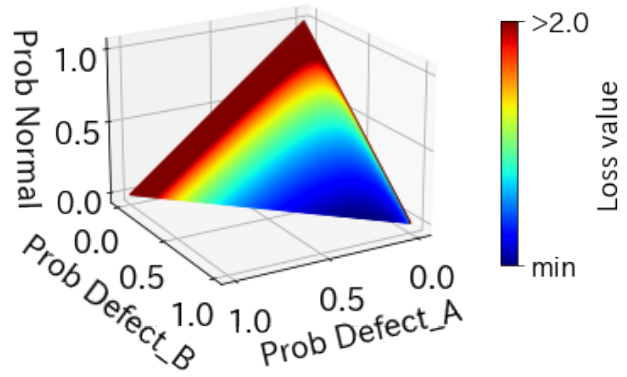
(a2) 正解ラベル
正解クラス：欠陥 A



(b2) 損失値の分布
正解クラス：欠陥 A



(a3) 正解ラベル
正解クラス：欠陥 B



(b3) 損失値の分布
正解クラス：欠陥 B

図 3.10 非対称ラベルスムージング手法における正解ラベルと損失値の分布

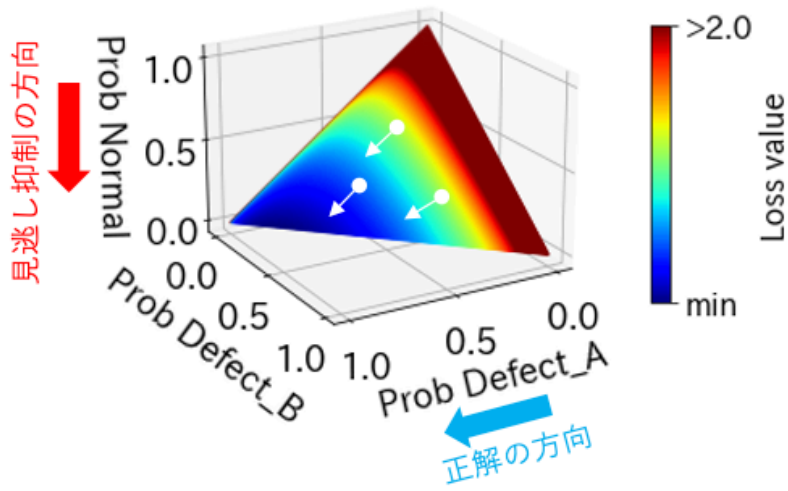
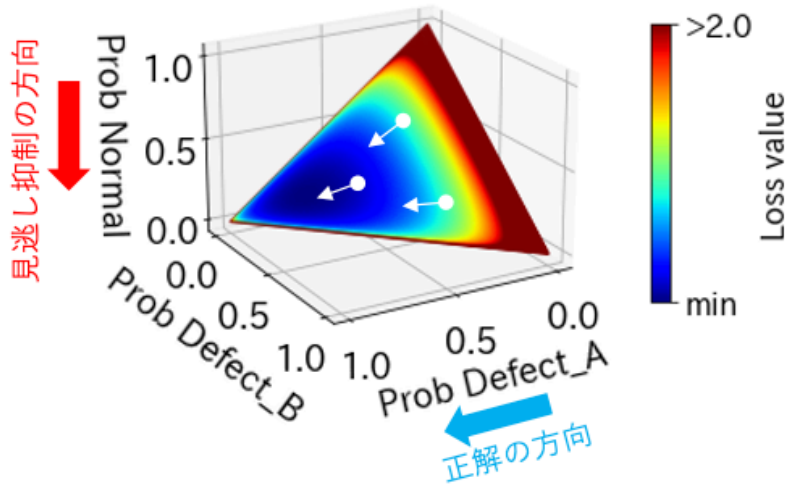
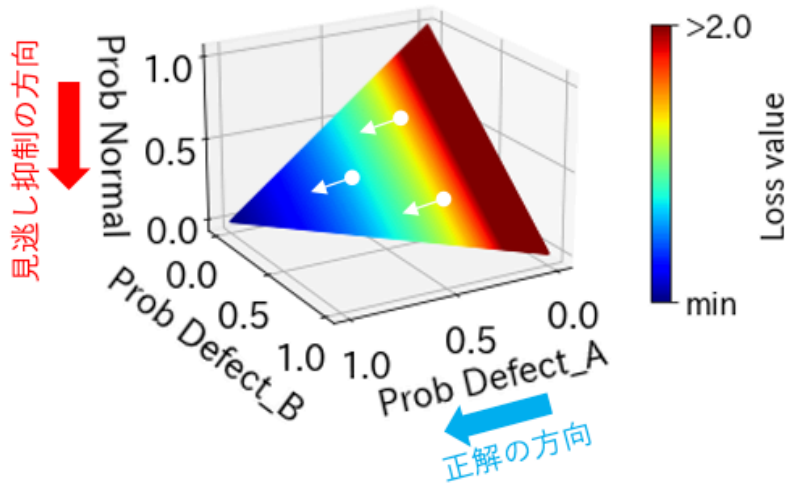


図 3.11 正解クラスが欠陥 A であるときの損失値勾配の比較

3.2 欠陥分類実験 4：開発手法 2 非対称ラベルスムージング手法

提案手法の有効性を確認するために実際の完成品検査を模擬した実験を行い、見逃し率と誤警告率の変化を観察した。実験では従来手法であるハードラベル手法、一様ラベルスムージング手法と提案手法である非対称ラベルスムージング手法の 3 つについて比較を行った。見逃し率は欠陥品と分類されるべきデータのうち、正常品と誤判定された割合、誤警告率は正常品と分類されるべきデータのうち欠陥品と誤判定された割合としている。どちらの指標も 0 に近いほど検出能力が高いことを意味する。

3.2.1 データセット

分類実験 4 では分類実験 3 と同じデータセットを用いて実験を行った。

3.2.2 学習モデル

分類実験 4 では分類実験 3 で用いたモデルの構成及びハイパーパラメータを変更せずそのまま学習モデルとして用いた。従来手法、一様ラベルスムージング手法、非対称ラベルスムージング手法の各モデルは、正解ラベルの付与方法のみが異なる。一様ラベルスムージング手法及び非対称ラベルスムージング手法を用いる際の平滑化パラメータ ϵ は、分類性能に影響を与えられられる。このモデルでの ϵ 最適値は不明確であるため、実験では $\epsilon \in \{0.1, 0.2, \dots, 0.9\}$ としてそれぞれの分類性能を列挙し比較する。ここで、 ϵ は 0 から 1 までの値を取りうるが、0 のときはハードラベル手法と等価であり、1 のときはスムージングされているクラス間での確率差がなくなり学習が収束しないため、実験からは除いている。なお、ハードラベル手法の実験結果は分類実験 3 の結果を用いている。

3.2.3 実験結果

正解ラベルの付与方法が異なる 3 つのモデルを用いてクラッチディスク完成品における外観欠陥の有無及び欠陥種の分類実験を行い、見逃し率と誤警告率について確認した。

まず、平滑化パラメータ ϵ の決定を行った。 ϵ を変化させた時の見逃し率の変化を図 3.12 に示す。ハードラベル手法は ϵ に依存しないが、比較のためその値を破線で示してある。一様ラベルスムージング手法では、 ϵ が大きくなるにつれてわずかに見逃し率が低くなる傾向がみられる。一方で非対称ラベルスムージング手法では、より顕著な低下傾向が確認でき、 $\epsilon=0.8$ のときに見逃し率が最も低くなった。ほとんどの試行で非対称ラベルスムージング手法は一様ラベルスムージング手法よりも低い値となっている。同様に誤警告率の変化を図 3.13 に示す。一様ラベルスムージング手法は横ばい、非対称ラベルスムージング手法は ϵ が大きくなるにつれて、緩やかな低下傾向となっている。どちらの手法もハードラベル手法よりも低い水準で推移している。これらの実験結果から ϵ の値には 0.8 を採用することとする。

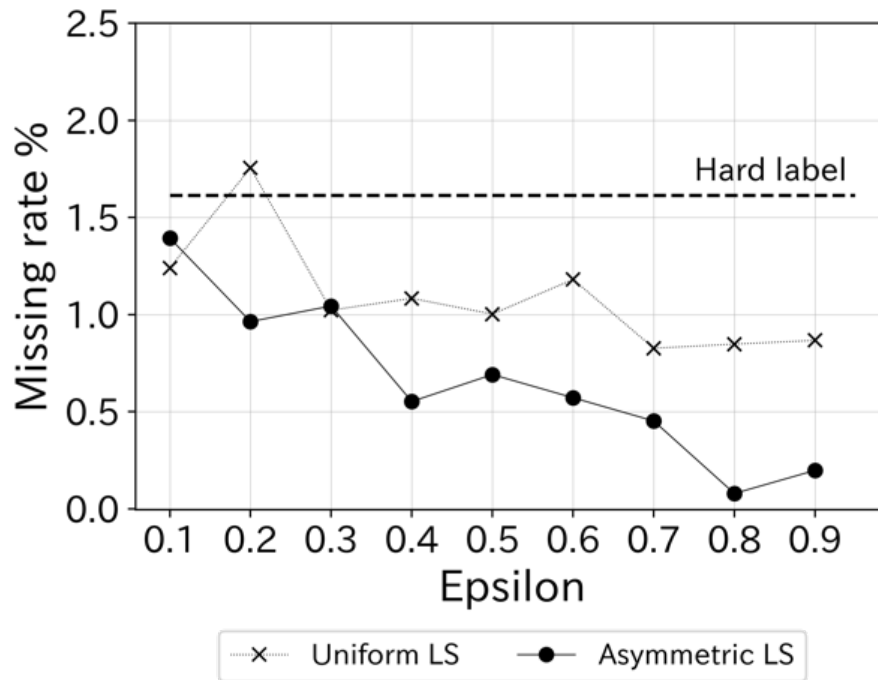


図 3.12 平滑化パラメータをシフトしたときの見逃し率の推移

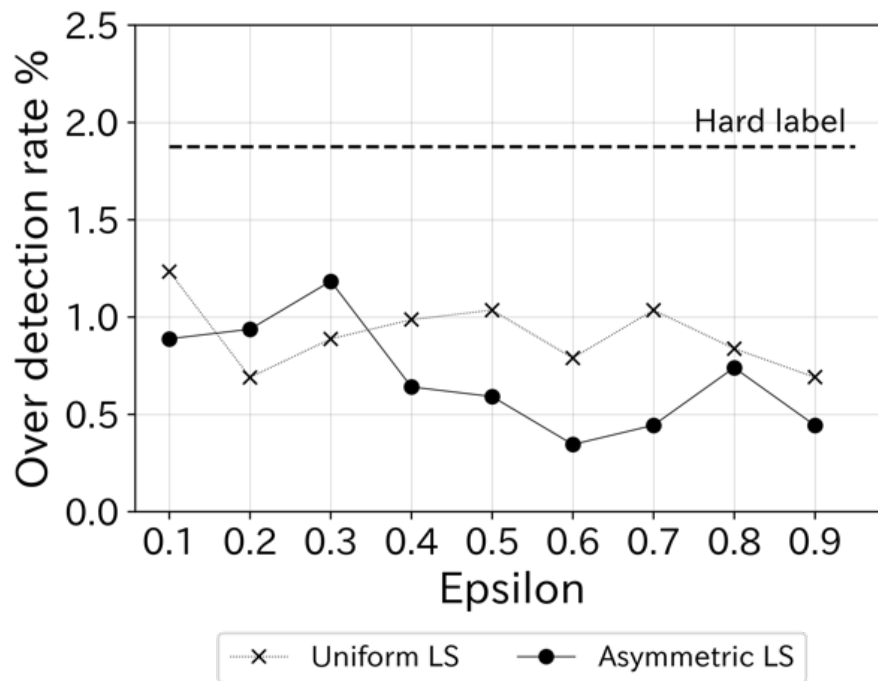


図 3.13 平滑化パラメータをシフトしたときの誤警告率の推移

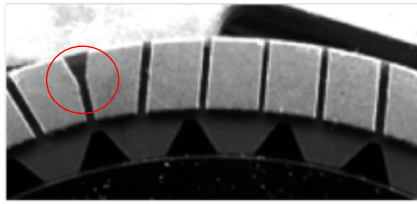
$\epsilon=0.8$ での各手法の見逃し率及び誤警告率を表 3.1 に示す。各手法での見逃し率は、ハードラベル手法 1.62%、一様ラベルスムージング手法 0.85%、非対称ラベルスムージング手法 0.08%となった。非対称ラベルスムージング手法によってハードラベル手法と一様ラベルスムージング手法に対して、見逃し率が各々1.54ポイント、および0.77ポイント改善し、同様に誤警告率は各々1.13ポイント、および0.10ポイント改善した。

表 3.1 $\epsilon=0.8$ における見逃し率と誤警告率

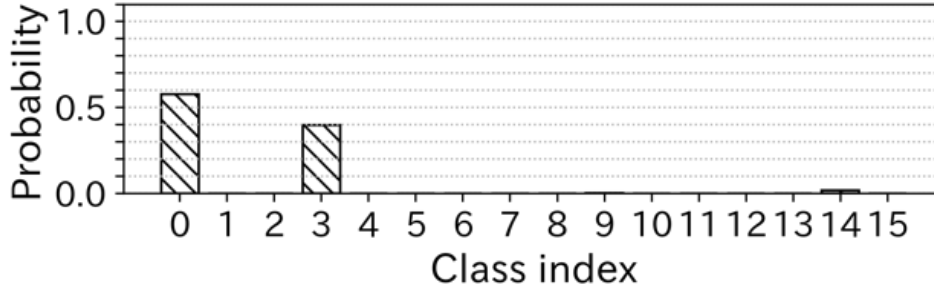
	見逃し率	誤警告率
ハードラベル手法	1.62% (5,075 件中 82 件)	1.87% (2,027 件中 38 件)
一様ラベルスムージング手法	0.85% (5,075 件中 43 件)	0.84% (2,027 件中 17 件)
非対称ラベルスムージング手法	0.08% (5,075 件中 4 件)	0.74% (2,027 件中 15 件)

本実験 ($\epsilon=0.8$) において欠陥品サンプル 5,075 件のうち、ハードラベル手法で見逃しとなったのは 82 件であった。同様に一様ラベルスムージング手法では 43 件、非対称ラベルスムージング手法では 4 件が見逃しとなった。非対称ラベルスムージング手法で見逃しとなった 4 件は、ハードラベル手法でも検出されなかった 1 件とハードラベルでは検出された 3 件に分けられる。ハードラベル手法では見逃しとなったが、非対称ラベルスムージングでは欠陥品として検出されたサンプルを個々に見ていくと、以下の 3 つに類型化できる。1)ハードラベル手法における予測ラベルで正解の欠陥クラスに確率を持ち、欠陥品が疑わしいと予測されながらも、正常クラスが最も高い確率を持つため、正常品と判断されたサンプル、2)欠陥品の可能性をほとんど持たずに見逃しとなったサンプル、3)複数の欠陥種で確率を分け合ったために正常品の確率が一番高くなったサンプルである。内訳は 1)が 57 件、2)が 23 件、3)が 1 件となった。1)のサンプル例を図 3.14 に示す。このサンプルの正解クラスは 3 で、ペーパー摩擦材の一部にオイル滲みが有る。この欠陥は目視でも発見し難いサンプルとなっている。ハードラベル手法では欠陥の可能性を持ちながらも正常と判断され、一様ラベルスムージング手法ではすべての欠陥種の中でオイル滲みの可能性が最も高いと予測しているが、わずかに正常の確率が高いために見逃しとなっている。これに対し、非対称ラベルスムージング手法では欠陥が疑わしいサンプルとして積極的に欠陥品として検出していることが確認できた。次に、2)のサンプル例を図 3.15 に示す。このサンプルの正解クラスは 10 で、ペーパー摩擦材の一部に剥がれがある。ハードラベル手法では欠陥の確率をほとんど持たず、見逃しとなっているが、非対称ラベルスムージング手法では欠陥として検出している。このことは、欠陥品の疑いがあるものを積極的に検出できるように学習することで、ほんのわずかな兆候も見逃さず、欠陥品として検出できるようになったためと推測される。

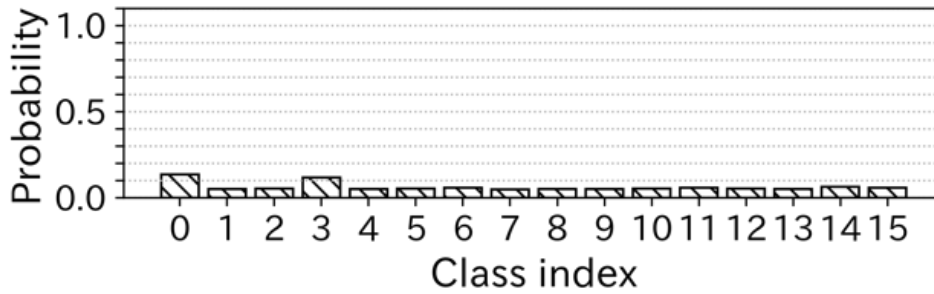
Ground truth:3



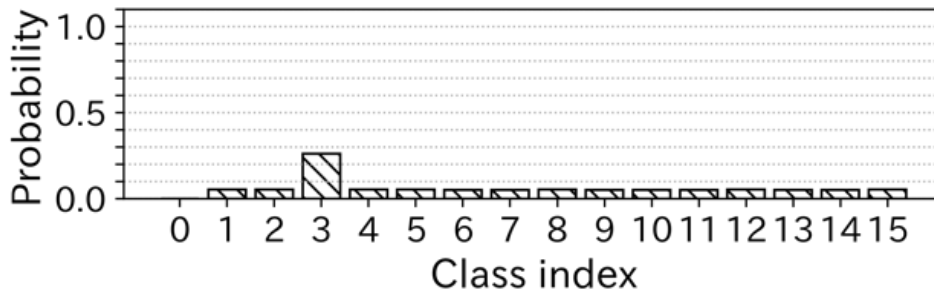
(a) サンプル画像



(b) ハードラベル手法での予測ラベル



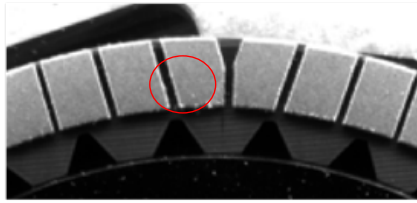
(c) 一様ラベルスムージング手法での予測ラベル



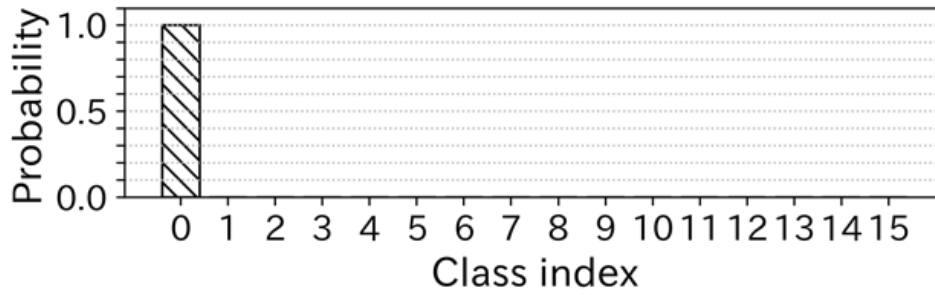
(d) 非対称ラベルスムージング手法での予測ラベル

図 3.14 サンプル画像と予測ラベル (1)

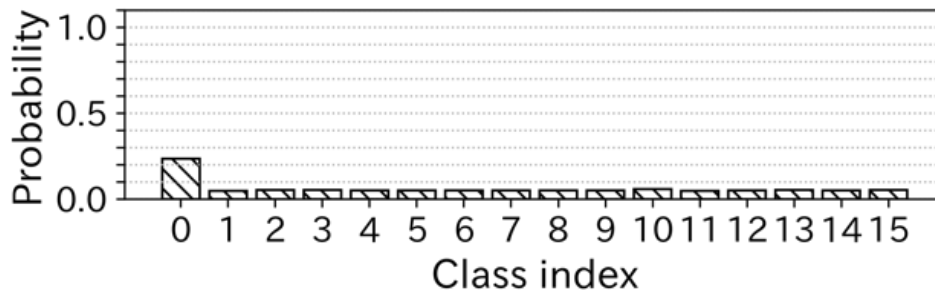
Ground truth:10



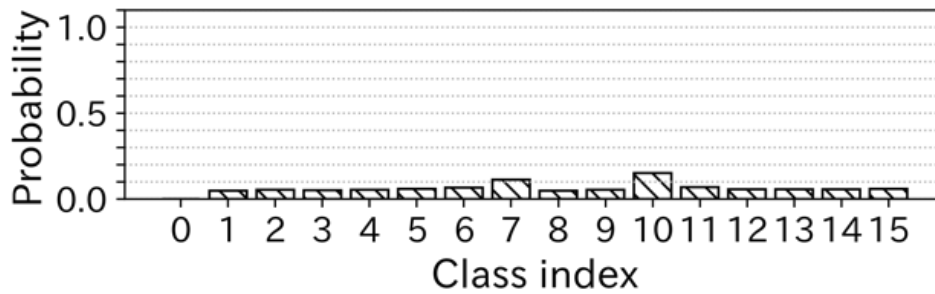
(a) サンプル画像



(b) ハードラベル手法での予測ラベル



(c) 一様ラベルスムージング手法での予測ラベル



(d) 非対称ラベルスムージング手法での予測ラベル

図 3.15 サンプル画像と予測ラベル (2)

各手法について、t-SNE 手法を用いてサブセット 4 における特徴量を 2 次元化し、可視化したものを図 3.16 (図 2.41 と同一)、図 3.17、図 3.18 として示す。ハードラベル手法では各クラスの分布に少しばらつきがある。特に、正常クラスの分布にまとまりがなく、オイル滲み、2 枚重ねクラスと分布の重なりがあることを確認できる。一様ラベルスムージング手法ではハードラベル手法よりも各クラスの分布がまとまっており、各クラスの特徴抽出がうまくできている様子が見て取れる。しかし、ハードラベル手法と同様に、正常クラスとオイル滲み、2 枚重ねクラスとの分布に重なりがある。一方で、非対称ラベルスムージング手法では、各クラスの分布のまとまりがさらに良くなっていると同時に、正常クラスとその他の欠陥クラス群の分布が区別されていることが確認できる。提案手法は、他の既存手法と同じデータで学習されているが、より正常クラスを明確に分離するように学習できている。このことは、予測ラベルの確率分布の考察と同様に、欠陥品の疑いがあるものを積極的に検出するために、欠陥クラスのわずかな特徴を逃さずに捉えようと学習した結果と考えられる。以上のように、非対称ラベルスムージング手法による外観欠陥の見逃し抑制効果が確認できた。

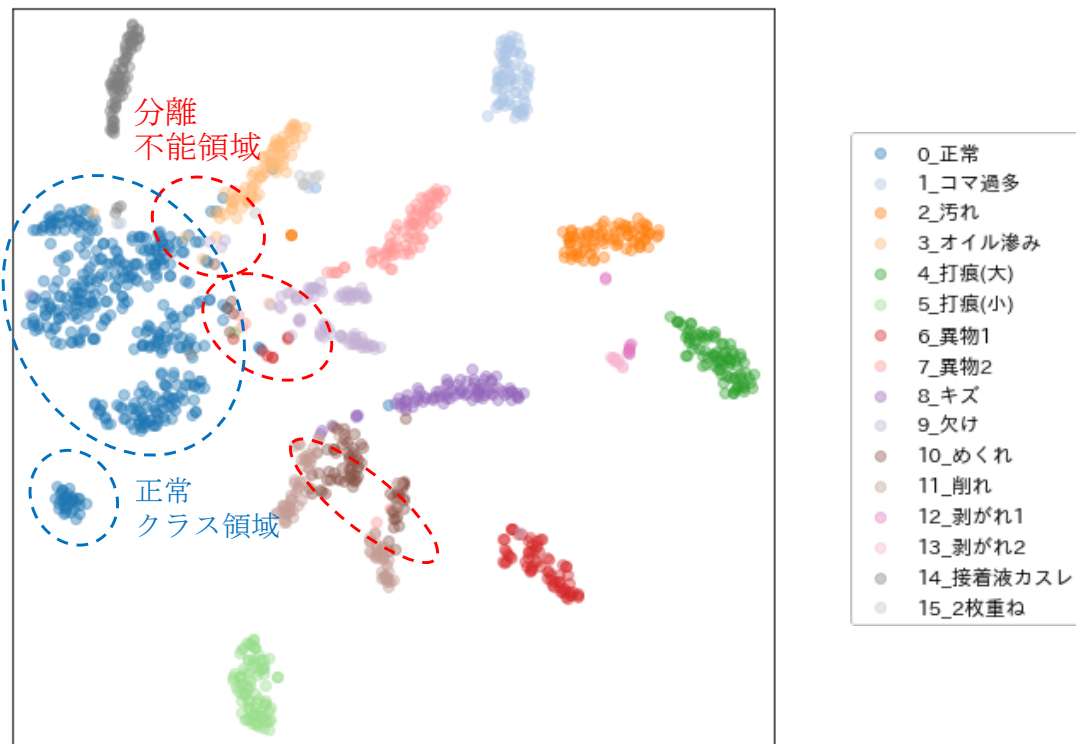


図 3.16 ハードラベル手法で学習したモデルの t-SNE による特徴量 2 次元マップ (サブセット 4)

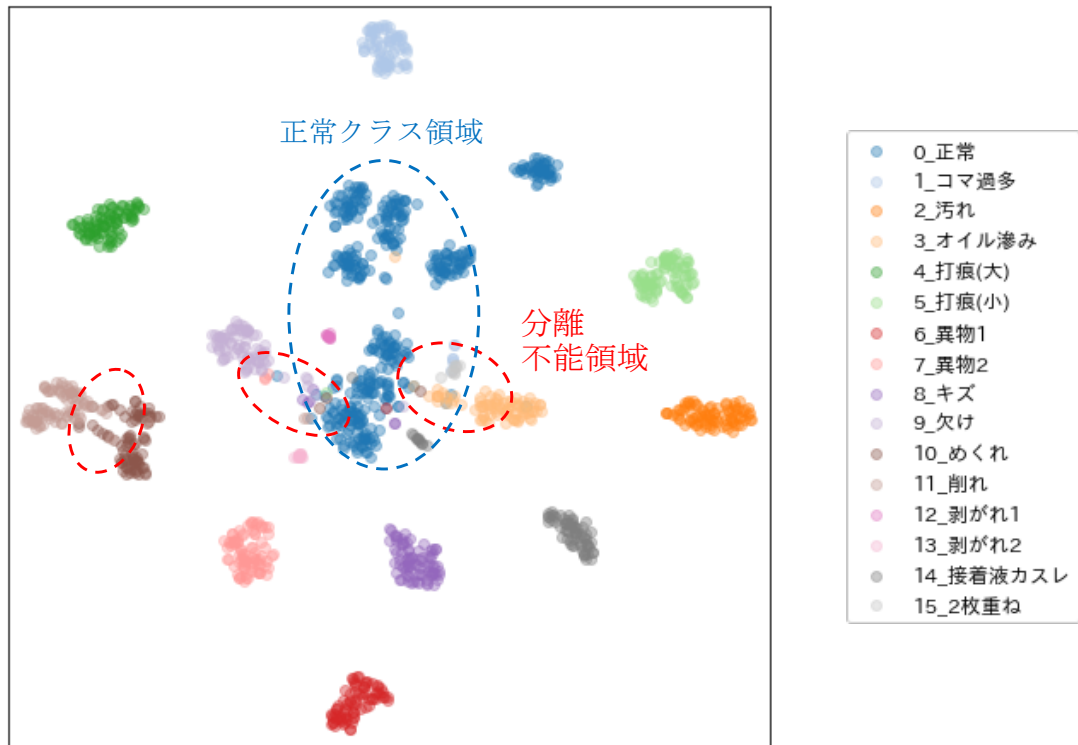


図 3.17 一様ラベルスムージング手法で学習したモデルの t-SNE による特徴量 2 次元マップ (サブセット 4)

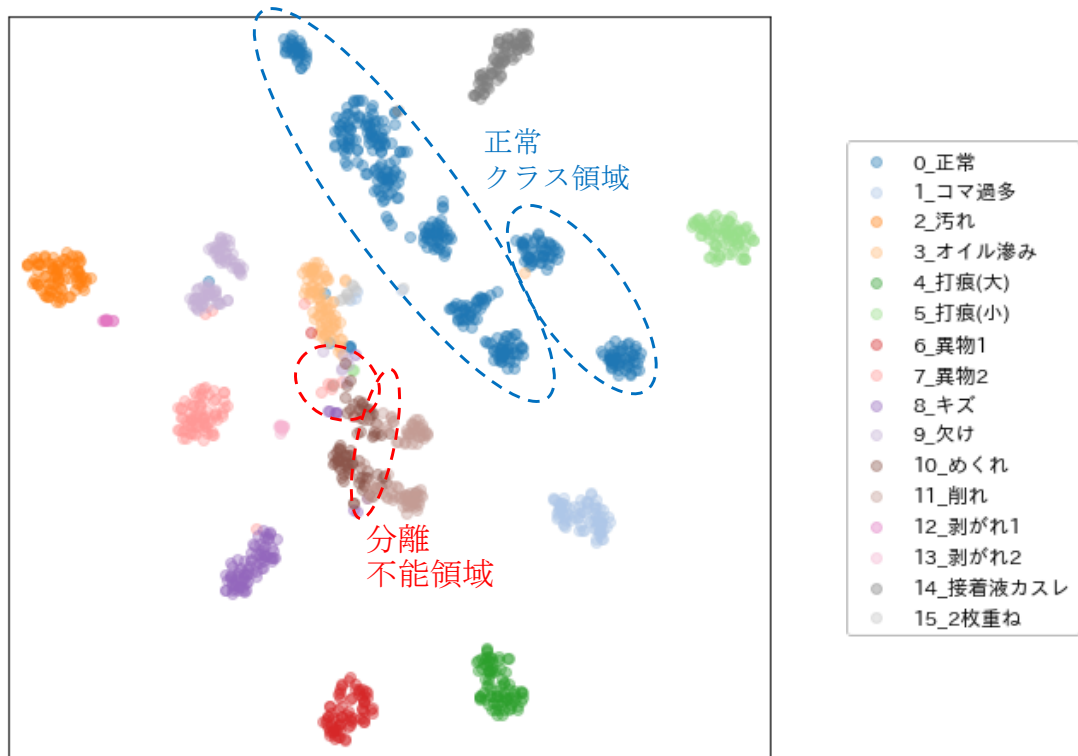


図 3.18 非対称ラベルスムージング手法で学習したモデルの t-SNE による特徴量 2 次元マップ (サブセット 4)

3.2.4 考察

非対称ラベルスムージング手法を用いた場合、平滑化パラメータ $\epsilon=0.8$ のときに見逃し率が最も低くなった。これは、先行研究[50]で用いられている 0.1 よりもかなり大きい。一様ラベルスムージング手法及び非対称ラベルスムージング手法を用いるとき、 ϵ の値を大きくすることは、正解クラスと正解クラス以外の確率を持つクラスとで確率差を小さくすることになる。そのため、正確な分類ができなくなり、見逃し率、誤警告率及び欠陥種内での誤分類率が高くなることが懸念されていた。しかし、実験の結果から ϵ が大きい範囲での見逃し率及び誤警告率は図 3.12、図 3.13 で示した通り、改善していく傾向にあり、欠陥種内での誤分類率については図 3.19 に示すように急激に高くなることはなく、懸念されていた性能の悪化は起こらないことが確認された。

ϵ が大きい範囲での正解ラベルの持つ確率差についてみると、本実験の 16 クラス分類の場合では、一様ラベルスムージング手法で $\epsilon=0.9$ としたとき、正解クラスに確率約 0.15、正解クラス以外に確率 0.05 が与えられ、約 3 倍の差がある。非対称ラベルスムージング手法で $\epsilon=0.8$ としたときは、欠陥品の正解クラスに確率約 0.25、正解以外の欠陥クラスに確率約 0.05 が与えられ、約 5 倍の差がある。このようにクラス数が多ければ ϵ が大きくても分類性能を保つための十分な確率差がついていると考えることができる。また、図 3.12 及び図 3.13 に示しているように一様ラベルスムージング手法で、その分類性能が ϵ の値に大きく左右されないことは、確率差の大きさの絶対値は分類性能に大きくは影響しないということを示している。一方で非対称ラベルスムージング手法の場合、 ϵ が大きくなるに従って、見逃し率が改善される傾向にある。その理由として、以下のようなことが考えられる。欠陥品教師データに与えられる正解ラベルは、一様ラベルスムージング手法の場合、図 3.6 (b) に示すように 2 段階であるのに対し、非対称ラベルスムージング手法の場合、図 3.7 (b) に示すように 3 段階の確率となる。 ϵ を大きくしていくと、欠陥クラス群のみで確率が平滑化されることで、正常クラスと欠陥クラス群の確率差は相対的に大きくなり、結果的に正常品か欠陥品かを分類する駆動力が大きくなる。別の見方をすると、正常クラスか欠陥クラス群かの 2 値分類と欠陥クラス群内での多クラス分類を同時に行っているとみることでもできる。

本実験では ϵ を正解の欠陥クラスと正解以外の欠陥クラスでの確率差が 3 倍から 5 倍となるように設定した場合に良好な結果が得られたが、最適値はモデル構成やデータセット、クラス数などによって異なる。 ϵ の決定方針に関しては、今後の課題として取り組みが必要である。

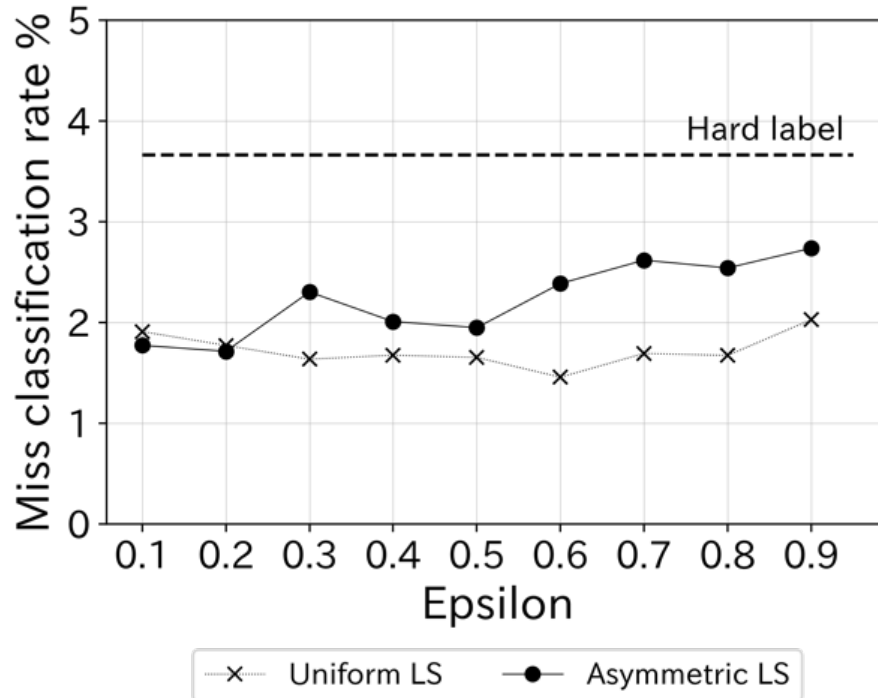


図 3.19 平滑化パラメータをシフトしたときの欠陥誤分類率推移

3.3 本章のまとめ

外観検査において欠陥品の見逃しと誤警告の抑制を同時に実現するため、CNN を用いた多クラス分類モデルに正常品と欠陥品の取り扱い非対称性の概念を取り入れた非対称ラベルスムージング手法を提案した。非対称ラベルスムージング手法を用いることによって欠陥が疑わしいサンプルを積極的に欠陥品として検出する性能を付与することができる。そのことを自動車用クラッチディスクの外観欠陥検出実験で実証した。このときの見逃し率は 0.08%、誤警告率は 0.74%であった。この値は従来のハードラベル手法と比較して 1.54 ポイント、および 1.13 ポイント改善されている。このように提案手法は実際の量産工程における欠陥検出精度を向上させ、欠陥の見逃しを抑制する有効な手段となることが期待できる。

本研究で開発した非対称ラベルスムージング手法は、従来の一様ラベルスムージング手法を改良することによって、検査工程の重要な機能である見逃し抑制を実現した。目視による検査では経験的に取り入れられている正常品と欠陥品の取り扱い非対称性の概念を正解ラベルの付与方法という形で CNN モデルに与えることで、CNN モデルの性格を操作することに成功した。このような先行研究はなく、新しい CNN の性能向上アプローチ手法として価値のある技術であると言える。また、ルールベース手法で行われていたような「特徴量をどう解釈して非対称性を与えるか」ではなく、「非対称性を取り入れた特徴量を抽出する」という点で大きく解決アプローチが異なり、エン

ジニアを困難な設定作業から開放できる点でも、開発した手法は従来ルールベース手法に対する優位性を持つものであり、高い技術的価値を有していると考えられる。

本研究で開発した非対称性を有する AI 構築手法は、完成品外観検査のみならず、画像診断、異常検出など、特定のクラスを積極的に検出したい場合に応用が可能な技術であると考えられる。また、教師あり学習で用いられる正解ラベルの付与を操作する手法に関して、その付与方法によって AI の性格を誘導する事が可能であるという新たな知見が得ることができた。

第4章 撮像方法変更による CNN 分類性能向上の取り組み

光技術は通信、農業、エネルギー、娯楽などさまざまな分野で応用されているが、製造業においても、加工、計測、画像認識をはじめとして、その応用が進んでおり、自動化を進めるための重要な技術となっている。筆者の所属企業においても、光技術はレーザーによる表面加工（加工）、製品寸法の測定（計測）、組み立て部品の欠品・取り付け位置・組み立て順序の確認、製品に刻印された 2 次元バーコードの読み取り（画像認識）など、様々な工程で活用されている。これらの機器性能は年々向上し、高精細化、高速化、小型化、低価格化が進んでおり、様々な工程の自動化・高効率化に貢献している。

クラッチディスクの完成品検査工程においては、画像認識、レーザー計測の光技術が導入され、社内ノウハウが蓄積されてきた。検査の対象となるペーパー摩擦材は、取り扱いが繊細で、測定具などが触れることによって、変形、キズの原因になる。また、クラッチディスクの生産ラインでは生産効率の向上が進んでおり、数秒のうちにクラッチディスク全体の検査を終える必要がある。そのため、非接触で高速であるという特徴を持つ光技術との相性がよく、応用が進んでいる。しかしながら、それらの検査システムは従来のルールベース手法を用いているために、完全な自動化には至っておらず、新しい自動化技術の開発が望まれている。

本研究では完成品外観検査の自動化を実現するため、CNN 多クラス分類モデルを用いた画像認識を用いている。CNN を用いた画像分類学習では、人間が指示することなく、CNN モデル自らが与えられた入力画像から特徴量を抽出し、分類のルールを作成する（図 4.1）。第 3 章では、CNN による特徴抽出および分類ルールを作成する「学習」における性能向上に取り組み、従来ルールベース手法を超える分類性能を達成した。一方で、学習の材料である「入力データ」について、分類性能を向上させるための詳細な検討は行われていない。本章では、入力データの「質」を向上させ、さらなる分類性能向上を実現するため、いくつかの異なる撮像方法によって作成されたデータを用いた学習の結果を比較することによって、光技術による CNN モデル性能向上の可能性を検討する。

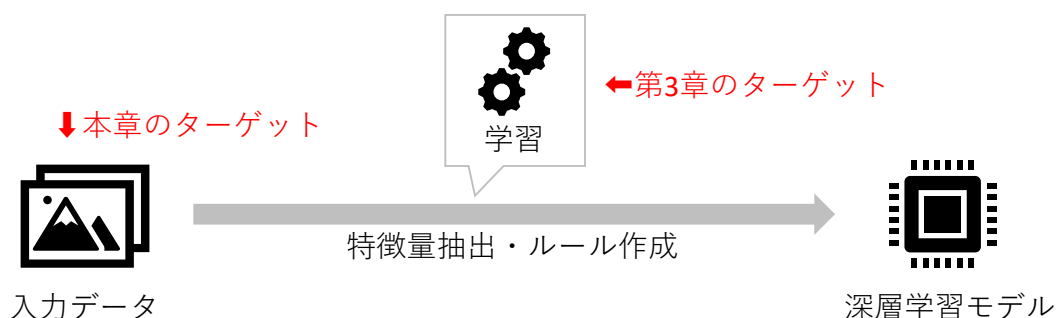


図 4.1 深層学習手法の画像判定モデル構成手順と開発ターゲット

4.1 欠陥分類実験 5 : 開発手法 3 画像データの質向上による分類性能向上手法

4.1.1 CNNにおけるデータの重要性

機械学習、深層学習モデルの学習における入力データは非常に重要である。CNN を始めとした深層学習モデルでは特徴抽出、ルール作成までを人手を介さず行うことができる。言い換えると、入力データの「質」が精度に大きく影響するということになる。データの質とは、ラベル付けの正確性、データ数はもちろんであるが、データに分類を行うために必要な情報が含まれていることまでを含む。本研究では CNN 多クラス分類モデルを教師あり学習することによって画像分類を行っている。入力データはモデルを学習するための教師データであり、正解ラベル付きの画像として与えられる。モデルの学習に使用した教師データにおいては正解ラベルの正確性をデータクレンジングで、データ数に不足がないことを事前実験の結果で確認している。データクレンジングは、エンジニアがすべての画像を確認し、教師データとして付与されているラベル（分類されているクラス）と検査結果として付与されるべき正解ラベル（分類されるべきクラス）が異なっていないかの確認をすることで行われた。また、画像では正確に判断できないサンプルに関しては、実物を確認し、検査基準に照らし合わせて、正しいラベルが付与されているかの判断を行った。データ数は事前の実験により、必要十分であることを確認している。学習に用いるデータ数を準備した教師データ 6,237 枚すべてとする場合を 100%として、データ数を 10%、20%、…と 10%ずつ増加させていったときの正解率を図 4.2 に、損失値の推移を図 4.3 に示す。すべての検証データセットは 1,550 枚から構成され、各試行で同じものを用いている。データ数が 50%を超えたあたりから正解率及び損失値の改善は僅かであり、100%のデータ数は性能を確認するのに十分であると言える。本研究で用いたデータセットは、ラベルの正確性、データ数について問題はないが、より分類を行いやすくする情報を含むデータを用いることによって分類の精度を向上させるため、撮像方法を工夫する余地がある。

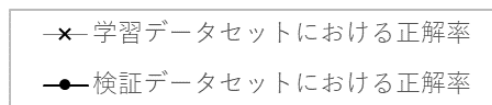
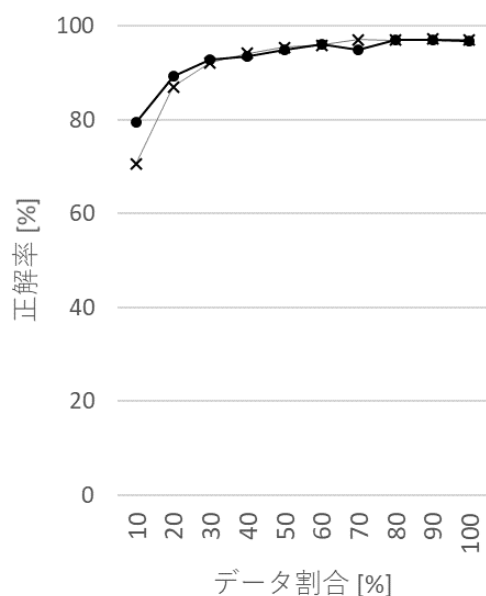


図 4.2 データ数による正解率の推移

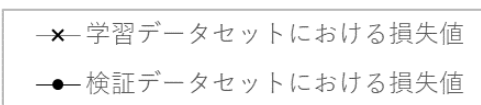
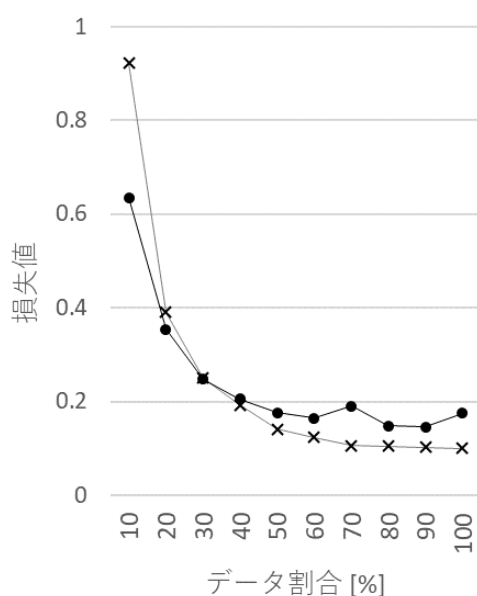


図 4.3 データ数による損失値の推移

4.1.2 撮像方法

学習に用いる入力画像の撮影時には光技術を応用し、欠陥部を強調したり、ノイズを低減したりすることが可能である。撮像方法によって分類性能がどのように変化するかの確認を行うための実験を行った。実験に用いる画像データセットは、第2章の分類実験1と同じサンプルと撮像環境(図2.17)を用い、5つの異なる撮像方法によって収集した画像データで作成した。画像は撮像方法のみが異なるように、1つのサンプルは設置後に、位置を変えることなく、5種の画像を取得するための撮像を行った。5種類の撮像は以下のように行った。

- i) モノクロ画像：モノクロカメラを用いて撮像した画像をモノクロ画像として用いた(図4.4)。分類実験1ではこのモノクロ画像データセットを用いた。
- ii) モノクロ偏光画像：LEDリング照明及びカメラレンズに直線偏光フィルタを設置して撮像した画像をモノクロ偏光画像として用いた(図4.5)。照明側偏光フィルタとカメラレンズ側偏光フィルタは偏光方向を直行させる向きで設置した。
- iii) カラー画像：カメラレンズに赤、緑、青のカラーフィルタを設置した3枚の画像を撮像した後、画像処理によって合成し、カラー画像として用いた(図4.6)。
- iv) カラー偏光画像：LEDリング照明に直線偏光フィルタを設置し、カメラレンズに偏光フィルタ及び赤、緑、青のカラーフィルタを設置した3枚の画像を撮像した後、画像処理によって合成

し、偏光カラー画像をして用いた（図 4.7）。照明側偏光フィルタとカメラレンズ側偏光フィルタは偏光方向を直行させる向きで設置した。

- v) エッジ抽出画像：LED リング照明からの照明光を遮光できるようなリングの 1/8 のみ開口させた遮光リングを用い、遮光リングの開口部を 45 度ずつ回転させながら 8 枚の画像を取得した後、画像処理によって合成した画像をエッジ抽出画像として用いた（図 4.8、図 4.9、図 4.10）。

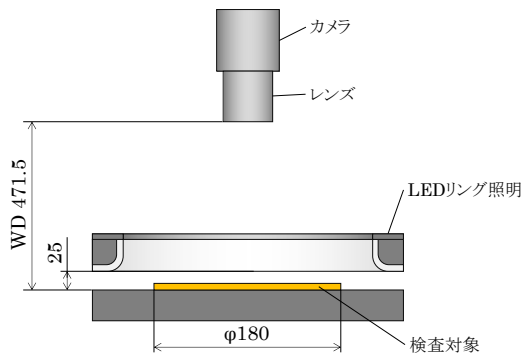


図 4.4 モノクロ画像撮像環境

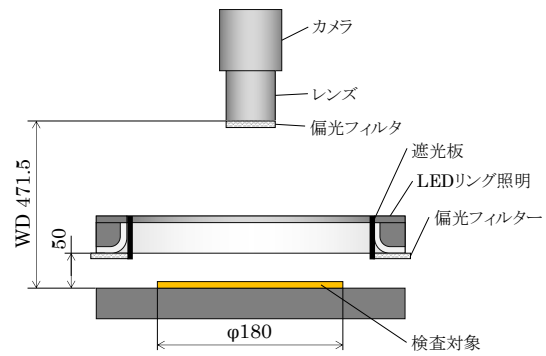


図 4.5 偏光モノクロ画像撮像環境

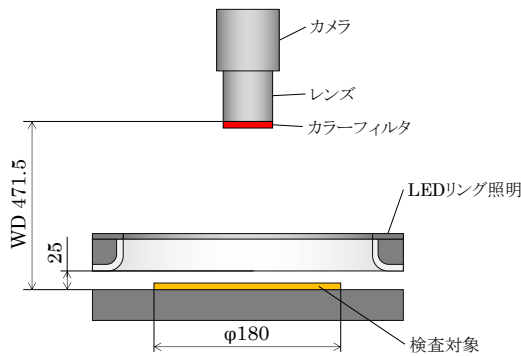


図 4.6 カラー画像撮像環境

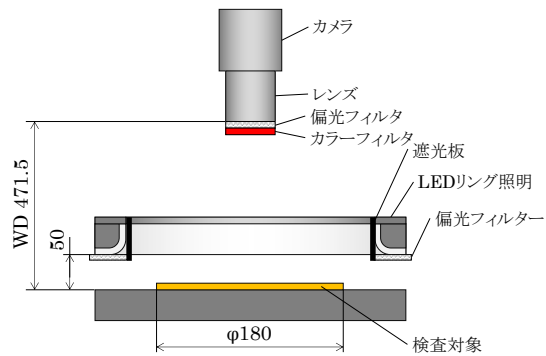


図 4.7 偏光カラー画像撮像環境

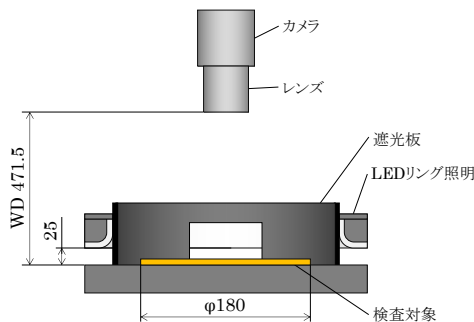


図 4.8 エッジ抽出画像撮像環境

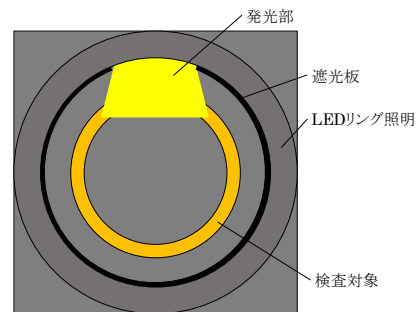


図 4.9 エッジ抽出画像撮像環境（平面図）

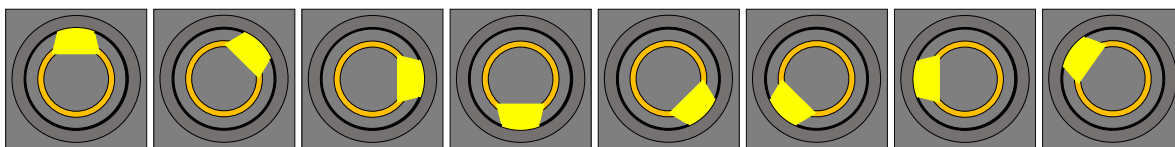


図 4.10 エッジ抽出画像撮像時の発光パターン

カメラや照明といった撮像機器は様々な機能を持つものが販売され、低価格化も進んでいることから、5種以外の異なる撮像方法を用いる検討ができたが、本実験では、量産を想定した場合に導入しやすい5種のみ絞って比較を行った。

撮像方法を工夫することによって CNN モデルの分類性能を向上させる研究に Aghaei らの研究 [49]がある。Aghaei らの研究では、オフラインの複数照明によって撮像された画像を学習することで、インラインの単一照明でも表面欠陥の検出性能向上を実現する手法を提案している。しかし、クラッチディスクの欠陥検出において、複数の撮像方法による精度を比較し、データの質向上による分類性能向上に取り組んでいる点で、先行研究とは異なる。

それぞれの撮像方法で取得された画像の例を図 4.11、4.12 に示す。図 4.11 は異物を含む画像、図 4.12 はキズを含む画像である。異物はアルミテープをカットし人為的に貼り付けたもので、光沢のある銀色をしている。キズは鋭利な金属で表面を擦ることでペーパー摩擦材の表面を毛羽立たせている。(a) モノクロ画像：異物部分がペーパー摩擦材部分よりも少し暗く写っているが、異物のエッジ部分は照明光が反射し、白く光っている。キズは散乱光が白く写っているが、正常部分との輝度差が小さく、正常部分との境目がはっきりとしない。(b) モノクロ偏光画像：異物部分がペーパー摩擦材部分よりも暗く写っている。異物のエッジ部が白く光っておらず、モノクロの場合にはペーパー摩擦材に対して明るい領域と暗い領域で構成されていた異物は、暗い領域のみで構成されている。加えて、芯板部分も暗く写っており、ペーパー摩擦材部分と異物のみを浮き上がらせるように写っている。キズに関してはキズ箇所のみが白く浮き上がっているために確認しやすく、モノクロ画像と比較して正常部分との境目がよりはっきりとしている。欠陥をよりシンプルに写し、欠陥以外の余分な情報を除去できている。(c) カラー画像：摩擦材は薄い茶色、異物部分は銀色をしており、色として異物を認識できる。モノクロ画像同様に、異物のエッジ部分が白く光って写っている。キズも異物同様に欠陥箇所を色として識別できる。モノクロ画像と比較して欠陥部と正常部分の境目がはっきりと認識できる。形状の変化を伴わず、色がはっきりとした欠陥は検出しやすいと考えられる。(d) カラー偏光画像：フィルタの影響で画像が暗くなってしまったが、ペーパー摩擦材は薄い茶色に、異物は黒く写っている。キズは正常部分と比較して欠陥部分が明るく写っている。(e) エッジ抽出画像：ペーパー摩擦材、芯板、異物それぞれのエッジが抽出できており、異物部は全周のエッジがしっかりと抽出できている。一方でキズは欠陥の多くの部分で認識するのが難しくなっている。凹凸がはっきりしている欠陥は検出しやすいが、凹凸のないキズや汚れなどの欠陥は検出が難しいと考えられる。

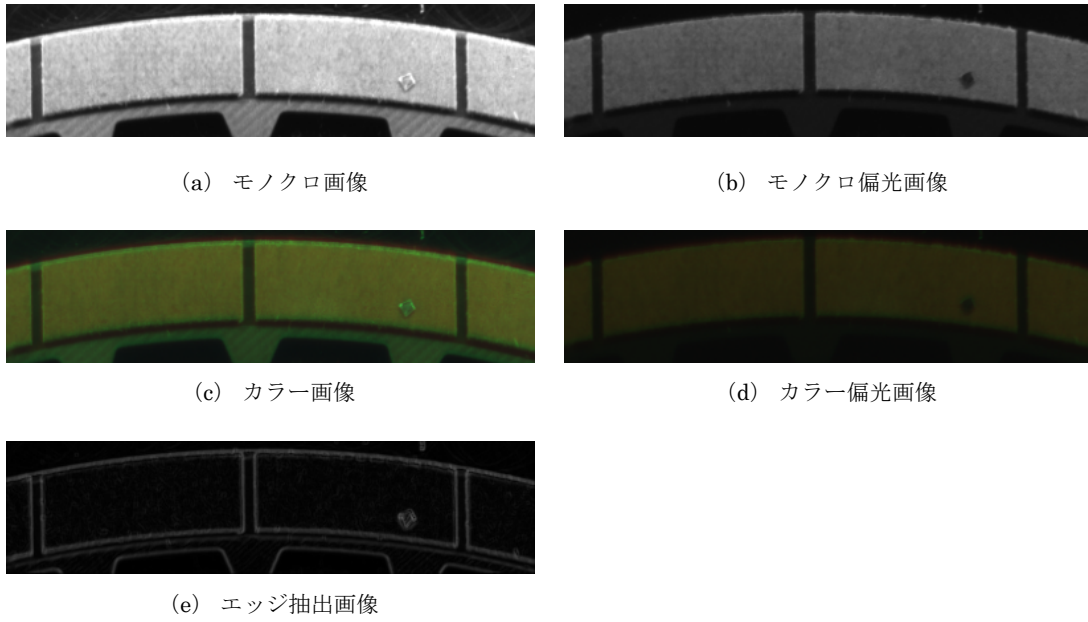


図 4.11 撮像方法別画像の例 欠陥種：異物

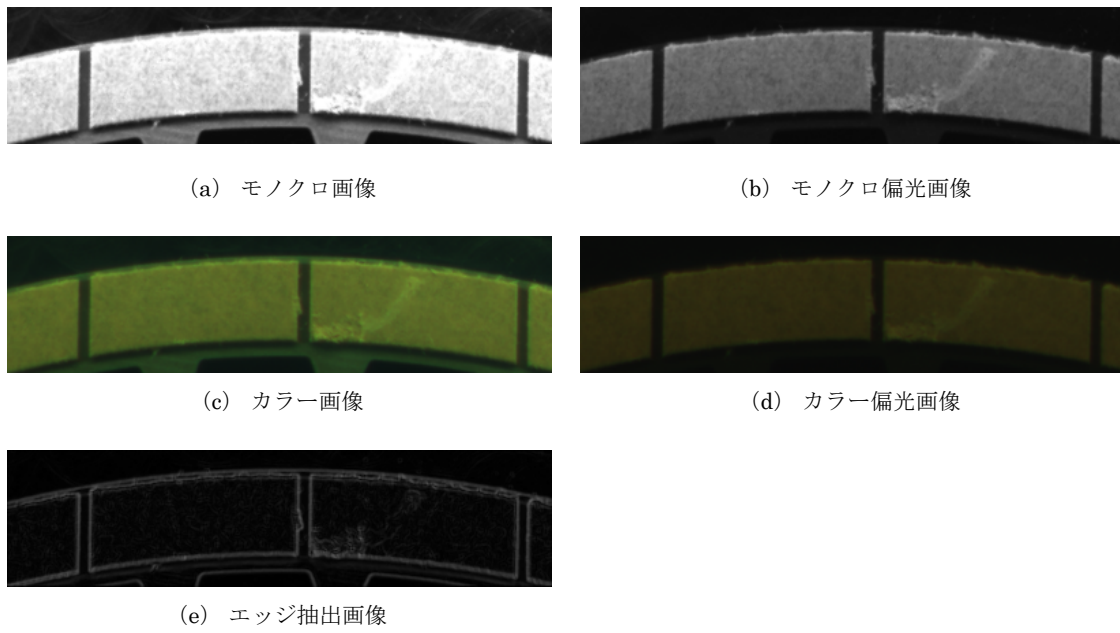


図 4.12 撮像方法別画像の例 欠陥種：キズ

4.1.3 データセット

データセットは 2 章で行った分類実験 1 と同様に、正常品クラスと 4 種の欠陥品クラスを含む 5 クラスのサンプル画像、合計 850 枚で構成された。各クラスから 20 枚ずつ取り分けた 100 枚の画像を検証データセットとし、残りの 750 枚を学習用の教師データセットとして用いた。また、(i)モノクロ画像のデータセットは 2 章の分類実験 1 で用いたものであるため、比較実験の結果には分類実験 1 の結果を用いている。

4.1.4 学習モデル

学習に使用したモデルは、画像がモノクロの場合とカラーの場合で入力サイズのみが異なるが、その他は分類実験 1 と同じ構成とした。

4.1.5 実験結果

5 つの異なる画像データセットを用いて分類を行った時の正解率を図 4.13 に示す。どのデータセットでも正解率は 90%前後となったが、(ii)モノクロ偏光画像データセットで学習したモデルの正解率が最も高く 92%となった。撮像方法による分類性能の変化は、検査対象や欠陥種などによって異なると考えられるが、本研究で対象とするクラッチディスクの検査においては偏光フィルタの利用が非常に有効であることが確認できた。

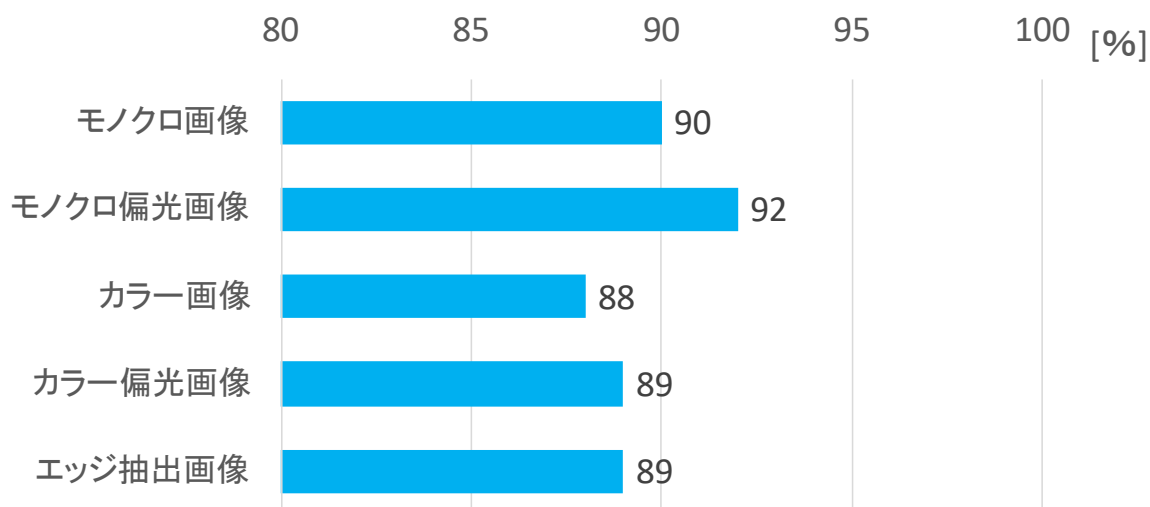


図 4.13 比較実験の結果

4.2 偏光撮像による画像の特徴

偏光とは特定方向に振動する光のことを言う。実験では偏光を得るために偏光フィルタを用いている。偏光フィルタとは、特定の方向に振動する光のみを透過し、それ以外の光を透過しない性質を持った素材である。そのため、偏光方向（透過方向）を並行にした 2 枚の偏光フィルタを重ねた

場合、光は透過されるが、直行させた場合には、光は遮蔽される（図 4.14）。本実験の偏光モノクロ画像撮像環境では照明側とカメラ側それぞれに偏光フィルタを設置しており、その偏光方向を直行させている。照明から発せられた無偏光の光は照明側偏光フィルタによって特定方向に偏光する。その後、被写体からの反射光はカメラ側偏光フィルタで再び偏光され、カメラの撮像素子によって受光され、画像が取得される。本実験で取得されている偏光モノクロ画像を観察してみると、通常モノクロ画像で確認できていた芯板に見られる縞模様、摩擦材下面の接着液、照明による影、異物エッジ部の反射は除去されており、背景や芯板部分は全体的に輝度が低く、暗く写っている。一方でペーパー摩擦材部分は輝度が高く、明るく写っている（図 4.15）。これは、芯板部の反射光は偏光向きが変わらないために、カメラ側の偏光フィルタで遮蔽されてしまい、光が撮像素子に届かないが（図 4.16 (a)）、ペーパー摩擦材部の反射光は、偏光の向きが変わっていることによって、カメラ側偏光フィルタを透過する成分があるため（図 4.16 (b)）と考えられる。本実験の撮像対象であるクラッチディスクの場合、偏光撮像を行うことによって、ペーパー摩擦材部分のみを浮き上がらせるような撮像が可能となる。

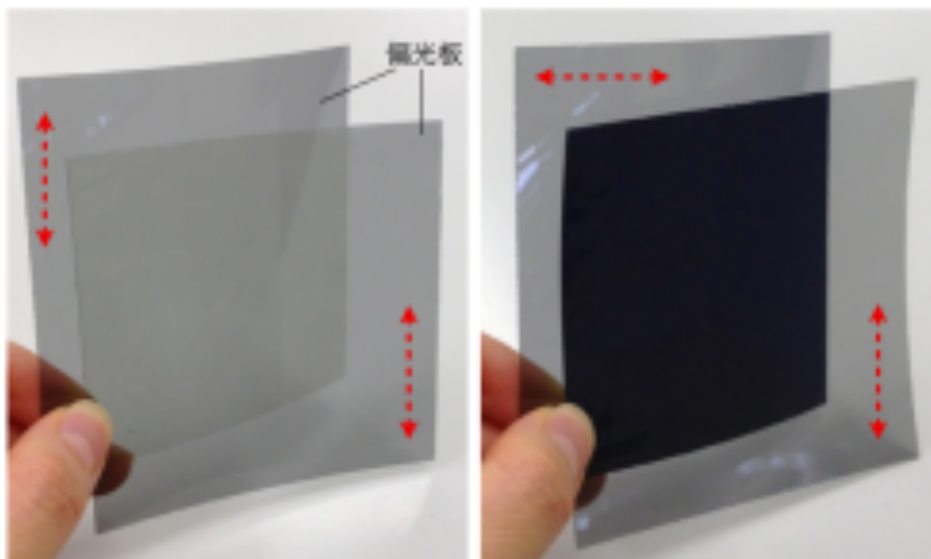


図 4.14 偏光フィルタを 2 枚重ねたときの様子（赤矢印は偏光フィルタの偏光方向）
 （引用：光範実物体と影の変形を用いた情報の秘匿と選択的視覚化 [57] 図 1）

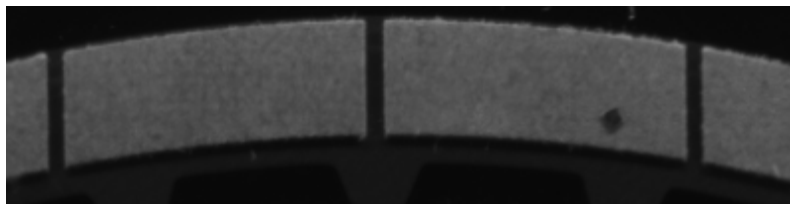
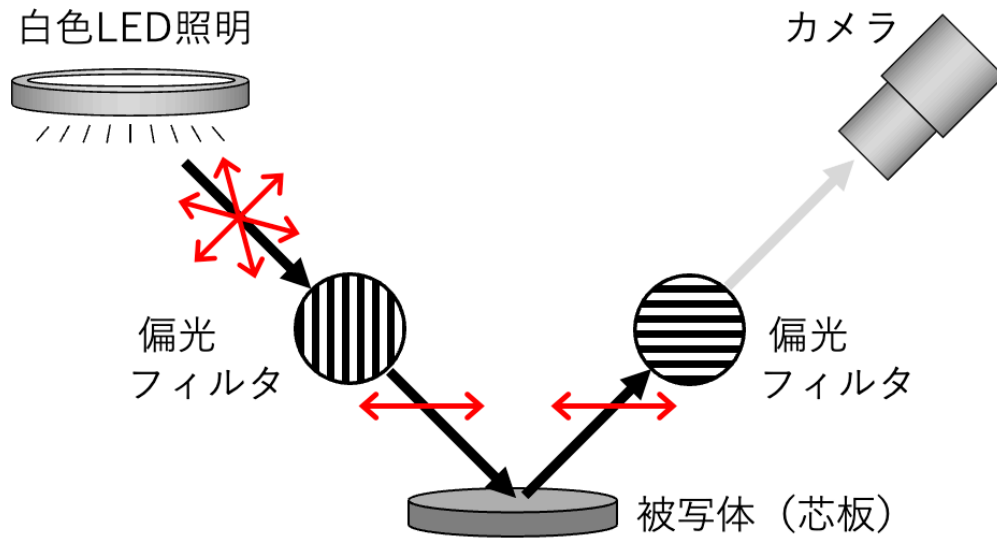
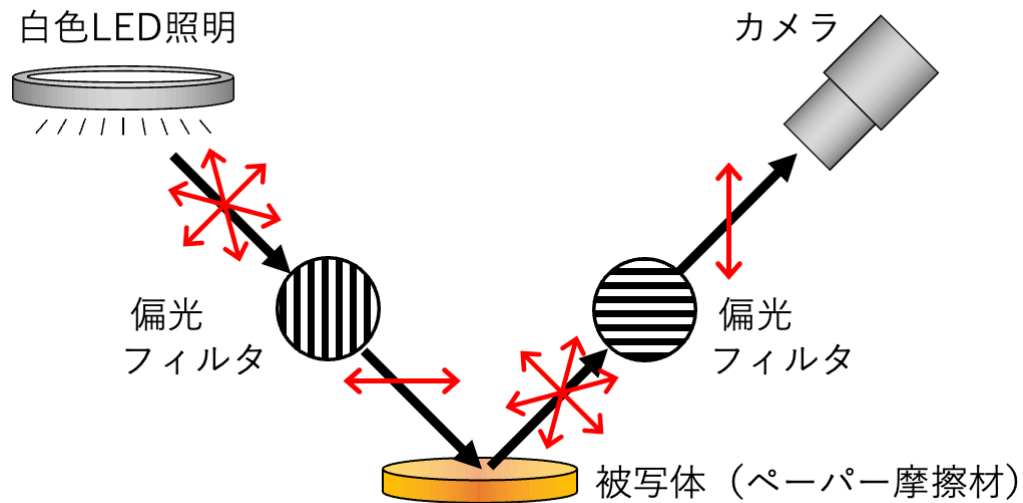


図 4.15 偏光フィルタの設置によりペーパー摩擦材のみが抽出された画像（図 4.10 (b) 再掲）



(a) 芯板部の反射光



(b) ペーパー摩擦材部の反射光

図 4.16 照明光と反射光の振動方向模式図（赤矢印は光の振動方向）

偏光撮像によって除去された情報は、欠陥検出においては不要なノイズであると考えられることができる。ノイズの多い画像は CNN の分類性能を低下させ、モデルの学習時間を増加させる[58]。画像からノイズを除去することは、画像処理における重要な課題であり、前処理として行われることが多い。画像処理によるノイズ除去の方法には、画像を一度ぼかし、その後にエッジを強調することで画像に含まれるノイズを抑制する方法などがある。しかし、ノイズ除去のための前処理は時間がかかる処理であるのに加え、CNN におけるノイズ除去単体での性能向上効果は限定的であるとの報告がある[59]。そのため、ノイズを含む画像に対して分類性能を向上させるための CNN アーキテク

チャが開発され、報告されている[58][59][60]。本実験における偏光画像は、取得した画像のノイズを抑制する処理を行うのではなく、画像取得時に不要なノイズを除去できている。つまり、ノイズを除去するための処理によって必要な情報の情報量削減が起きることがなく、より「質」の高いデータが取得できていると考えることができる。この事により、通常のノイズ除去方法のみでは分類性能向上が難しかった先行研究とは異なり、分類性能の向上が確認できていると考えられる。一般的な画像認識においては偏光画像を用いることは困難であるが、工業製品検査のような限定的な状況で画像を取得する場合、分類性能の向上に有効な手段となることが期待できる。

4.3 開発手法との組み合わせによる精度向上の効果

4.3.1 データセットと学習モデル

予備実験にて分類性能が高かった偏光モノクロ画像と開発した非対称ラベルスムージング手法を組み合わせる実験を行うため、3章の分類実験3に用いたサンプルを再度撮像し、偏光モノクロ画像データセットを作成した。データクレンジングを行った後の画像は正常品クラス画像が 2,045 枚、14種の欠陥品クラス画像が合わせて 5,092 枚、合計 7,137 枚となった。分類実験3同様、5つのサブセットに分割し、そのうちの4つのサブセットを教師データセットとして学習に用い、残りの1つのサブセットを検証データセットとして検証を行う試験を、組み合わせを変えて5回行った後、5回の試験結果を平均し、評価する交差検証を行った。モデルは入力画像データセットのみが異なり、モデルの構成やハイパーパラメータはすべて同じとした。

4.3.2 実験結果

平滑化パラメータ $\epsilon \in \{0.1, 0.2, \dots, 0.9\}$ としたときの見逃し率の推移を図 4.17 に示す。偏光モノクロ画像においても非対称ラベルスムージング手法によって見逃し率を低減できることが確認できた。見逃し率が最小になったのは通常のモノクロ画像と同様に $\epsilon=0.8$ の時に、その値もモノクロ画像と同じ 0.08% (5,092 件中 4 件) となった。 ϵ が小さい範囲で見ると、見逃し率は通常モノクロ画像と比較して低く推移している。同様に、誤警告率の推移を図 4.18 に示す。通常のモノクロ画像で学習したモデルと比較して同等の数値であることが確認できる。実験の結果より、偏光モノクロ画像を用いることによって分類性能が向上し、データセットの「質」を変えることによって分類性能が向上できることを示唆している。

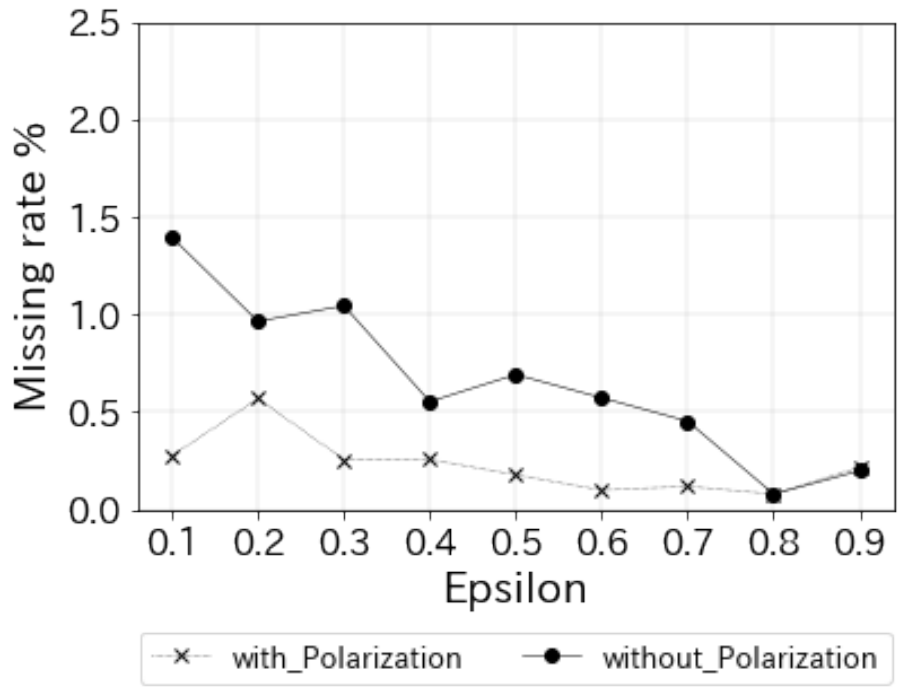


図 4.17 モノクロ偏光画像と非対称ラベルスムージング手法の組み合わせ結果（見逃し率）

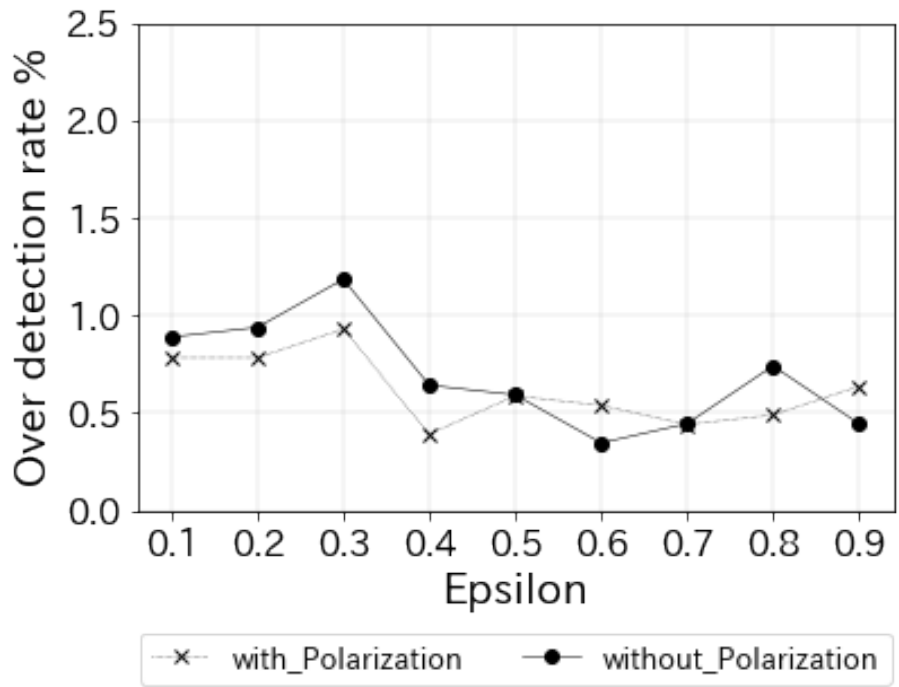


図 4.18 モノクロ偏光画像と非対称ラベルスムージング手法の組み合わせ結果（誤警告率）

4.4 本章のまとめ

CNN を用いた学習におけるデータの重要性に注目し、光技術を応用したデータ取得による分類性能向上に取り組んだ。5つの撮像方法で取得した画像データを用いて分類性能を比較した結果、偏光フィルタを用いた画像を教師データとすることで分類性能を向上させることができた。研究の対象であるクラッチディスクにおいては、偏光フィルタを設置することによって、背景や角部のテカリなどを抑え、ペーパー摩擦材部分のみの輝度を高く保つことができる。このことによって、学習時の特徴抽出に必要な欠陥の特徴のみを残し、余分な情報を削減することができるため、分類性能の向上が実現したと考えられる。第3章で開発した技術と組み合わせた実験では、欠陥検出工程における重要な指標である見逃し率で比較した場合、最も低い値はモノクロ画像と同じであったが、モノクロ偏光画像を用いて学習したモデルでは、平滑化パラメータ ϵ が小さい範囲で、さらなる性能向上の効果が確認できた。本実験結果より、CNNの性能に影響するデータの「質」向上には、光技術が有用であることを確認し、光技術によってCNNの性能をより引き出すことに成功したと言える。

実験は量産を想定した上で実用化可能性が高い5種類の異なる撮像法によって得られた画像を用い、全体最適化を目指して行ったが、クラッチディスクに現れる欠陥は多様であり、特徴を顕在化するための最適な撮像方法は、欠陥種や状態によって異なることが想定される。今後は特定の欠陥に対して最適な撮像方法を見つけ、組み合わせることでの検査性能向上や特定欠陥の高精度検査などの研究を行っていく必要がある。

第5章 組織内での新技術導入活動

AI は DX、スマートファクトリーのキーテクノロジーであると同時に、各種の自動化、さらには新しいサービスの実現にも不可欠な技術となっており、その活用への期待が高まっている。しかし、その導入状況を見てみると、現実には、国内企業の 80%が AI 導入に関心を持ちながらも、実際に導入している企業は全体の 20%程度、実証実験（Proof of Concept）を行っている企業を入れても 30%にも満たないと報告があり（図 5.1）、関心と導入率の間には乖離が見られる。日本企業における AI 導入率は 2019 年の 3.1%、2020 年の 4.2%から 2021 年の 20.5%と大幅に増加しており、乖離の程度は低くなってきているものの、米国企業の導入率 44.2%、実証実験を行っている企業を含めた 60.5%とは大きな差があり（図 5.2）、さらなる導入の促進が必要となっている。筆者の所属企業でも、AI 導入の検討は行われているが、実証実験や実際の導入に至ったケースは少なく、新しい技術の導入を加速するため、あらたな活動を始める必要がある。

AI 導入を検討するにあたっての課題は様々あるが、IPA によるアンケート結果の上位に注目すると、2つに類型化できる（図 5.3）。一つは AI 人材不足や社内の AI 理解不足からなる導入基盤の課題、もう一つは効果の見積もりが難しい、費用が高い、手軽なサービスがないことなどからなる導入の障壁の高さである。AI 導入を促進するためには、これらの課題にアプローチしていく必要がある。

本章では筆者自身が大学院大学の学生であるという環境を利用して専門スキルを習得することで AI 開発技術者となり、社内の AI 理解度向上を目指した取り組みを行いながら、AI 導入の障壁を下げる活動を行ったことによって、AI 導入を促進し、開発案件獲得の成果を得た実践活動内容について述べる。

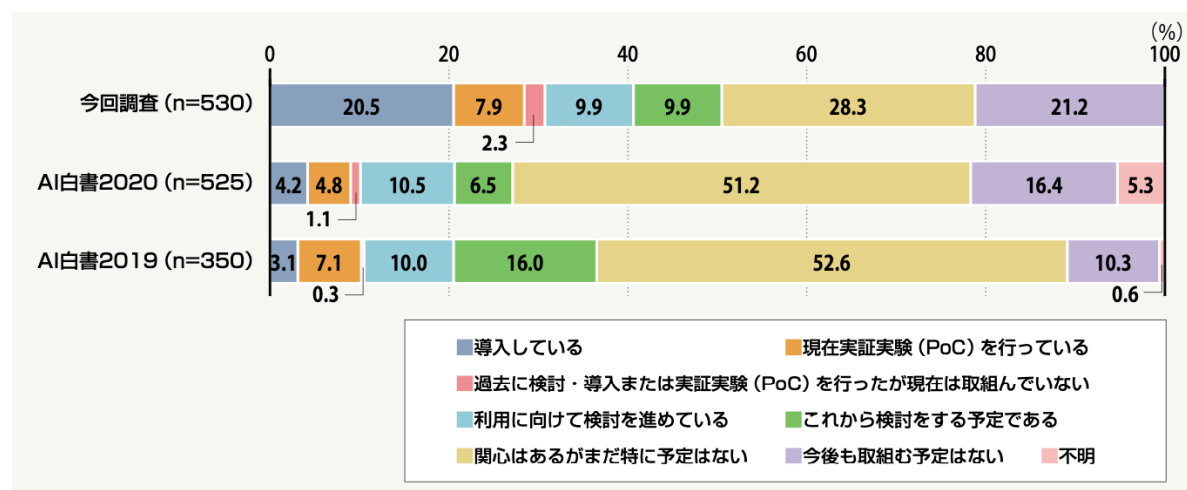


図 5.1 国内の AI 技術の活用状況
(出典：DX 白書 2021 [5])

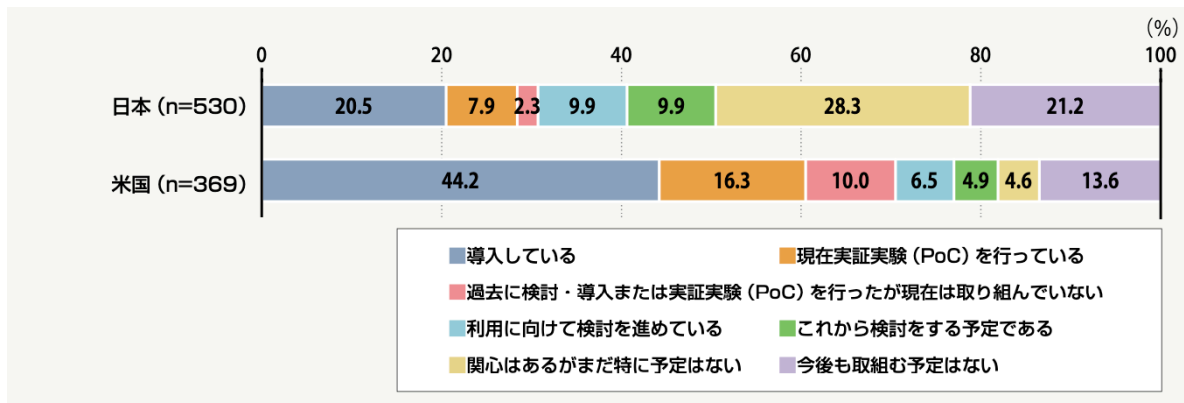


図 5.2 AI 技術の活用状況の比較
(出典：DX 白書 2021 [5])

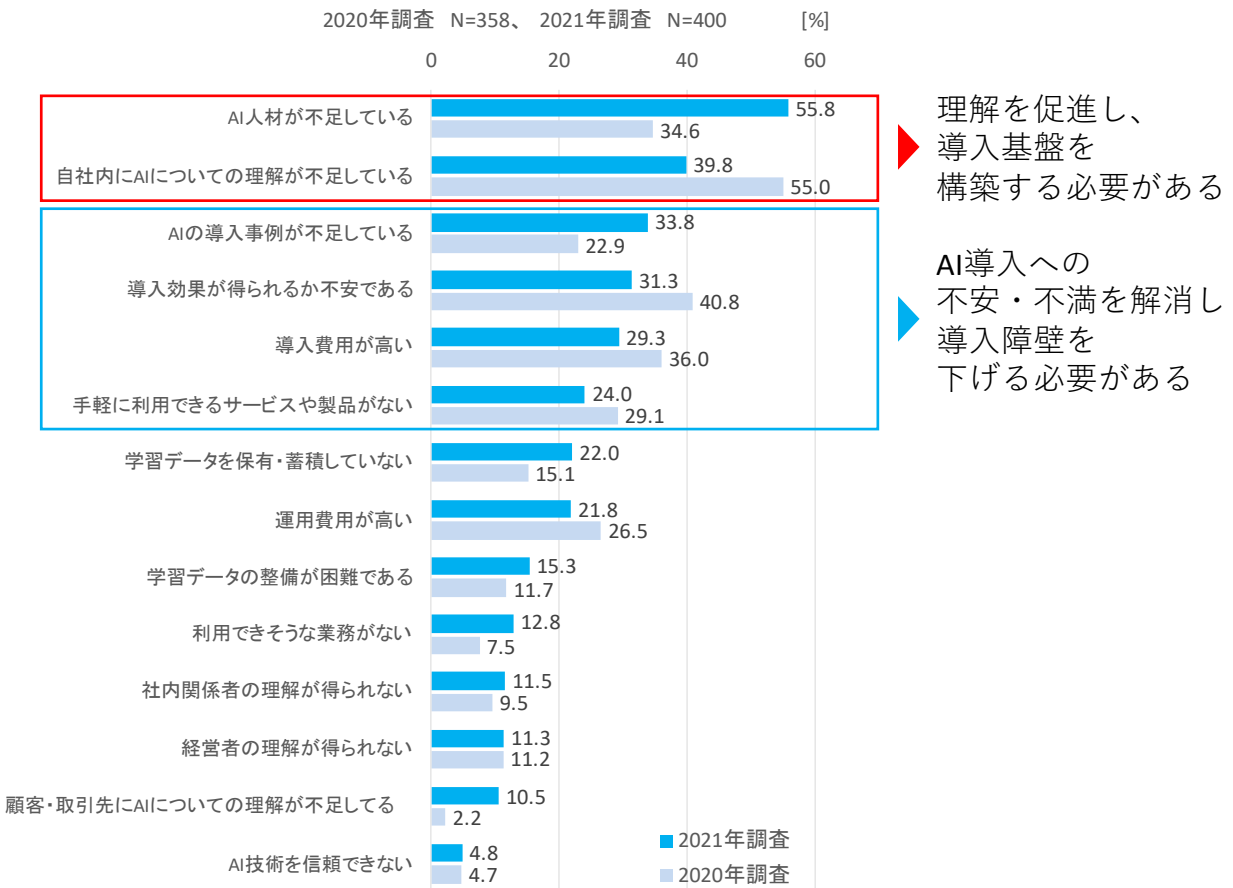


図 5.3 AI 導入検討するにあたっての課題の類型化
(引用：AI 白書 2020 [13] p306 図 3-4-21 および DX 白書 2021 [5]p283 図 42-53 をもとに作成)

5.1 新技術導入基盤の構築

AI 導入に関心を持つ企業が多くあるのに対して、導入や実証実験を行えない課題はどこにあるかについて、IPA によるアンケート結果によれば、導入が進まない一番の原因は、「AI 人材が不足している」こととなっている（図 5.3）。AI 導入の基盤を構築し DX 推進を加速させるためには、その技術開発のみならず、企業内で AI に対する理解度を高め、AI エンジニアを育成する取り組みが必要となっている。

そこで、本研究では、自社内において、AI 活用による DX において中心的役割を担う生産技術部のエンジニアを対象に AI に対する理解度向上を図る取り組みを行った。社内 AI 人材育成が進まない原因の一つに、AI に実際に触れ、知識や技術を習得する機会が極めて少ないことが挙げられる。そこで、AI に対する理解度の向上を目的とした講習会を開催する。この講習では、知識学習だけではなく、実際にプログラミングを行い、外部機器との連携を行うという実践的な演習も実施する。そして受講後に、参加者に対してアンケート調査を行い、調査結果をもとに、AI に対する理解度向上の効果と、導入事例創出のための課題について考察する。

5.1.1 研究方法

生産技術エンジニアは、システムの導入を推進し、維持管理することが求められる。そのため、システムの操作方法を知るだけでなく、背後にあるプログラミングの基礎的な知識や仕組み、考え方の理解が必要となる。本研究では生産技術エンジニアに対し、AI 活用の基礎となるプログラミングの講習を行った。そして、参加者へのアンケート調査の結果をもとに、AI を始めとしたデジタル技術に対する理解度向上の効果と、導入事例を創出するための課題について考察を行った。

講習を行うにあたっては、アクティブ・ラーニングの視点を取り入れた。文部科学省[61]によれば、アクティブ・ラーニングは「主体的・対話的で深い学び」とされ、学習過程の改善手法として取り上げられている。「新たな未来を築くための大学教育の質的転換に向けて生涯学び続け、主体的に考える力を育成する大学へ(答申)」[62]の用語集[63] (2012) では「教員による一方向的な講義形式の教育とは異なり、学修者の能動的な学習への参加を取り入れた教授・学習法の総称。学修者が能動的に学修することによって、認知的、倫理的、社会的な能力、教養、知識、経験を含めた汎用的な能力の育成を図る。発見学習、問題解決学習、体験学習、調査学習なども含まれるが、教室内でのグループ・ディスカッション、ディベート、グループ・ワーク等も有効なアクティブ・ラーニングの方法である」(p37)と記載されている。今回の講習は問題解決学習、体験学習、調査学習、グループ・ディスカッションの要素を含んでいる。プログラミング教育にアクティブ・ラーニングを取り入れた先行研究に田中らの報告[64]、高岡らの報告[65]がある。これらの報告では、アクティブ・ラーニングの視点を取り入れ、講習を実践的、実用的な内容としたことで、プログラミングの知識やスキルの習得を促進し、理解度や意欲の向上の効果があつたと述べられている。先行研究では事前学習や事前知識が求められており、そのような学習時間は含まれていないが、田中らは 45 時間程度、高岡らは 13 時間程度と、どちらも比較的短時間での講習の報告であり、その点、本講習と類似しており、アクティブ・ラーニングを取り入れた実践事例として、同様の効果が期待できる。本研

究では、プログラミング初学者を対象にした短期講習にアクティブ・ラーニングを取り入れることで、理解度向上の促進を狙った。

5.1.2 講習の概要

講習は2020年8月、2020年9月、2021年4月の3回開催した。講習では毎回生産技術エンジニアからAI利活用の人材候補である4名が選抜され、合計12名に対して講習を実施した。講習は14時間のカリキュラムを2日間に分けて行った。講習は講義形式で基礎知識や仕組みなどの言語化されているものを「知識」として学習するだけではなく、実際に機器を操作し、プログラミングを行い、外部機器との連携を行う実践的な演習を経験することによって「スキル」を習得する内容とした。講習時間の3割は座学による「知識」学習に、7割は演習による「スキル」習得に割り当てた。アクティブ・ラーニングの観点から、スキル演習では課題を伝えた後、講師から答えを教えるのではなく、一緒に悩みながら、受講者が能動的に問題解決に取り組むように促した。講習資料作成、環境構築、当日の講師などに関しては、社内IT部門有識者の協力を得て行われた。

講習で用いるプログラミング言語はpython（パイソン）とした。pythonはAI開発によく用いられる言語で、可読性が高く、初心者にも扱いやすいという特徴がある。プログラミングの環境を構築するための機器には教育用の小型コンピュータであるRaspberryPi（ラズベリーパイ）を用いた（図5.4）。RaspberryPiはコンピュータ、コンピュータサイエンスなどの教育を目的として作られた小型コンピュータで、複雑な環境設定をすることなくpythonを実行可能である。加えて、汎用の入出力ポートを備えているため、実際の生産設備を模擬した信号のやり取りが手軽に行えるという利点がある。課題とした外部機器との連携は、生産ラインでよく用いられる物体検出センサとカメラを用い、物体を検出した場合にカメラで画像を撮像するという、実際の設備を模擬した内容になっており、生産技術エンジニアにとっては実務での活用がイメージしやすい内容となっている。演習のシステム構成を図5.5に、実際に使用した機器を図5.6に示す。



図 5.4 RaspberryPi（ラズベリーパイ）
（出典：ラズベリーパイ HP）

<https://www.raspberrypi.org/products/>

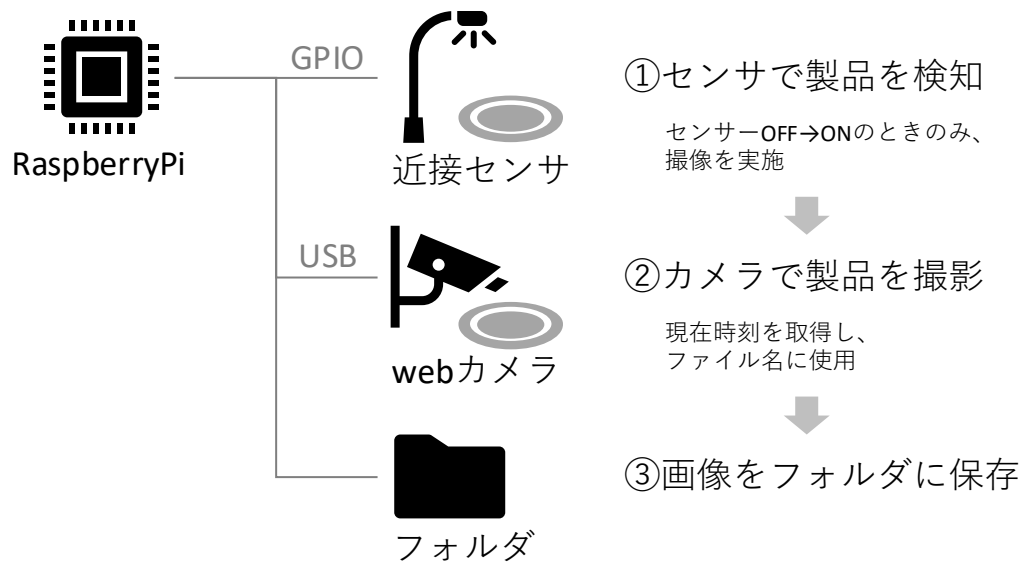


図 5.5 実際の設備を模擬したシステムの構成図

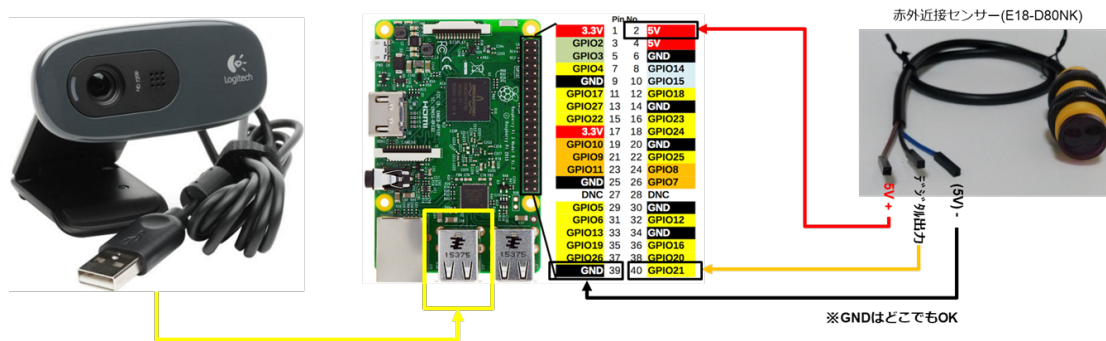


図 5.6 講習で使用した機器の写真

5.1.3 アンケート調査

アンケートは講習による AI、IoT、デジタルに対する理解度向上の効果と講習で得た「知識」と「スキル」の業務への活用状況、活用ができていない場合の理由などを調査する目的で行った。調査対象は講習の受講者 12 名とした。アンケートは調査対象 12 名のうち 10 名から回答を得た。

調査は受講から 4 ヶ月～12 ヶ月を経過した 2021 年 8 月に実施した。アンケートはあらかじめ用意した設問に web アプリケーションを用いて回答する方式で行われた。調査項目の内、必須回答とした項目を表 5.1 に示す。必須回答項目については 5 段階の選択式もしくは複数選択可能な選択式とした。各設問のあとには任意で理由などをコメントできる回答欄を設けた。

表 5.1 アンケート内容

1. 講習の内容には満足いただけましたか。
2. 講習前、ご自身の AI、IoT、デジタルへの理解度はどの程度だったと思いますか。
3. 講習を終えて、AI、IoT、デジタルへの理解度はどの程度だと感じますか。
4. 講習前、AI、IoT、デジタルについての興味はどの程度でしたか。
5. 講習を終えて、AI、IoT、デジタルについての興味はどの程度ですか。
6. 講習前、AI、IoT、デジタルの取り組みに対して不安感・不信感がありましたか。
7. 講習後の今、AI、IoT、デジタルの取り組みに対して不安感・不信感がありますか。
8. ご自身の業務に講習で得た知識が活用できそうだと感じますか。
9. ご自身の業務に講習で得たスキルが活用できそうと感じますか。
10. 講習は継続して開催したほうが良いと思いますか。
11. 知識やスキルの活用はどんなフォローがあればできそうですか。
12. 関連講習があるとしたらどのような内容があると嬉しいですか。

5.1.4 調査結果と考察

①理解度の調査：理解度に関する調査結果を図 5.7 に示す。講習前の AI、IoT、デジタル技術に対する理解度について、6 名が「全然理解していなかった」と回答したのに対し、講習後「全然理解できていない」と回答したのは 1 名のみとなっている。受講後は 8 名が「少し理解できている」と回答しており、講習を通して理解度を向上できたと考えられる。受講者の「講習前よりは理解は深まったと思うが、まだ知識不足だと感じる」や「概要を把握することが出来た」といったコメントが示す通り、「よく理解している」、「おおむね理解できている」まで理解度を大きく向上させることはできていない。

②興味の調査：興味に関する調査結果を図 5.8 に示す。講習前から受講者らは AI、IoT、デジタル技術への興味の程度が高く、10 名中 8 名が「すごくあった」もしくは「少しあった」と回答した。講習後にはさらに興味の程度が高まり、10 名中 10 名が「すごくある」もしくは「少しある」と回答した。講習を通して興味の程度が向上したことが確認できる。受講者からは「いろんなことにチャレンジしたいと思った」、「もっと知識・経験を深められたら良い」、「理解して運用が出来れば日頃の生活、仕事が効率的に出来ると思う」といった意欲的なコメントが得られた。

③知識・スキル活用状況の調査：講習で得た知識の活用状況に関する調査結果を図 5.9 に示す。自身の業務に知識が活用できそうと回答したのは7名であった。そのうち、1名は実際に活用できていると回答した。同様に、スキル活用状況に関する調査結果を図 5.10 に示す。自身の業務にスキルが活用できそうと回答したのは4名で、そのうち1名が実際に実務で活用できていることが分かった。一方で、活用できていない理由として「自身の業務に活用できるものが見当たらない」や「従来技術からの切り替えがイメージできない」といった意見が得られ、知識やスキルを活用できる機会の創出や、先行事例の創出が必要であることが明らかになった。知識やスキルを活用するためにどのようなフォローを望むかについての調査結果を図 5.11 に示す。もっと基礎的な講習や復習、スキルアップを望む声が多く、実際に活用が進んでいない状況であるためか、OJT（On the Job Training）を望む声はなかった。

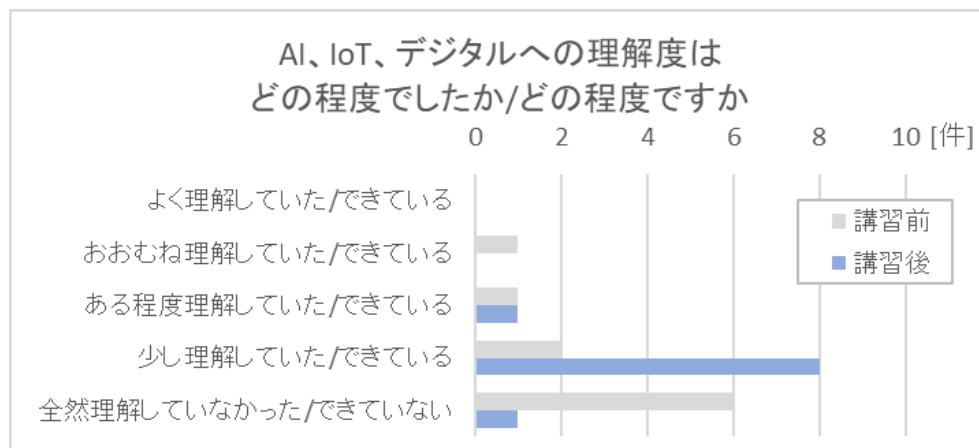


図 5.7 理解度の調査結果

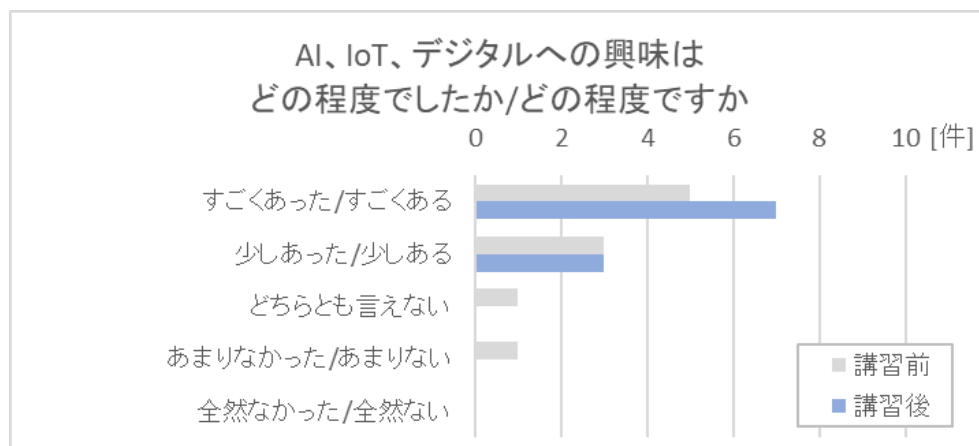


図 5.8 興味の調査結果

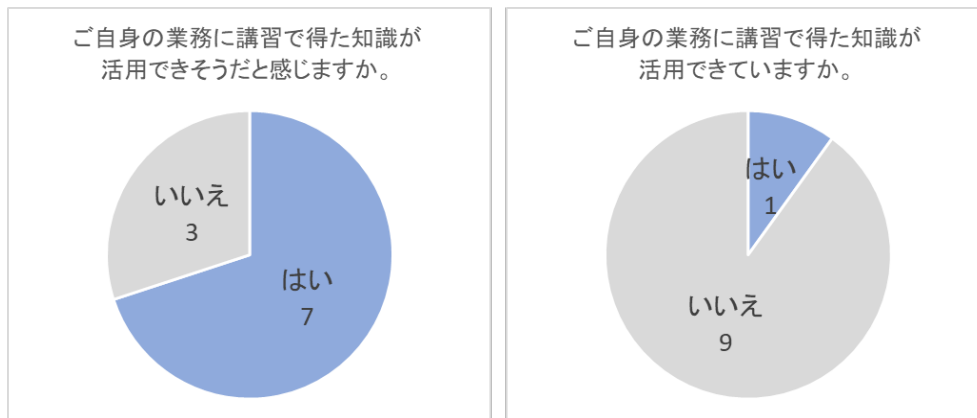


図 5.9 知識活用状況の調査結果

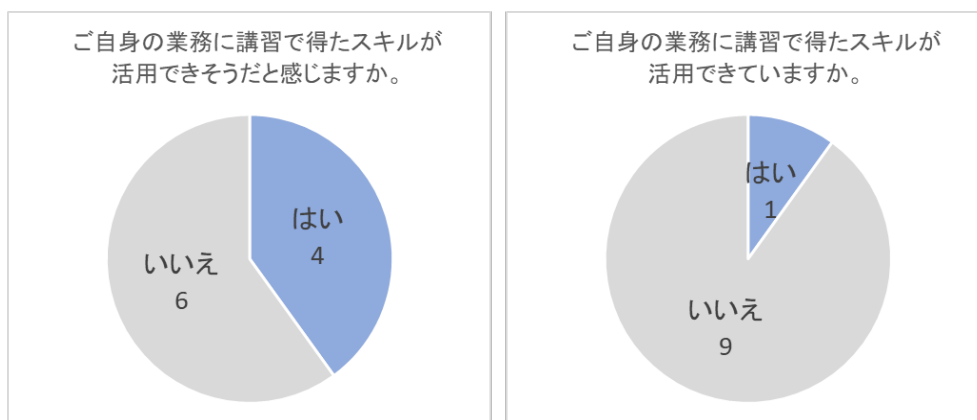


図 5.10 スキル活用状況の調査結果

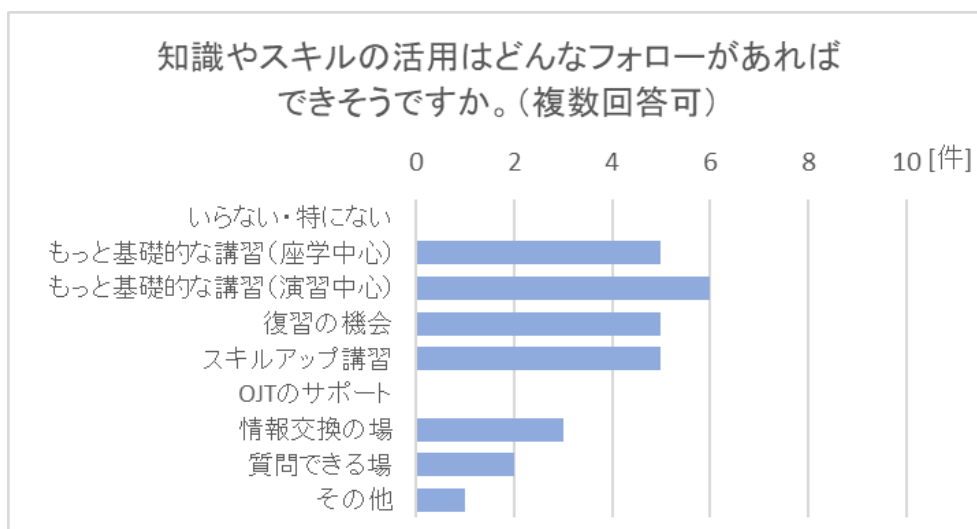


図 5.11 知識・スキルの活用に必要なフォローの調査結果

5.1.5 新技術導入基盤構築活動のまとめ

講習の実施によって理解度と興味の程度を向上する効果があることが確認できた。受講者の多くは講習後に理解が向上したと回答しており、一定の効果が確認できた。しかし、講習は短期間での基礎講習であり、AI や IoT の入り口に触れるにとどまっている。そのため、理解度を大きく向上させることはできなかったと考えられる。興味の程度は講習前から高く、意欲的な人材が多いことが分かった。導入に対して意欲的なコメントも得られている。講習にアクティブ・ラーニングを導入し、実際の作業を行い、自らの手で演習を行ったことも理解度向上に寄与していると考えられる。実際の AI 開発ではただ机に向かってコードを書くだけではなく、課題解決のためにトライアンドエラーを繰り返し、分からないことを調査し、協働者・ユーザーとのディスカッションを行う必要がある。アクティブ・ラーニングを取り入れた講習は実務に沿った形であり、講習内容の理解促進のためのみならず、実務における問題解決力を養う意味でも、AI を開発するエンジニアの育成にも適した学習方法であると考えられる。

知識・スキルの活用状況については、1 つの活用事例を創出できた。反面、講習から 4 ヶ月～12 ヶ月経過しても自身の業務に AI、IoT の活用ができていないのはほんの僅かであることが明らかになった。その原因として挙げられたのは、自身の業務で活用する機会がない、使えそうな業務が見当たらないというものである。各企業は DX をはじめとしたデジタル推進活動を進めている。しかしながら先にも述べた通り、国内企業の多くは未だ AI の導入に至っていない。このように AI 導入が進んでいない状況では知識やスキルを活用できる場面が限られていることが原因と考えられる。

5.2 新技術導入に関する導入障壁へのアプローチ

AI 導入が進まない一番の原因となっている自社内の AI 知識不足を解消するために、生産技術エンジニアに対してプログラミングの講習を行い、理解度向上の成果を得た。しかしながら、講習後のアンケート調査では、講習で得た知識やスキルが実務に活用できているケースは限られていることが明らかになっている。このことは、AI 導入が進んでいない状況では知識やスキルを活用できる場面が限られていることが原因と考えられる。企業の多くが AI 導入を検討しながらも、なかなか進められない状況は、その他の要因がいくつも絡んでいる複雑な問題であることを示しており、AI 導入を進めるためにはそのような問題に対して多面的にアプローチしていく必要がある。

5.2.1 システム内製化の必要性

企業が導入や実証実験を行えない課題はどこにあるかについて、一番の原因は、「自社内に AI についての理解が不足している」であったが、それ以外の主な課題には以下が挙げられている。これらは、AI 導入の障壁となっており、その解決が望まれている。

- 導入効果が得られるか不安である
- 導入費用が高い
- 手軽に利用できる AI のサービスがない
- 運用費用が高い

このような課題が上がってくる背景の一つにシステム開発の内製化率の低さがあると考えられる。これまで多くの日本企業は、自社システムの開発・運用・保守をベンダーに委託することが多かった。ベンダーにシステム開発を委託することは、自社で高度な技術者を育成、雇用する必要がない、開発環境の設置、維持管理のコストがかからない、短期間でシステム開発が可能、システムが一定の品質を持って納入されるなど、社内開発と比較してメリットが多い。しかし、ベンダーへ「依存」する傾向が強くなってしまいあまり、導入効果を確認するための簡単な実証実験や比較的小規模なシステムの開発、軽微なシステム改修でもベンダーに依頼しなければならないケースが一般化し、その度に費用が発生することや、調整工数を要するなどのデメリットが目立つようになっている。特に新しい技術は不確実性が高く、効果の見通しが得にくいため、検証を十分に行えない状況では高額なシステムの導入に消極的になる企業は多いと考えられる。AI の導入についても同様に、ベンダーへのシステム開発依存体質が導入効果への不安を生み、導入促進の障壁の一つになっていると考えられる。

先に挙げた課題について、実際にベンダーを活用したシステム導入を検討している側にたって考えてみると、以下のように解釈できる。

- 導入効果が得られるか不安である：様々な導入事例の紹介はあるが、自身の課題に対して有効なシステムであるかの判断ができない。少量のテストサンプルのみで行われた簡易実験の結果だけでは、自身の課題が解決できるかわからないため、システム購入に踏み切れない。
- 導入費用が高い：購入するシステムは汎用性が高く、様々な機能を備えているが、導入費用が高い。実際に導入した場合には限定的な機能しか使用しないため、限定的な機能を安く導入したい。反対に、汎用的なシステムでは機能が十分ではない場合に、購入したシステムをベースにメーカーオプションやカスタマイズを行う場合が多い（図 5.12）が、その場合には追加で費用が発生してしまう。
- 手軽に利用できる AI のサービスがない：導入効果を確認するための実証実験を行ってみたいが、そのためには費用が発生し、契約を結ぶ必要がある場合もある。また、比較的小規模なシステムの開発を行いたいが、汎用的なシステムは、様々な機能を備えており、導入費用が高いため、実験的な導入が難しい。

- 運用費用が高い：システムの更新を行わなくても保守費用が発生するケースが有る。また、対象の変更や僅かな設定変更、軽微なシステム改修でも、費用が発生する。

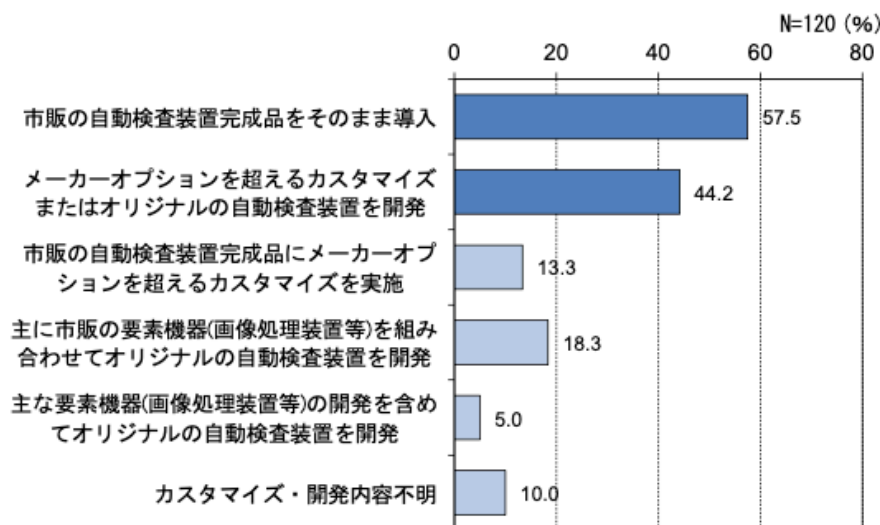


図 5.12 導入している自動検査装置
(出典：ものづくり企業の生産現場における検査の自動化促進可能性調査 [8])

このような状況を改善するため、自社でシステム開発や保守・運用を行うシステム内製化に取り組んだ。システムを内製化するには、高度な技術、知識、ノウハウを持った技術者を育成、雇用し、開発環境を整える必要があるなどの問題点もあるが、長期的に見れば上記の課題を解決できる。AI 導入企業のうち、約 6 割がコスト削減やノウハウ蓄積の観点から内製化を検討もしくは内製化に取り組んでおり (図 5.13)、非導入企業を含めた内製化検討率よりも非常に高くなっている。AI に限らない、多くの IT システムで内製化が求められる背景には、市場変化スピードの加速に伴う柔軟なシステム開発、データの重要性増大、社内人材育成・ノウハウの蓄積が挙げられる。不確実性の高い開発が多い現代において、内製化はその重要度を増しており、多くの企業でその取り組みが始められている (図 5.14)。

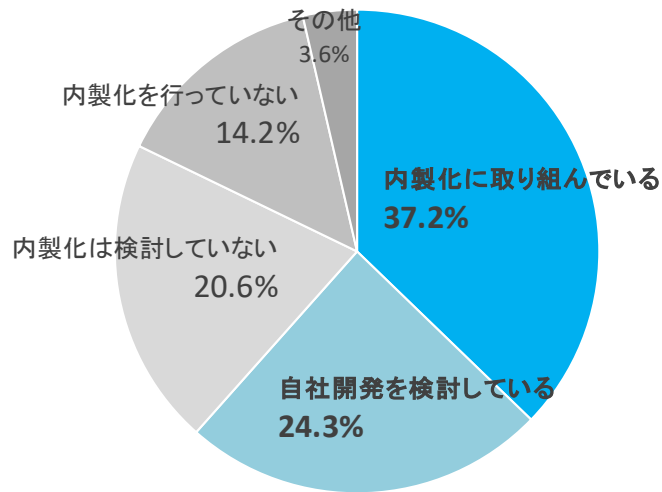


図 5.13 AI 導入企業の AI 開発内製化率
 (出典：AI inside 株式会社プレスリリースをもとに作成)
<https://inside.ai/news/2021/12/21/ai-researchresults/>

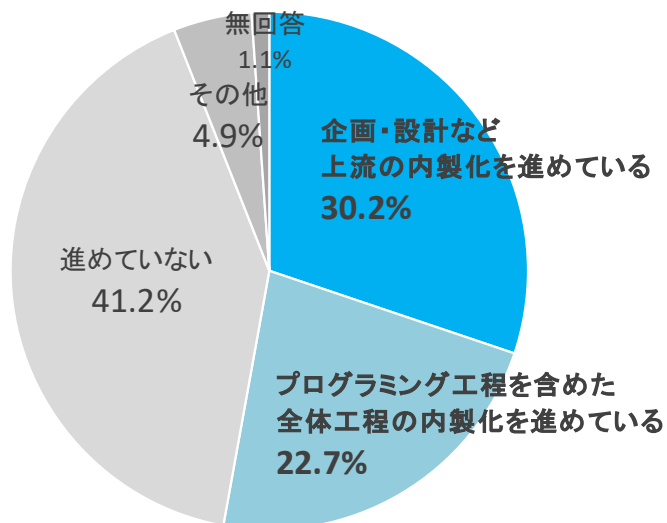


図 5.14 ユーザー企業が社内に IT のスキルを蓄積・強化するための内製化状況
 (出典：IT 人材白書 2020 [66] をもとに作成)

5.2.2 システム内製化の手段

システム内製化の手段として以下の 2 つについて取り組みを行った。

① シンプルなシステムの開発に特化する：AI システムを内製化するにあたり、ベンダーとの事業ポジションの違いを明確にしておく必要がある。内製 AI システム開発にあたっては、多機能高価格なベンダーシステムに対して、単機能低価格のポジションをとることで、差別化が可能となる（図 5.15）。ベンダーシステムは汎用性が高く、高機能なものが多い。しかし、実際の現場で運用を開始した場合に、使用する機能を変更することはほとんどなく、付帯している機能の多くが使われないことがある。また、製造の工程は一点一様であり、必要とされる機能も一般的なものと異なること

が多い。汎用的なシステムには必要な機能がなく、最初から組み込まれている機能のみで運用することができないため、システムをカスタマイズすることになり、多額の費用が発生するケースも少なくない。一方でシステムを内製化できれば、開発段階で顧客ニーズに特化したシンプルなシステムの開発を行うことができ、技術開発費、システム開発費を抑えることができる。また、検査項目や検査時間などの機能を限定することで、計算量を削減し、安価な計算機を採用できることから、ハード面での低価格化も期待できる。AI システム内製化は機能を絞ったシンプルなシステムとすることで実現可能性が増し、ベンダーとの差別化も行うことができる。

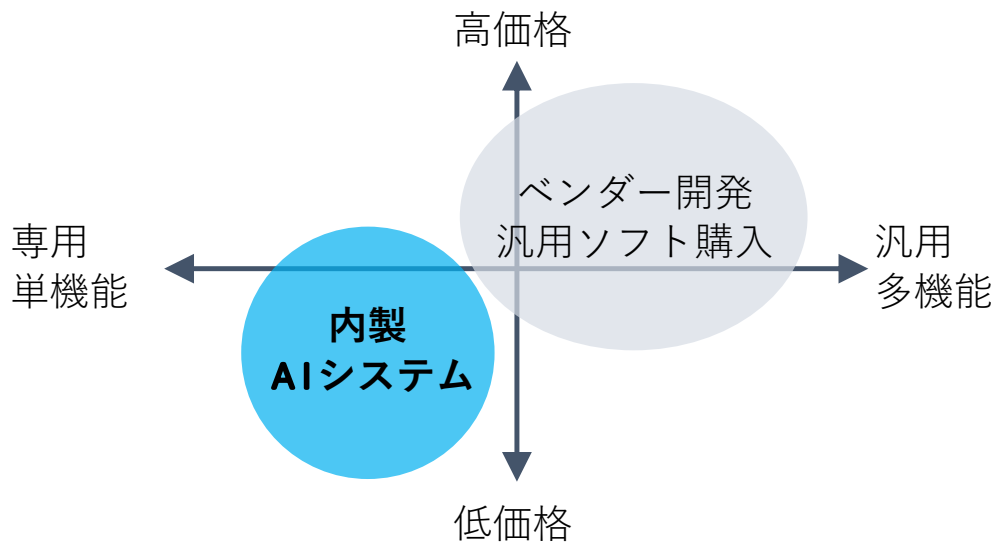


図 5.15 ベンダーシステムと内製システムの位置づけ

本研究の対象となっているクラッチディスクの完成品検査では摩擦材位置などの寸法検査が必要となっているが、内製 AI システム開発にあたっては、寸法検査を切り捨て、外観検査項目にターゲットを絞った。こうすることで既存検査装置（ベンダー依存型設備）との差別化（低価格化）を実現する。しかし、実際の工程では寸法検査の必要があるため、既存の寸法検査装置に AI による外観検査機能の追加を行える提案とする必要があった。そこで既存の画像検査装置と検査項目を棲み分け、同時に 2 つの検査を行う仕様とした（図 5.16）。設備費用を抑えるために、AI システム側のカメラには比較的安価な産業用カメラを用い、照明は既存の検査装置のものを共用した。AI システムを稼働させるための機器には安価な小型 PC である JetsonXavierNX（以下、Jetson）を用いた。Jetson は AI 推論用のデバイスではあるが、大規模な学習をするには能力が足りない。そこで、学習用の環境として、GPU を搭載した計算能力の高い学習用の計算機を用意し、オフラインで学習を行った。システムフローを図 5.17 に示す。実際の設備に設置された AI システムで画像を撮像したあと、データセットを生成し、学習用計算機にてモデルの学習を行った。完成したモデルを設備の AI システムにインストールし、推論に用いた。

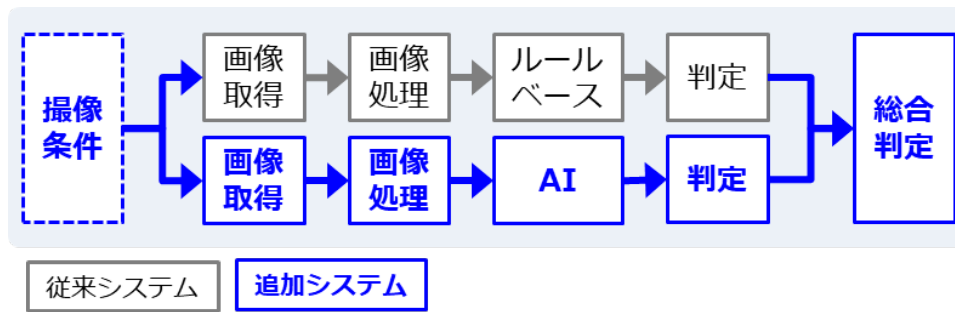


図 5.16 既存画像検査システムへの AI による外観検査機能の追加

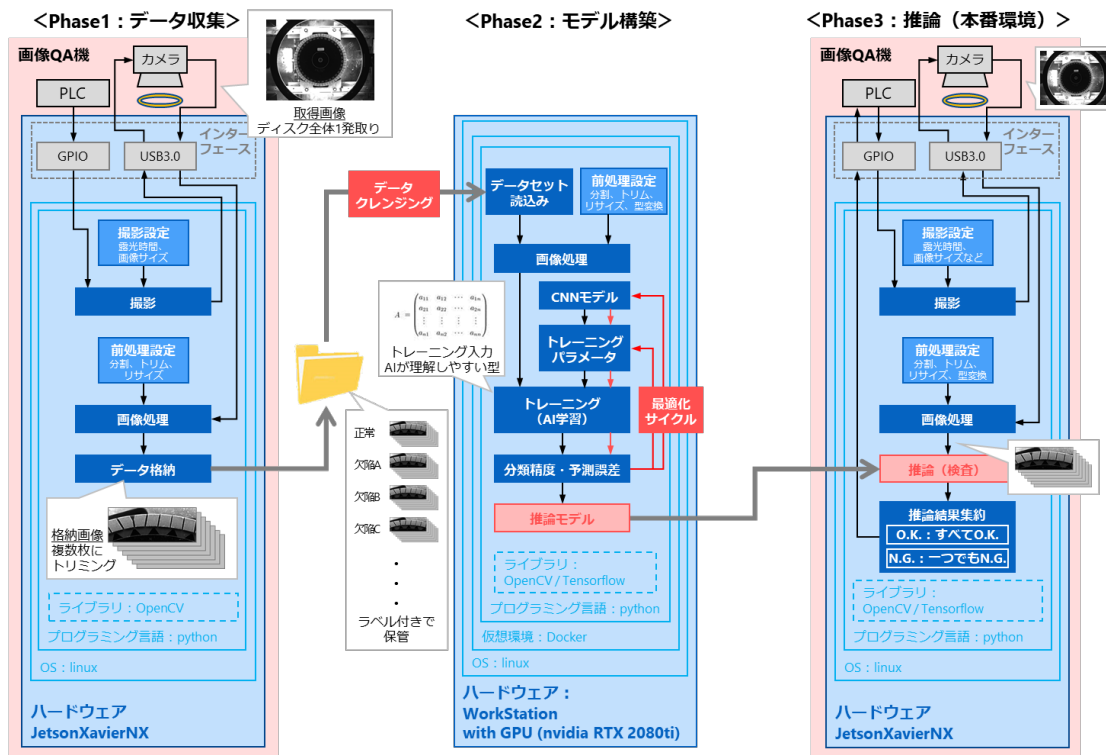


図 5.17 システムアーキテクチャ

本研究では内製化の実現のため、寸法検査を切り捨て、外観検査項目にターゲットを絞った。このように機能を絞り込むことによってシステムを単純化することができ、設備側に高価な PC を用いず、安価なエッジ AI デバイスによる実装が可能となる。また、システムの物理的なサイズもコンパクトとなるため、既存設備への追加設置が可能となり、ベンダーとの差別化が行える。

② 高度技術者の育成：内製 AI システムを構築するためには、当然、AI をプログラミングできるスキルを習得している必要がある。しかし、実際に求められるスキルはそれだけではない。実際の AI システムのうち、コードで構成されているのはごく一部で、周辺基盤には、データ収集、特徴抽出、データ検証、モデル評価、環境構築などを含み、膨大で複雑である（図 5.18）。筆者の所属企業では、AI 活用の検討が始まったばかりで、このような基盤は存在せず、当然、専門的なスキル保有

者は不在で、開発の環境も整っていなかった。すべての基盤に対してスキル保有者を確保し、開発環境を整備した後にシステム開発に取り組もうとしたのでは時間と費用が掛かりすぎる。本ケースでは前述のようにシステムをシンプルにしたことで、それぞれの必要スキル、環境は大幅に削減できる。しかし、これらをマネジメントできる技術者が必要であることに変わりはなく、そのような人材の育成は容易ではない。

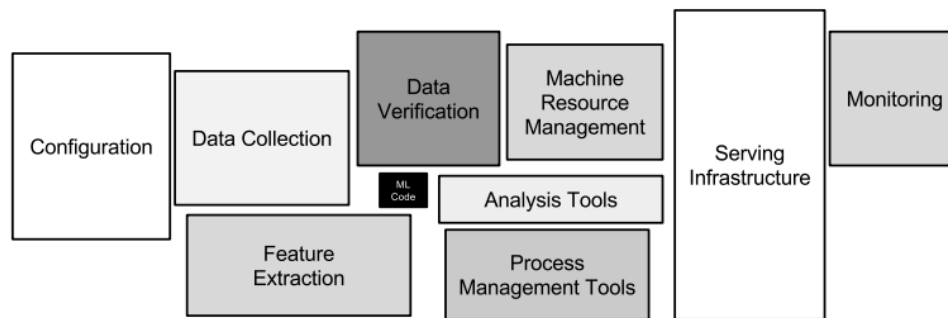


図 5.18 AI システムを構成する周辺基盤
 (引用：Hidden Technical Debt in Machine Learning Systems [67] Figure 1)

本研究では、高度技術者の育成に関して、大学院大学へ通学していることが大きな成功要因となった。通常、業務を遂行するのに必要なスキルは、企業内の OJT で習得することが基本である。しかし、全く新しい技術は基本的に OJT の機会がないうえ、社内にノウハウがないため、技術的課題を乗り越えることが難しい。セミナーや講習などに参加することが有効となることもあるが、本ケースのように周辺基盤を含めた開発が必要で、短期的にスキルを習得することが難しい場合には、大学院大学で長期的に学ぶことによるメリットが大きい。プログラミングスキルの向上、実験結果の深い考察はもちろんのこと、周辺基盤を構築するにあたってのノウハウ習得、顧客志向の開発マインドをもった開発を行うことができた。学びのモチベーションを維持する上でも非常に大きな役割を持っている[68]。社内ノウハウのない事例に対し、社外の環境で長期的に学ぶ機会を得たことで、AI 開発のスキルを習得し、周辺基盤の整備を行い、システム内製化スキルを手の内化することに成功した。

5.2.3 システム内製化のまとめ

AI を始めとした新しい技術の導入事例を創出するためには、自社組織内でシステムを内製できることが有効であると考え、AI システムの内製化に取り組んだ。ベンダーへの委託開発でも導入事例の創出は可能だが、新しい技術は不確実性が高く、効果が見通しにくいために、高額なシステム導入には消極的な企業が多いと考えられる。本研究では、必要機能に絞ったシンプルなシステムとすることで、内製化を加速させた。同時に、大学院大学への通学によって技術者を育成し、システム内製化スキルの手の内化に成功した。システム内製化を実現したことで、低コストのシステムが提供できるようになり、短期間で投資回収が見込まれる導入計画を策定することが可能となった。ま

た、実証実験を手軽に行える環境を整え、実際に動作する AI を目の当たりにできるようになった。これらの内製化効果によって、先に挙げた新技術導入の障壁の高さについて、以下のように改善することができた。

- 導入効果が得られるか不安である：様々な先行事例を参考に、自社内で実証実験を行い、自身の課題に対して有効なシステムであるかの判断ができる。導入効果が得られるようなシステムの構成を選択することができる。
- 導入費用が高い：現実の問題に合わせて、必要最低限の機能に絞ったオリジナルの安価なシステムを開発・提供できる。
- 手軽に利用できる AI のサービスがない：特定の工程に特化した単機能の小規模システムが開発可能なため、手軽に導入の検討ができる。
- 運用費用が高い：対象の変更や僅かな設定変更、軽微なシステム改修が社内に対応が可能となるため、保守契約費用などを削減することができる。

5.3 新技術導入活動実践の成果と実用化に向けた取り組み

5.3.1 新技術導入活動実践の成果

AI などの新しい技術はその有効性を感じながらも、技術に関する知識不足や効果の見積もりが難しいことから、なかなか導入が進まない。そこで、筆者が大学院大学への通学を経て、AI 開発スキルを習得し、組織内部へ理解者を生むためにエンジニアを対象とした技術講習を行い、組織全体の理解度を向上させると同時に、システム内製化により不確実性に対応できる体制を整えた。通学開始から導入案件獲得までの実績を図 5.19 に示す。AI 開発スキルの習得を行いながら、講習の開催及びシステム内製化を行い、AI 導入の体制が整った時期に、社内向け生産設備の特定機能のみをタスクとするシンプルな AI システム開発の依頼を複数件受けており、現在開発を推進中である。開発依頼の獲得に関しては、導入基盤を構築したこと以外にも、実際のシステムが低コストであり、短期間で投資回収可能な費用対効果の高いシステムであるという実務的なインパクトを持っていたことが大きな要因となっている。このように、本研究における新技術導入活動によって導入事例創出の成果を得ることができている。



図 5.19 AI 導入案件獲得に至るまでの実績

5.3.2 開発した技術の実用化に向けたロードマップ

スマートファクトリー構築の重要な取り組みに人作業の自動化があり、従来技術では自動化が難しい外観検査をターゲットとして、2章から4章まで技術開発を進めてきた。また、5章ではAIという新しい技術の導入基盤を構築するための活動に取り組んだ。活動の成果として得られた開発案件には完成品の外観検査システムが含まれており、開発した技術の実用化に向けて取り組みが始まっている。

外観検査システム実用化に向けては、大きく「導入」、「熟成」、「安定運用」の3つのフェーズを設定し、取り組んでいく。「導入」フェーズでは、検査システムの肝となる検査性能の整合や運用側へのAI理解促進が主な取り組みとなる。外観検査において、判定基準を定めることは容易ではない。導入段階で量産において発生する可能性のある全ての欠陥を準備することは不可能であるうえ、明確な基準を定めることが難しい欠陥も存在する。また、AI手法は外観欠陥の位置や大きさを定量的に評価することが難しい。この点については、運用側との整合を重ね、判定基準をできるだけ明確にしていくほか、テストサンプルを用いてAI判定と検査員判定の整合性検証などを繰り返し、性能を評価していく必要がある。また、AI手法は、例えば欠陥の見逃しがあつた場合に、なぜその見逃しが発生したのかの説明が難しいなど、従来のルールベース手法とは異なるAI特有の特性がある。実際に実用化を進める中では、このようなAI特性を理解し、受け入れてもらう活動が必要になる。しかし、検査性能をはじめとしたシステム要件を細部まで煮詰め、AIに対する理解を十分に得るには時間を要する上、実際に運用開始した場合に要件の変更・追加なども想定される。例えば欠陥検出においては、実際に流れてくる欠陥品すべてを想定できるわけではなく、運用しながら都度、最適化していく必要がある。そこで、優先度の高い要件を満たした時点で、工場への導入、運用を開始し、「熟成」フェーズとして検査システムと目視検査と併用してもらう期間を設ける。併用期間を設けることで品質保証能力を確保しながら、AIシステムの機能を監視し、検査員の判定との整合性を検証することができる。この期間にシステムをブラッシュアップし、運用を安定化させていく。また、この過程を通して、運用側から理解を得られていくと考えている。「安定運用」のフェーズでは、プログラム保守や機器メンテナンスができる要員を育成し、徐々に運用側に引き継いで、自立させ、水平展開できる体制を整えていく。

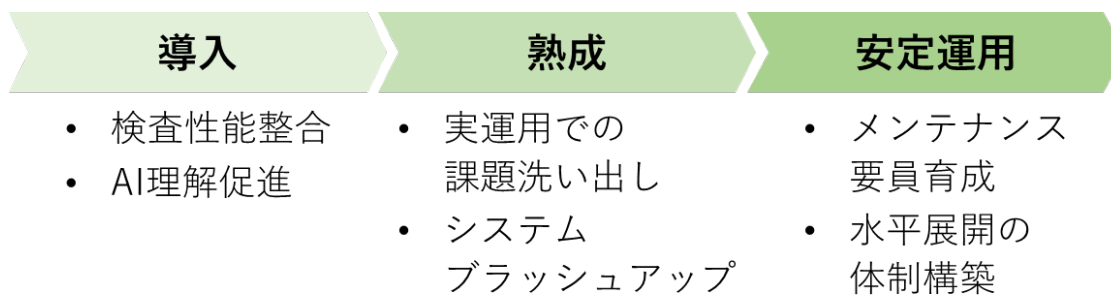


図 5.20 システム実用化のロードマップ

製造業における品質保証能力の確保は非常に重要な課題である。そこに AI のような従来技術とは異なる特性を持つ技術を導入し、実用化を進めていくためには、いくつもの課題に取り組んでいく必要がある。これらの 3 つのフェーズを経て、システムの実用化を実現し、継続的な改善を行っていくためには、運用側との連携が重要であり、欠かせない。開発側と運用側で協調し、新しい技術の実用化に向けての取り組みを進めていく。

5.4 本章のまとめ

多くの企業が AI 導入を検討しながらも、なかなか進められない状況は、いくつもの要因が絡んでいる複雑な問題であることを示しており、AI 導入を進めるためにはそのような問題に対して多面的にアプローチしていく必要があった。まず、デジタル進化基盤構築活動として生産技術エンジニアを対象に AI に対する理解度の向上を目的として、アクティブ・ラーニングを取り入れた講習会を開催した。講習によって AI に対する理解度向上の成果を得ることができた。次に効果の見通しが得にくいなど AI 導入に関する不安・不満を解消し、障壁を取り除くために、システム内製化の体制を整えた。実証実験を手軽に行い、実際に動作する AI に触れる機会を創出し、シンプルなシステムが提供できる体制を整えた。これらの活動によって、システム開発の依頼を複数件受けており、導入事例創出の成果を得ることができている。また、本研究の実務的なインパクトとして、省力化における費用対効果の高さがある。本システムを複数ラインに導入することで、要員削減が見込まれているが、製造ラインへの導入の際には、短期間で投資回収が可能な費用対効果の高いシステムが求められる。本システムは内製の低コストシステムであるため、短期間での投資回収が可能であり、このことが案件獲得の大きな要因となっている。

このような組織内への新技術導入を促進するための活動によって、導入を加速させることに成功し、DX、スマートファクトリー実現への足がかりを得ることができた。スマートファクトリー構築のためには単独の工程や小規模なシステムではなく、工場全体や工場をまたいだ大規模な相互連携システムが必要になる。この活動をきっかけに、小さな導入事例を重ねていながら、有りたい姿に向けての活動を継続する必要がある。また、本研究が取り組んだ AI 導入に関する課題は、製造業、サービス業、流通業、金融業を含む複数業種へのアンケート結果を基にしており、その課題の解決に取り組んでいる点で、様々な産業領域への展開が可能であると考えられる。また、AI に限らない新しい技術の導入に関しても、多くの企業が同様の問題を抱えていると考えられる。本研究の事例はそれらの課題解決手法の一つとして貴重な事例になり得る。

第6章 結論

本研究は、人に頼った作業の自動化を実現するため、光技術と AI を用いた完成品外観検査技術開発と、スマートファクトリーの構築を加速するための新しい技術を受け入れる組織体質造りを目的として取り組み、以下の4つを成果として得た。4つの成果はそれぞれ、欠陥検出としての意義、AI 開発としての意義、検査装置としての意義、新技術導入方法論としての意義を持つ。

① 欠陥品の見逃し抑制手法の開発

目視検査で経験的に取り入れられている欠陥品と正常品の判定過程の違いに着目し、完成品外観検査で最も重要な、欠陥品を正常と誤認する「見逃し」の抑制を実現する手法を開発した。検査員による目視検査では、見逃しを発生させないために、いずれの欠陥種にも属さず確実に正常品と判断されるサンプルである場合のみ正常品として分類する一方、欠陥の存在が疑われるサンプルは積極的に欠陥品として分類することがよく行われている。この判定過程の違いを非対称性と定義し、その機能を AI に実装した「非対称ラベルスムージング手法」を開発した。これにより、AI を用いた完成品外観検査に、欠陥品を積極的に検出する機能を付与することに成功した。

② AI の性格を誘導する手法の開発

AI 学習の指針となる正解ラベルを操作することにより、AI の性格を誘導する手法を考案した。学習過程では、教師データに付与されている正解ラベルと AI の予測した予測ラベルとを比較して損失値が算出される。モデルはこの損失値を最小とするように学習を進める。学習の重要な要素である正解ラベルの付与方針を各クラスで別にすることで、特定のクラスを積極的に検出する機能を持たせることに成功した。考案した手法は、AI が人の与えた指針に沿って特徴量を抽出するという点で、一般的な学習手法とは異なる。このことは AI の性格を誘導し性能を向上させる新たなアプローチであることを示唆している。

③ 実用化につながる検査性能の実現

欠陥品見逃し抑制のために開発した「非対称ラベルスムージング手法」を用いることによって、ルールベースの画像検査装置ではトレード・オフの関係にあった見逃し抑制と誤警告抑制を両立することに成功した。クラッチディスクを対象とした完成品外観検査を模擬した実験で、開発手法は一般的なハードラベルを用いた CNN 多クラス分類手法では正常品と判断した欠陥品の多くを検出することに成功し、見逃し率を 1.62% から 0.08% まで、誤警告率を 1.87% から 0.74% まで低減した。これにより、生産効率を犠牲にすることなく、エンジニアを困難な設定作業から開放し、より信頼性の高い検査工程の構築を実現できる。

④ 新技術導入の基盤を構築

在来技術と考え方が大きく異なる新技術を導入する際の課題として、新技術に対する理解の不足と、導入の障壁の高さの2点に着目し、解決に取り組んだ。企業内のAI理解不足を解消するため、アクティブ・ラーニングを取り入れた体験型の社内講習会を実施し、AIエンジニア候補の理解向上活動を行った。導入障壁を取り除くためには、システムの内製化に取り組み、実証実験を手軽に行い、実際に動作するAIに触れる機会を創出し、シンプルなシステムが提供できる体制を整えた。これらの活動により、新技術の導入基盤を構築し、開発案件獲得の成果を得た。

以上のように、本研究では外観検査自動化のための画像検査システム、特にクラッチディスク完成品における外観欠陥を検査するシステムの自動化技術開発とその技術導入を促進するために組織構築の実践を行った。本システムは製造業の量産工程における課題を解決するためのシステムであり、スマートファクトリーの実現やエンジニアの手助けに貢献できる。本研究の成果は人の視覚を代替えし、人を超える性能を実現する「光×AI」テクノロジーとして、製造業に限らずさまざまな分野での活用が期待され、光産業の創成に寄与することが期待できる。

参考文献

- [1] 内閣府: 統計表一覧(2021年10-12月期1次速報値), (2020).
https://www.esri.cao.go.jp/jp/sna/data/data_list/sokuhou/files/2021/qe214/gdemenuja.html
(参照 2022.6.1)
- [2] 経済産業省・厚生労働省・文部科学省: 2020年版ものづくり白書, (2020).
- [3] 一般社団法人日本自動車工業会: 日本の自動車工業 2020, (2020)
- [4] 経済産業省: DXレポート, (2018).
- [5] 情報処理推進機構: DX白書 2021, (2021).
- [6] 経済産業省: DXレポート2 中間とりまとめ, (2020).
- [7] 公益社団法人 精密工学会画像応用技術専門委員会 感察工学研究会, 公益財団法人ちゅうごく産業創造センター: 周辺視目視検査法の理解と導入のためのヒント, (2017).
- [8] 公益財団法人ちゅうごく産業創造センター: ものづくり企業の生産現場における検査の自動化促進可能性調査, (2016).
- [9] 秦清治: 外観検査の自動化現状と展望, 精密工学会誌 56, 8, pp1366-1370, (1990).
- [10] 須川成利: 撮像素子の高機能化と小型化, 光学, 42 巻, 3 号, pp351-356, (2013).
- [11] 秋田純一, 井口義則, 池辺将之, 丸山基之, 関根寛, 黒田理人, 香川景一郎, 小室孝, 頼田崇, 船津良平, 西山円, 廣瀬裕, 藤澤大介, 杉山行信: 情報センシングの研究開発動向, 映像情報メディア学会誌, vol.74, No.5, pp845-856, (2020).
- [12] パテント・インテグレーション レポート: 画像処理技術 (特許分析レポート・日本).
<https://patent-i.com/report/jp/theme/5B057/> (参照 2022.6.1)
- [13] 情報処理推進機構: AI 白書 2020, (2020).
- [14] 人工知能学会: 人工知能の FAQ.
<https://www.ai-gakkai.or.jp/whatsai/AIfaq.html> (参照 2022.6.1)
- [15] John McCarthy: WHAT IS ARTIFICIAL INTELLIGENCE? (2007).
<http://www-formal.stanford.edu/jmc/whatsai/whatsai.html> (参照 2022.6.1)
- [16] JIS X0028:1999:情報処理用語—人工知能—基本概念及びエキスパートシステム.
<https://jis.eomec.com/jisx00281999> (参照 2022.6.1)
- [17] 内閣府: 科学技術基本計画, (2016).
- [18] Hinton, Geoffrey E., and Ruslan R. Salakhutdinov: Reducing the dimensionality of data with neural networks, science 313.5786, pp504-507, (2006).
- [19] 株式会社野村総合研究所: ICT の進化が雇用と働き方に及ぼす影響に関する調査研究, (2016).
- [20] 総務省: 令和元年版情報通信白書, (2019).
- [21] 山下和: 学生用 GPU 計算サーバの導入とパフォーマンス測定, 富山大学総合情報基盤センター広報, vol.17, pp28-31, (2020).

- [22] 人工知能学会: AI マップ β 2.0, (2020).
https://www.ai-gakkai.or.jp/pdf/aimap/AIMap_JP_20200611.pdf (参照 2022.6.1)
- [23] 総務省: AI ネットワーク社会推進会議 (報告書 2018 別紙 2-1), (2018).
https://www.soumu.go.jp/main_content/000564151.pdf (参照 2022.6.1)
- [24] 内閣府: 人間中心の AI 社会原則, (2019).
- [25] K. Fukushima and S. Miyake: Neocognitron: A new algorithm for pattern recognition tolerant of deformations and shifts in position, *Pattern Recognit.*, vol.15, no.6, pp.455–469, (1982).
- [26] V. Nair and G.E. Hinton: Rectified linear units improve restricted boltzmann machines, *Proc. ICML*, pp.807–814, (2010).
- [27] K. He, X. Zhang, S. Ren, and J. Sun: Delving deep into rectifiers: Surpassing human-level performance on imagenet classification, *Proc. ICCV*, pp.1026–1034, (2015).
- [28] D.A. Clevert, T. Unterthiner, and S. Hochreiter: Fast and accurate deep network learning by exponential linear units (ELUs), *Proc. ICLR*, pp.1–14, (2016)
- [29] Lecun, Y., Boser, B., Danker, J. S., Henderson, D., Howard, R. E., Hubbard, W. and Jackel, L. D.: Backpropagation applied to handwritten zip code recognition, *Neural Computation*, Vol.1, No.4, pp.541-551, (1989)
- [30] LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P.: Gradient-based learning applied to document recognition, *Proc. IEEE*, (1998)
- [31] A. Krizhevsky, I. Sutskever, G. Hinton: ImageNet classification with deep convolutional neural networks, In *NIPS*, pp.1106, (2012).
- [32] Zeiler, Matthew D., and Rob Fergus: Visualizing and understanding convolutional networks, *European conference on computer vision*. Springer, Cham, (2014).
- [33] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich: Going deeper with convolutions, In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (2015).
- [34] He, Kaiming, et al.: Deep residual learning for image recognition, *Proceedings of the IEEE conference on computer vision and pattern recognition*, (2016).
- [35] K. Simonyan, A. Zisserman: Very deep convolutional networks for large-scale image recognition, *arXiv*, 1409.1556, (2014).
- [36] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov: Dropout: A simple way to prevent neural networks from overfitting, *The journal of machine learning research*, 15, 1 pp1929, (2014).
- [37] Tieleman, T. and Hinton, G.: Lecture 6.5-RMSProp: Divide the gradient by a running average of its recent magnitude. *COURSERA: Neural networks for machine learning*, 4(2), pp26-31, (2012).
- [38] Choi, Dami, et al.: On empirical comparisons of optimizers for deep learning, *arXiv*, 1910.05446 (2019).

- [39] L. Perez, J. Wang: The Effectiveness of Data Augmentation in Image Classification using Deep Learning, arXiv, 1712.04621, (2017).
- [40] A.M. Hafiz, G.M Bhat: Multiclass Classification with an Ensemble of Binary Classification Deep Networks, arXiv, 2007.01192, (2020).
- [41] Maaten, L. and Hinton, G.: Visualizing data using t-SNE, Journal of Machine Learning Research, pp1-48, (2008).
- [42] Mukherjee, Sourabrata: t-SNE based feature extraction technique for multi-layer perceptron neural network classifier, International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICT). IEEE, (2017).
- [43] Perez H, Tah JHM: Improving the Accuracy of Convolutional Neural Networks by Identifying and Removing Outlier Images in Datasets Using t-SNE. Mathematics. 8(5), (2020).
- [44] Lin, Min, Qiang Chen, and Shuicheng Yan: Network in network, arXiv preprint arXiv, 1312.4400, (2013).
- [45] E. Hoffer, N. Ailon: Deep metric learning using triplet network, arXiv, 1412.6622, (2014).
- [46] B. Staar, M. Lutjen, M. Freitag: Anomaly detection with convolutional neural networks for industrial surface inspection, Procedia CIRP, vol79, pp 484-489 (2019).
- [47] X. Tao, D. Zhang, W. Ma, X. Liu, D. Xu: Automatic metallic surface defect detection and recognition with convolutional neural networks, Applied Sciences, vol 8, no 9, pp 1575 (2018).
- [48] 半澤雄希, 池田泰之, 栗田真嗣, 長谷川友紀: AI 技術による外観検査の自動化, OMRON TECHNICS, Vol51, 1, pp21,(2019).
- [49] M. Aghaei, M. Bustreo, P. Morerio, N. Carissimi, A. D. Bue, V. Murino: Complex-Object Visual Inspection via Multiple Lighting Configurations, arXiv, 2004.09374, (2020).
- [50] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna: Re-thinking the inception architecture for computer vision, In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp2818, (2016).
- [51] R. Müller, S. Kornblith, G. Hinton: When Does Label Smoothing Help?, NeurIPS2019, vol32, pp4696-4705, (2019).
- [52] Michal Lukasik, Srinadh Bhojanapalli, Aditya Krishna Menon and Sanjiv Kumar: Does label smoothing mitigate label noise?, arXiv, 2003.02819, (2020).
- [53] C. B. Zhang, P. T. Jiang, Q. Hou, Y. Wei, Q. Han, Z. Li, M. M. Cheng: Delving Deep into Label Smoothing, arXiv, 2011.12562, (2020).
- [54] Russakovsky, Olga, et al.: Imagenet large scale visual recognition challenge, International journal of computer vision, 115.3, pp211-252, (2015).
- [55] G. Hinton, O. Vinyals, and J. Dean: Distilling the knowledge in a neural network, in Adv. Neural Inform. Process. Syst. Worksh., (2015).
- [56] T. DeVries, G. W. Taylor: Improved regularization of convolutional neural networks with cutout, arXiv, 1708.04552, (2017).

- [57] 阪口紗季, 東納ひかり, 松下光: 範実物体と影の変形を用いた情報の秘匿と選択的視覚化, 芸術科学会論文誌, Vol.15, No.4, pp.157-166, (2016)
- [58] Momeny Mohammad, Sarram Mehdi, Latif Alimohammad, Sheikhpour Razieh, Zhang Yu-Dong: A Noise Robust Convolutional Neural Network for Image Classification, Results in Engineering, vol.10, June2021, (2021).
- [59] Yim Jonghwa, Kyung-Ah Sohn: Enhancing the performance of convolutional neural networks on quality degraded datasets, International Conference on Digital Image Computing, Techniques and Applications, IEEE, (2017).
- [60] Dodge, Samuel, Lina Karam: Understanding how image quality affects deep neural networks, eighth international conference on quality of multimedia experience (QoMEX). IEEE, (2016).
- [61] 文部科学省: 新しい学習指導要領の考え方, (2017).
- [62] 文部科学省: 新たな未来を築くための大学教育の質的転換に向けて～生涯学び続け、主体的に考える力を育成する大学へ～ (答申) , (2012).
- [63] 文部科学省: 新たな未来を築くための大学教育の質的転換に向けて～生涯学び続け、主体的に考える力を育成する大学へ～ (答申) 用語集, (2012).
- [64] 田中宏和: アティブラーニングによるプログラミング教育法の提案～基幹系業務システム開発演習の実践事例, 日本情報経営学会誌, vol.36, No.2 (2015).
- [65] 高岡詠子, 山内崇裕, 滑川敬章: 高等学校における実用的プログラミンの教育実践, 情報処理学会論文誌, vol.2, No.2, pp37-52, (2016).
- [66] 独立行政法人情報処理推進機構 (IPA) 社会基盤センター: IT 人材白書, (2020).
- [67] Sculley, David, et al.: Hidden technical debt in machine learning systems, neural information processing systems, 28th, vol2, pp2503-2511, (2015).
- [68] 三好きよみ, 川田誠一: 産業技術大学院大学における横断型人材育成とキャリアアップ, 第4回横幹連合コンファレンス, (2011).
- [69] 経済産業省: DX 推進指標, (2020).

謝辞

本論文をまとめるにあたり、終始親切丁寧な指導をいただきました光産業創成大学院大学 花山良平先生に心より感謝を申し上げます。花山良平先生には、入学後の研究テーマが定まらないうちから適切な助言を賜り、光技術、情報システム、機械学習の分野を基礎からご指導いただきました。また、研究活動を通して物事を深く考え、多面的に捉える大切さをご指導いただき、研究者のあり方を学ばせていただきました。学会発表や論文執筆の方法に関しても一から熱心にご指導いただきました。ありがとうございました。

光産業創成大学院大学 増田靖先生には社会科学分野でのご指導を頂き、数多くのご助言を賜りました。定期的にゼミを開催していただき、常に相談できる環境を準備してくださいました。論文の執筆にあたっては、何度も推敲していただき、完成度を高めていただきました。深く感謝申し上げます。

博士課程在学中、光産業創成大学院大学 坪井昭彦先生、石井勝弘先生、沖原伸一郎先生には指導教官として、多大なるご指導をいただき、有益なご教示を数多く賜りました。深く感謝いたします。森清和先生、楠本利行先生には勉強会などを通じて、機械学習に関するご指導、アドバイスいただきましたことに深く感謝いたします。

光産業創成大学院大学 瀧口義浩学長、江田英雄先生、宇佐美健一先生、長谷川和男先生、藤田和久先生、横田浩章先生、内藤康秀先生、森芳孝先生、姜理恵先生、林寧生先生には、ゼミおよび講義を通して有益な議論をいただきました。深く御礼申し上げます。

光産業創成大学院大学事務局の皆様には、本研究を進めるにあたり、常に研究に集中できる環境を整えていただきました。深く御礼申し上げます。

日頃から熱心に討論してくださいました光情報・システム分野の諸氏、先端光産業経営分野ゼミのメンバー、14期生の諸氏に感謝の意を表します。

所属企業である株式会社 エフ・シー・シーの社長、生産技術センター長、第1技術ブロックリーダーの皆様には、光産業創成大学院大学に入学し、研究を行う機会を与えていただきました。深謝申し上げます。所属企業の職場の皆様には長期間に渡り業務の調整など、研究に専念できるようにご尽力いただきました。心より感謝申し上げます。管理ブロックの皆様には、手続きや事務処理など、様々な面で研究をサポートしていただきました。深く御礼申し上げます。

本研究を進めるにあたり、ここに挙げた方々の他にも、多くの方々からご指導、ご支援、ご協力をいただきました。皆様に深く感謝申し上げます。

最後に、通学に理解を示し、4年間いつも自分を支え、励ましてくれた妻 彩加、いつも笑顔で応援してくれた長女 咲葉、長男 梓留に深く感謝いたします。ありがとうございました。

2022年9月

岡崎元樹

業績目録

1. 論文

- [1] 岡崎元樹, 花山良平: 畳み込みニューラルネットワークを用いた外観検査における非対称検出手法の開発, 精密工学会 (2022年9月出版)

2. 国際学会発表

- [1] Motoki OKAZAKI, Ryohei HANAYAMA: Defect detection method for clutch disk inspection using AI, IoT-SNAP2021, (2021)

3. 国内学会発表

- [1] 岡崎元樹, 花山良平: AI を用いたクラッチディスク完成品検査における欠陥検出方法, OPJ-2020, (2020)
- [2] 岡崎元樹, 花山良平: AI を用いたクラッチディスク欠陥検出手法の開発, 2021年度精密工学会秋季大会, (2021)
- [3] 岡崎元樹, 増田靖: AI 導入を推進するリテラシー向上活動に関する考察, 日本情報経営学会第82回全国大会, (2021)

4. 特許

- [1] 特願 2021-142656, 「欠陥検査装置、欠陥検査方法および予測モデル生成方法」, 発明者 岡崎元樹