

正常画像パターンが複数ある教師なし異常検知 Unsupervised Anomaly Detection for Multiple Normal Image Patterns

南 俊輔* 堂 蘭 浩*
Minami Shunsuke Dozono Hiroshi

1. はじめに

大量生産における不良品検知では、目視検査によってすべてを判断することは、効率や精度面において最適であるとは言えない。近年では、この問題を解決するために画像分析を用いて異常検知を行うことが求められている。本研究では、画像データの中の未学習の画像(異常)を検出することを目的として、オートエンコーダ、One-Class SVM, MNIST データベースを用いて行った。[1],[2]

工業製品を画像データとした異常検知では正常データに大差は存在せず、従来の異常検知手法で対応が可能であるが、食料品などといった、正常品の中に大きな幅がある場合の異常検知手法はあまり提案されていない。そこで本研究では正常データの中に幅のある場合にも適した異常検知手法の提案を試みた。

今回用いる画像データセットである MNIST データベースは、「0」から「9」までの手書き数字で構成されたデータセットであり、正常データの種類を複数選択可能である。また、各数字において複数パターンの手書き数字が存在するため、本研究の目的である正常画像の中に複数パターンの画像が存在する場合の異常検知に適していると考え、MNIST データベースを用いることとした。

2. CAE (Convolutional Autoencoder)

本節では CAE について説明する。CAE とは、Convolutional Autoencoder の略であり、畳み込みオートエンコーダである。オートエンコーダは、入力層、中間層、出力層で構成され、前半部分がエンコーダ、後半部分がデコーダとして機能するシステムとなっている。エンコーダとは入力画像を圧縮する層、デコーダとは圧縮した画像を復元する層である。学習段階では、入力画像と出力画像が一致するように学習される。オートエンコーダの概略図を図 1 に示す。

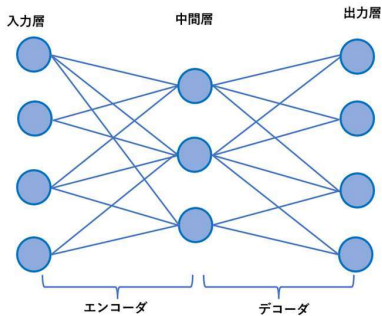


図 1: オートエンコーダの概略図

本研究で用いた、畳み込みオートエンコーダについて説明する。畳み込みオートエンコーダとは、エンコーダ、デコーダ部分に畳み込みニューラルネットワークを用いたオー

* 国立大学法人佐賀大学 Saga University

トエンコーダである。畳み込みニューラルネットワークとは、AI が画像分析や音声分析を行う際に用いる学習方法の一つである。人間の視覚感覚を参考にしたアルゴリズムであり、入力データの特徴量を捉える畳み込み層と、捉えた特徴の重要な情報だけを保存したまま入力された画像サイズの縮小を行うプーリング層を加えたニューラルネットワークで構成されている。また、畳み込み層により、空間的に意味のある特徴を抽出できるため、画像に対して用いられることが多い。

3. One-Class SVM (Support Vector Machine)

本節では、機械学習の分類アルゴリズムである One-Class SVM について説明する。One-Class SVM とは、2 クラス分類問題の代表的な手法である SVM(Support Vector Machin) を教師なし学習へと拡張したものである。正常データとして 1 つのクラスを学習させ、識別境界を決定する。その後、その境界を基準に外れ値を検出する。異常検知を行う場合、一般的に正常データは多数入手できるが、異常データを入手することは困難な場合が多い。One-Class SVM は大多数が正常であるデータセットをもとに関数を学習し、未知のデータが正常であるかどうかを判別することが可能であるため、今回の条件に有効な外れ値検知手法と考えた。

4. 実験手法

データセットについて説明する。まず、異常とする数字を任意で選択し、その数字を含まない画像データで学習データを構成する。具体的には「9」を異常画像とした場合には学習データは「0-8」の数字で構成される。また、正常画像の数字が 1 種類の場合と複数種類ある場合(5 種類, 9 種類)で構成する。テストデータについては、学習データに異常とした数字を加えた画像データで構成する。

学習方法としては、学習データを用いて、一つの畳み込みオートエンコーダに対して入力画像と出力画像の再構成誤差が小さくなるように学習させる。この際、学習データの中に含まれる正常画像の種類は 1 種類, 5 種類, 9 種類の場合で実験を行った。

異常検知の方法としては、学習済みモデルのオートエンコーダにテストデータを入力し、入力画像と出力画像の再構成誤差を計算する。その際に、すでに学習されている画像、つまり、正常画像に関しては、再構成が可能となる。しかし、未学習の画像(異常データ)については十分に再構成がされない。この時生じる、再構成差の値に閾値を設け、正常・異常の判断を行う。再構成誤差の評価方法としては、入力画像と出力画像との差の 2 乗誤差を用いた。この平均 2 乗誤差が小さいほど、入力画像に近い画像であり、正常画像と判断される。

また、学習データに含まれる正常画像の種類が複数ある際には、畳み込みオートエンコーダに学習させる前にデータ間の類似度に基づいてクラス分けを行うクラスタリング

を行う。その後、クラスタリングをしたクラス数と同じ数の畳み込みオートエンコーダに学習させ、異常検知を行う。

この時クラス分けを行うクラス数は 5, 9, 15 とした。概略を図 2 に示す。

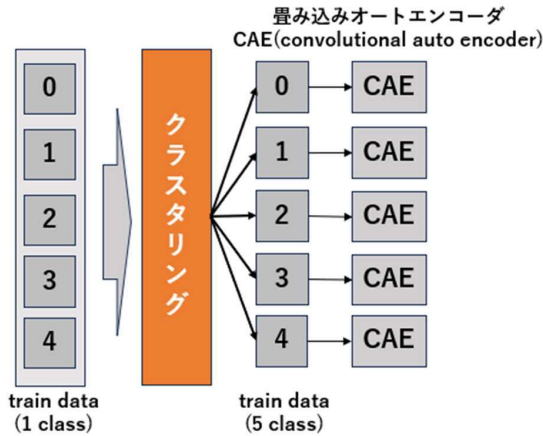


図 2: クラスタリングを用いた CAE の学習手法

次に、One-Class SVM を用いた異常検知手法について説明する。学習段階については、先ほど同様に正常データを用いて正常画像と出力画像が等しくなるように学習させる。その後、学習済みオートエンコーダのエンコーダ部分を用いて、特徴量を抽出し、その特徴量を入力データとし、入力値が正常データであれば+1, 異常データであれば-1 を返すように One-Class SVM を学習する。

異常検知の方法としては、学習済みオートエンコーダを用いて、テストデータの特徴量を抽出し、One-Class SVM を用いて異常検知を行った。[3]

5. 実験結果

畳み込みオートエンコーダの学習に使用した正常データが 1 種類の場合は、表 1 に示すように、全体的に高い精度での異常検知が可能となった。表 2 には正常データが 9 種類の場合の実験結果を示す。正常データが複数ある場合、正常データが 1 種類の時と比べ、精度が低下する結果となった。表 3 には、畳み込みオートエンコーダに学習させる前にクラスタリング(クラス数 5)を行った場合の結果を示す。すべての条件で精度が向上したわけではないが、精度が向上する結果となったものもみられた。表 4 に正常データが 9 種類の場合において、畳み込みオートエンコーダを特徴抽出器として、One-Class SVM を用いた異常検知の結果を示す。異常検知精度が最も高い結果となったが、正常検出率が低くなる結果もみられた。

表 1: 正常データ 1 種類

Normal data	Normal dection rate(%)	Abnoemality detection rate(%)
0	89	97
1	95	98
2	72	82
3	61	82
4	74	82
5	81	81
6	88	88
7	93	86
8	83	76
9	81	73

表 2: 正常データ 9 種類(クラスタリング無し)

Abnormality data	Normal dection rate(%)	Abnormality detection rate(%)
0	71	78
1	2	9
2	63	85
3	60	87
4	43	52
5	84	59
6	63	73
7	27	14
8	70	77
9	40	58

表 3: 正常データ 9 種類(クラス数 5)

Abnormality data	Normal dection rate(%)	Abnormality detection rate(%)
0	66	93
1	79	0
2	66	76
3	70	59
4	73	51
5	69	63
6	69	69
7	71	32
8	69	77
9	71	28

表 4: 正常データ 9 種類(One-Class SVM)

Anormal date	Normal dection rate (%)	Abnormality detection rate(%)
0	60	86
1	30.6	100
2	89	93
3	38	99
4	53	99
5	78	81
6	55.6	100
7	81	75
8	68.5	73
9	33	99

6. まとめ

本研究では、畳み込みオートエンコーダを用いた異常検知、畳み込みオートエンコーダを特徴抽出器として用いた One-Class SVM で異常検知を行う手法を提案した。結果として、正常画像パターンが複数ある場合には、畳み込みオートエンコーダを特徴抽出器として用い、One-Class SVM で異常検知を行うことが異常検知の精度向上には有効であることが分かった。しかし、異常検出率が高いが、正常検出率が低い結果となる場合が多く見られた。異常検知の現場では良品を不良品と判断することは望ましいことではないため、異常検出率を保った状態で正常検出率を向上させることを今後の課題としたい。

7. 参考文献

- [1] Sebastian Raschka, Vahid Mirjalili /Python 機械学習
- [2] 岡谷 貴之 / 深層学習
- [3] 清水竜樹, 永田寅臣, 有馬滉宜, 加藤博久, 渡辺桂吾: 「VAE を特徴抽出器とした 1 クラス SVM の提案と欠陥検出への応用, 産業応用工学会全国大会 2022 講演論文集。」 P45-46