

# 機械学習を用いた画像判別に関する研究

上 菌 剛\*, 谷山清吾\*

## Study on Image Distinction using Machine Learning

Tsuyoshi UEZONO and Seigo TANIYAMA

工業製品（鋳造加工部品）の外観検査を自動化することを目的として、機械学習技術の一つであるDeep learningによる良否判別の研究を実施した。対象製品の鋳造加工部品にドリルで $\phi 0.5\text{mm}$ ,  $1.0\text{mm}$ ,  $1.5\text{mm}$ の貫通穴を開けて疑似不良を作り、製品の全体像ではなく部分的な画像で良否判別を検討した。その結果、多量の良品画像と少量の不良品画像を用いた深層距離学習を適用することにより、貫通穴 $\phi 1.0\text{mm}$ および $1.5\text{mm}$ の大きさの疑似不良について判別可能な外観検査モデルを構築できた。

**Keyword** : Deep learning, 機械学習, 距離学習, 外観検査, 鋳造

### 1. 緒 言

近年、AI (Artificial Intelligence: 人工知能) という言葉が当たり前のように使われている。辞書<sup>1)</sup>には「学習・推論・判断といった人間の知能の持つ機能を備えたコンピューターシステム」とあり、人間の振る舞いを模倣するようなシステムを指す大きな概念のことである。この人工知能における研究課題の一つに機械学習 (Machine learning) がある。機械学習とは、「データから反復的に学習し、そこに潜むパターン (特徴) を見つけ出すこと」であり、アルゴリズムとしてサポートベクターマシーン (SVM) ニューラルネットワーク (NN), 深層学習 (Deep learning), 決定木 (DT), ナイブベイズ (Naive Bayes) など様々な提案がなされている。その中で、現在のAIブームの火付け役となった技術がDeep learningである。Googleの研究者ジェフリー・ヒントン氏が2012年に論文発表<sup>2)</sup>し、同年に開催されたAI系画像認識コンテストで、初めてDeep learningを利用したシステムが使われ、圧倒的な認識精度を発揮した。

Deep learningでは、コンピューターが自ら学習データから特徴量を抽出し、予測モデルを構築することが可能であるため、人間が介在して特徴量設計を行う必要がなくなり、以前よりはるかに容易に機械学習を行うことができるようになった。更にコンピューターの性能向上や、2010年代からビックデータを扱うようになり、膨大な計算リソースを獲得したこと、またインターネットの普及による広範囲で迅速な情報の共有、交換など、技術開発に必要な要素が十分にそろったことから、AI技術は飛躍的に向上し本格的な活用が始まることとなった。AIの身近な活用事例として、パンの画像識別を行うベーカリースキャン、キュウリ

の画像選別システム、医療分野の画像診断、音声認識によるAIスピーカー、自然言語処理による文字変換システムなどがあり、既に多くの分野で利用が進んでいる<sup>3)~7)</sup>。

本研究では、工業製品の外観検査の自動化を目的に Deep learning技術を適用して良品、不良品の判別を行う技術について研究を実施した。

### 2. 実験方法

#### 2. 1 学習と評価に用いる画像

実験に供する工業製品 (以下、サンプル) は、県内企業が目視検査で月産7万個程出荷する鋳造加工部品 (ナット : M36) を対象にした。一般的にDeep learningに使用する画像データは膨大な量が必要であることが想定されたので、サンプル個々の全体画像ではなく、1個のサンプルから多数の画像を得ることとした。具体的には鏡を利用してサンプル内側加工面のねじ山部分を撮像して内面展開図を得て、そこから小さな画像を多数切り出すこととした。撮像に供するサンプルは4個とし、疑似不良として $\phi 0.5\text{mm}$ ,  $1.0\text{mm}$ ,  $1.5\text{mm}$ の3種類の貫通穴をドリルで開けた。使用したサンプルの一つを図1に示す。穴径は、最も小さな $\phi 0.5\text{mm}$ が製造現場で不良と判断する下限値として設定している数値で、これを基準に3種類の大きさを設定した。穴は、サンプル1に $\phi 1.5\text{mm}$ を、サンプル2に $\phi 1.0\text{mm}$ を、サンプル3に $\phi 0.5\text{mm}$ を、サンプル4には3種類全ての大きさの穴を開けた。実際の不良品の画像と、ドリル貫通穴の画像を同じ縮尺で比較した図を図2に示す。ほぼ同様の画像が得られており、ドリル穴による疑似不良でも代用可能と思われる。

\* 生産技術部

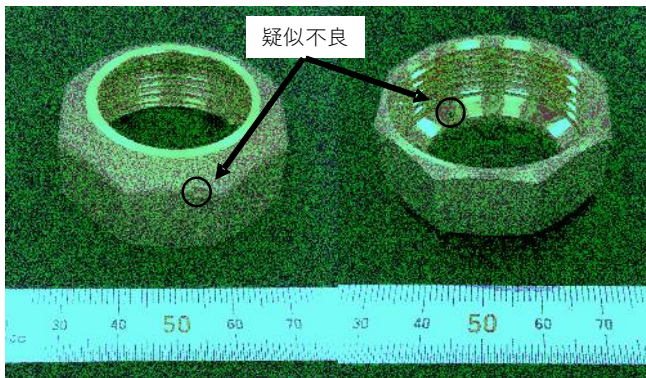


図1 鋳造加工部品（ナット）



図2 不良画像の比較

2.2 AIアルゴリズム

工業製品の外観検査において、一般的には不良品の発生はそれほど多くないため、不良品の情報が無くても学習できることが望ましい。この解決策としては、教師無し学習でアプローチすることになる。教師無し学習に対して教師有り学習というものがある。これは、多量の良品データと多量の不良品データを用意して、両者の特徴をそれぞれ学習させた後、評価において、入力データがどちらの特徴に近いかで判別を行う手法である。一方、教師無し学習は、良品のみを学習データとして利用し、良品の特徴のみを学習させる。評価において、入力されたデータが良品の特徴を持っているか否かで判断する。このような異常を検知する分野で、現在主流となっている技術は、Deep learningに距離学習を付加した深層距離学習（Deep metric learning）である。距離学習とは、データ間の計量（距離や類似度など）を学習する手法である。直感的には、特徴が似ているデータは近く、そうでないデータは遠くなるように計量を学習する。本研究では深層距離学習で取り組むこととした。

データを入力して何かしらの出力を得るAIアルゴリズムを一般的にモデルという。本研究におけるモデルの全体像を図3に示す。入力画像をDNN（Deep Neural Network）に入力し、次に距離学習を適用する処理を付加し、最終的に2つの出力（良品か不良品）を得る。DNNには、MobileNet V2（2018年にGoogleが開発）を用い、これにImageNet（データセットと呼ばれる多種多様な画像データ群）で学習したパラメータを適用しベースに用いた。

一般的にDNNは多数の層構造となっており、入力層から中間層を経て出力層に向けて情報が流れる。層間ではパラメータ（重みやバイアス）により情報の変換がなされる。

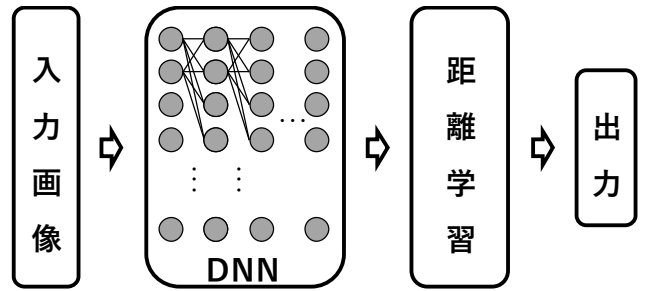


図3 モデルの全体像

MobileNetV2は、入力層が1層、中間層が155層、出力層が1層の157層で構成される。入力層は224(縦)×224(横)×3(色)=150,528個で、出力層の個数は任意に設定が可能である。

次に、距離学習の機能は、今回はL2 Softmax Lossという損失関数<sup>8)</sup>を使用した。損失関数とは、正解との誤差を表す指標で、L2 Softmax Lossの損失関数は、式1、式2で示される。Mはバッチサイズで、一度に学習させる画像の枚数であり、例えば16の場合16枚学習させる毎にパラメータを1回更新することになる。Xiは、モデルに入力する画像のデータで、関数f(Xi)はDNN部分からの出力を指す。Cは、出力のカテゴリ数で2（良品と不良品）となる。WとBはモデルのパラメータで、それぞれ重みとバイアスである。式1は、一般的なSoftmax Lossの式で、画像を入力した際に出力される結果と正解との誤差を表しており、この出力が小さくなるように、モデルのパラメータが更新され学習が進む。式2は、本損失関数の特徴で、DNNの出力の大きさを任意の定数αにスケールすることを表している。式1の損失関数は、パラメータが適切に更新され、精度が上がることに伴い小さくなるが、精度が上がっていてもDNNの出力が小さくなることでも小さくなる。つまり損失は小さくなっているが、認識精度が上がらないという状況が生まれる。これを抑制するために式2の制約を加えている。L2 Softmax LossはDNNの最終出力層の直前の情報を利用するため、最終出力層の手前に処理を追加し学習モデルを完成する。評価の際は、L2 Softmax Loss直後の出力を基にLocal Outlier Factor（LOF：データの集まりの中から外れた値を見つけ出す外れ値検知アルゴリズム）を実施し、その値を異常スコアとして判定に利用した。

$$-\frac{1}{M} \sum_{i=1}^M \log \frac{e^{W_{y_i}^T f(x_i) + b_{y_i}}}{\sum_{j=1}^C e^{W_j^T f(x_i) + b_j}} \dots (1)$$

$$\|f(x_i)\|_2 = \alpha, \forall i = 1, 2, \dots, M \dots (2)$$

- M : バッチサイズ
- Xi : 入力画像のデータ
- f (Xi) : DNN部分の出力
- C : 出力のカテゴリ数 (2)
- W, B : パラメータ (重みとバイアス)
- α : 任意の定数

3. 結果及び考察

3. 1 画像の作成

撮像する内側加工面は、カメラによる直写像ではねじ山の死角により不良が撮像できない可能性があることから、ミラーを用いて反射像を撮像する構成とした。撮像の構成図を図4に、構築した撮像システムを図5に示す。サンプルは、開口部の広い方を下向きに設置し、その中心位置に円錐ミラーを設置した。照明は、偏りの無い光とするために面照明を用い、下方から照射する構成とした。カメラは上部に設置し、円錐ミラーに写り込む像を撮像した。取得した画像において、ミラー頂点部から一定距離の同心円上にある全ての画素を、円から線分に座標変換した。サンプルを移動させ、ミラーに写り込む画像を1ドット分ずらしして同様の座標変換を繰り返し、得られた複数の線分を順番に並べることで内側加工面の展開画像(図6)を得た。

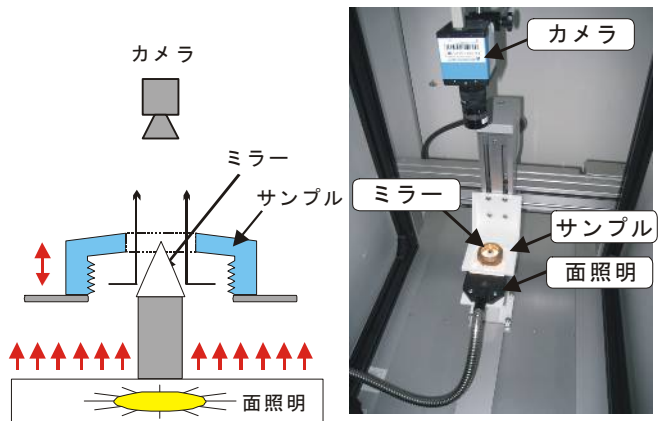


図4 内側加工面の撮像

図5 構築したシステム

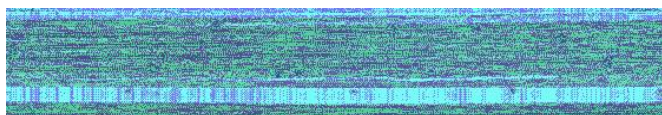
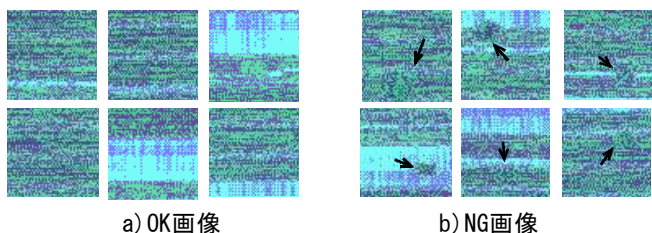


図6 内面の展開画像

次に、展開画像から小さな画像を複数作成する。展開画像は1,080×180ドットの画像である。ここから56×56の小さな画像を左上から順番に1ドットずつシフトさせながら切り出して作成する。予め、不良の場所を指定しておくことで、切り出した画像の中に不良部位が含まれているものは不良画像(以下、NG画像)として、含まれないものは正常画像(以下、OK画像)としてラベルを付与する。作成した画像を図7に数例示す。画像は4つのサンプルそれぞれ



a) OK画像

b) NG画像

図7 実験に供する画像

OK画像とNG画像を50,000枚ずつ(合計400,000枚)作成した。

3. 2 OK画像のみでの学習

使用する画像は表1のとおりで、バッチサイズは16, 32, 64, 128, 256, 512の6通り、エポック数は50回までとした。エポックは学習の回数で、1エポックで9,000枚の画像を学習させることになる。バッチサイズとエポックの組合せは全て行い、OK画像とNG画像の異常スコア平均値の差が最も大きい条件で評価した。

結果は、OK画像のみの学習では、全ての場合においてOK画像とNG画像を判別することは全くできなかった。評価結果の一例を図8に示す。

表1 使用画像数 (単位:枚)

サンプル	1	2	3	4
学習画像 OK画像	—	3,000	3,000	3,000
学習画像 NG画像	—	—	—	—
評価画像 OK画像	1,000	1,000	1,000	1,000
評価画像 NG画像	1,000	1,000	1,000	1,000

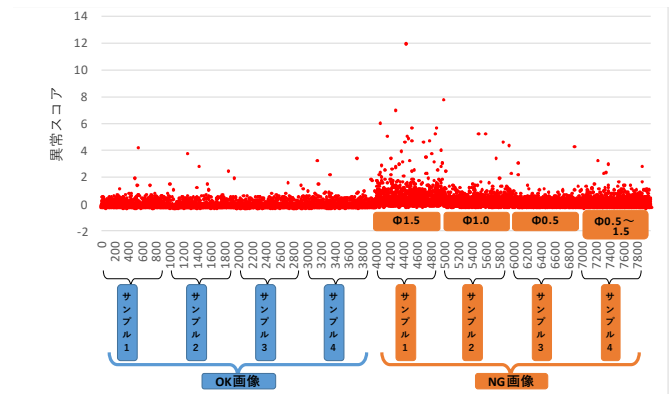


図8 評価結果 (OK画像のみで学習)

横軸は評価画像で8,000枚分、左半分がOK画像、右半分はNG画像の評価結果でサンプル順に並んでいる。縦軸は、OK画像と判断できる特徴量を基準(0付近)として、そこからの特徴量の違いの大きさを示している。期待すべき結果としては、NG画像のプロットが縦軸の大きな位置(OK画像の特徴からの違いが大きい)に出現することであるが、OK画像とNG画像がほぼ同じ特徴量となっており、うまく判別できていない。これは、NG画像の不良領域(穴部分)の全体に占める割合が1~4%程度と小さく、ほとんどの場所でOK画像と同等の様相であることが原因と思われる。そこで、この原因を確認するために、色調、模様異なるネコの絵(図9)から切り出した画像をNG画像として、OK画像のみで学習した上記モデルに入力した。その結果を図9に示す(縦軸は対数表示)。左側はOK画像1,000枚で、右側はNG画像1,000枚である。別な画像がベースであるため、NG画像の異常スコアが総じて大きくなっており、OK画像とは違うものであると判断していることが判る。ただ、OK画



像とNG画像の異常スコアがオーバーラップする領域があり、閾値で明確に分離することはできない。

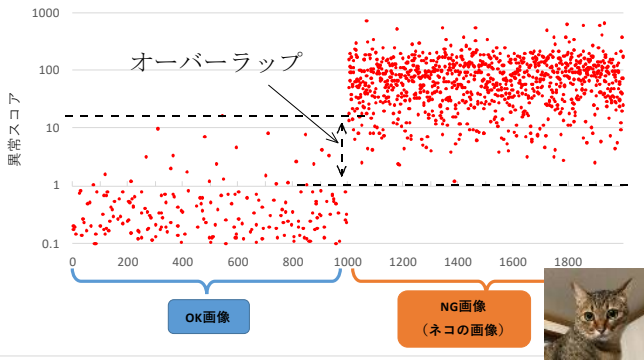


図9 評価結果 (別画像で検証)

3.3 OK画像+少量のNG画像での学習

距離学習は、似ているもの(同じカテゴリ)は固めてプロットし、違うもの(別なカテゴリ)は離れた場所にプロットするという特徴がある。工業製品において少なからず不良は発生するので、これをNG画像として学習させることでNG画像が持つ特徴を学習し、異常スコアを大きくすることができると考えられる。

本実験の条件は、NG画像(サンプル4)100枚を前述の9,000枚の学習画像の中にランダムに加えたこと以外同じ条件で実施した(表2)。評価結果の一例を図10に示す。評価方法は前述と同じである。

表2 使用画像数 (単位:枚)

サンプル		1	2	3	4
学習画像	OK画像	—	3,000	3,000	3,000
	NG画像	—	—	—	100
評価画像	OK画像	1,000	1,000	1,000	1,000
	NG画像	1,000	1,000	1,000	1,000

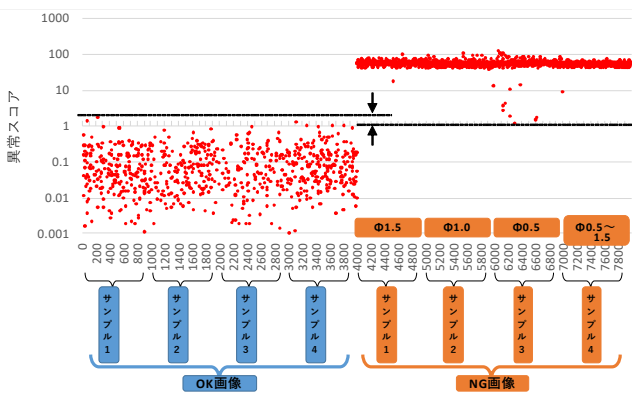


図10 評価結果 (少量のNGを追加して学習)

右半分のNG画像の異常スコアが総じて大きくなっており、OK画像とは違うものであると判断していることが判る。サンプル1(φ1.5mm)とサンプル2(φ1.0mm)はOK画像とNG画像のオーバーラップが無く、はっきり区別できているが、サンプル3(φ0.5mm)とサンプル4(φ0.5~1.5mm)の最小異常スコアは1.2で、OK画像の最大異常スコア

の1.9より小さく、少しオーバーラップが確認できた。この結果から、検出限界はφ0.5mm~1.0mmの間となり、製造現場の規定であるφ0.5mmはクリアできなかった。対策としては、より最適なモデルを構築することにより、検出精度を向上させるか、不良部位の全体に占める割合を高める(=切り出す画像のサイズを小さくする)ことで対応が可能と思われる。ただ、実際の製造現場において100%の判別は必ずしも必要ではないと考える。OK画像の99%は異常スコアが0.4以下であり、NG画像の最小値より低い。つまり99%はOK画像として判別でき、残りの1%は目視により検査するという運用で、検査コストを大幅に軽減することができることになる。

3.4 ラベリング間違いの影響

本研究では、OK画像とNG画像の両方を使用する。つまりそれぞれの画像に対応するラベルを付与する必要がある。その際に、OK画像にNGのラベルを付与したり、またその逆であったり、間違ったラベリングをしたデータが混入する可能性が少なからずある中で、その影響について実験を行った。

学習データとしてOK画像9,000枚と、NG画像100枚と、間違ったラベリング(NG画像にOKラベルを付与)をしたもの10枚を追加、および100枚を追加したものを使用して実験を行った(表3)。評価結果を図11, 12に示す。10枚程度(全体の約0.1%)の間違い混入では、判別に大きな影響は無いが、100枚(全体の約1%)の混入ではOK画像の中心領域(0付近)だけの領域が存在せず、現場での利用には適さなくなる。ラベリングには注意が必要である。

表3 使用画像数 (単位:枚)

サンプル		1	2	3	4
学習画像	OK画像	—	3,000	3,000	3,000
	NG画像	—	—	—	100 + α
評価画像	OK画像	1,000	1,000	1,000	1,000
	NG画像	1,000	1,000	1,000	1,000

(αはOK画像として学習)

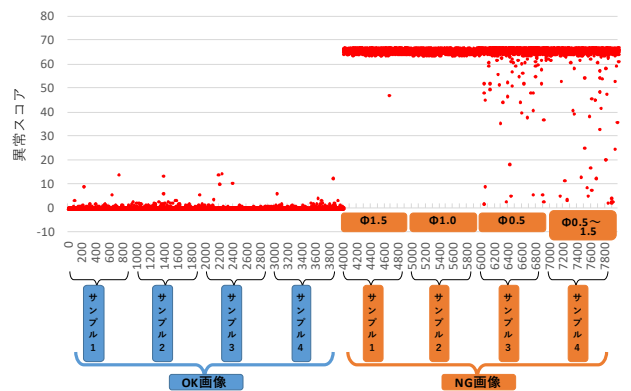


図11 評価結果 (0.1%の間違い混入で学習)

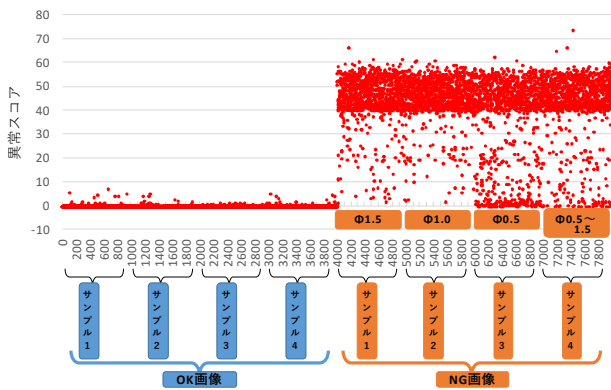


図12 評価結果（1%の間違い混入で学習）

#### 4. 結 言

外観検査の自動化が実現していない鋳造加工部品の良否判定を対象に、Deep learningによる機械学習の適用を実施し、以下の結果を得た。

- (1) 円錐ミラーを用いて反射像を撮像するシステムを構築し、サンプル内面を正確に撮像することが可能となった撮像した画像から、Deep learningに供する多量の画像（400,000枚）を取得することができた。
- (2) 深層距離学習という手法を用いることによりOK画像の特徴を学習したモデルを作成できた。しかし、微小な不良部位を持つOK画像によく似たNG画像は判別できなかった。そこで、少量のNG画像を学習サンプルに追加することで $\phi$ 1.0mm及び1.5mmのNG画像を100%判別できた。しかし $\phi$ 0.5mmのNG画像は、異常スコアがOK画像とオーバーラップしており、100%の判別はできず、製造現場の規定はクリアできなかった。ただし、OK画像の99%はNG画像の最小異常スコアよりも小さく判別が可能であるため、99%は一次スクリーニングが可能となり、残りの1%を目視検査するという運用が可能である。

(3) 誤ったラベル付けを行った学習サンプルが多くなると、正しい判断ができなくなる。ラベル付けには注意が必要である。

今後は、本手法をベースにして、モチーフの全体像を撮像し、製造現場に適用可能な外観検査システムの構築を進める予定である。

#### 参 考 文 献

- 1) 三省堂：“大辞林 第四版”（2019） p.1397
- 2) A. Krizhevsky, I. Sutskever and G. EHinton: “Image Net Classification with Deep Convolutional Neural Networks”, In Proc., NIPS2012 (2012)
- 3) BakeryScan製品ホームページ  
<https://bakeryscan.com/>
- 4) キュウリをAIで判定するシステムを農家が自ら開発、足りない部品は3Dプリンターで作成も  
<http://digital-innovation-lab.jp/kyuri-ai/>
- 5) 「医用画像解析ソフトウェアEIRL aneurysm」のサービス提供を開始  
<https://lpxel.net/news/press-release/2020/10026/>
- 6) AIスピーカーの仕組みとは？【徹底解説】  
<https://nissenad-digitalhub.com/articles/ai-for-speaker/>
- 7) 様々なシーンで活用できる自然言語処理ソリューションTrueText  
<http://www.textmining.jp/case.html>
- 8) R. Ranjan, CD Castillo, R. Chellappa: “L2-constrained softmax loss for discriminative face verification” arXiv preprint arXiv:1703.09507 (2017)