

機械学習における良品・不良品の識別精度に及ぼす

訓練データのサンプルサイズの影響

森 大樹*

Effect of the Sample Size of Training Data on the Accuracy of Discriminating Good and Defective Products in Machine Learning

Daiki MORI*

When machine learning being applied to visual inspections in manufacturing, the relationship between accuracy of discriminating good/defective products and the sample size of training data is an important issue to be investigated beforehand. In this report, a machine learning system was applied to a visual inspection of drilled hole parts in order to identify good/defective products while increasing the sample size of training data. Then the discrimination accuracy was evaluated by applying the trained model to some test data, and the effect of bias in the ratio of the sample size of good to defective products in the training data was also confirmed: the ratio between good to defective products was varied from 1:9 through 5:5 (referred to as no-bias) to 9:1, and the deviation from 5:5 was referred to as the bias. As a result, the accuracy tends to increase with the sample size of training data, and to decrease with the bias in the good/defective product ratio.

Keywords: Machine Learning, Sample Size of Training Data, Accuracy, Visual Inspection, Transfer Learning

1. はじめに

製造現場における製品の外観検査では、様々な検査項目があるため、目視による検査に頼ることが多く¹⁾、その自動化が求められている。外観検査の自動化において、画像処理技術を応用した機械学習の活用が注目されている一方で、導入が進んでいないのが現状である。その理由として、導入効果に関する不安や訓練データの保有・蓄積がないこと、訓練データの整備が困難であることなどが挙げられている²⁾。

導入効果に関する不安を解決するためには、安定した高い精度で機械学習を行うことが望ましく、そのためには適切な数の訓練データを用いること

が必要とされている。必要となる訓練データ数は、解決したい問題によって異なると言われていたが³⁾、訓練データの保有・蓄積や整備にも相応のコストがかかることから、その数は少ない方が望ましい。したがって、解決したい問題に応じて、導入効果が期待できる精度を得るための訓練データ数の目安について、あらかじめ把握することが重要である。

そこで、本報では、多くの製造現場で行われる加工の一つである、穴加工部品における穴の有無や異物の付着などの外観検査を取り上げ、機械学習における良品・不良品の識別精度と訓練データのサンプルサイズの関係について評価を行ったので、その結果を報告する。

* ものづくり研究課

2. 実験方法

2.1 訓練データの準備

2.1.1 検査項目

本報では、外観検査を模擬した穴加工部品として、長さ 60 mm、幅 25 mm、厚さ 1 mm、中心に $\phi 6$ mm の穴加工があるアルミニウム製の板材を用いた。外観検査の項目は、板材における穴加工の有無と異物の付着を対象とし、2 つの検査項目における良品・不良品の二分類問題とした。穴加工の検査では、同形状の板材に穴がないものを不良品とした。異物の検査では、板材表面に黒色のカラーサンドを付着させたものを不良品とした。そのカラーサンドの粒径は 0.03~0.15 mm, 0.2~0.8 mm, 1~1.7 mm の 3 種類であり、順に異物(小・中・大)の不良品とした。

2.1.2 撮影

2.1.1 項で用意した板材に対し、暗室環境下で外観を撮影した。黒色のテーブル上に板材を設置し、テーブル面に垂直上方に照明とカメラを設置した。カメラは Canon 製 EOS KissX4、レンズは EF-S18-55 mm f/3.5-5.6IS を用いた。照明は外径 470 mm、内径 430 mm のリングライトを用い、照度 8460 lx の白色光で板材を照らした。板材からカメラまでの距離を 320 mm、照明までの距離を 400 mm とした。カメラの撮影条件は、画質 5184×3546 画素(px)、露光時間 0.3 s、F 値 F36、ISO 感度 100 とし、マニュアルフォーカスで撮影を行った。撮影により得られた板材の外観画像 (Table1 参照) を訓練データとした。

2.2 機械学習

2.2.1 学習モデル

機械学習には、CNN 手法⁴⁾をベースとした 13 層の畳み込み層と 5 層のプーリング層、3 層の全結合層で構成される VGG16 モデル^{5,6)}の転移学習⁷⁾を用いた。Fig.1, Table2 に VGG16 モデルの構造を示す。ここでは、学習の高速化を図るため、撮影画像は 1/16 にリサイズして学習させた。

VGG16 の畳み込み層では、種類の異なる複数の特徴フィルタ (フィルタサイズ 3×3 px) を、1 px のストライドで計算し、特徴を抽出した画像 (以下、特徴マップとする) を得た。また、この処理では、特徴マップのサイズが変化しないように、予め入力画像の周囲を 1 px 分、画素値 0 で埋める処理を行った。

VGG16 のプーリング層では、特徴マップから 2×2 px 内の最大画素値に置換していく処理を行うことにより、特徴マップのサイズを 1/2 となるように圧縮した。

VGG16 の全結合層では、圧縮された特徴マップを特徴ベクトル (数値データ) として取り扱い、その特徴をニューラルネットワーク (Fig.2 参照) で学習することで、良品・不良品を分類した。なお、ニューラルネットワークの学習では、ノードを不活性化させて過学習を抑制するドロップアウト⁷⁾を隠れ層に 50% の割合で行った。

学習回数は 10 epoch とし、畳み込み層は VGG16 のパラメータをそのまま用い、全結合層の重みを 1epoch ごとに更新することにより、板材の訓練データへのファインチューニングを実施した。

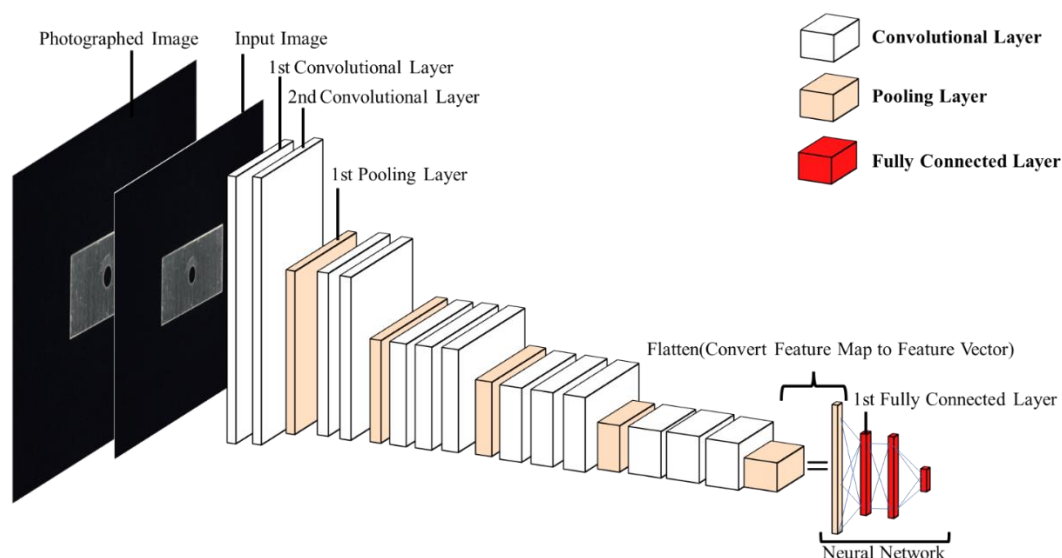


Fig.1 VGG16 Transfer Learning Model

Table1 Photographed Images for Each Inspection

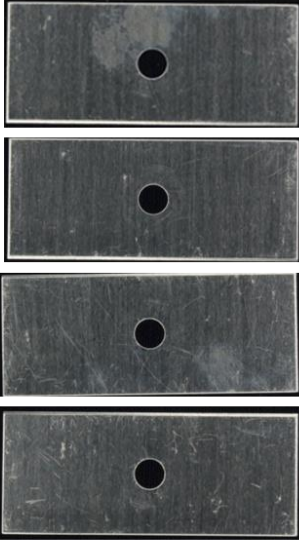
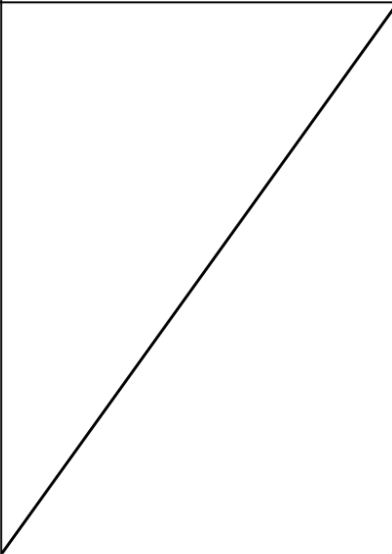
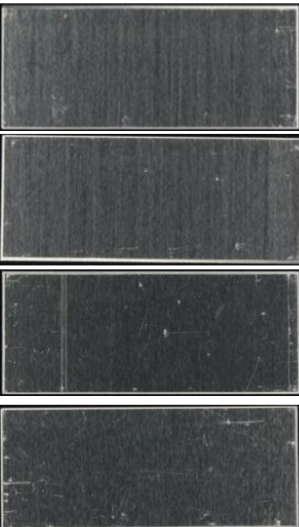
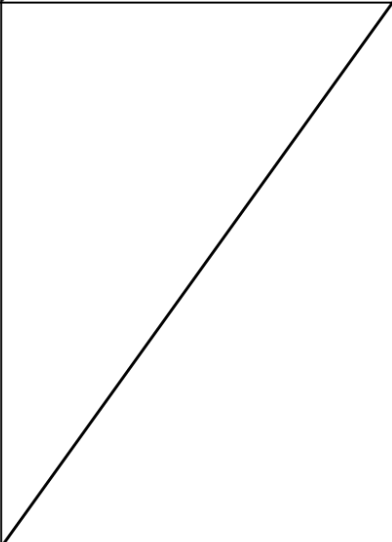
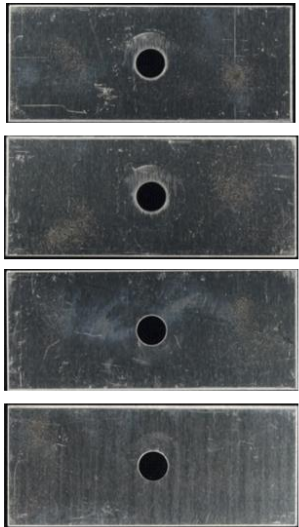
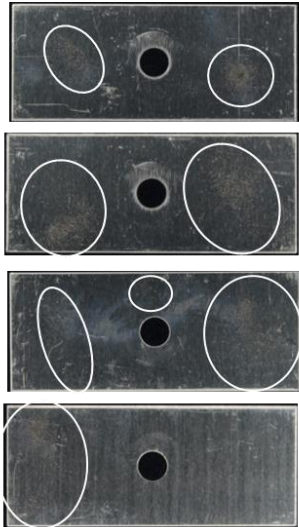

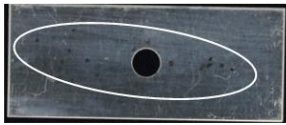

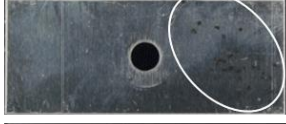

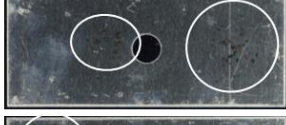


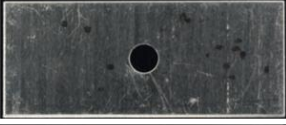
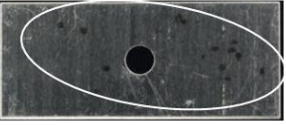






Inspection	Image	Defective Areas
Good Product		
No Hole		
Foreign Objects (Size:Small)		

Table1 Photographed Images for Each Inspection <Continued>

Inspection	Image	Defective Areas
Foreign Objects (Size:Medium)		
		
		
		
Foreign Objects (Size:Large)		
		
		
		

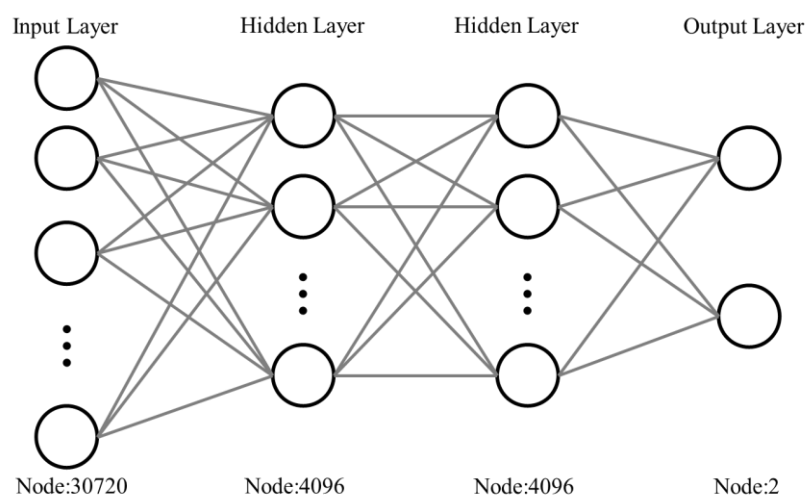


Fig.2 Neural Network

Table2 VGG16 Transfer Learning Model's Layer

Layer Name	Output Data
Photographed Image	Image:5184 × 3456 (Horizontal Pixel × Vertical Pixel)
Input Image	Image:324 × 216
1st Convolutional Layer	Feature Map:324 × 216 × 64 (Horizontal Pixel × Vertical Pixel × maps)
2nd Convolutional Layer	Feature Map:324 × 216 × 64
1st Pooling Layer	Feature Map:162 × 108 × 64
3rd Convolutional Layer	Feature Map:162 × 108 × 128
4th Convolutional Layer	Feature Map:162 × 108 × 128
2nd Pooling Layer	Feature Map:81 × 54 × 128
5th Convolutional Layer	Feature Map:81 × 54 × 256
6th Convolutional Layer	Feature Map:81 × 54 × 256
7th Convolutional Layer	Feature Map:81 × 54 × 256
3rd Pooling Layer	Feature Map:40 × 27 × 256
8th Convolutional Layer	Feature Map:40 × 27 × 512
9th Convolutional Layer	Feature Map:40 × 27 × 512
10th Convolutional Layer	Feature Map:40 × 27 × 512
4th Pooling Layer	Feature Map:20 × 13 × 512
11th Convolutional Layer	Feature Map:20 × 13 × 512
12th Convolutional Layer	Feature Map:20 × 13 × 512
13th Convolutional Layer	Feature Map:20 × 13 × 512
5th Pooling Layer	Feature Map:10 × 6 × 512
Flatten (Input Layer)	Feature Vector:30720
1st Fully Connected Layer (Hidden Layer)	Feature Vector:4096
2nd Fully Connected Layer (Hidden Layer)	Feature Vector:4096
3rd Fully Connected Layer (Output Layer)	Feature Vector:2

2.2.2 評価方法

(1) 精度と訓練データのサンプルサイズ

精度と訓練データのサンプルサイズの関係について、訓練データの機械学習を終えたモデル(以下、学習済みモデルとする)に、学習に使用しなかったデータ(以下、テストデータとする)を識別させた際の損失で評価した。損失の算出には、式(1)に示す損失関数を用いた。ここで、テストデータのサンプルサイズは 100 とし、良品・不良品の個数が同数になるように割り当てた。なお、損失関数はモデルの予測値と真の値との間の誤差を数値化するために使用するものであり、値が小さいほど優れたモデルと言える。本報では二値交差エントロピーと呼ばれるものを用いた。これは、

2 つのクラス分類問題における予測値と真値の間の交差エントロピーを計算するものである。

機械学習に用いた訓練データのサンプルサイズは 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 とし、総数 1000 の訓練データから無作為に良品・不良品のデータを同数取り出して学習させた。各サンプルサイズの条件の学習済みモデルに対し、テストデータを良品・不良品に識別させた結果から、損失と訓練データのサンプルサイズの間を確認した。なお、訓練データの取り出しと学習は、各サンプルサイズの条件で 10 回ずつ行い、得られた 10 個の学習済みモデルをテストデータに識別させた結果の平均値を評価値とした。

$$L_{BCE}(y, \hat{y}) = -\sum_{i=1}^N [y^i \log \hat{y}^i + (1 - y^i) \log(1 - \hat{y}^i)] \tag{1}$$

L_{BCE} : Loss(Binary Cross Entropy) y : Probability Distribution of True Values

\hat{y} : Probability Distribution of Predicted Values N : Sample Size

(2) 精度と訓練データの偏り

良品・不良品に偏りが生じた際における、訓練データのサンプルサイズが精度に及ぼす影響についても評価した。訓練データのサンプルサイズの偏りについての実験では、訓練データのサンプルサイズが 50, 500 の条件で、良品・不良品のデータの比 (偏り) を、1:9, 2:8, 3:7, 4:6, 5:5, 6:4, 7:3, 8:2, 9:1 まで変化させて、学習させた。ここでは、5:5 を「偏りなし」とし、学習済みモデルの評価方法は(1)と同様の方法で実施した。

3. 結果と考察

3.1 精度と訓練データのサンプルサイズ

各検査項目における学習済みモデルをテストデータで評価し、得られた損失を Fig.3 に示す。

穴加工の有・無に関する検査では、訓練データのサンプルサイズが 100 で損失 0.02 以下となり、学習済みモデルの識別が正確であることが確認できた。

異物 (小) の検査では、サンプルサイズが 300 まで損失が小さくなる傾向にあり、以降は 0.1 前後で安定した。異物 (中) の検査では、サンプルサイズが 100 から 300 までの間に、損失が約 0.2 減少し、以降も緩やかに低下する傾向となった。異物 (大) の検査では、サンプルサイズが 400 まで損失が低下する傾向にあり、以降は 0.1 前後で安定した。以上より、今回扱った異物の検査では、粒径が異なる条件下であっても、その損失はサンプルサイズが 300 から 400 程度で安定すると考えられる。

3.2 精度と訓練データの偏り

学習済みモデルをテストデータで評価し、得られた損失を Fig.4 に示す。

すべての検査項目において、サンプルサイズに偏りが生じるほど、損失が大きい傾向が確認できた。また、サンプルサイズが 500 の条件 (Fig.4 (b)) は 50 の条件 (Fig.4 (a)) に比べ、サンプルサイズに偏りが生じた際に生じる損失が小さくなる傾向となった。これらのことから、訓練データのサンプルサイズは良品・不良品間で偏りなく (均等) かつできる限り多く確保することが望ましいと考えられる。

なお、Fig.4(b), Fig.5(b)より、良品・不良品の訓練データのサンプルサイズの比率が 1:9 の条件

と 9:1 の条件を比較すると、9:1 に偏った条件の方が、損失が小さい傾向となった。すなわち、訓練データ数が良品に偏った方が、不良品に偏るよりも、精度・損失の結果が良くなる傾向が確認できた。

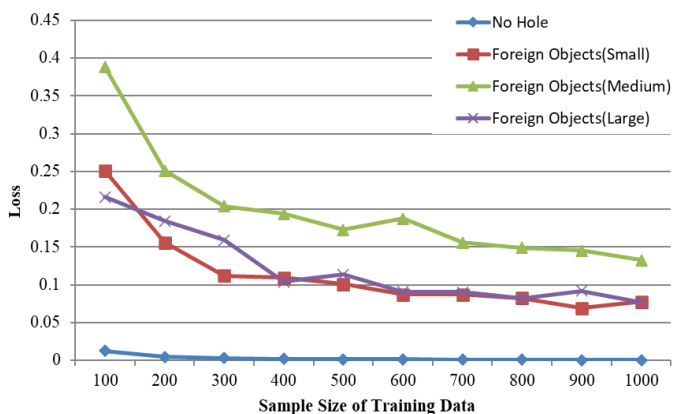
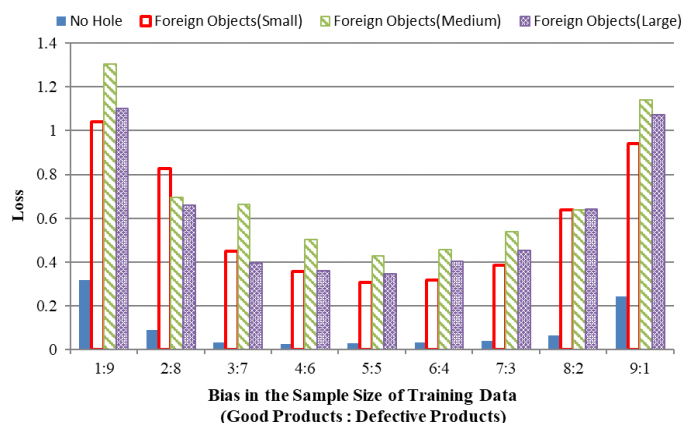
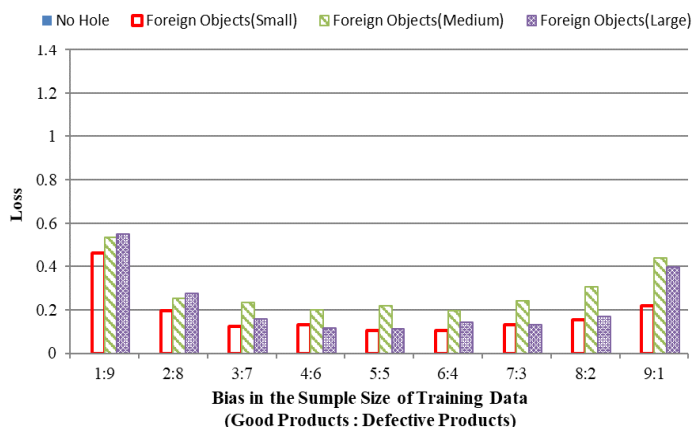


Fig.3 Relationship between Loss and Sample Size of Training Data



(a) Sample Size = 50



(b) Sample Size = 500

Fig.4 Relationship between Loss and Bias in Sample Size of Training Data

このことについては、以下のように推測される。まず、基本的に、良品であろうと、不良品であろうと、それぞれの特徴空間内の分布はサンプルサイズが大きいかほど凝集する。良品と不良品を比較した場合、良品は、不良品に比べ、外観の変化が小さいことから、特徴の分布はより凝集していると考えられる。例えば、ここで良品と不良品の比 1:9 の条件と 9:1 の条件で比較する。それぞれの比を保ちながら訓練データのサンプルサイズを増やした場合には、良品 9:不良品 1 の条件の方がその特徴をより正確に捉えられる。その結果、良品 9:不良品 1 の条件における損失が小さくなったと考えられる。

4. 結論

機械学習の精度に及ぼす訓練データ数の影響について、外観検査を模擬した訓練データを元に、検査項目ごとで評価を行った結果、以下のことが明らかとなった。

穴加工の検査では、100 のサンプルサイズで正確な識別ができる。

異物の検査では、すべての粒径において、300 から 400 程度のサンプルサイズで損失が安定した。このことから、今回扱った異物の検査に類似する項目であれば、同様の損失と訓練データ数の関係を指標にできると推察される。

サンプルサイズに偏りが生じた際、偏りが大きいほど損失が大きい傾向が得られた。また、サンプルサイズが多い条件下で偏りが生じた場合は、この傾向が小さくなることから、集めるべき訓練データは、できる限り均等かつ多く確保することが望ましい。

謝辞

本研究にあたり、研究助成をいただいた公益財団法人岡三加藤文化振興財団様に謝意を表します。

参考文献

- 1) 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 報告書”. p8-12(2011)
- 2) AI 白書編集委員会：“AI 白書 2023”. 角川アスキー総合研究所, p6-7 及び p209-216(2023)
- 3) 国立研究開発法人産業技術総合研究所：“機械学習品質評価・向上技術に関する報告書”. p10-11(2022)
- 4) 岡谷貴之：“画像認識のための深層学習の研究動向 —畳み込みニューラルネットワークとその利用法の発展—”. 人工知能, 31(2), p169-179(2016)
- 5) Karen Simonyan, Andrew Zisserman: “Very deep convolutional networks for large-scale image recognition”. published as a conference paper at ICIR 2015, p1-14(2015)
- 6) 吉岡智和：“深層学習による被害写真を用いた RC 方立壁の損傷度識別器の生成”. 日本建築学会技術報告集, 26(64), p1252-1257(2020)
- 7) Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, Ruslan Salakhutdinov: “Dropout: A simple way to Prevent Neural Networks from Overfitting”. J. Mach. Learn. Res 15, p1929-1958(2014)