

# 畳み込みニューラルネットワーク (CNN) の転移学習を用いた ラップフィルムの不良品検出

－ InceptionV3 の転移学習による CNN の設計と評価－

有馬 滉宜 (山口東京理科大学工学部機械工学科), 中島 健斗, 三木 康平

永田 寅臣 (山口東京理科大学大学院工学研究科), 渡辺 桂吾 (岡山大学大学院自然科学研究科)

## Detection of Defective Wrap Roll Product Using Transfer Learning of Convolution Neural Networks

－ Design and evaluation of CNNs by transfer learning of InceptionV3 －

Koki ARIMA, Department of Mechanical Engineering, Sanyo-Onoda City University

Kento NAKASHIMA, Kohei MIKI, Fusaomi NAGATA (Graduate School of Engineering,  
Sanyo-Onoda City University)

Keigo WATANABE (Graduate School of Natural Science and Technology, Okayama University)

Abstract : The authors are developing an application that make it easy to design convolutional neural networks (CNNs) and support vector machines (SVMs). In this study, the authors try to detect defects that occur in the manufacturing process of wrap roll products by using a transfer learning based CNN model. In advance, template matching is applied to all the images of the wrap roll products to efficiently extract only the film part. InceptionV3 is used as the trained CNN model for transfer learning, in which the fully connected layers are replaced according to the number of classifications, i.e., OK or NG. After the structure of the CNN model is newly designed, additional training is done using a large number of non-defective and defective images increased by image augmentation technique in order to enhance the generalization ability. The CNN model obtained by the transfer learning of InceptionV3 is evaluated through the classification experiment of test images including defective products.

### 1. 緒言

様々な工業製品の検査工程においては一部で自動化が進んでいるものの、それぞれの製品の品質管理に精通した検査員の目視検査に頼るところが大きい状況である。最近、深層学習の技術を画像認識に特化させた CNN を製品の欠陥検出に応用しようとする試みがなされている。例えば、寺野らは VGG16 アーキテクチャーを用いた転移学習によるコンクリート構造物のひび割れ抽出及び判別方法に関する研究を行っている [1]。また、Perez らは建造物の欠陥・劣化の自動検出と位置特定ができるように VGG16 の転移学習を用いて 4 カテゴリー分類用に新たな CNN を設計し、分類性能を評価する研究を行っている [2]。

著者らは現在、図 1 に示すようなラップフィルム製品の不良品検出の研究を行っている。透明なフィルムは、はみだしや光の反射などにより、市販の画像認識システムを用いても十分な検出結果が得られないという課題がある。これまでに、オリジナルで設計した CNN を特徴抽出器として用いた SVM [3] と、転移学習により新たに設計した CNN [4] による欠陥部分の検出の基礎研究に取り組んできた。本研究では、より高い分類性能を得るために図 2 に示す開発済みの CNN&SVM 設計ツールを用いて、学習済み

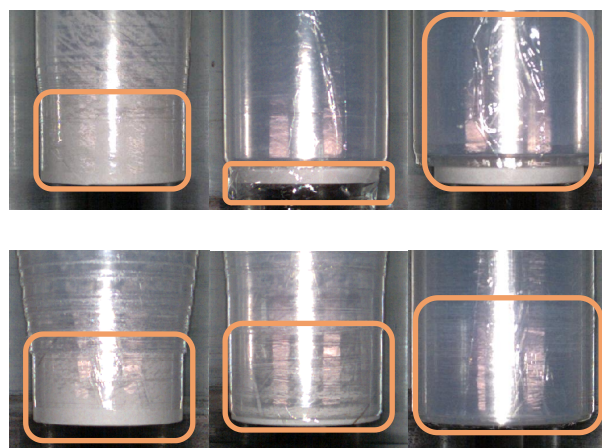


Fig. 1: Image samples of wrap film products with several kinds of defects after template matching.

の CNN である InceptionV3 [5] の転移学習により新たな CNN を設計し、分類性能を評価する。

まず、治具を含むラップフィルム製品の全体画像に対してテンプレートマッチングを適用し、ラップフィルム部分のみを抽出する画像処理を行う。次に学習済みの CNN には InceptionV3 を用い、2 クラス分類用に終段にある

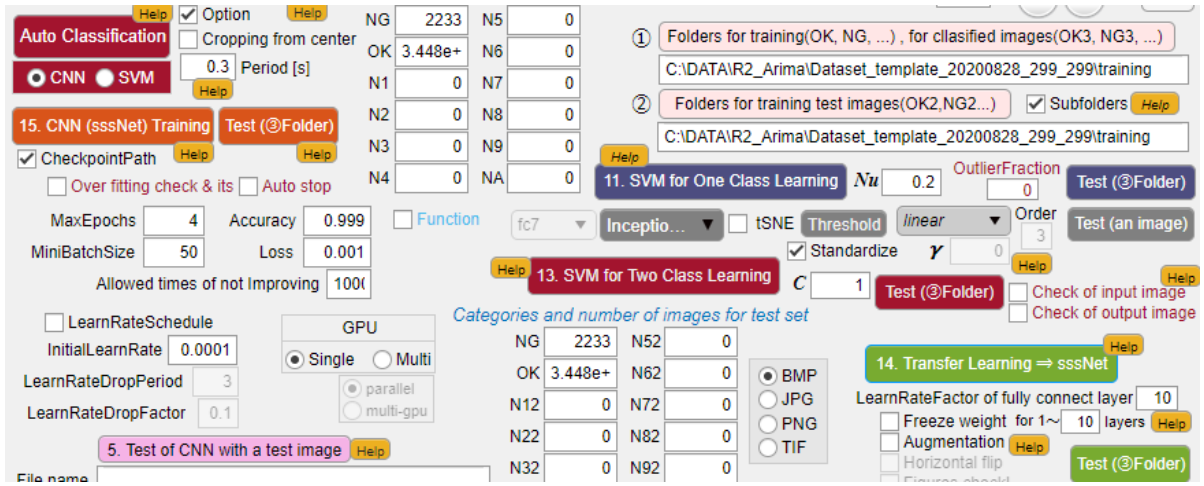


Fig. 2: A part of main dialog developed on MATLAB system to user-friendly design original CNN[6].

全結合層部を置き換えた転移学習用の新たな CNN を設計する．この CNN に対して全層を追加学習させた CNN (IncA) と，終段にある全結合層部のみを追加学習させた CNN (IncB) を用いて分類性能を比較する．汎化性を発揮できるように画像拡張（画像オーギュメンテーション）を施した多数の良品と不良品の画像を訓練に用いる．テスト画像の分類実験によって設計した IncA と IncB を比較評価する．

## 2. テンプレートマッチングによる訓練画像のダウンサイジング

様々な課題に対する画像処理で広く利用されているテンプレートマッチングは，撮影されたワークの中で欠陥が含まれやすい領域を抽出するためにも非常に有効である．これから設計する CNN への入力画像のサイズを大幅に減らすことができ，計算コスト，メモリ占有コストを軽減することができる．このため，開発したメイン，オプション及びオーギュメンテーションのダイアログの中ではテンプレートマッチング機能を利用できるようにしている．

$(M, N)$  のサイズのテンプレートをターゲット画像内でラスタースキャンさせて相関係数の高い位置を検出する場合，周辺領域でもマッチング評価ができるように図 3 に示すようにパディング処理を行う．テンプレートとパディングにより拡張されたターゲット画像内の同面積の領域との相関係数  $\alpha(u, v)$  は，次式から計算される．

$$\alpha(u, v) = \frac{s_{it}(u, v)}{s_i(u, v)s_t(u, v)} \quad (1)$$

$$\sum_{y=v}^{v+N-1} \sum_{x=u}^{u+M-1} \{f(x, y) - \bar{f}_{u,v}\} \{t(x-u, y-v) - \bar{t}\} \quad (2)$$

$$s_{it}(u, v) =$$

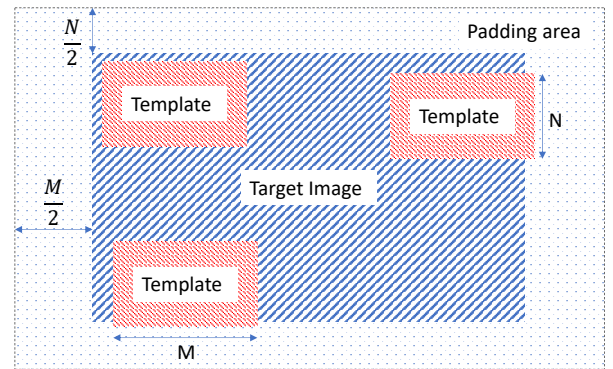


Fig. 3: Configuration among a target image, padding area and template image with the size of  $(M, N)$ .

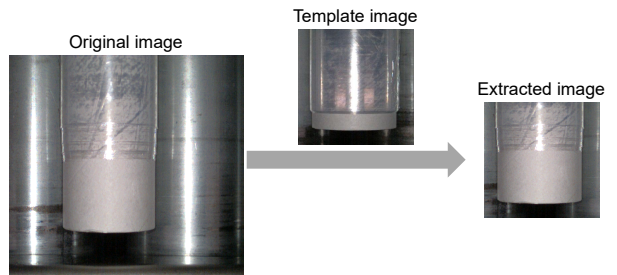


Fig. 4: An example of extracted image using the template matching technique.

$$s_i(u, v) = \sqrt{\sum_{y=v}^{v+N-1} \sum_{x=u}^{u+M-1} \{f(x, y) - \bar{f}_{u,v}\}^2} \quad (3)$$

$$s_t(u, v) = \sqrt{\sum_{y=v}^{v+N-1} \sum_{x=u}^{u+M-1} \{t(x-u, y-v) - \bar{t}\}^2} \quad (4)$$

ここで， $(u, v)$  はターゲット画像内におけるテンプレ

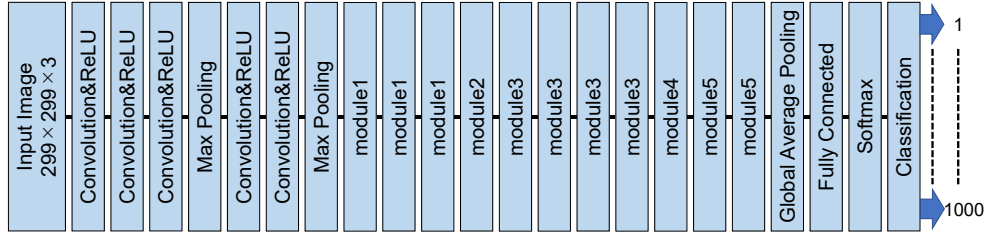


Fig. 5: Network architecture of original InceptionV3 .

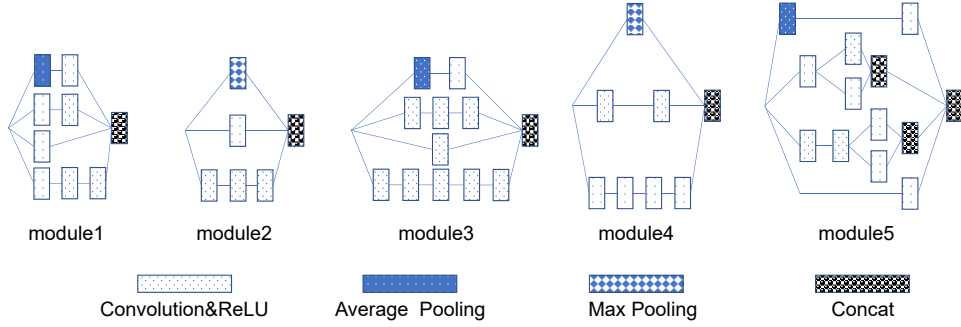


Fig. 6: Each module included in original InceptionV3 network shown in Fig. 5.

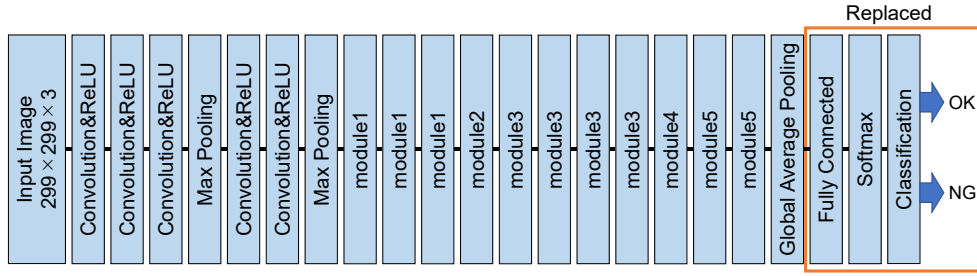


Fig. 7: Reconstructed InceptionV3 architecture for binary classification designed by application shown in Fig. 2.

ト左上コーナーの座標である． $s_t(u, v)$  と  $s_i(u, v)$  はそれぞれ，テンプレート内とターゲット内比較領域の標準偏差であり， $s_{it}(u, v)$  は共分散である． $f(x, y)$  は拡張された画像内の  $(x, y)$  におけるグレースケール 256 階調値を正規化した値である． $t(x - u, y - v)$  はテンプレート内の  $(x - u, y - v)$  における同様の値である． $M$  と  $N$  はそれぞれ，テンプレートの幅と高さである． $\bar{t}$  と  $\bar{f}(u, v)$  はそれぞれ，テンプレート内のグレースケール値の平均値と，ターゲット画像内のテンプレート真下の領域のそれである．式 (1) で与えられる相関係数  $\alpha(u, v)$  は，テンプレートをターゲット画像内の左上から右下までラスタースキャンさせることで計算される．ラスタースキャン後，テンプレートと最もマッチする領域，すなわち最も大きな値  $\alpha(u, v)$  を持つ領域が抽出される．図 4 には，ラップフィルム品のオリジナル画像からテンプレートで抽出された画像の例を示している．今回の実験では，図 1 のようにテンプレート

マッチングにより  $640 \times 480$  の解像度を持つラップフィルム製品画像から目標領域を正方形  $347 \times 347$  の解像度で切り出し，ダウンサイジングした．

### 3. InceptionV3 による欠陥検出

#### 3.1 InceptionV3 の転移学習について

本研究では図 2 のアプリケーションを用いて InceptionV3 の転移学習により 2 クラス分類用の CNN を新たに設計する．図 5, 6 にはそれぞれオリジナルの InceptionV3 と各モジュールの構造を示している．InceptionV3 は，ILSVRC2014 で優勝した InceptionV1(GoogLeNet) の畳み込みフィルタのサイズ変更とバッチ正規化を導入したアーキテクチャである [5]．これは ImageNet データベースの 100 万枚を超える画像で事前に学習されており，入力

画像のサイズは  $299 \times 299 \times 3$  である。

転移学習の手法の 1 つとして Fine-tuning がある。Fine-tuning は異なるタスクに対して学習済みの CNN のモデルの重みを初期値とすることで、適用先のデータセットで学習を進める手法である [7]。図 7 には、この方法に基づき、InceptionV3 の終段にある全結合層部を 2 クラス分類用に新たに設計した構造を示す。Kandel らは InceptionV3, VGG16, VGG19 それぞれに対して、幾つかのブロックごとに Fine-tuning を適用し、組織病理画像をより高い精度で分類できる方法についての研究を行っている [8]。この研究では InceptionV3 のネットワーク全体を Fine-tuning させることにより、最も良い結果が得られたと報告されている。

本研究では、全層を Fine-tuning させたものを IncA、終段の全結合層部のみを Fine-tuning させたものを IncB とし、2 種類の CNN を設計した。このため、学習後の IncB については、全結合層部以外の全ての層の重みは InceptionV3 のそれと同値を持つことになる。共通の学習条件として、訓練用データセットは良品画像 34,482 枚と不良品画像 2,233 枚で構成した。また、最大エポック 4, ミニバッチサイズ 50, 学習率を 0.0001 に設定し、訓練を行った。なお、図 2 で紹介したアプリケーション内の右下のチェックボックスで重み更新を行わない層の設定を可能にしている。

### 3.2 分類実験

訓練後の汎化性能を評価するためにテスト用データセットとして良品画像 4,035 枚と不良品画像 21 枚の計 4,056 枚を用いて IncA と IncB による分類実験を行った。表 1 には IncA の全体的な認識結果を表す混合行列を示す。誤認識された合計枚数は 3 枚で、その内 2 枚の不良品画像は良品画像として誤認識され、1 枚の良品画像は不良品画像として誤認識されていた。この 1 枚の画像を観察すると、不良品の画像に類似したグレーな状態であったことが確認された。表 1 と同様に表 2 には IncB の全体的な認識結果を表す混合行列を示す。誤認識した合計枚数は 22 枚で、その中には IncA が誤認識していた画像も含まれていた。

### 3.3 比較結果

次に 4 つの評価基準である認識率、精度、再現率、及び  $F$  値について説明する。認識率はテスト画像データセットに対する正解率を、精度は不良品と判断された画像のうち実際に不良品であった割合を、再現率は実際に不良品である画像が正しく不良品と判断された割合を、さらに  $F$  値は精度と再現率の調和平均の値をそれぞれ表す。これらの

Table 1: Confusion matrix classified by IncA (row: true labels, column: predicted labels).

	Normal	Anomaly
Normal	4034	2
Anomaly	1	19

Table 2: Confusion matrix classified by IncB (row: true labels, column: predicted labels).

	Normal	Anomaly
Normal	4017	4
Anomaly	18	17

Table 3: Statistical evaluation results of IncA and IncB.

CNN	Accuracy	Precision	Recall	F-measure
IncA	0.998	0.791	0.904	0.844
IncB	0.994	0.485	0.809	0.607

指標をもとに IncA と IncB の汎化性能を評価する。全体のデータの中で不良品を不良品として分類したものを  $T_P$ 、良品を良品として分類したものを  $T_N$ 、不良品を良品として誤分類したものを  $F_N$ 、さらに良品を不良品として誤分類したものを  $F_P$  としたとき認識率  $A_c$ 、精度  $P_r$ 、再現率  $R_e$ 、及び  $F$  値  $F$  はそれぞれ、次式で与えられる。

$$A_c = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \quad (5)$$

$$P_r = \frac{T_P}{T_P + F_P} \quad (6)$$

$$R_e = \frac{T_P}{T_P + F_N} \quad (7)$$

$$F = 2 \times \frac{P_r \times R_e}{P_r + R_e} \quad (8)$$

表 3 には表 1, 2 の結果をもとに算出した IncA と IncB の認識率、精度、再現率、及び  $F$  値を示している。IncA の認識率と再現率は、それぞれ 0.998, 0.904 であった。再現率が高いということは、取りこぼしなく不良品を検出できているということになる。また、精度は 0.791 と低い値であったが、これはテスト画像に含まれる良品のサンプル数が 4,056 であり、不良品の 21 に対して多かったことに起因していると考えられる。製造ラインにおける不良品の未検出は流通後の商品回収、リコールなど重大な問題に発展する。このため、不良品を良品として誤認識する画像の枚数を減らすために精度よりも再現性の値が重要視される。

表 1, 2 から確認できるように IncB は IncA より誤認識枚数が多かったため、表 3 のように 4 つの評価基準となる

数値は IncA に比べて低い結果となった。これらの結果から、全層を Fine-tuning させることでより高い分類性能を得られることが確認できた。

## 4. 結言

本研究では、開発済みの CNN&SVM 設計ツールを用いて InceptionV3 の転移学習による CNN を構築し、ラップフィルム製品の製造過程で発生する欠陥検出を試みた。まず、治具を含むラップフィルムの全体画像に対してテンプレートマッチングを適用し、ラップフィルム部分のみを抽出する画像処理を行うことで、良品と不良品のターゲット領域のオリジナル画像を採集した。次に、これらのオリジナル画像と、それらを画像拡張して増やした多数の画像を加えてデータセットを構成し、転移学習の手法の 1 つとして層の重みを軽微に更新する Fine-tuning を用いた。全層を Fine-tuning させた CNN と、終段の全結合層部のみを Fine-tuning させた CNN をそれぞれ作成し、不良品の画像を含むテスト画像のデータセットを用いた分類実験により比較したところ、全層を Fine-tuning させた CNN を用いた場合がより良い分類性能が得られた。これは、Kandelらの組織病理画像を用いた研究と同様の結果となり、ラップフィルム製品の分類でも全層を Fine-tuning させた方が効果的であると確認できた。

今後の展開として、他の製品にも転移学習ベースの CNN による欠陥検出法を適用し、Fine-tuning の適用方法とともに有効性を評価していきたい。

## 参考文献

- [1] 寺野 聡恭, 出水 享, 古賀 掲維, 中島 賢哉, 松田 浩, “深層学習を用いたコンクリート構造物のひび割れ抽出・判別方法に関する研究,” 長崎大学大学院工学研究科研究報告, pp. 71–76, 2020.
- [2] H. Perez, J. H. M. Tah, A. Mosavi, “Deep Learning for Detecting Building Defects Using Convolutional Neural Networks,” *Sensors*, Vol. 19, No. 16, pp. 3556–3579, 2019.
- [3] 中島 健斗, 永田 寅臣, 渡辺 桂吾, “畳み込みニューラルネットワーク (CNN) とサポートベクターマシン (SVM) を用いた微小な欠陥を持つ不良品検出の基礎研究,” ロボティクス・メカトロニクス講演会 2019 講演論文集, 2A1-Q05(1–4), 広島国際会議場, 2019.
- [4] K. Nakashima, F. Nagata, H. Ochi, A. Otsuka, T. Ikeda, K. Watanabe, M. K. Habib, “Detection of Minute Defects Using Transfer Learning-Based CNN Models,” *Procs. of 25th International Symposium on Artificial Life and Robotics*, pp. 871–875, 2020.
- [5] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna “Rethinking the Inception Architecture for Computer Vision,” *Procs. of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2818–2826, 2016.

- [6] F. Nagata, K. Tokuno, K. Mitarai, A. Otsuka, T. Ikeda, H. Ochi, K. Watanabe, M. K. Habib, “Defect Detection Method Using Deep Convolutional Neural Network, Support Vector Machine and Template Matching Techniques,” *Artificial Life and Robotics*, Vol. 24, No. 4, pp. 512–519, 2019.
- [7] 中山 英樹 “深層畳み込みニューラルネットワークによる画像特徴抽出と転移学習,” 信学技報, pp. 55–59, 2015.
- [8] I. Kandel and M. Castelli, “How Deeply to Fine-Tune a Convolutional Neural Network: A Case Study Using a Histopathology Dataset,” *Applied Science*, Vol. 10, No. 10, pp. 1–20, 2020.