

# 畳み込みニューラルネットワークを用いた ラップロール製品の不良品検出

Detection of Defective Wrap Roll Product Using Convolution Neural Networks

○ <sup>1</sup> 中島 健斗      <sup>1</sup> 永田 寅臣      <sup>2</sup> 渡辺 桂吾  
○ <sup>1</sup>Kento NAKASHIMA    <sup>1</sup>Fusaomi NAGATA    <sup>2</sup>Keigo WATANABE  
<sup>1</sup> 山口東京理科大学大学院      <sup>2</sup> 岡山大学大学院  
<sup>1</sup>Sanyo-Onoda City University      <sup>2</sup>Okayama University

**Abstract:** Although the automation of inspection processes for various kinds of industrial products has progressed, the situation seems to be largely depending on visual inspection ability of inspectors who are familiar with the quality control of each product. Recently, not a few attempts have been tried to apply convolutional neural networks (CNNs) specialized in deep learning technology to image recognition for product defect detection. The authors have developed an application that can design and train CNNs and support vector machines (SVMs). In this paper, the application is tried to be applied to the defect detection in the manufacturing process of wrap roll product. Firstly, a template matching technique is used to extract only the target film areas from the entire images of wrap roll products. Next, a CNN named sssNet consisting of 15 layers is originally designed so as to classify input images into defective or not, then trained using a large number of original and augmented images to enhance the generalization ability. Finally, the trained sssNet is evaluated through classification experiments of test images. The usefulness of the developed application with a promising function of defect visualization is also assessed through this test trial.

## 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。これまでも畳み込みニューラルネットワーク (CNN) を製品の欠陥検出に応用した研究がある。例えば、多田らは多段階型 CNN を用いた自動車用タイヤの内面における欠陥識別の研究を行っている [1]。Young らはコンクリートの亀裂損傷検出に 8 層で構成された CNN を用いた研究を行っている [2]。この CNN では、第 1, 3, 5 層の畳み込み層の後に Batch Normalization が配置されており、さらに第 6 層には Drop Out を配置するという工夫が見られる。

筆者らは、現在、図 1 に示すようなラップロール製品の不良品検出を具体的な課題としているが、透明なフィルムのはみ出しや光の反射などにより、市販の画像検出器を用いても十分な不良品検出の性能が得られないという問題がある。この課題を解決するために、AlexNet などの優れた 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 それぞれによる微小な欠陥検出の基礎研究に取り組んできた [4]。しかしながら、既存の CNN を用いた場合、それらが既に実際の製品とは異なる多くのカテゴリの画像で学習されていることから、各層の特徴マップをもとにした欠陥部分の可視化を十分に行えなかった。本研究では、図 2 に示す MATLAB 環境で開発済みのアプリケーションを用いてラップロー

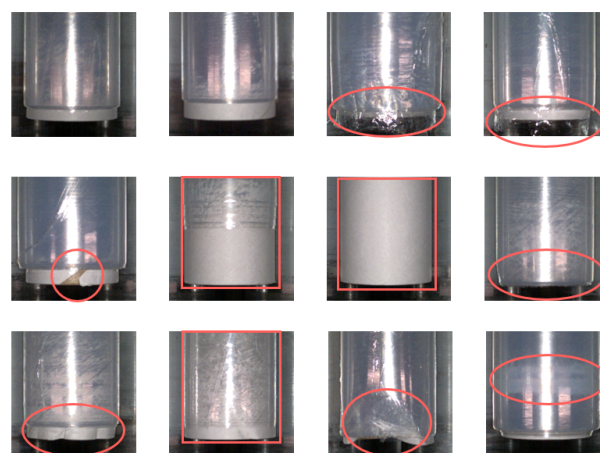


Fig. 1: Image samples of wrap roll with several kinds of defects after template matching.

ルの製造工程で発生する欠陥の検出を試みる。まず、治具を含むラップロール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、ラップロール部分のみを抽出する画像処理を行う。次に 15 層から成る CNN(sssNet) を設計し、汎化性を発揮できるようにオリジナルの良品と不良品の画像に加えて、画像オーギュメンテーションを施した多数の画像を用いて訓練する。不良品を含むテスト画像の検出実験によって設計した sssNet を評価する。また、最終的な分類結果に影響を及ぼした特徴部分の視覚化機能を含め、開発済みのアプリケーションの有用性を確認する。

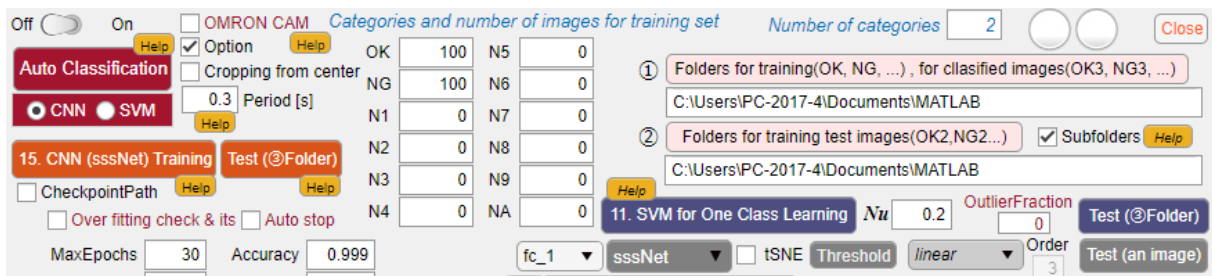


Fig. 2: A part of main dialog developed on MATLAB system to user-friendly design original CNN [4].

## 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

### 2.1 相関係数に基づくマッチング領域の抽出

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。

さて、 $(M, N)$  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。

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

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

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

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

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

ここで、 $(u, v)$  はターゲット画像内におけるテンプレート

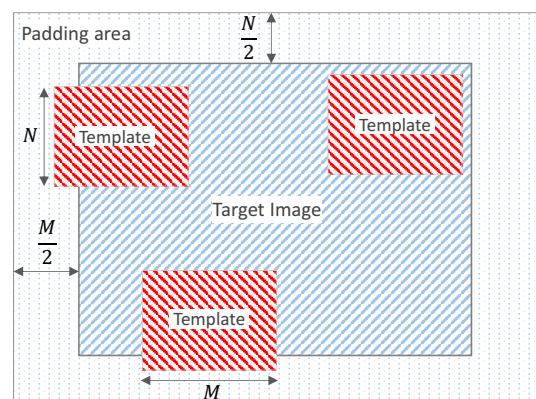


Fig. 3: Configuration among a target image, padding area and template image whose size is  $(M, N)$ .

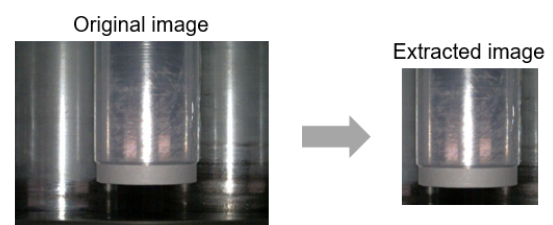


Fig. 4: An example of extracted image using the template matching technique.

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

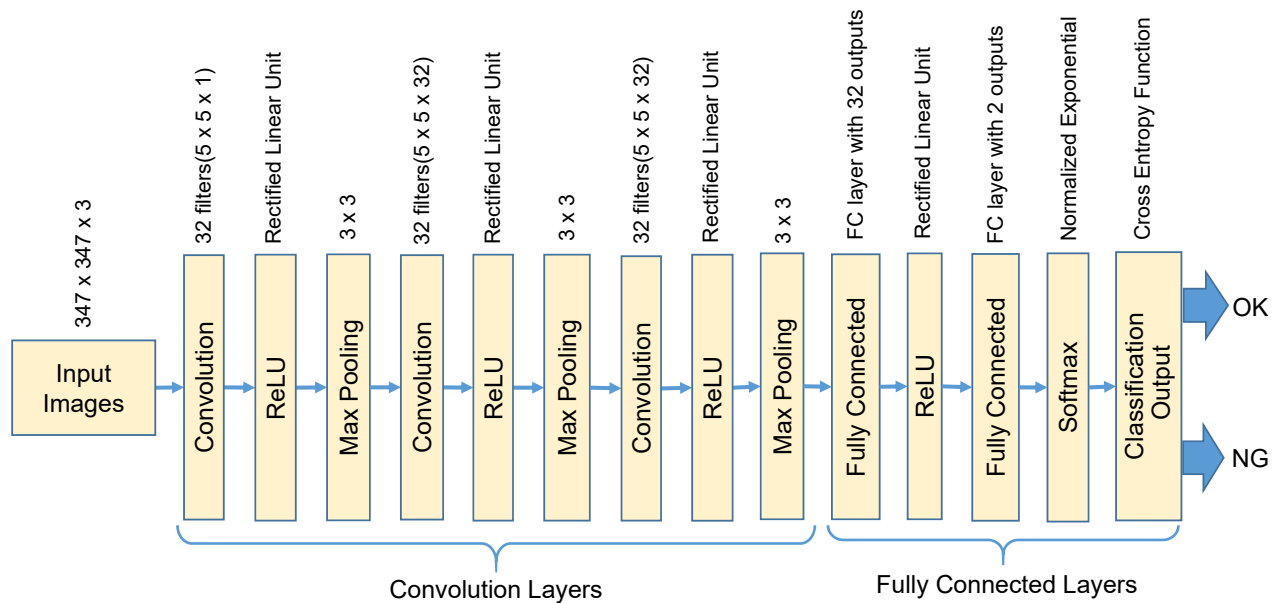


Fig. 5: Our original CNN named sssNet designed using the proposed application shown in Fig. 2.

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、テンプレートで抽出されたラップロール品の画像の例を示している。今回の実験では、テンプレートマッチングによりラップロール製品の画像解像度を  $640 \times 480$  から  $347 \times 347$  にダウンサイジングした。

### 3 ターゲット画像で訓練した CNN による欠陥検出

AlexNet などの既存の CNN を転移学習に用いる場合、入力画像の大きさが  $227 \times 227$  のように限定されるために、解像度の大きな訓練画像を用意できても画像サイズの縮小に伴う解像度の低下を余儀なくされる。微小な欠陥の場合、画像のダウンサイジング時にその特徴が失われてしまう恐れがあるために、もともとの訓練画像の解像度を落とさなくてもよいようにオリジナルで設計した CNN (sssNet) を用いて欠陥検出を行うこととした。

#### 3.1 CNN(sssNet) の設計

図 5 には 15 層で構成されたオリジナルの CNN である sssNet の構造を示す。画像を良品または不良品に分類できるバイナリ分類用に設計しており、図 1 に示すような画像を使用して、ランダムに初期化された重みで学習させた。訓練では、図 1 に示す入力画像が入力層に与えら

れ、正規化される。第 2, 5, 8 番目の畳み込み層は、それぞれ  $5 \times 5 \times 1 \times 32$  (縦×横×チャネル数×フィルタ数)、 $5 \times 5 \times 32 \times 32$ , および  $5 \times 5 \times 32 \times 32$  の構造のフィルタを持つ。畳み込み層では、特徴抽出要素として作用するフィルタをストライドで指定されたピクセル数ずつ左上から右下にシフトさせながら特徴マップを生成していく。第 3, 6, 9, 12 層には正規化線形ユニット (ReLU) を配置し、さらに第 13 層の最終的な判断を行う全結合層を経て第 14 層では正規化指数関数である softmax 関数により各カテゴリに対する確率であるスコアを出力できるようにしている。

#### 3.2 sssNet による学習

訓練用データセットとして良品 29,483 枚と不良品 2,167 枚の画像を用いた。学習条件は最大エポックを 30, ミニバッチサイズを 50, 学習率を 0.0001 に設定して sssNet の訓練を行った。次にこの訓練済みの sssNet の汎化性能を評価するためにテスト画像として良品 5,001 枚と不良品

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

	Normal	Anomaly
Normal	4982	4
Anomaly	19	29

Table 2: Statistical evaluation results.

CNN	Accuracy	Precision	Recall	F-measure
sssNet	0.995	0.604	0.879	0.716

33 枚の画像を用いて分類実験を行った。表 1 には sssNet の全体的な認識結果を表す混同行列 (横軸: 実際のクラス, 縦軸: 予測クラス) を示す。表 1 より, 良品のデータセットに含まれていたにもかかわらず不良品として認識された画像は 19 枚であった。これら 19 枚の画像のほとんどが不良品の画像にも類似したグレーな状態であったからだとと思われる。

次に 4 つの評価基準である認識率, 精度, 再現性, 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, 再現性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  $F$  値は精度と再現性の調和平均の値をそれぞれ表す。これら 4 つの指標をもとに sssNet の汎化性を評価する。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  $P_r$ , 再現性  $R_e$ , 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 2 には表 1 をもとにしたこれらの計算結果を示す。認識率と再現性は, それぞれ 0.995, 0.879 であった。再現性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。次に精度は 0.604 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 5,001 であり不良品の 33 に対して多かったことに起因している。実際の製造ラインにおいては不良品が良品のロットに含まれることは流通後の商品回収など重大なトラブルにつながる恐れがあるため注意を要する。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。また,  $F$  値が低いということは精度と再現性のバランスが悪いということを表しているが, 原因としては, 今回の実験では良品と不良品のテスト画像の枚数が大きく異なり,  $P_r$  の値が 0.604 と小さかったことが考えられる。

#### 4 特徴マップの可視化

次に, 図 6 に示す畳み込み層を例にして活性化領域の可視化のための処理について述べる。この層では  $N$  枚のフィルタにより,  $w \times h$  のサイズの特徴マップが  $N$  枚生成されている。 $N$  枚のフィルタはそれぞれ, 手前側の畳み込み層で生成された特徴マップと同数のチャンネルを持つ。各特徴マップ  $^k \mathbf{m} (k = 1, \dots, N)$  は  $w \times h$  のサイズを持つ二次元配列であり, その要素である各ピクセル  $^k m_{ij}$  は活性化状況に応じたスカラー値を持つ。したがって, 全ての特徴マップを重ね合わせて得られる活性化マップ  $^{all} \mathbf{m}$  は次式から得られる。

$$^{all} \mathbf{m} = \sum_{k=1}^N ^k \mathbf{m} \quad (9)$$

さらに, 活性化状況をグレー諧調画像で視覚化するために, すなわち, 強く活性化している領域をより明るい白色で表示できるようにするために  $^{all} \mathbf{m}$  の要素  $^{all} m_{ij}$  の最小値  $m_{min}$  と最大値  $m_{max}$  をもとに, 各要素を次式のように正規化する。

$$^{all} \tilde{m}_{ij} = \frac{^{all} m_{ij} - m_{min}}{m_{max} - m_{min}} \quad (10)$$

なお, 各特徴マップの大域的 (全体的) な活性化状況を表す定量値  $^k v_{act}$  は次式で与えられる。

$$^k v_{act} = \sum_{i=1}^w \sum_{j=1}^h m_{ij} \quad (11)$$

式 (9) は全ての特徴マップをもとにした活性化マップであるが, 式 (11) で与えられるマップ毎の活性化値を用いることで活性化状況のランク付けが可能となり, 例えば, 活性化状況の高い上位 5 枚の特徴マップを畳み込んだ活性化マップを生成させることなども可能となる。図 7 には sssNet の第 10 層にある Max Pooling 層から抽出した活性化状況の高い上位 32 枚を平均化することで得られた  $42 \times 42$  ピクセルの特徴マップを示す。図 7 より, ラップロール製品のはみ出した透明フィルムの部分が白く活性化しているため, sssNet が透明なフィルムの部分まで特徴として捉えていることが分かった。

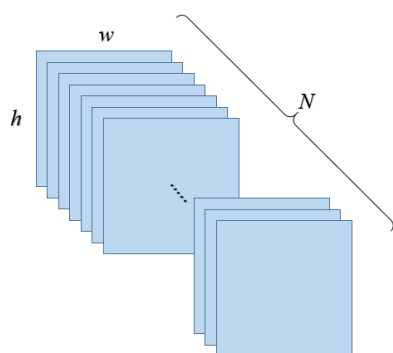


Fig. 6: Example of convolution layer, in which feature maps are generated as many as the number of filters. The filters have channels as many as the number of feature maps in the previous layer.

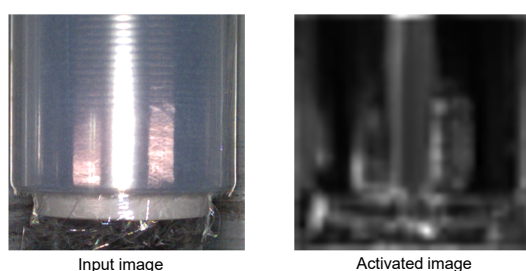


Fig. 7: Generated feature map of a wrap roll product.

## 5 結言

本研究では、開発済みの CNN&SVM 設計アプリケーションを用いてラップロール品の製造工程で発生する欠陥の検出を試みた。まず、ラップロール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、ラップロール部分のみを抽出する画像処理を行った。次に  $347 \times 347$  ピクセルの画像を入力として利用できるように 15 層から成る CNN (sssNet) を新たに設計し、汎化性を高めるために実際に撮影した良品と不良品のオリジナル画像に対して、画像オーギュメンテーションを施した多数の画像を加えたデータセットを用いて訓練した。不良品の画像を含むテスト画像のデータセットを用いて sssNet を評価したところ、認識率と再現性の 2 項目においてそれぞれ、0.995, 0.879 という結果が得られた。また、最終的な分類結果に影響を及ぼした特徴部分の視覚化が可能になり、開発したアプリケーションの有用性を確認することができた。

## 参考文献

- [1] 多田 拓太郎, 杉浦 彰彦, “多段階型 CNN を用いた自動車用タイヤ内面部における欠陥識別,” 電気学会論文誌 C (電子・情報・システム部門誌), pp. 1586–1594 2018.
- [2] Young-Jin Cha, Wooram Choi, Oral Buyukozturk, “Deep Learning-Based Crack Damage Detection Using Convolutional Neural Networks,” *Computer-Aided Civil and Infrastructure Engineering*, pp. 1–18, 2017.
- [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] Kento Nakashima, Fusaomi Nagata, Hiroaki Ochi, Akimasa Otsuka, Takeshi Ikeda, Keigo Watanabe, Maki 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.

