

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

— VGG19 の転移学習による CNN の設計と評価 —

○岩田 賢洸 (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 Film Using Transfer Learning of Trained Convolutional Neural Network (CNN)

— Design and evaluation of CNN model obtained by transfer learning of VGG19 —

○ Takahiro IWATA (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 have developed an application that can carry out transfer learning of Convolutional Neural Networks (CNNs). In this paper, we design two kinds of CNN models transferred from VGG19, i.e., with and without weights freezing in convolutional layers. The CNN models are trained using a large number of two categorized images of wrap film products, i.e., with and without defects. Before training process, our implemented template matching is applied to all images to extract only the film portion, so that the load regarding image processing can be largely reduced. The performance of the two kinds of CNN models are quantitatively evaluated and compared.

1. 緒言

これまでも畳み込みニューラルネットワーク (CNN) を製品や構造物の欠陥検出に応用した研究がある。例えば、浅田らは道路舗装の路面画像に対して CNN を適用し、ひび割れとパッチングを検出する研究を行っている [1]。

本研究では学習済みの CNN モデルの転移学習により、ラップフィルムの製造工程において発生する欠陥を持つ不良品の検出を行うことができるシステムを提案する。学習済みの CNN には VGG19 を用いる。図 1 には今回の不良品検出で対象とするラップフィルムの良品と不良品の画像の例を示す。

本研究ではまず、治具を含んだラップフィルムの画像に対してテンプレートマッチングを適用し、不良品検出のターゲットとなるフィルム部分のみを抽出する画像処理を行う。次に、VGG19 の全結合層部を 2 クラス分類用に置き換えて新たな CNN を設計する。画像オーギュメンテーションを施した多数の訓練画像を用いて追加学習 (ファインチューニング) を行う。この場合、畳み込み層の重み更新を軽微に行いながら訓練した CNN と重み更新を行わずに訓練した CNN を準備する。不良品が含まれたテスト画像の分類実験によって設計した 2 種類の CNN の性能を評価し、比較する。

2. VGG19 の転移学習による CNN を用いた不良品検出

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。

図 2 に示すように (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。

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

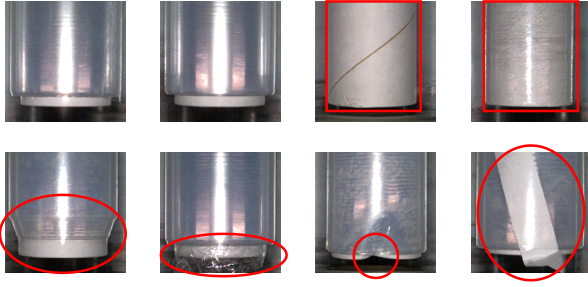


Fig. 1: Image samples of wrap films with and without several kinds of defects.

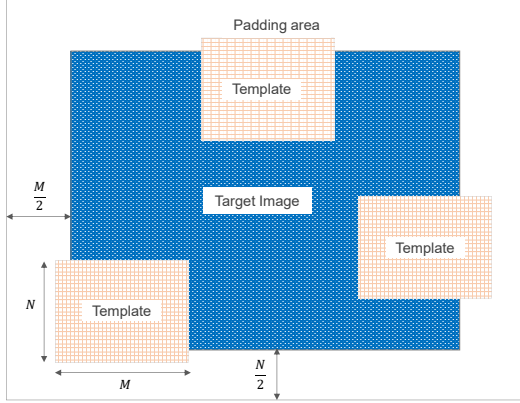


Fig. 2: Configuration among a target image, padding area and template image whose size is (M, N) .

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

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

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

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。

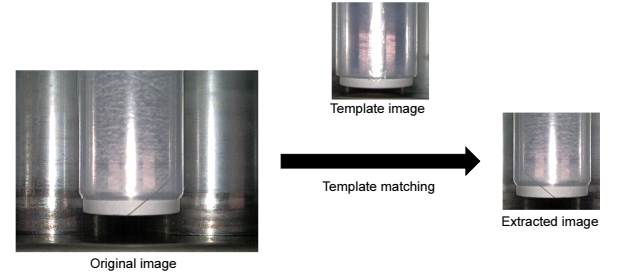


Fig. 3: An example of extracted image using the implemented template matching technique.

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 3 には、テンプレートで抽出されたラップフィルム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フィルムの画像解像度を 640×480 から 347×347 にダウンサイジングした。

2.2 VGG19 の転移学習による新たな CNN の設計

転移学習には学習済みの CNN のモデルの重みを初期値として用い、目的のタスクに応じた学習データセットに合うように全結合層を置き換え、追加学習を行うファインチューニングという手法がある。本研究では VGG19 の転移学習により設計した CNN に対してファインチューニングを行う。図 4 に示すように VGG19 は 16 の畳み込み層と 3 つの全結合層を含む合計 47 層から構成されている。最初の畳み込み層は、 $224 \times 224 \times 3$ のサイズの入力画像に対して、縦横サイズとチャンネル数が $3 \times 3 \times 3$ の 64 枚のフィルタにより 64 枚の特徴マップを生成する。第 2 の畳み込み層は最初の畳み込み層の出力を受け取り、 $3 \times 3 \times 64$ の特徴マップを 64 枚、第 3 の畳み込み層は、 $3 \times 3 \times 64$ のサイズのフィルタを 128 枚、第 4 の畳み込み層は、 $3 \times 3 \times 128$ サイズのフィルタを 128 枚、第 5 の畳み込み層は $3 \times 3 \times 128$ サイズのフィルタを 256 枚、第 6 から第 8 の畳み込み層は $3 \times 3 \times 256$ サイズのフィルタを 256 枚、第 9 の畳み込み層は $3 \times 3 \times 256$ サイズのフィルタを 512 枚生成する。第 10 から第 16 の畳み込み層は $3 \times 3 \times 512$ サイズのフィルタを 512 枚有し、フィルタと同数の特徴マップを生成する。全結合層は、 $4,096 \times 4,096 \times 1,000$ 個の重みを有する。オリジナルの VGG19 は、ImageNet データベースの 130 万枚の画像を用いて訓練されており、入力画像をキーボード、マウス、鉛筆、多くの種類の動物などを含む 1,000 種類のオ

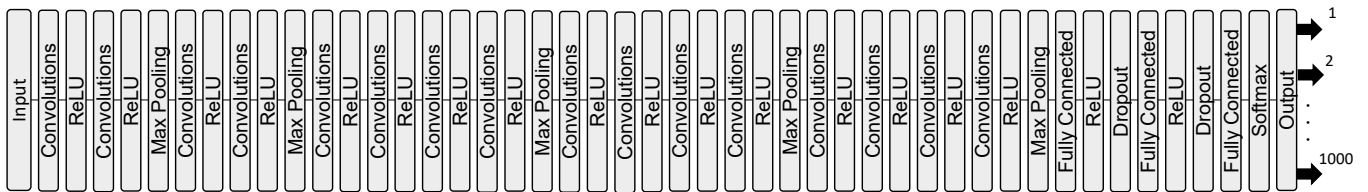


Fig. 4: Architecture of original VGG19.

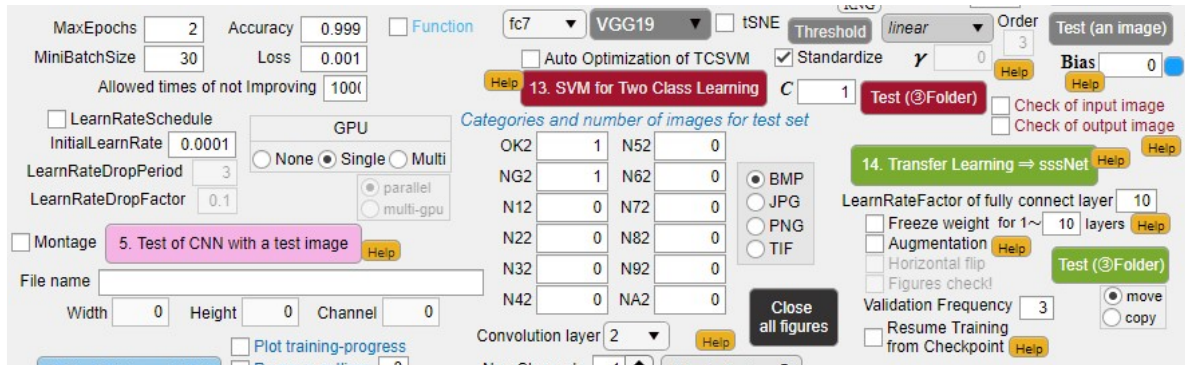


Fig. 5: A part of main dialog developed on MATLAB system to user-friendly design original CNNs.

ブジェクトクラスに分類できる [2].

2.3 分類実験

図 5 には MATLAB 上で開発した CNN の設計及び転移学習のためのソフトウェアを示す [3]。本研究ではこのソフトウェアの転移学習機能を用いて 2 クラス分類用の VGG19.2 を設計した。図 6 には設計した VGG19.2 の構造を示す。訓練用データセットとして良品 34,482 枚と、不良品 2,233 枚、計 36,715 枚を用いた。共通の学習条件として、最大エポックを 2、ミニバッチサイズを 30、全結合層の学習率を 0.001 に設定した上で、畳み込み層の学習率を 0.0001 とした場合と 0 (畳み込み層の重み更新、すなわち学習を行わない) とした場合の 2 つの条件で訓練を行った。ここで、学習率を 0.0001 とすることで畳み込み層の重み更新を軽微に行いながら訓練した CNN を VGG19.2A、学習率を 0 とすることで畳み込み層の重み更新を行わずに訓練した CNN を VGG19.2B とする。

次に訓練した VGG19.2A, VGG19.2B の汎化性能を評価するために未学習のテスト画像として良品 4,035 枚と、不良品 21 枚用意し、分類実験を行った。表 1 と表 2 にはそれぞれ VGG19.2A, VGG19.2B の全体的な認識結果を表す混合行列 (縦軸：実際のクラス, 横軸：予測クラス) を示す。テスト画像に対して誤認識した画像枚数は、表 1 により VGG19.2A を用いた場合は 6 枚、表 2 により VGG19.2B を用いた場合は 4 枚となった。

Table 1: Confusion matrix classified by VGG19_2A (row: true labels, column: predicted labels).

	Normal	Anomaly
Normal	4033	4
Anomaly	2	17

Table 2: Confusion matrix classified by VGG19_2B (row: true labels, column: predicted labels).

	Normal	Anomaly
Normal	4031	0
Anomaly	4	21

3. 比較結果



GG19.2A と VGG19.2B の性能を比較するために、前章までの分類結果をもとに認識率、精度、再現性および F 値の 4 つの基準で評価を行った。

認識率はテスト画像の全データのうち正しく分類された割合を、精度は不良品と分類された画像のうち実際に不良品である割合を、再現性は実際の不良品データのうち正しく不良品と分類された割合を、さらに F 値は精度と再現性の調和平均をそれぞれ表している。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分類したものを F_N 、さらに良品を不良品として分類したものを F_P としたとき、認識率 A_c 、精度 P_r 、再現性 R_e 、F 値 F はそ

