

# 畳み込みニューラルネットワーク(CNN)・畳み込みオートエンコーダ(CAE)・サポートベクタマシン(SVM)のための設計支援ツールの開発

## Development of Design and Evaluation Tool for CNN, CAE and SVM

山口東京理科大学大学院 永田寅臣

**Abstract:** 製造業が抱える目視による品質検査の自動化に関する課題を解決するために、プログラム開発や画像処理の初心者であっても高性能な不良品検出用 AI を構築できるように CNN(Convolutional Neural Network), SVM(Support Vector Machine), CAE(Convolutional Auto Encoder)などの設計, 訓練および評価を効率的に支援できるアプリケーションの開発を行っているので紹介する。

### はじめに

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。最近では深層学習の技術を画像認識に特化させた CNN(Convolutional Neural Network)を製品の欠陥検出に応用しようとする試みが活発に行われている。筆者らは製造業が抱える目視による品質検査の自動化に関する課題を解決するために、プログラム開発や画像処理の初心者であっても高性能な不良品検出用 AI を構築できるように CNN, CAE(Convolutional Auto Encoder), SVM(Support Vector Machine)などの設計, 訓練および評価を効率的に支援できるアプリケーションの開発を行っている<sup>1)</sup>。これにより、CNN, SVM といった分類器に対して転移学習、追加学習などの学習アルゴリズムを適用し、静止画や動画に含まれる工業製品の欠陥を検出するためのシステム開発の支援が可能になってきた。実際、地域のものづくりの現場では作業者が環境設定から運用までを容易に行えるシステムへのニーズが高いため、試用を通じた普及と利用者からのフィードバックに基づく機能改善を図っている。開発環境は MATLAB であり, Neural Network Toolbox, Parallel Computing Toolbox for GPU, Deep Learning Toolbox, Statistics and Machine Learning Toolbox<sup>2)</sup>などを用いている。

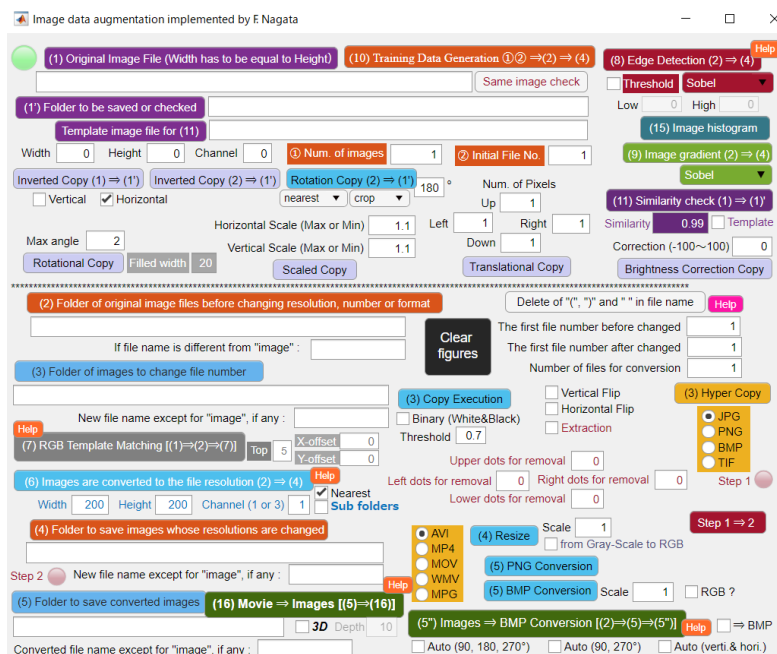
### 開発内容

CNN・CAE・SVM 設計支援ツールは、画像オーギュメンテーション（データ拡張）ダイアログ、メインダイアログ、パラメータダイアログ、オプションダイアログなどから構成される。現場ユーザへの分かり易い操作インタフェースとともに提案するために、様々な対話型コンポーネントが提供された開発環境である App Designer を用いている。

### 画像の編集と拡張用ダイアログ

図 1 には訓練用画像を効率的に生成させるために開発した画像オーギュメンテーションのため

のダイアログを示す。製造現場では例えば、図 2 のような樹脂成形品の不良品の画像の収集に苦慮するケースが多いようであるが、限られた数のオリジナルの良品画像と不良品画像をもとに生成させたい枚数を入力すれば、ワンクリック操作で多数の訓練用画像を生成できる機能を持たせている。左右回転、上下左右反転、拡大縮小、明暗変化、トリミングなどの変化を与えながら、PNG, JPG, BMP, TIF など指定した画像フォーマットと解像度で自動生成させる機能である。このほか、全ての訓練画像に対して一括でテンプレートマッチングを適用し、欠陥の特徴が現われやすい領域のみを抽出する機能や、微分係数にソーベル近似、プレウィット近似、ロバーツ近似などを適用したエッジ検出で輪郭情報を抽出する機能により、画像サイズに関する計算コストを大幅に軽減しながら重要な特徴を含む新たな訓練画像を効率的に生成できるようにしている。これまでの経験から、訓練用の画像データの内容が訓練後の CNN や SVM の分類性能及び汎化性能に大きく影響することがわかっている。

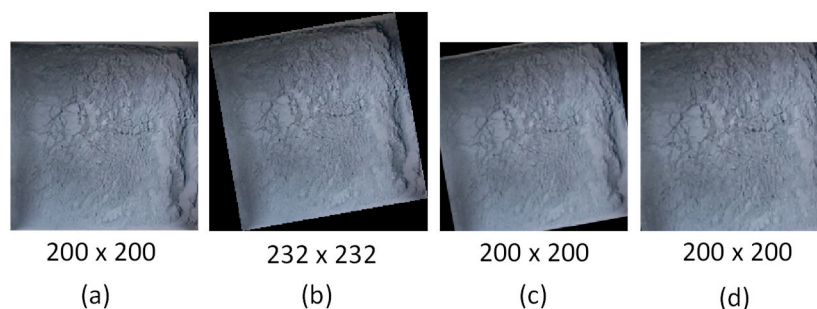


第 1 図 画像オーギュメンテーションダイアログ



第 2 図 樹脂成形品に発生する 6 種類の欠陥の例

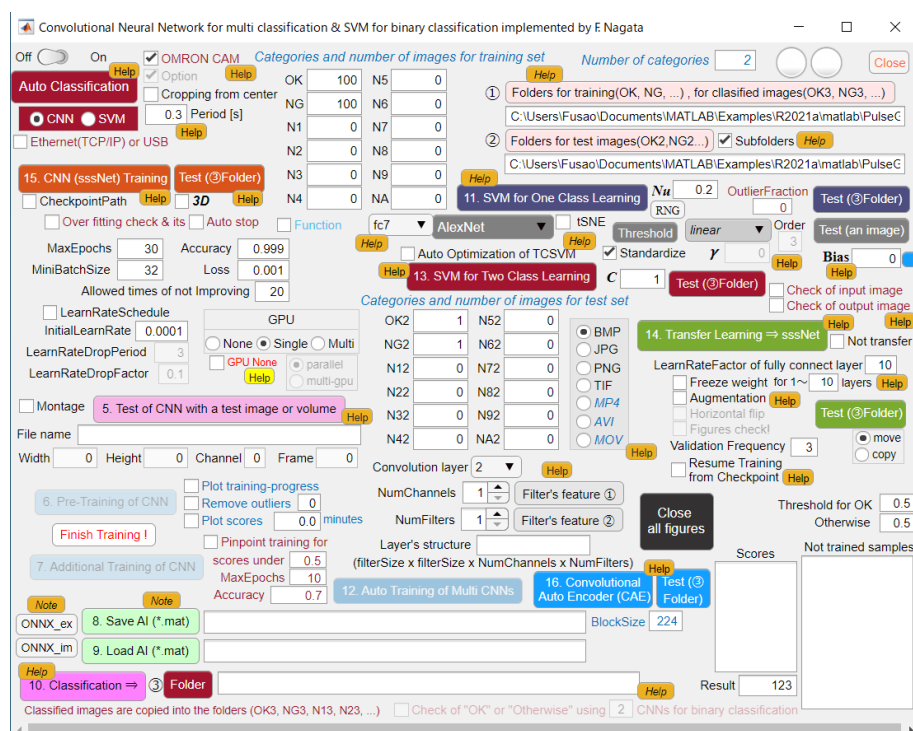
工業製品によっては不良品の発生頻度が低いにもかかわらず、その視覚的な種類の違いが多岐にわたるものがあり、限られた枚数のオリジナル画像からそこに含まれる欠陥の特徴を損なわず拡張させる機能に対するニーズが高い。図 3 にはそのようなニーズに対応して作成した回転による拡張機能を紹介する。(a)は工業材料のオリジナル画像であり、その解像度は  $200 \times 200$  である。(b)は imrotate 関数により反時計回りに  $10^\circ$  回転させた画像であるが、ワーク部分の解像度を維持させたままの回転に伴い画像そのものの解像度が大きくなっているだけでなく、周囲には単色による埋め込みが行われており、これにより無駄な特徴が加わってしまうことになる。(c)は(b)の中心から  $200 \times 200$  のサイズでクロッピングした画像であるが、単色の埋め込み部分はまだ残っており、訓練データとして加える場合には問題となる。このような画像回転による拡張を行った場合の問題点を解決するために、(d)には 115% に拡大後にオリジナルサイズである  $200 \times 200$  でクロッピングした画像を示す。画像拡張ダイアログには画像の回転角度に応じて徐々にサイズを拡大させながら指定サイズでクロッピングできる機能などを持たせている。



第 3 図 工業材料の画像拡張の例

## CNN, CAE 及び SVM 構築用のメインダイアログ

図 4 に示す設計ツールのメインダイアログではシリーズ型の CNN の設計と学習に加えて、有向非循環グラフ(Directed Acyclic Graph: DAG)ネットワークを含む既に公開済みの CNN の後半部にある全結合層部を新たな目的のクラス分類用(例えば、OK と NG)に置き換え、転移学習による新たな CNN の設計などを可能にしている。転移学習時、転移元の CNN の畳み込み層の重み調整は学習率を全結合部分のそれと比べて 10 分の 1 程度に小さく設定することで、値があまり変化しないように進めることが重要である。

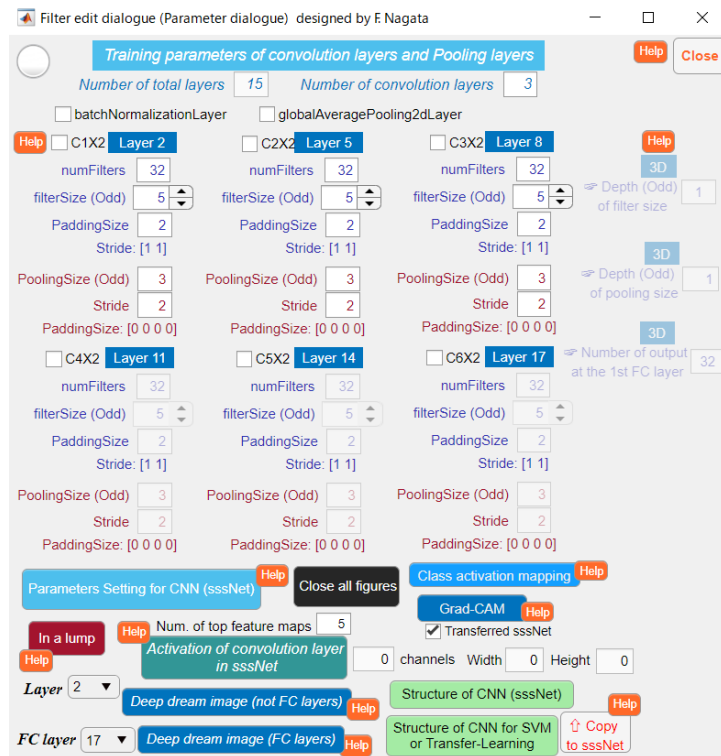


第 4 図 CNN と SVM 構築用のメインダイアログ

オリジナルな CNN の設計時には図 5 に示すパラメータ設定用のサブダイアログを表示させて、畳み込み層、ReLU 層、Pooling 層などを含む層数、各畳み込み層のフィルタ数とサイズ、パディング、ストライド、Pooling 層の各パラメータなどの細かな構造に加えて、バッチ正規化層、Global Average Pooling 層の使用の有無を設定できる。CNN 単体の学習では、分類するカテゴリ数、各カテゴリの訓練用及びテスト用画像のあるフォルダ、学習率、ミニバッチサイズ、マックスエポック、目標の認識精度などの設定や、初期化されたフィルタの重みで学習を始める事前学習、既に訓練されている重みを使って学習を再開させる追加学習を選択できるなどの機能を持つ。このほか、MP4 や AVI などの動画に含まれる連続した複数のフレーム画像から構成する 3D ボリュームを訓練データセットに利用できる 3D CNN の設計が可能になっている。

また、バイナリクラスの分類を行うことができる SVM の学習では、様々なシステムでその性

能が認知されているシリーズ型ネットワークの AlexNet, VGG16, VGG19, DAG ネットワークである GoogleNet, Inception-v3, IncResNetv2、さらに最近注目されている NASNet large (Neural Architecture Search Net large), NASNet mobile に加えてオリジナルで設計した CNN(sssNet)を特徴抽出器として選択できる機能に加えて、良品 1 クラスのみの画像データで学習を行う教師なし学習、良品と不良品の 2 クラスの画像データで学習を行う教師あり学習などを選択できるようにしている。NASNet が従来のニューラルネットワーク設計と違う点は、NAS はニューラルネットワークのアーキテクチャ自体を最適化できる点にある<sup>3)</sup>。

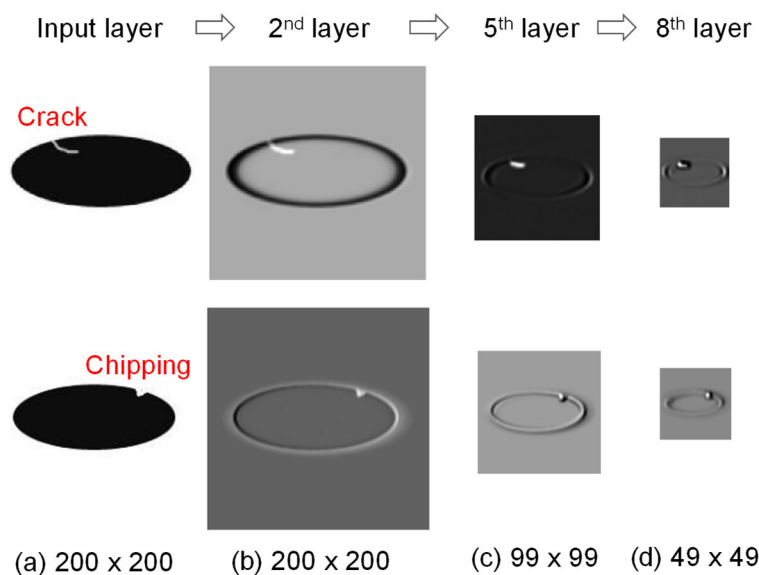


第 5 図 パラメータ設定用のサブダイアログ

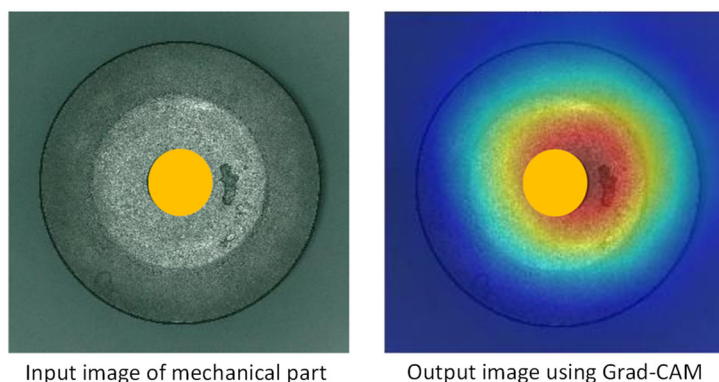
さらに、評価したい画像が CNN に入力された際、図 6 のように指定された畳み込み層で最も活性化した特徴マップとそのマップ内で最大活性化値を持つ領域をより明るいグレー色で視覚化できるようにしている。これにより、入力画像から最終的に決定されたラベルに至ったその根拠となった特徴部分を確認できるようになった。この他、勾配加重クラス活性化マッピング (Gradient weighted Class Activation Mapping: Grad-CAM)<sup>4)</sup>により、分類結果に影響を及ぼした領域を可視化できる機能も持たせている。例として、図 7 には欠陥が含まれた機械部品が CNN により不良品と判定されたときに Grad-CAM により生成されたマップを示している。

一般的な欠陥検出は不良品を分別する 2 クラス分類に帰着できることから、(1)オリジナルでシンプルな 2 クラス分類用の CNN を設計し、実際の良品と不良品の画像（データ拡張を含む）を用いて訓練する方法、(2)AlexNet や VGG19 など既に学習済みの優れた CNN の転移学習により、2 クラス分類用の CNN を新たに設計する方法、(3) (1)の CNN を特徴抽出器として用いた 2 クラス分類用の SVM を設計する方法、(4)AlexNet や VGG19 など既に学習済みの優れた CNN を

特徴抽出器として用いた 2 クラス分類用の SVM を設計する方法などが選択肢となるため、今後も比較評価を進めて行きたい。なお、(2)(3)(4)の場合もデータ拡張を含む実際の良品と不良品の画像をもとに訓練することとなる。



第 6 図 各畳み込み層における最も活性化している特徴マップの例



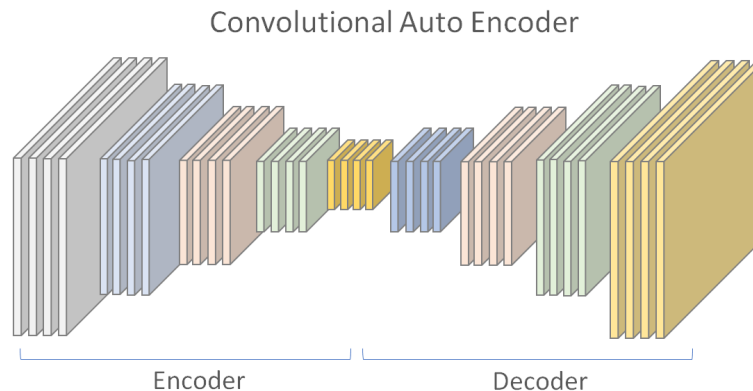
第 7 図 Grad-CAM による分類結果に影響を及ぼした領域の可視化

### 畳み込みオートエンコーダ(CAE)による欠陥検出

メインダイアログでは畳み込みオートエンコーダ(Auto Encoder: AE)の設計も可能になっている。一般的な 3 層あるいは 5 層程度の構造を持つオートエンコーダの場合、入力層と出力層は同じ数のニューロンを持ち、隠れ層（中間層）はそれより少ない次元のニューロンを持つ構造となる。このネットワークの学習では入力パターンがそのまま出力されるように重みが調整されるため、中間層では入力データが持つ有効な特徴が抽出されることとなる。ここではこのオートエンコーダのコンセプトを CNN に応用した図 8 のような畳み込みオートエンコーダ(Convolutional



Auto Encoder : CAE) の設計機能について述べる。



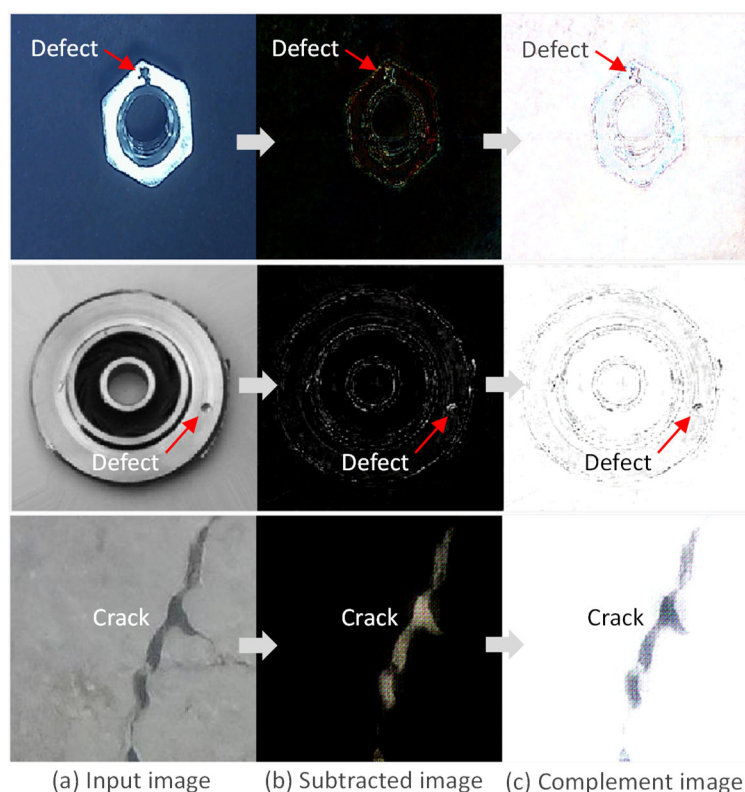
第 8 図 畳み込みオートエンコーダ(CAE)の設計

良品のみの画像を訓練データとして用いて CAE を学習させると、テスト画像 A が良品画像の場合にはそれに非常に似た画像 A' を出力できるようになるため、差分の絶対値  $|A - A'|$  を求めて表示させると全てのピクセル要素がほぼ 0 となり、黒色単色に近い画像として得られる。これに対して、キズやクラックなどの欠陥が含まれた不良品の画像 B が入力された場合、CAE は訓練データにはなかった欠陥部分の特徴を再構成できず、結果的に出力画像 B' には欠陥部分を正確に再現できない部分が現われることとなる。従って、 $|B - B'|$  を表示させると再現できなかった欠陥部分が黒より明るい傾向の色で表示されることとなる。

例として VGG16 ベースのセマンティックセグメンテーションネットワークである SegNet を取り上げ、それをもとにした CAE の設計方法について紹介する。まず、SegNet では、図 8 のように VGG16 の後部にある全結合層の部分が削除され、代わりに前部の畳み込み層と対称構造をもつアップサンプリング用の転置畳み込み層を接続する。Pooling 層を持つ手前側をエンコーダ、アップサンプリングを行う層を持つ後半側がデコーダである。SegNet では入力画像に対してピクセル毎のクラスラベルを出力できるようになっているため、最終の 90 層と 91 層にあるソフトマックスとピクセル分類レイヤをそれぞれ削除した後、CAE では入力画像を出力側の教師データとして訓練させるために新たに回帰出力層を加える構造となっている。

ここでは公開されているナット、ブレーキロータおよびコンクリート表面の良品画像のデータセット<sup>6,7,8)</sup>をそれぞれ用いて学習させた 3 つの CAE による欠陥部分の可視化実験の結果を紹介する。図 9(a) はキズあるいはクラックなどの欠陥が含まれた入力画像であり、図 9(b) には再現できなかった欠陥部分を黒以外の明るい色で表示させた  $|B - B'|$  を示す。さらに画像を構成する全ピクセルの値を補数で置き換えると図 9(c) のように欠陥部分の特徴が含まれた領域を視覚化できることがわかる。今回はナット、ブレーキロータおよびコンクリート表面の良品画像の枚数が、100 枚あるいは 1000 枚程度であったがより多くの訓練画像を用いることで欠陥部分の可視化時の鮮明さを高めることができるものと期待される。

現在開発中の CNN, CAE, SVM 設計ツールを使って CAE を設計する場合、VGG16 あるいは VGG19 をベースとした SegNet の出力部分をピクセル毎のクラス名の出力から、ピクセル毎の再帰出力(ピクセル値を出力)へと置き換える方法と、オリジナルで設計した CNN (sssNet) をベースとして新たな CAE を設計する方法が利用可能である。sssNet の場合は MATLAB 2021a からのサポートされている 2 次元転置畳み込み層(transposedConv2dLayer)を使って特徴マップをアップサンプリングさせることで CAE としての設計を可能にしている。このほか、学習後の CAE においては入力画像と出力画像の差分を定量的に解析することで CNN, SVM と同様に欠陥検出器としての利用も可能となる。



第 9 図 CAE を用いた欠陥部分の可視化実験

## おわりに

本稿では、プログラム開発や画像処理の初心者であっても高性能な不良品検出用 AI を設計できるように現在開発中の CNN, CAE 及び SVM のための設計支援ツールを紹介した。現在、このアプリケーション内に操作方法と基本用語の説明を組み込みながら、山口県内外のものづくり企業数社との共同研究を通して実ラインへの適用研究に取り組んでいる。そこでは本稿で紹介した AI 設計支援ツール、作業座標系とカメラ座標系との間のビジュアルフィードバック制御、複数のロボットの協調制御などのソフトウェア技術に加えて、それぞれの製造工程が対象とするワーク、PLC(Programmable Logic Controller)や産業用ロボットといった機械装置、ピッキングのためのエンドエフェクタなどハードウェアとの統合を考慮したシステム開発が中心となっている。なお、紹介したアプリケーションの利用、読者皆様からの評価実験実施の依頼については永田研究室



(nagata@rs.socu.ac.jp)までいただければ幸いです。

## 参考文献

- 1) 永田, 渡辺, 不良品検出のための畳み込みニューラルネットワークとサポートベクタマシン設計支援ツール, システム/制御/情報, Vol. 64, No. 8, pp. 304–309, 2020.
- 2) <https://jp.mathworks.com/>
- 3) B. Zoph, V. Vasudevan, J. Shlens, Q.V. Le, “Learning Transferable Architectures for Scalable Image Recognition,” arXiv preprint arXiv:1707.07012v4, 2018.
- 4) R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra. “Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization,” *Proc. of the IEEE Conference on International Conference on Computer Vision (ICCV)*, pp. 618–626, 2017.
- 5) V. Badrinarayanan, A. Kendall and R. Cipolla, “SegNet: A Deep Convolutional Encoder-decoder Architecture for Image Segmentation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 39, No. 12, pp. 2481–2495, 2017.
- 6) <https://www.mathworks.com/matlabcentral/fileexchange/64068>
- 7) <https://www.kaggle.com/ravirajsinh45/real-life-industrial-dataset-of-casting-product>
- 8) <https://www.kaggle.com/arunrk7/surface-crack-detection>