AI 技術による外観検査の自動化

Deep learning による事前学習型欠陥検出について

半澤　雄希・池田　泰之・栗田　真嗣・長谷川　友紀

本稿では外観検査自動化システムにおいて、一様背景（ヘアライン、梨地、等）ワークにおける多種多様な欠陥を検出する技術を提案する。現在、ものづくり領域において、「人材不足」および「嗜好の多様化」が進み、目視検査工程の自動化ニーズが高まっているが、既存の画像センサでは多種多様なワーク・欠陥種に対する自動化は実現できていない。そこで我々はDeep Learning技術を活用し、多種多様な欠陥状態を事前に学習することで、人のような検査を複雑な設定なく誰でも自動で実行することが可能な欠陥検出技術を構築した。また、GPUのような高コストな装置を導入することなく、CPUで高速動作可能な工夫を導入することで、既存の画像センサに搭載可能にした。これにより、一様背景ワークの外観検査自動化が可能となる。

AI Visual Inspection System

Defect Detection using Deep Learning

Yuki Hanzawa, Yasuyuki Ikeda, Masashi Kurita and Yuki Hasegawa

In this paper, we propose a technique to detect various defects in uniform background objects (hairline, satin, etc.) in visual inspection system. "Human resources shortage" and " diversification of consumer needs" have emerged recently in the field of manufacturing, and the demands for the autonomous visual inspection system are increasing. The existing inspection system, however, cannot adapt to a wide variety of objects / defect types. We built Deep Learning model that learns a variety of defect types in advance, thereby constructing a defect detection system that enables anyone to automatically perform inspection like human eye without any complicated setup. In addition, without introducing high-cost computing resources such as GPU, it can be feasible on existing vision system by highly optimizing for modern CPU.

1. まえがき

1.1　背景　ものづくり領域において、「人材不足」および「嗜好の多様化」が進み、人による目視検査工程の自動化ニーズが高まっている。しかし、既存の画像センサでは検査工程の自動化は一部のみしか実現できていない。その要因は、「つくるものの多様化により、検査ワークの材質や形状の多様性に対応しきれない」「高度な専門知識を保有しなければ調整できない」などの課題がある。そこで、我々は検査工程を自動化するために、「多種多様なワーク・欠陥種に対応可能」かつ「誰でも簡単に設定可能」という要件を満たす外観検査自動化技術の実現を目指す。

多種多様なワーク・欠陥種に対応するためには、欠陥を検出するアルゴリズム、欠陥を強調する入力（照明・撮像）技術、ワーク形状に対応するための駆動技術が必要であるが、その中で、本稿では、多様な欠陥を検出可能にする欠陥検出アルゴリズムを提案する。

1.2　Deep Learningを活用した事前学習型欠陥検出の提案

多種多様な画像解析においてDeep Learningが成果を上げている中、ネックとなっているのは、画像の収集である。外観検査でのDeep Learningの実用化にあたり、学習に使用する画像の収集が、現場の作業者にとって大きな負担となり、製品ラインの立ち上げ時に十分な枚数の学習画像を確保することは困難であることが想定される。本稿ではこの課題を解決するために、ラインごとに学習画像を準備する必要のない、事前学習型のアルゴリズムによる外観検査の自動化を提案する。

外観検査においては多種多様なワーク表面・欠陥状態の微小な差から欠陥のみを検出する必要性がある。外観検査の対象ワーク・欠陥状態は表1のように分類できる。この分類の中で、既存の画像センサで実用化がなされているのは、一様ワーク上の顕著な欠陥の検査のみである。それに対して、本稿で提案するアルゴリズム、一様ワーク上の微小欠陥を対象としている。一様ワークであれば異なる製品ライン間でも似た特徴を示すため、事前学習型のアルゴリズムでも十分に対応可能である。

連絡先：半澤　雄希　yuki.hanzawa@omron.com

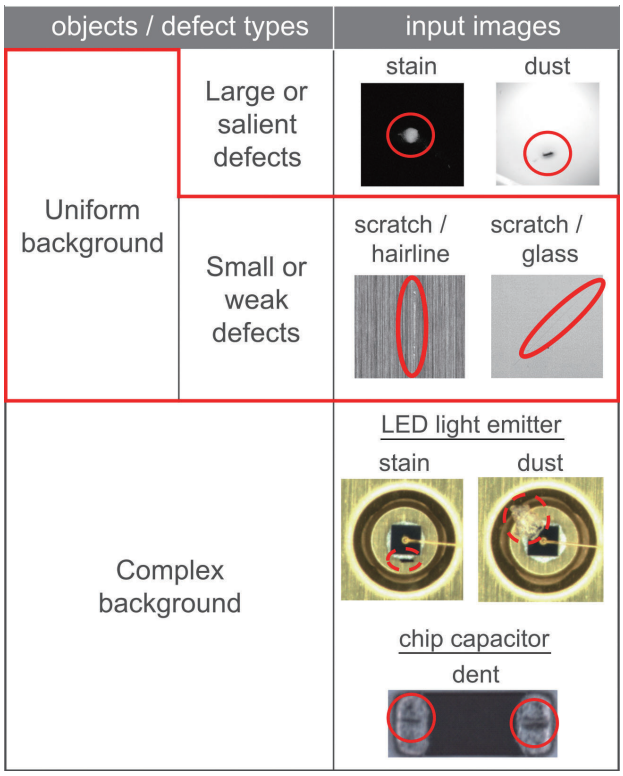


表1　表1 ワーク･欠陥状態の分類

2. システム概要

検出処理の概要を図1に示す。入力画像を事前学習型のCNN（Convolutional Neural Network）によって処理することで、欠陥らしさを表すFeature Mapを生成する。そして、得られたFeature Mapを2値化し、欠陥領域として抽出する。

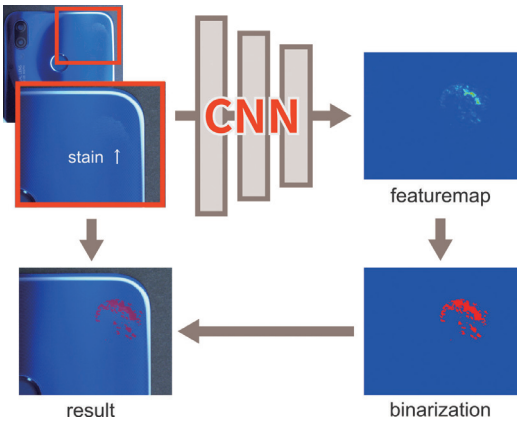


図1　検査システム概要

このようなタスクを行う場合のDeep Leaningの研究事例としては、一般にはObject detectionや、Semantic segmentationといった処理結果を出力するアルゴリズムがベースとなる[1]ことが多いのに対して、本稿で提案するアルゴリズムでは、Feature Mapを最終出力としている。この理由は、実際の現場においては、生産ラインごとに製品の欠陥を不良品とするか良品として許容するかの判定基準が異なり、ラインごとの閾値を設定する余地を残す必要があるためである。実際の運用形態としては、画像センサに搭載されたうえで、提案アルゴリズムによって欠陥部分のみを強調したFeature Mapを生成した後に、単純な画像処理の2値化やラベリングを適用し、欠陥の位置や大きさを用いて良否判定を行う、といった一連の検査フローの一部として活用されることを想定している。

3. 欠陥検出画像の生成アルゴリズム

製造現場での欠陥検査では、欠陥以外のものが画像に映りこむことも考えられ、後段処理で欠陥とそれ以外の分類が必要な場合がある。そのために、「欠陥らしさ」を画像化したFeature Mapにより、位置や大きさを特定できるようにする。Feature Mapの作成は下記2ステップから成り立つ（図2）。

（1）検査画像に欠陥が含まれる確率を推定
（2）推定した欠陥の位置を特定し、画像化

まず、（1）では検査画像をCNNに入力し、画像内に欠陥が存在する可能性を0～1の確率で出力する。CNNは事前に大量の欠陥画像を用いて学習しておくことで、検査画像に欠陥に近いパターンが含まれると高い値を出力するようになっている。

次に、（2）では（1）が推定した欠陥確率が、検査画像内のどの場所に由来するものかを求める。CNNでは欠陥位置を表す情報は、各中間層の計算結果に含まれることが知られており[2]、これらを用いて、欠陥確率への寄与度を検査画像の画素単位で算出できる。最後に、画素毎の寄与度に適切な倍率をかけてFeature Mapとする。

Feature Mapの例を図3に示す。この図では欠陥確率への寄与度が高いほど色が青から赤に変化する。検査画像中で欠陥が存在する箇所が高い値になっていることがわかる。

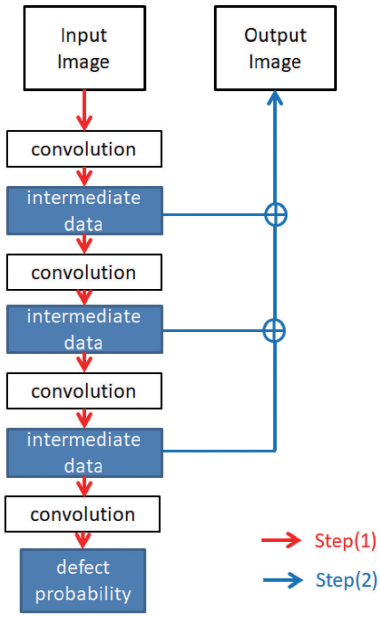


図2　Feature Mapの生成フロー

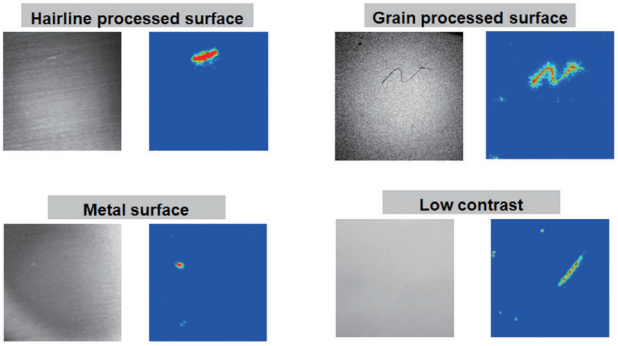


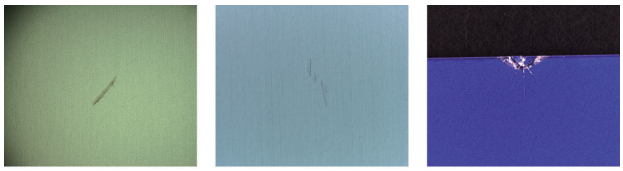
図3　Feature Mapの例

4. 学習画像DB

事前学習型のアルゴリズムで十分な性能を発揮するには、現場で生産される多種多様なワークを網羅する、大規模な学習画像DBを開発時に構築する必要がある。しかし、実ラインのデータは常に保存されているわけではなく、企業秘である場合もあるため、容易には開発用途での収集はできない。特に欠陥を含む画像データは絶対数が少なく、良品とのデータ数のバランスがとれないという問題もある。

学習に使用可能な画像の調達が困難な条件においてのDB構築方法として、擬似的にCG（Computer Graphics）やGAN（Generative Adversarial Network）で生成した画像を学習画像として使用する手法が提案されている[3) 4)]が、実環境で実際に有効性が示されるケースは稀である。

そこで、本稿で提案するアルゴリズムの開発においては、実際の現場で発生が想定される欠陥種別、位置、サイズ、色、背景素材、光源設定等の組み合わせを網羅したパターンのワークを実際に作成、撮影し、学習画像DBを構築した。図4は実際に作成した組み合わせパターンの例である。



(a) metal (matte) + stain　(b) metal (hairline) + scratch　(c) glass + crack

図4　学習画像の例

学習処理においては、撮影した各画像に対してクリッピングやノイズ付加等のオーギュメンテーション処理を加え、800万枚に増強した画像を使用している。

この学習DBを使用することにより、本稿で提案するアルゴリズムは、現場で多種多様なワークや欠陥形状に対応することが出来る。

5. 高速化の検討

一般にDeep Learningを実行するにはGPUなど潤沢な計算資源を用いるが、現場で使う画像センサの場合、コスト等の問題によりこれらの活用が困難である。そこでCPUのみの処理にて高速性を改善するために、CNNの処理時間の多くを占める畳み込み層に着目し、ネットワーク構造とコードの実装形態を画像センサのHW構成に最適化している。これにより、入力画像の解像度次第で、100ms ～ 600msの処理時間を実現している。

5.1　ネットワーク構造の最適化　高精度なDeep Leaningのネットワーク自体は数多く提案されているものの、GPUの使用を前提としたものが多く、画像センサのCPUでの処理時間は1000msを超えてしまうものがほとんどである。本稿で提案するアルゴリズムで採用しているネットワークは、一般物体認識のタスクにおいて高い精度で高速なネットワークとして代表的なResNet[5)]や、Inception[6)]といったCNNのネットワークをベースとしつつ、ネットワークの各レイヤーを、多種多様な欠陥、背景パターンの組合せに対応することを考慮して、Effective Receptive Fieldの多様性を重視した構成と[7) 8)]し、さらに、様々な高速化構造[9)]を取り入れることで、速度と精度の両立を図った。

5.2　カーネルの近似　畳み込み演算は、カーネルのサイズに比例して計算量が増加するため、カーネルは小さいほうが高速である。一方、CNNでは小さなカーネルの畳み込み層を多段で使うことで、大きなカーネルの畳み込み層と同等の効果があることが知られている[10)]。そこで我々はカーネルサイズNxNxCの畳み込み層を1xNxCおよびNx1xCの2段の畳み込み層に分けることで、全体の計算時間を短縮する。（図5）

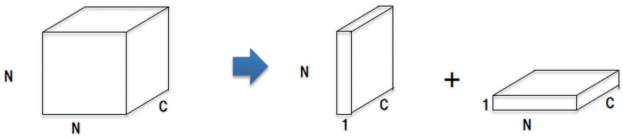


図5　畳み込み層のカーネルの小型化

5.3　固定小数点化と並列演算命令の活用　畳み込みは浮動小数点のデータ列に対する積和演算である。一方、画像センサで採用するCPUでは、複数のデータに対する積和演算を並列で実行する命令セット（Single Instruction Multiple Data：SIMD）が使用できる。また、SIMDではデータを表現するbit数が少ないほど多の並列実行ができる。我々は畳み込みの計算において、入力と出力を8bit固定小数点化することで、最大32並列で演算実行するようにする。

6. 性能評価

本稿で提案したアルゴリズムを、466枚の検査画像（良品62枚、不良品404枚）に対して適用し、性能を評価した。比較対象とした従来手法の精度は、画像センサに搭載されている、コントラスト強調や、エッジ検出といった一般的なフィルタをワークごとに最適化し、欠陥領域を抽出することで算出している。また、評価に使用した画像は、本稿

で採用した学習画像とは撮影環境も撮影対象も異なる画像である。

性能評価の結果を表2に示す。表中の見すぎは良品画像からの誤検出、見逃しは不良品画像に対する未検出である。

表2　性能評価結果

	見逃し	見すぎ
従来手法	3.2%	6.7%
提案手法	0.9%	3.4%

見逃し、見すぎともに、提案手法が高い精度を示した。また、従来手法がワークごとに複数のフィルタの選択と多数のパラメータの調整を要するのに対し、提案手法を使用する上で調整が必要なパラメータはFeature Mapに対する閾値のみであるため、現場での調整の手間の削減も期待できる。

図6、7に提案アルゴリズムによる欠陥部位の抽出結果の例を示す。図の1段目は入力画像で、2段目は比較用に画像センサに実装されている従来手法によって入力画像から欠陥部位の抽出を行った結果である。そして、3段目が提案アルゴリズムによる欠陥部位の抽出結果である。図6に示したワークは学習画像に類似したパターンの画像を含む場合の結果で、図6（a）は梨地のアルミにキズが付いているワークである。入力画像の解像度が約2000 × 2000に対してキズの幅は約4pixと細く、かつコントラスト比も低いため、従来手法によって処理を行った場合は全く抽出出来ていない。これに対して、提案アルゴリズムによって処理した結果では、従来手法では不可能だったキズ部位の抽出が出来ており、かつ、キズ部位と隣接した画像上ではより顕著なコントラスト比を示す照明の影を無視できている。図6（b）は汚れのついたプラスチックフィルムである。フィルムには細かなヘアラインが入っており、従来手法によって処理した場合、それらに起因するノイズがワークの広範囲に発生している。それに対し、提案アルゴリズムによって処理した結果では、ヘアラインの影響を受けずによりクリアに欠陥部位の抽出が出来ている。

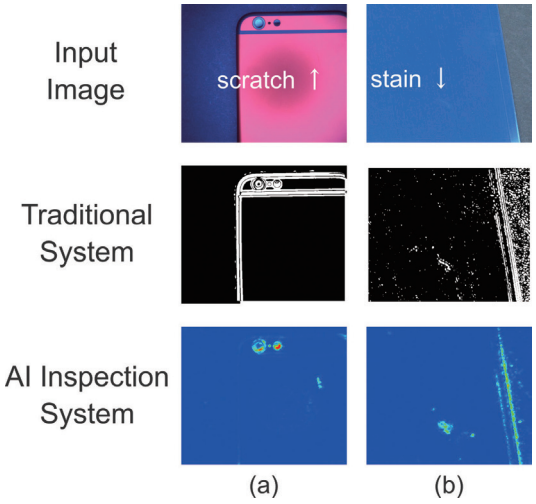


図6　学習画像と類似したワーク

図7は学習画像との差異の大きく、かつ背景部分の情報量が多いワークに対する出力結果である。図7（a）はフェライトコアにクラックが発生している画像で、学習画像にはこのような形状の情報が含まれる画像は存在しない。この画像に対して従来手法によって処理を行った場合、クラック部分の抽出こそ行えているものフェライトコアそのもの形状や表面荒れに起因するエッジも同様に抽出されている。対して、提案アルゴリズムによって処理した結果では、クラック部分のみが顕著な欠陥部位として出力されている。図7（b）は木製のスマートフォンカバーにホコリが乗っている画像である。学習画像には、木製のワークの画像はなく、ホコリを欠陥として含むパターンも今回の開発におけるターゲットとしていなかったため、存在しない。このワークに対して従来手法によって処理を行った場合、ホコリだけでなく、木目の部分も同等の強度で抽出されてしまっている。これに対して、提案アルゴリズムによって処理した結果では、木目パターンもホコリも学習していないものの、木目を無視してホコリ部分が欠陥として出力できている。

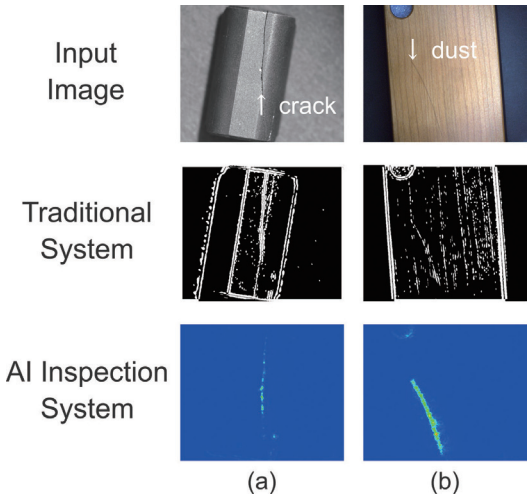


図7　学習画像と差異の大きいワーク

これらの結果から、本稿で提案するアルゴリズムは、事前学習型でありながら、学習画像に存在しない、未知パターンに対して高い対応能力も持つといえる。

7. むすび

本稿では、多種多様なワーク、欠陥種別に対応可能な事前学習型の欠陥検出アルゴリズムを提案した。提案手法により、未知のパターンのワーク、欠陥に対しても対応できることが確認できた。

今後の展望としては、より複雑な入力（照明・撮像）技術、ロボットの駆動技術との連携や、より複雑なデザインのワークに対応するための、オンライン追加学習を検討している。

参考文献

1) Hiroya Maeda; Yoshihide Sekimoto; Toshikazu Seto; Takehiro Kashiyama; Hiroshi Omata. Road Damage Detection Using Deep Neural Networks with Images Captured Through a Smartphone. arXiv:1801.09454
2) B Zhou; A Khosla; A Lapedriza; A Oliva; A Torralba. Learning Deep Features for Discriminative Localization. CVPR (2016)
3) César Roberto de Souza; Adrien Gaidon; Yohann Cabon; Antonio Manuel López Peña. Procedural Generation of Videos to Train Deep Action Recognition Networks. CVPR (2017)
4) Ashish Shrivastava; Tomas Pfister; Oncel Tuzel; Josh Susskind; Wenda Wang; Russ Webb. Learning from Simulated and Unsupervised Images through Adversarial Training. CVPR (2017)
5) Kaiming He; Xiangyu Zhang; Shaoqing Ren; Jian Sun. Deep Residual Learning for Image Recognition. CVPR (2016)
6) C. Szegedy; V. Vanhoucke; S. Ioffe; J. Shlens; Z. Wojna. Rethinking the Inception Architecture for Computer Vision. CVPR (2016)
7) Wei Xiang; Dong-Qing Zhang; Heather Yu; Vassilis Athitsos. Context-Aware Single-Shot Detector. arXiv:1707.08682
8) Songtao Liu; Di Huang; Yunhong Wang. Receptive Field Block Net for Accurate and Fast Object Detection.ECCV (2018)
9) Mark Sandler; Andrew Howard; Menglong Zhu; Andrey Zhmoginov; Liang-Chieh Chen. MobileNetV2: Inverted Residuals and Linear Bottlenecks. arXiv:1801.04381
10) K Simonyan; A Zisserman. Very Deep Convolutional Networks for Large-scale Image Recognition. ICLR (2015)

執筆者紹介

半澤　雄希　Yuki Hanzawa
技術・知財本部　センシング研究開発センタ
専門：画像処理

池田　泰之　Yasuyuki Ikeda
インダストリアルオートメーションビジネスカンパニー　商品事業本部
センサ事業部　第2開発部
専門：画像処理
所属学会：情報処理学会

栗田　真嗣　Masashi Kurita
技術・知財本部　センシング研究開発センタ
専門：画像処理

長谷川　友紀　Yuki Hasegawa
技術・知財本部　センシング研究開発センタ
専門：画像処理
所属学会：精密工学会