

検査工程における水栓ラベル認識システムの構築

永見 英臣^{*1}, 水谷 真也^{*2}, 森田 亮介^{*3}, 伊藤 聡^{*4}

Construction of a label recognition system on water valves in the factory inspect process

Hideomi NAGAMI^{*1}, Shinya MIZUTANI^{*2}, Ryosuke MORITA^{*3} and Satoshi ITO^{*4}^{*1} Graduate school of Natural Science and Technology, Gifu University, 1-1 Yanagido, Gifu-shi, Gifu 501-1193, Japan^{*2} Mizutani Valve Company limited. 194 Tominaga, Yamagata-shi, Gifu 501-2257, Japan^{*3,*4} Faculty of Engineering, Gifu University, 1-1 Yanagido, Gifu-shi, Gifu 501-1193, Japan

Received: 11 April 2019; Revised: 2 August 2019; Accepted: 8 October 2019

Abstract

This paper presents a label recognition system aiming at being introduced to the inspection process in the water valve factory. This inspection process checks whether all the labels are pasted to the right place on the products. The numbering management that has been adopted in the current factory lines can help us with finding the missing or double pasting of the labels, but it never notifies the wrong-place pasting. Furthermore, this requires much time to prepare the same number of the parts including the labels. To solve these problems, we here introduce an image-based recognition method utilizing the deep-learning. Adopting the object recognition algorithm, not only the existence but also the pasted place of five pre-learned labels are inspected using several numbers of the frames within 1 second. The learning from about 3,000 sample data resulted in more than 95% correction rate in three of five labels though it fell nearly 80% for the others.

Keywords : Machine learning, Image recognition, Object detection, Image processing, Factory automation, Productivity improvement

1. 緒 言

工場の生産現場で安全性の次に常に求められていることの一つに生産性の向上がある。長年勤務した労働者は比較的高い技術力を持っているが、高齢化が進み高い技量をもつ技術者が減少していく。一方、少子高齢化の影響で生産年齢人口そのものが減少しつつある。したがって、技術力のある労働力を安く手に入れることは、今後いっそう困難となる。つまり、技術力を維持したまま労働力をいかに確保していくかが、これからの大きな課題になると考えられる。

生産での技術とは、精緻な作業を素早く正確に行なう職人技ばかりではない。単純な作業でもミスなく効率的に実施することも生産性の向上には重要な技術の一つであり、むしろ代替可能な労働力に対しては、こちらを常に維持できることの方が有用である。

作業をミスのなく効率的に実現するための一つとして、同じ作業を確実に繰り返し実行できる機械を導入した生産の自動化がある。しかし自動化には設備投資が必要で、容易に導入できるわけではない。生産している商品の需要が少ない場合には自動化してもコストに見合わず、機械化導入の見極めも重要である。また自動化した生

産システムの維持・管理にはコストばかりでなく、それなりの知識とノウハウが必要であり、労働力減少に対応できるものの、担当する管理者を相応の教育により育成する必要があるという新たな問題も生じる。

しかしながら機械のもつ高い再現性や、高速性・正確性は、生産現場のニーズに合致するように導入できれば、高い利用価値を生み出す。そこで、本稿では、完全自動化ではなく、「機械の支援のもとで作業を実施することで効率性や信頼性を向上させる」というスタンスをとり、部分的な自動化で生産現場の一改善に取り組む。

対象とするのは、水栓を製作・梱包し、出荷を行なっている生産現場である。水栓は砂型に“湯”と呼ばれる溶解銅を注入して本体を製作する鑄造工程から始まる。鑄造で大枠が作製された本体は、中子取り出し、型ばらしにより取り出され、バリ取りを経て、加工工程に送られる。加工工程ではねじ切りを中心に工作機械による加工がおこなわれる。その後、サンドペーパーおよび布による研磨が行われ、最後に Ni - Cr のメッキが施される。完成した本体は水没検査の後、水管などの他の部品と組み立てられ、通水検査を経て、梱包・出荷となる。

この中で自動化により効果が期待できる箇所として、検査工程の一部である「ラベル検査」に着目した。例えば、湯水混合水栓には「湯水」と「冷水」それぞれ専用の水管があり、住宅・工場での現場で反対に設置されないように、それぞれ赤と青のラベルが水管の先方に貼り付けられる。もし誤りがあると、水と違って熱湯がでるようなケースも想定され、安全面からも重要な表示の一つである。もちろん作業時にラベルの確認をしているが、ヒューマン・エラーを零にすることは難しく、数年に1回は貼り忘れを含めたクレームが届いている。このラベルの貼り付けを機械で確認させ、ヒューマン・エラーをさらに減少させること、そして機械がチェックすることで確認作業への精神的負担を減少させ、安心して作業に取り組んでもらえる作業環境を提供することを本稿のねらいとする。

ラベルの検査には、近年注目されつつある機械学習の導入を試みる。機械学習とは与えられたデータと課題から統計的に最適化を行う手法のひとつである。近年では、大量のデータが扱えるようになったばかりでなく、深層学習技術の発展によって、より複雑な問題にも適用可能となっている。機械学習は画像認識分野において有力な手法のひとつとされており、画像を取り扱う今回の検査システムに導入することにした。画像から得られる情報を活用する研究として、画像から文字情報を読み上げる (Yi et al., 2014) ものや、標識を読み取る (Rong et al., 2016) ものがあり、画像情報から別の情報を取り出すことによって新たな問題解決が可能になる。本稿では画像内の物体位置を利用して製品の検査に活用する。

2. ラベル検査システムの構成

2.1 ラベル検査と定数管理

最初にラベルの貼付け作業についての現状を述べる。水栓が完成すると、付属部品、包装用のビニール、取扱説明書など梱包に必要なものが作業台付近に集められ、作業者が確認しながら手作業で梱包を行っている。ラベルは水栓本体の検査が完了し出荷可能なものだけに貼り付けるため、ラベルの貼り付け作業および確認は、梱包時の作業の一つとして行われている。ここで注意すべきことは、ラベルの確認はその時に行わないと、梱包された後では実施ができないことである。これは梱包されるもの全てについていえることで、もう一度確認するには再度開梱するしかない。

梱包時のミスを取り除くために現場で導入されているのが「定数管理」と呼ばれる手法である。定数管理では事前に梱包時に必要な全ての種類の物品を一定数ずつ正確に用意し、全種の物品が同時になくなることで入れ忘れないことを担保する。もし過不足が生じたときには、用意した個数内の製品でミスが生じているはずである。このため再検査が必要な梱包済みの製品は、定数管理内の範囲に限定できる。

ところが、定数管理は間接的な検査であり実際に確認を行うわけではないため、例えばある箱には2つあってある箱には入っていなかった、といった最終的な数が合致してしまうミスの検出は不可能である。また、ラベルは貼り付け位置が指定されていることが多く、貼り間違いは定数管理で検出することもできない。さらに定数管理では前もって各部品数を整える必要がある。生産は部品数が整わないと開始できないため、実際にはその準備に多くの時間を注いでいる。

したがって、ラベル検査を機械で自動的に実施し貼り付けを確認できれば、信頼性の向上につながるばかりでなく、定数管理の廃止で不要となる準備時間を作業に当てることで生産性の向上につながる。

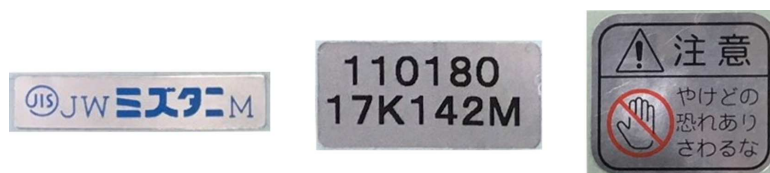


Fig. 1 Three types of silver label.



(a) Whole shape.



(b) Labels at the hose tips.

Fig. 2 Blue and red labels on twin hoses.

2.2 検査対象ラベル

製造されるすべての水栓には、図 1 のように、製造会社名の記載されたラベル (MIZ) と製品管理のためのロット番号が示されたラベル (LOT) が貼り付けられる。水栓には図 2(a) のように湯水・冷水の両方を取り扱えるものもいくつかあり、湯水を扱うものには、熱湯に注意するように促す注意ラベル (WAR) も取り付ける。このような水栓では、内部には湯水・冷水それぞれに対し専用の水管が用意されている。取り付け時に間違いがないよう水管の先方には図 2(b) のように、湯水側には赤のラベル (Red) を、冷水側には青のラベル (Blue) を貼り付ける。本研究では以上の 5 種類のラベルを今回の検査対象とした。

2.3 システムの設計仕様

本稿ではラベルの確認を機械に行わせることで、生産性および製品の信頼性の向上を図る。このラベル検査では画像を用いることとする。ラベル検査ではラベルをセンシングすることが第一である。しかし、ラベルに IC チップを取り付けられるわけでもなく、重量で有無を判別できるほどの質量もない。ラベルを人間は目視で確認できているのであるから、視覚と同じような機能を持つカメラ画像という手法を選択した。またラベルには色が意味をもつことも、色を扱える画像を選択した理由である。

実際に生産現場で利用することを踏まえ、次の 3 点を特に意識した。

- その場で検査結果が判明する (リアルタイム性)
- 検査に必要な手順を増やさない (低負荷性)
- 判別の誤りが少ない (高再現性)

1 つ目の項目に関しては、作業手順の工夫により、一連の梱包作業の 1 つとして本検査処理を組み込むため、現場では 30 秒以内に処理ができれば十分である。2 つ目の項目に関しては、検査部品を検査場所へ配置することは可とするが、撮影箇所を一定とするために治具で固定することは避けることとする。また、3 つ目の項目としては、画像処理を初めて導入したこともあり、最低 80% の再現率を確保する目標で判別器を構成することとする。

これに基づき、構成するシステムの目標動作を以下のように設定した。

1. 検査者はラベルが貼られる予定の箇所を撮影する。もしくは、指定場所に設置することで自動的に撮影が行われる
2. 画像に基づきラベルの有無を検査する。
3. 検査結果を音により検査者に通知する。

検査結果はラベルにより以下のように定義する。図1の銀ラベルに関しては、3つのうち1つでも確認できなければ不可とする。図2の赤・青ラベルに関しては、湯水・冷水の各水管に対応した色のラベルが貼り付けられているか判断し、誤りがあればその誤りを検査者に知らせる。このとき、2本の水管に対しともに青、ともに赤の場合も想定され、その場合には同色が2つと判定することが望ましい。したがって、ラベルが存在することだけでなく、複数あることを検知できる必要もある。

これにより、本研究のラベル検査の仕様として、1. ラベルの有無、2. ラベルの種類、3. 水管とラベルの対応、の3つを判定することと定義する。

2.4 画像の認識

ニューラル・ネットワークを用いた深層学習が、学習の実装法、GPUの計算利用、学習データのデータベース化などいくつかの面でのブレークスルーを経て、その認識率を飛躍的に増大させた(原田, 2017)。深層学習に基づく画像認識ネットワークが、2015年のコンペティションで人間の認識精度を超えたといわれた(中山, 2017)。この性能を発揮した課題はクラス分類であり、1枚の画像に対して1つの認識結果が返されるものである。クラス分類問題では、高精度化のために一つのクラスに多くの画像を用意する。ILSVRCというコンペティションでは1000クラスに対して120万枚の画像を使用している(Russakovsky et al., 2015)。

当初、この認識率の高いクラス分類ネットワークを利用して、ラベル検査システムを構成しようと考えた。すなわち、例えば赤・青ラベルの場合、2つのラベルが正しく貼り付けられた画像と逆に張られた画像をそれぞれ正例・反例として学習させて判別させる方法である。

しかし、このようなクラス分類ネットワークでラベル検査システムの構築を試みたところ次のような問題が生じた。

- ラベルの種類が増加で、その有無の組み合わせが指数的に増加する。
検査するラベルが n 個のとき、その有無の組み合わせは 2^n となる。どのような誤りが起きているかまで知るには全ての場合を学習させる必要があるため、学習に必要なデータ数も指数的に増加する。したがって、学習データの作成により労力が必要となる。
- 赤・青ラベルなど位置関係が必要ラベルには、位置の組み合わせも生じる。
例えば、湯水/冷水の赤・青ラベル2枚に関しても、1枚しかないケース4通り、2枚貼られている場合4通り、何も貼られていない場合1通りの9通り存在し、正解はそのうちの1通りしかない。製品には色ラベルを3つ貼り付けるものもあり、その全て場合について学習が必要となる。すなわち、一つの画像内でのラベルの位置関係も学習の対象となり、学習が複雑になる。
- 画像内に映る製品(水栓)の変更にもなって、ラベルの認識に影響が生じる恐れがある。
クラス分類問題では、画像全体に対して1つのクラスが学習される。本稿の場合、検査対象はラベルだが、クラス分類で様々な水栓画像の学習を行った場合、ラベルより水栓そのものの方を認識対象と判断されてしまう可能性がある。

このような問題を回避するため、ここではラベルそのものを学習させ、それを「検出する」というアプローチを採用する。すなわち画像認識の問題設定の中で「物体検出」という手法を利用してシステムの構築を行う。

物体検出は画像内に認識対象が複数存在する場合でも同時に複数検出することができ、ラベル有無の組み合わせの問題が解決される。また、製品は変わっても貼り付けるラベルは同一である場合が多い。この場合、ラベル自身が検出対象となっていれば、過去に利用したラベルの画像データをそのまま学習データとして利用可能となるほか、過去の学習をそのまま適用できるという再利用性も高い。

2.5 物体検出の特徴を利用した機能実現

物体検出には深層学習を用いた高速モデルが提案されており、高性能な GPU を用いると 50Frames/sec (FPS) で 70% を超える平均認識率となる高速処理も実現可能である (Liu et al., 2016) . カメラの撮影画像を 1 回処理してラベルの検査を行ってもよいが、高速処理が可能ならば複数の画像を用いて判断した方が信頼性は向上する。また物体検出では、認識対象の画像内での座標が取得可能なため、座標情報をラベル貼り付け位置の正誤判定に利用できると思われる。

物体検出の高速性と位置検出能力を利用して、次の 3 点を実現する。

1. 複数回の画像判定による信頼性向上

物体検出では対象を複数枚の画像内から探す性質上、存在しない対象を誤検出することがある。このため一度きりの検出結果で判定を行うと誤る確率が高くなる。そこで物体検出を高速に繰り返して複数回実行し、直近中の検出割合が一定数以上のラベルを「貼られている」と判定する。これにより瞬間的な誤検出が防がれ、信頼性向上が見込まれる。

2. ラベルなしの判定機能

(1) の方法では、ラベルが検出される撮影画像が連続すれば、実際にラベルが水栓に貼られていると判断し、検査は合格として終了できる。しかし、ラベルが検出されない場合は検出処理を終了できず、検出割合が上がるまでいつまでも検査を試みることとなる。そこで、ある制限時間を設け、その間にラベルを検出できなければ、ラベルは存在しないものと判断することとした。

3. ラベル位置の判定機能

湯水管に赤ラベル、冷水管に青ラベルが貼られているかを検査する。本来、同一の水栓では湯水・冷水各水管の配管は決まっており、例えば正面から撮影した場合、得られた画像上での赤・青ラベルの位置関係は一意に定まる。しかし、検査台上での左右逆向き設置や水管のねじれなどが存在すると、赤・青ラベル位置関係が撮影画像では逆になってしまう。そこで撮影画像上で赤と青の位置関係が一定となるように、作業手順の方に「検査台には水管をねじることなく、一定の向きで設置する」ことを取り決めとする。具体的には、水栓を水管が右側となるように横向きに置き、蛇口部が上側となるよう回転させた後固定する。ラベルが正しく貼り付けられている場合には、この状態で青ラベルが上で赤ラベルが下となる。つまり青が赤より上に存在するとき正常、青が赤より下のときは誤りと判定する。

2.6 システム構成

2.6.1 ハードウェア

画像撮影には、ラベル認識には十分な解像度をもつ Web カメラ (ロジクール C270, 解像度: HD720p) を採用した。また、画像処理には、GPU (NVIDIA GTX1060 6GB) を搭載した PC (CPU: Intel core i7-7800X, OS: Linux, Ubuntu 16.04) を使用した。判断結果を効果音で知らせることとし、スピーカ (サンワサプライ MM-SPL7UBK) を接続した。

2.6.2 ソフトウェアとアルゴリズム

本システムの検査アルゴリズムのフローチャートを図 3 に示す。開発プログラミング言語は Python 3.6 である。この検査アルゴリズムでは、いくつかのチェックフラグを導入している。初期設定ではそのチェックフラグは 0 であるが、直近の N 回の検査で正常検出の割合が設定した値 α を越えたら、そのチェックフラグに 1 をセットする。銀色ラベルでは 3 種類のすべてのラベルが検出されればよいので、それぞれ有無を示すチェックフラグ 3 ビットに加え、3 種類がすべて検出されたこと (Complete) を示すフラグの計 4 ビットを用意した。結果的には、銀色ラベルの場合 complete フラグ (CMP) がセットされればよいこととなる。一方、赤・青ラベルに対しては貼り付け位置が 2 カ所あり、その組み合わせも考慮し、Blue(青 1 枚), Red(赤 1 枚), OK(赤青位置正常), NG(赤青位置反対), BB(青 2 枚), RR(赤 2 枚) の 6 ビットを導入した。6 つフラグの中で OK のみが検出されればよく、他は異常時の原因判別に利用する。また Blue, Red は 1 枚でも検出されればセットされるため、他の 4 種類の内 1 つでもセットされている場合は必ず 1 となる。フラグがセットされるのに必要な条件をフラグ条件 (flag condition) と呼ぶことにする。また、各フラグ条件に対し、繰り返される撮影でこれまで満たされた回数を記憶するカウンタを

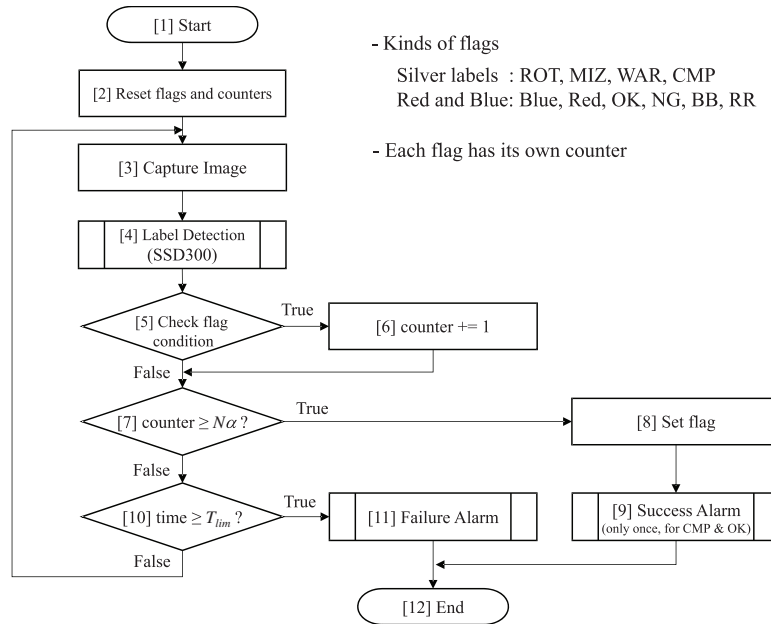


Fig. 3 Flowchart of our inspection algorithm. It includes ten flags to memorize the inspection result over several sequential images. Counters for each flag stores the number of time for the successful detection over the recent N images.

準備しておく。

以下にラベル検出の手順の詳細を述べる。以下の文中の数字は図3のフローチャート中の番号に対応している。検出対象となる水栓をカメラの撮影範囲にセットし、ラベル検出をスタート [1] する。初期化 [2] 後画像を撮影 [3] すると、物体検出により学習させたラベルがその位置情報とともに認識 [4] される。その結果に応じて各フラグ条件をチェック [5] し、もし満たされていれば、対応するカウンタ値を+1 [6] する。赤と青のラベルが見つかった場合に限り、その位置情報により上下判定を行い、OK か NG かを区別してカウントする。上下の比較には、検出された矩形の上下方向中央値を利用した。カウンタの値が αN を越えた [7] とき、検出の条件が満たされたと考え、チェックフラグに 1 をセット [8] する。CMP フラグもしくは OK フラグがセットされた場合には正常を示すピープ音を鳴らす [9]。一方、設定時間 T_{lim} 内検査を繰り返しても正常判定が得られない [10] 場合、異常すなわちラベルに問題ありと判定し、それを示すピープ音を鳴らし [11]、ラベル判定を終了 [12] する。

物体検出には Wei Liu らの SSD300 (Liu et al., 2016) を採用する。SSD300 は現在の物体検出アルゴリズムのうちで高速な動作速度と高認識精度を両立しており、同時に多物体の検出に比較的優れているため、検査工程での利用に適していると考えられる。認識エンジンには特に指定はなく、デフォルトで VGG16 モデル (Simonyan and Zisserman, 2014) が利用される。SSD300 ではもともとの VGG16 の出力に加え、それに畳み込み演算を繰り返し異なるサイズの同時認識を可能とする層を 5 層に分けて作成している。これら 6 層と出力層を結合し、その結合を重みを学習することで、さまざまなサイズと位置の画像を同時に判別することが可能となっている。このとき、一つの対象物が、位置やサイズさえもことなる複数の領域で認識させることもあるが、jaccard overlap という指標を用いて重なりが少ない BoudingBox を生成している。この同時判別できるネットワーク構造が SSD での判別速度の向上に大きく貢献している。

なお、SSD300 の Keras プログラムは、GitHub で公開 (Rykov, 2019) されており、実装は容易である。カスタマイズの方法 (@slowsingle), Processing を用いた学習データの作成方法 (AI coordinator) もインターネットで情報提供されている。

3. 評価実験

3.1 目的

今回構築したシステムについて、認識精度と処理時間について評価する。

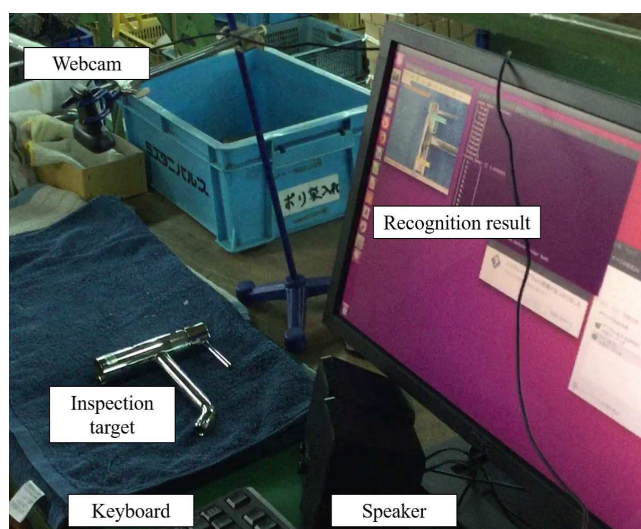


Fig. 4 Snapshot of the experimental environment.

3.2 物体検出モデルの学習

距離や角度を変えながらラベルと水栓本体の動画撮影を行い、動画から静止画を切り出して学習データとした。学習データ作成の途中で学習結果を確認したところ、金属光沢による誤認識がみられたため、照明や背景などの周囲の環境を変えて画像を用意することで認識精度向上を図った。各撮影画像内の検査対象ラベルを矩形で囲い、ラベルの種類を関連付けて教師データを作成する。総計 3092 枚の画像を利用した教師データでネットワークの学習を行った。内訳は銀ラベルの 3 クラスが 1435 枚、赤と青のラベルが 503 枚、今回の検査対象ではない水栓に貼られたラベルの画像が 1154 枚である。赤と青のラベルは 1 枚の画像内に同時に写っていることを考慮し、各クラス 500 枚程度になるように学習データを与えた。

実装プログラムとして SSD の Keras モデル (Rykov, 2019) を使用した。学習では、PASCAL VOC2007 (Everingham et al., 2007) のデータによって学習済みの重みを流用し、最上層のみファインチューニング (Fine Tuning) (Girshick et al., 2014) を行った。学習は 200epoch 行い、そのうち学習時のテスト誤差 Loss が最も小さかった重みを使用した。2.6.1 節に示したシステムでの学習時間は約 4 時間であった。

3.3 実験環境

実験は、工場の生産ラインの隣の空いたスペースを借用して行った。開始前の調整においては、認識精度が大きく変化しないよう、カメラは検査対象から 35cm ほど離して鉛直真上に固定した。実際の生産ラインでは製品保護のためにタオル敷く。今回の実験でも同様のタオルを敷き、実際の作業環境と近くなるよう整備した。実験の様子を図 4 に示す。

3.4 実験方法

正常・異常すべてが正しく検出できるかを確認するため、想定される組み合わせのテスト用製品を用意する。具体的には、3つの銀色ラベルでは、各ラベルの有無の 8 種類を、赤・青ラベルでは正常 (OK)、ラベル反対 (NG)、青 1 枚のみ赤無し (Blue)、赤 1 枚のみ青無し (Red)、両方青 (BB)、両方赤 (RR)、ラベル無し (NO) の 7 種類のテスト用水栓を作成し、検査にかける。

なお、作成するテスト用水栓の組み合わせが多くなるため、銀ラベル (実験 1) と赤・青ラベル (実験 2) のように分けて実験を実施した。全種類のテスト用水栓を 1 回ずつ認識させる実験を 1 セットとし、これを各実験で 50 セット繰り返す。各セット内で検査順序は一定とするが、その順序は最初にランダムに設定した。

実験では作業者はテスト用水栓をカメラの下に用意したタオルの上の決められた方向に設置し、以降は検査結果に影響を与えないよう触れずに制限時間内放置する。実験 1 では、ラベルが設置した水栓の反対側にも貼り付けられているため、検査時間が過ぎたら上下反転させ裏側の検査を行う。

Table 1 Recognition performance of all five labels. The recall rate over 0.8 was achieved in four of five labels.

	Precision	Recall	F-measure
LOT	0.908	0.990	0.947
MIZ	0.994	0.820	0.899
WAR	0.994	0.770	0.868
Blue	0.995	1.000	0.998
Red	1.000	0.995	0.997

Table 2 The result of three silver labels. Rows denote the kind of tested labels while columns denote detected results.

		Tested label			
		LOT	MIZ	WAR	nothing
Detected result	LOT	198	1	15	4
	MIZ	0	164	1	0
	WAR	0	0	154	1
	miss	2	36	46	—

実験での検査時間 T_{lim} は 5 秒と設定した。ただし、銀色のラベルは水栓を上下反転させて裏側を検査するため、制限時間は倍の 10 秒となる。したがって、すべての対象の検査時間は等しくなっている。物体検出アルゴリズムで検出と判定するクラスの確信度閾値を 0.6 とし、瞬間的な誤検出の防止のため直近 $N=10$ 回の判定のうち 5 回以上 ($\alpha=0.5$) 認識したときラベルが存在すると判定するようにした。

3.5 実験結果

実験 1・2 の結果から求められた各ラベルに対する精度、再現率、F 値を表 1 に示す。すべてのラベルに対し精度は 90% 以上であったが、一方で再現率は 80% を満たないラベルが存在した。

そこで、各ラベルがどのラベルと誤って認識されたかを調査した。実験 1 の詳細を表 2 に示す。

表の列はテスト用水栓に実際に貼られているラベルで、nothing はラベルが存在しないにもかかわらずラベルが検出されてしまったものである。なお実際に貼られたラベルを正常に認識しなかった回数を miss とし、偽陰性の数として扱った。

数値は検査で認識したら 1 とし、各ラベルは 200 回ずつ検査にかけられているため、各セルの取りうる最大値は 200 である。また各列の合計が 200 とならないのは、見逃しおよび複数ラベルの検出を計上しているためである。

実験 1 では LOT ラベルの見逃しが 2 回で再現率が最も良かったが、WAR ラベルと判断する誤認識が 15 回あり、精度は最も低かった。一方で WAR ラベルの検出については再現率が最も低かった。

表 3 は実験 2 の識別結果をまとめたものである。サンプルに対するラベルの認識とアルゴリズムによる分類で 2 段に分けた。

上段の数値は実験 1 と同様に、ラベルが認識された検査の回数で、ラベルの認識された枚数ではない。例えば BB のサンプルでは B ラベルが 2 枚貼られているが、B があるかないかの判定を行い、その検査で 1 枚でもあると判定されれば 1、ないと判定されれば 0 とする。このため各セルの取りうる最大値は各サンプルの検査回数となっている。

下段は上段での認識結果を利用したアルゴリズムによるサンプルを分類した結果であり、いずれかのクラスに分類されるようになっている。この表記によってどのように誤ったかが明らかとなる。今回の結果では、OK のサンプルは赤の見逃しによって Blue と分類されている。また、RR は片方の R の見逃しによって Red と分類され、nothing は B の誤検出によって Blue と誤って分類されたことがわかる。すなわち検出時のミスを除き、アルゴリズムではすべてのサンプルを正しく分類を行うことができた。

また実験 1・2 で同一のアルゴリズムを利用したため、実験 1 の銀ラベルの検査では R が 3 回、実験 2 の赤青

Table 3 The result of blue and red labels. The upper parts presents the number of times that each of blue and red flag was detected at least once. The lower parts represents what kind of test samples were recognized to which classes.

		Test sample						
		Blue	Red	OK	NG	BB	RR	nothing
number	B	50	0	50	50	50	0	1
	R	0	50	49	50	0	50	0
Detected result	Blue	50	0	1	0	0	0	1
	Red	0	50	0	0	0	1	0
	OK	0	0	49	0	0	0	0
	NG	0	0	0	50	0	0	0
	BB	0	0	0	0	50	0	0
	RR	0	0	0	0	0	49	0
	nothing	0	0	0	0	0	0	49

ラベルの検査では銀ラベルである MIZ が 3 回認識された。しかし、実験 1 で赤や青のラベルを認識しても、必要な銀色ラベルの種類数にはカウントされていない。また実験 2 で銀ラベルを認識してもアルゴリズムによる分類結果に影響はない。以上を踏まえて、これらの他種ラベルの誤認識はそれぞれの検査において影響がないため、今回の実験結果からは除外している。

各ラベルの認識に要した平均時間は、実験 1 では LOT 0.73sec, MIZ 7.79sec, WAR 1.07sec, 実験 2 では Blue 0.38sec, Red 0.39sec であった。MIZ を除けば概ね 1 sec 程度で認識が行われており、十分な高速性を備えているといえる。MIZ ラベルは水栓の最初の設置で裏側に貼られている。したがって 5 秒経過後に反転したあとで撮影範囲に収まっているため、認識が遅れている。

3.6 考察

実験 2 と比較して実験 1 の認識精度が低かった理由として、反射や映り込みといった影響が考えられる。今回の実験で使用した水栓はメッキ加工によって銀色の金属光沢があり、また 3 種のラベルのいずれも銀色で光沢が見られる。ラベルが本体と同色のため紛れて見えにくい上、光沢によってはラベル自体が見えないこともある。さらに水栓はくすみのない銀色であるため、鏡のように周囲が写り込んでしまう。これらの要因が銀色のラベルの検出を困難にし、誤検出の発生や認識精度の低下につながったのではないかと考えられる。

学習に関しては、学習データ数が比較的少数の 3000 枚程度であった。学習データが少なく済んだ理由は、Fine Tuning を利用したことがある。それに加え、ラベルは車のようにさまざまな色形があるわけではなく、また動物のようにさまざまな姿勢・体勢が存在するものではない。均一の大きさ、形状、色彩を持った型にはまったものであったため、パターンとしての揺らぎが少なかったことが考えられる。また認識環境が工場内で、変動が少ない環境であったことも要因に考えられる。特に赤・青ラベルは、撮影画像内には他に存在しない特徴的な色であるため識別が容易であり、実際に認識精度も高かった。

4. 現場への導入に向けて

4.1 改善案

実験の結果と考察を踏まえ、認識精度向上と効率化の観点から次の 3 点の改善が考えられる。

- (1). 検査場所で撮影した学習データを用意する。
 - (2). 検査環境を整え、反射や映り込みを減少させる。
 - (3). 複数のカメラで多方向からの同時検査を行う。
- (1) では、学習データを大学の研究室や企業の会議室内で撮影したため、現場の環境での学習データにより認識率の向上が見込まれる。(2) は作業の妨げにならない程度に照明条件を整えることは許されると考えるが、設置されるラインの状況にも依存する。(3) では、銀色ラベル検査で行っている上下反転の作業を省略できたり、取り付け



Fig. 5 The inspection Test using three webcams.

位置が離れている銀色ラベルと赤・青ラベルを同時に安価なカメラで大きく撮影できたりすることが可能となり、大幅な効率アップが期待できる。

(2)については、天候および朝昼夕での彩光の時間的変動や労働者の動きによる照明器具からの光の遮断などの影響が相対的に少なくなるように、それらの変動幅よりも強い光量を検査時に供給することを考えている。具体的には、生産ラインにおいてカメラの近くに照明を設置する予定であり、その効果を検証する計画である。

(3)に関し、実際にカメラを3台接続して実験を行い、実行速度を中心に評価してみた。図5は実際に3台のWebカメラを同時に使用して動作確認を行った例である。水栓を寝かせると裏側になって見えなくなってしまうが、今回は動作確認が目的ということで複数の水栓を使用した。画面左上に表示されている数値は検出の結果を表しており、ここからカメラ3台分の検査結果を統合して判定できていることが確認できた。またカメラを1台使用して検査を行うとき、動作速度はFPS20前後であったが、3台同時に使用したときはFPS8程度であり、動作速度の低下が見られた。しかし、3箇所を順次検査するよりは10秒ほどの短縮となる。ただし、1つの水栓のすべての検査箇所を同時に撮影するためには、設置方法を工夫する必要がある。

4.2 現場環境での学習データによる認識

現場環境で撮影した学習データを用いると、表1以上の性能が得られることを期待し、ライン導入計画中の図6(a)に示す水栓を対象に追加実験を行った。この水栓に対しては、白・緑・銀3種類のラベル貼り付け検査が必要である。ラインで導入を予定している30cm前後でカメラの高さを振り、図6(b)に示すように実際にラインに送られてくる梱包済み製品を撮影し、その覗き窓から出荷直前の状態を検査する。実際に撮影された覗き窓からの水栓ラベルの画像を図6(c)に示す。実際のラインでは、貼ってあるラベルを見逃して警告を出すことよりも、存在しないラベルを存在すると判断される方を防止することが望まれる。そこで、物体検出の検出閾値をデフォルトの0.6から0.7に引き上げ、後者の防止を重視するパラメータに設定後、再現率を比較する。学習用画像は7518枚、442epochの学習を行い、最もlossの小さい学習重みを使用した。

実際のラインは生産で使用中であったため、ライン横の空きスペースを利用して実験を行った。作業台から30cm

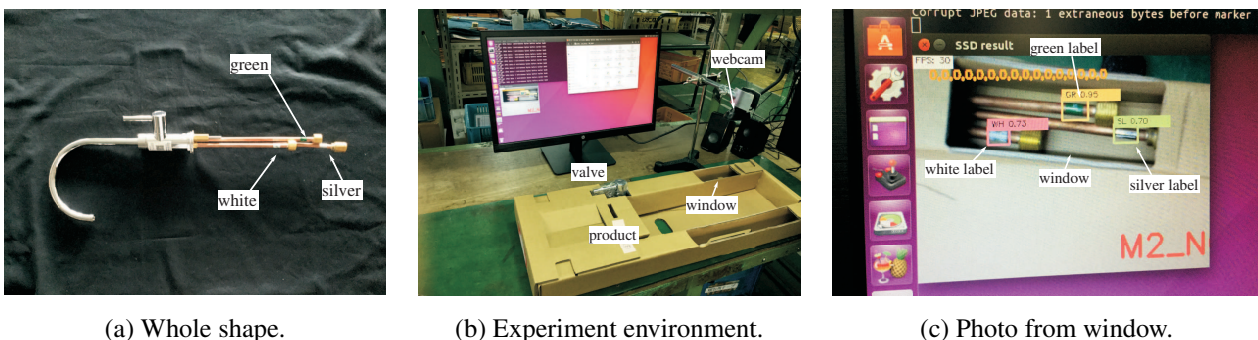


Fig. 6 Photos of the additional experiment.

Table 4 Result of the additional experiments with the white, green and silver labels. Compared with table 1, the performance of these three labels is improved.

	precision	recall	F-measure
White	1.000	0.930	0.964
Green	0.982	0.810	0.888
Silver	0.966	0.995	0.980

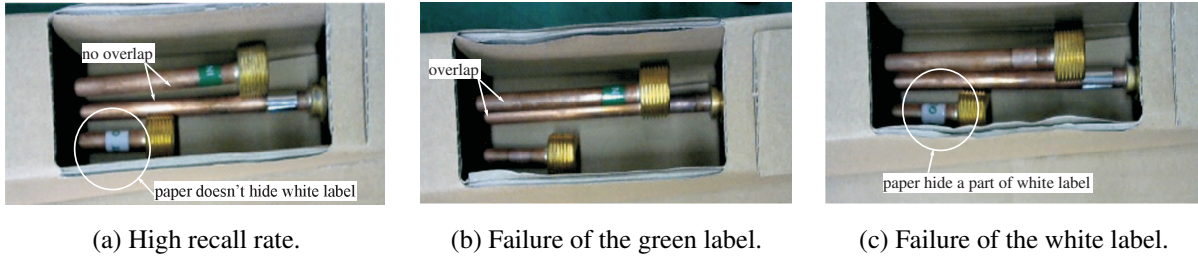


Fig. 7 Photos from package window.

の高さにカメラを固定し、学習データと同様、梱包積み製品をラベルが目視できる覗き窓からテストした。検査時間は5秒と設定した。3色ラベルのそれぞれの有無で計8種類の検査用サンプルを用意し、8種類を1回ずつ提示する実験を1セット、合計50セットを同じ順序で提示した。各ラベル有り無しが計400回検査されることとなる。

3色のラベルの検査結果を表4に示す。学習データの数の違いもあるが、現場環境での学習データを用いた今回の結果の方が、表1の銀色ラベル3種（LOT, MIZ, WAR）よりも再現率やF値で最高・最小・平均すべてで上回っている。しかし、学習データが多いのにもかかわらず、赤青ラベルほどの性能は得られなかった。今回の追加実験では、現場環境での撮影というよりも、別の要因が影響してしまった可能性もある。

実験をしてみると、検出されやすいサンプルとそうでないサンプルがあった。検出されやすいサンプルを図7(a)に示す。再現率が最も低い緑ラベルを調べたところ、緑のラベルが貼り付けられている4種のサンプルのうち、取りこぼしは特定の2サンプルで89%を占めていた。図7(a)と比較すると、そのサンプルでは図7(b)のように緑の配管の奥に別の配管が重なっている特徴があった。実験終了後に奥の配管の位置を手で写真下方に曲げて重ならないような状態にして物体検出にかけてみたところ、緑ラベルの検出が向上した。また、白ラベルについても、図7(c)に示す特定のサンプルが全見逃しの93% (13/14)であった。見逃した画像の特徴は、白ラベルに段ボール覗き窓の窓枠が一部掛かっていることであった。こちらも実験後に、段ボールの窓枠を手で下方に移動させ、白ラベルの配管が隠れないように撮影したところ、白ラベルの検出を示す検出枠が容易にあらわれた。

したがって、取りこぼした画像に記録された特徴は学習データに含まれていない可能性が高く、これらを集めて教師データを作成し再学習を行うと、より性能の高いラベル検出システムが構築できると考えられる。また、本来は現場環境で撮影した学習データを用いることで性能が向上することを示すのが目的であったが、それを実証するには銀色ラベルもしくは赤・青ラベルで他の条件は揃えて実験を行うべきであった。

4.3 機械学習の利用について

本稿でラベル認識に採用しているSSD内では、認識エンジンに学習済みVGG16モデル (Simonyan and Zisserman, 2014) を使用しており、出力層直前の重みのみを水栓ラベルが識別できるようファイン・チューニングで再学習を行って利用している。

このような機械学習の手法を取らずに、特に赤・青ラベルなど現場にあまり見かけない色彩の場合は、色やその面積情報のみからでもラベル判別も十分可能であると考えられる。この場合、数百枚以上もの学習データは必要なく、撮影環境さえ規定してしまえば必要パラメータ値も数枚の画像から設定できるという利点が考えられる。識別結果もラベルの色情報・面積といった統計量に基づく理論的検証が得られたものとなる。

しかし、そのためには、現場でプログラムについて知識を持った技術者が、プログラム内の該当箇所のみを適

切な値に修正する必要がある。また、赤・青の他にも種類があること、銀ラベルの識別のようにどのような統計量を取ればよいか直ぐには見当がつかないものもあることもあり、それらに対応するためにもプログラムそのものに機能を付け加えていく必要がある。

一方、機械学習による手法は、プログラムの修正は必要なく、学習結果を反映させれば認識対象の変化に対応できる。ただし、数多くの学習データが必要となるが、このデータ作成はプログラム等の知識が要らない単純作業で、その手順は予めレジュメとして作成できる。機械学習を用いない方法では、パラメータ調節や時には新たな統計量計算機能の追加が必要で、その場での試行錯誤も必要となり予め手順として記述できない。したがって、必然的にプログラミングの知識に基づいた現場対応が必要となる。

実際に製造業の場合には、プログラミングに精通した技術者が企業に存在しているとは限らない。したがって、再現率が同じもしくは少し落ちたとしても、専門的知識なしでアプリケーション・ソフトで学習データが簡単に作成でき（数は必要である）、認識対象をカスタマイズできる機械学習を用いる手法の方が、システムの保守・運用面で有利となり、早期の導入が可能であると考えられる。

5. 結 言

本稿では、画像認識によるラベルの検査システムの構築を行った。仕様として多種のラベルを同時に検査する必要があった。この仕様を物体検出により各種ラベルの有無や重複の有無、検出した物体の座標を利用して水管との対応を判断できるようにすることができた。システムの構築にあたり3つの仕様を設定した。ラベルの有無とそのラベルの種別の判定という2つ仕様は、物体検出という画像認識手法を利用することによって解決した。物体検出によって画像内にあるすべてのラベルを探し出すため、全5種類のラベルそれぞれの有無が判定可能である。また貼り付けられている水管とラベルとの対応は、検査時の水栓の姿勢を統一することにより、赤と青のラベルの上下関係を調べることで判定可能になった。上下関係を調べる手段として、検出された赤と青のラベルの画像内位置座標を利用することとし、ラベル位置の上下を比較することで判定を行った。

検証の実験の結果より、必要な判定を行えていることがわかり仕様を満足できていることが確認できた。再現率に関しては、目標のすべてのクラスでの再現率80%を達成することができず、十分とは言えない結果であった。しかし5つのうち3つのラベルにおいて再現率99%を達成しており、今後の改善によっては十分実用可能であるといえる。

今回とりあげたラベル検査は、水栓に限らず工業製品では広く見られる作業であり、その検査の効率化も一般的な問題である。したがって本システムによる効率化が期待できる分野は広いと考えている。

今後は実際に現場で試用実験を行い作業手順などを含めてシステムを再調節を行う。最終的には現場導入を目指し、作業の効率化と作業の信頼性向上を行っていききたい。

謝 辞

工場内での実験にご協力をいただきました、ミズタニバルブ株式会社の嶋井 敦氏、森元 翼氏に感謝いたします。

文 献

@slowsingle, 物体検出アルゴリズム (SSD : Single Shot MultiBox Detector) を学習させてみる, available from <<https://qiita.com/slowsingle/items/64cc927bb29a49a7af14>>, (参照日 2019 年 7 月 21 日).

AI coordinator, 物体検出用 SSD.Keras で使える学習モデルの作成方法, available from <<https://ai-coordinator.jp/ssd-keras-train>>, (参照日 2019 年 7 月 21 日).

Everingham, M., Van Gool, L., Williams, C., Winn, J. and Zisserman, A., The PASCAL visual object classes challenge 2007 (VOC2007) results (2007).

Girshick, R., Donahue, J., Darrell, T. and Malik, J., Rich feature hierarchies for accurate object detection and semantic segmentation, 2014 IEEE Conference on Computer Vision and Pattern Recognition (2014), pp.580–587.

原田達也, 画像認識 (2017), p.153, 講談社.

- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y. and Berg, A. C., SSD: single shot multibox detector, European Conference on Computer Vision (2016), pp.21–37.
- 永見英臣, 水谷真也, 森田亮介, 伊藤聡, 物体検出の位置座標を利用した水栓ラベル認識システムの構築, 第36回日本ロボット学会学術講演会予稿集 (2018), RSJ2018AC2E2-07.
- 中山英樹, 画像解析関連コンペティションの潮流, 電子情報通信学会誌, Vol.100, No.5 (2017), pp.373–380.
- Rong, X., Yi, C. and Tian, Y., Recognizing text-based traffic guide panels with cascaded localization network, European Conference on Computer Vision (2016), pp.109–121.
- Russakovsky, O., Deng, J., Su, H., Krause, K., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C. and Fei-Fei, L., Imagenet large scale visual recognition challenge, International journal of computer vision, Vol.115, No.3 (2015), pp.211–252.
- Rykov, A., Port of single shot multibox detector to Keras, (online), available from <https://github.com/rykov8/ssd_keras>, (accessed on 18 March, 2019).
- Simonyan, K. and Zisserman, A., Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556 (2014).
- Yi, C., Tian, Y. and Ardit, A., Portable camera-based assistive text and product label reading from hand-held objects for blind persons, IEEE/ASME Transactions On Mechatronics, Vol.19, No.3 (2014), pp.808–817.

References

- @slowsingle, Try learning an object detection algorithm (SSD: Single Shot MultiBox Detector), available from <<https://qiita.com/slowsingle/items/64cc927bb29a49a7af14>>, (accessed on 21 July, 2019) (in Japanese).
- AI coordinator, How to create a learning model that can be used with object detection SSD_Keras, available from <<https://ai-coordinator.jp/ssd-keras-train>>, (accessed on 21 July, 2019) (in Japanese).
- Everingham, M., Van Gool, L., Williams, C., Winn, J. and Zisserman, A., The PASCAL visual object classes challenge 2007 (VOC2007) results (2007).
- Girshick, R., Donahue, J., Darrell, T. and Malik, J., Rich feature hierarchies for accurate object detection and semantic segmentation, 2014 IEEE Conference on Computer Vision and Pattern Recognition (2014), pp.580–587.
- Harada, T., Image recognition (2017), p.153, Kodansha (in Japanese).
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y. and Berg, A. C., SSD: single shot multibox detector, European Conference on Computer Vision (2016), pp.21–37.
- Nagami, H., Mizutani, S., Morita, R. and Ito, S., Construction of recognition system for the labels on water valves in factory test process, The 36-th Annual Conference of the Robotics Society of Japan (2018), RSJ2018AC2E2-07 (in Japanese).
- Nakayama, H., Trends on competitions of image analysis and related areas, The Journal of the Institute of Electronics, Information and Communication Engineers, Vol.100, No.5 (2017), pp.373–380 (in Japanese).
- Rong, X., Yi, C. and Tian, Y., Recognizing text-based traffic guide panels with cascaded localization network, European Conference on Computer Vision (2016), pp.109–121.
- Russakovsky, O., Deng, J., Su, H., Krause, K., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C. and Fei-Fei, L., Imagenet large scale visual recognition challenge, International journal of computer vision, Vol.115, No.3 (2015), pp.211–252.
- Rykov, A., Port of single shot multibox detector to Keras, (online), available from <https://github.com/rykov8/ssd_keras>, (accessed on 18 March, 2019).
- Simonyan, K. and Zisserman, A., Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556 (2014).
- Yi, C., Tian, Y. and Ardit, A., Portable camera-based assistive text and product label reading from hand-held objects for blind persons, IEEE/ASME Transactions On Mechatronics, Vol.19, No.3 (2014), pp.808–817.