

AI を活用した 画像認識技術への取り組み ～機器情報の読み取り手法について～

■ 高田 晃平
Kohei Takada

■ 高野 裕基
Hiroki Takano

東光高岳では、AI の手法である機械学習、およびその発展形である深層学習を活用した画像認識技術の適用に関し、各種機器の管理に用いられている銘板や、管理用シールに記載された内容を読み取るための検討を進めている。本稿では、電力量計および管理シールの写真を対象として、その撮影条件にかかわらず、電力量計の銘板上にある電力量や、管理用シールの内容が記載された領域を検出できたことを報告する。

1 はじめに

近年、AI（人工知能）に関する研究が各所で盛んに進められており、製造・医療・金融など、さまざまな分野での活用も始まっている。東光高岳においても、業務効率化や、製品・システムへの適用に向け、AI に関する調査・研究を積極的に実施している。

AI の発展により、大きく進展している分野の代表例が、画像認識である。従来の画像認識の場合、見本となる写真との比較などによって認識を行う。そのため、工場の生産ライン上で使われる画像検査装置のように、照明や認識対象に対するカメラの位置関係といった撮影条件を整える必要があった。それに対し、AI による高度な画像認識では、撮影条件の異なる複数の画像から、共通する特徴をコンピュータが学習し、認識を行う。これにより、照明やカメラの位置、撮影場所や撮影者などのさまざまな撮影条件に左右されにくい画像認識が可能となった。

製品や部品を管理するために、銘板やシールを用いる方法がある。銘板や管理用シールの、さまざまな撮影条

件の写真から、必要な情報を読み取ることができれば、製品や部品の製造工程や在庫管理、保守にかかわる管理の効率化に応用できる。

東光高岳では、かかわりの深い機器である電力量計（図 1）や、その管理用シール（図 2）の写真から、電力量（指針値）を含む諸元や管理用シールの内容を読み取ることを目的とした検討を進めている。

本稿では、それらの写真に対して AI を活用した画像認識技術を適用することで、その諸元やシールの内容が記載された領域（以下項目領域）を検出する手法について報告する。

2 AI について

2.1 AI と機械学習と深層学習

「AI」は、認識・推論・創造などの知的活動をコンピュータに行わせるための情報科学の分野を指す。その中に「機械学習」があり、機械学習の一部に「深層学習（ディープ・ラーニング）」がある（図 3）。

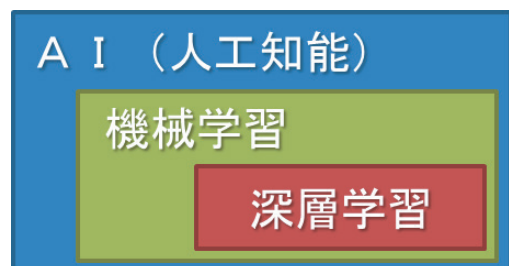


図 3 AI と機械学習と深層学習の関係



図 1 電力量計の写真



図 2 管理用シールの写真

2.2 機械学習について

機械学習は、入力されたデータに対してコンピュータが「学習」をすることで、コンピュータによる知的活動を実現する。その学習方式は、「教師あり学習」と「教師なし学習」に大別される。

教師あり学習では、データとそれに対する答えをセットで用い、これを教師データとよぶ。教師データをもとに、データと答えとの関係性を機械学習のアルゴリズム

が学習する。それにより、新たなデータが与えられた際に、コンピュータはそのデータがどの答えと紐づいているかを予測できる。

一方、教師なし学習では、教師データを用意しない。答えのないデータから、その中にひそむ法則性を機械学習のアルゴリズムが抽出する。このため、コンピュータによるデータの分析手法として用いられることが多い。

以降、本稿で扱う画像認識は、教師あり学習である。

2.3 深層学習の特徴

本来、深層学習を含めて機械学習とよぶが、以降、深層学習を除いた機械学習を指して「機械学習」と記述する。本節では機械学習と深層学習の違いについて述べる。

機械学習では、特徴量^{注1)}計算のアルゴリズムと予測のアルゴリズムが独立しており、予測に適した特徴量計算のアルゴリズムを人間が設計する必要があった。

一方、深層学習は、特徴量計算のアルゴリズムと予測のアルゴリズムを内部に有しており、その双方を相互に最適化するように学習することができる。

教師あり学習において、機械学習と深層学習の性能を比較すると、一般に、予測精度は深層学習のほうが高い。ただし、深層学習は、機械学習と比べて多くのデータを必要とし、学習および予測に高性能なコンピュータを要するなど、欠点も存在する。よって、条件に合わせて、機械学習と深層学習を選択する必要がある。

3 物体検出アルゴリズム

図1に示した電力量計の諸元を読み取るために、文字認識ソフトウェアによる文字読み取りを行った場合、諸元の記載された銘板以外の背景や、銘板内の模様を文字として誤認識する可能性が高くなり、文字読み取りの精度が著しく低下する。

これに対して、物体検出アルゴリズムを用いて、画像中から諸元の記載された領域のみを検出し、その領域に対して文字読み取りを行う。これにより、背景や模様を文字として誤認識しなくなるため、諸元の読み取り精度が向上する。

3.1 物体検出と画像分類

「物体検出」は、画像中のどこの位置に、どの大きさで、何の物体が写っているかを予測する教師あり学習のアルゴリズムである。図4に例を示す。この例では、猫と犬が、画像中でそれぞれどこの位置に、どの大きさで存在するかを予測する。

一方、画像に写っている物体が何であるかを予測する教師あり学習のアルゴリズムが「画像分類」である。図5に例を示す。この例では、画像に犬と猫のどちらが

写っているかを予測する。画像分類は物体検出の処理の一部としても用いられる。

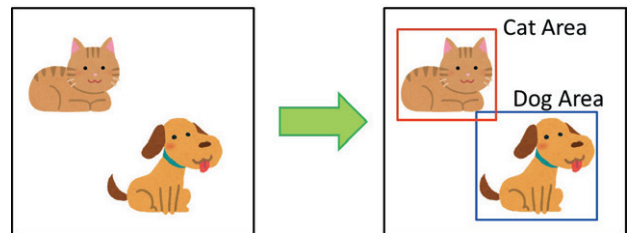


図4 物体検出の例

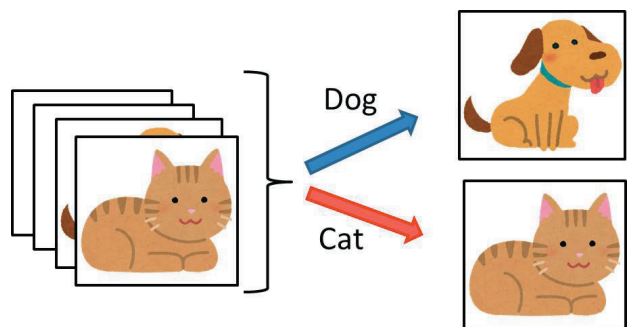


図5 画像分類の例

3.2 物体検出アルゴリズムの仕組み

本節では、機械学習を用いた物体検出アルゴリズムについて述べる。その概要を図6に示す。なお、深層学習を用いた物体検出アルゴリズムの詳細については、参考文献⁽¹⁾を参照されたい。

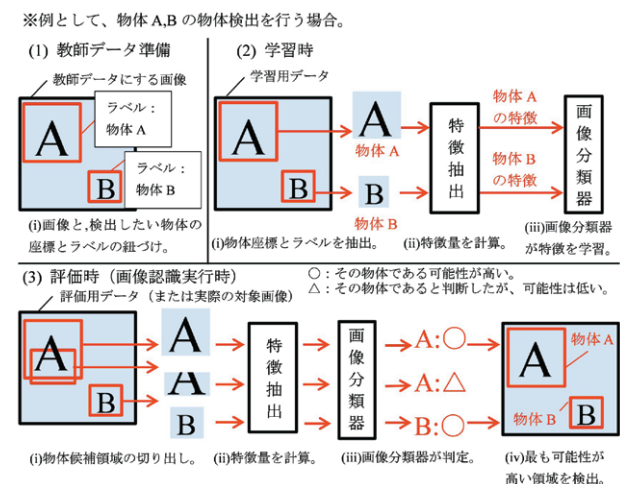


図6 機械学習による物体検出アルゴリズムの概要

(1) 教師データ準備

物体検出では、学習用に検出対象が写った画像と、その画像内の検出対象の座標 (位置および大きさ) とラベル (検出対象の分類先) を紐づけたデータ (教師データ) が必要であり、これを作成する。作成した教師データは学習データと、評価データに分割する。

(2) 学習時

- (i) 学習データから検出対象の座標とラベルの情報を抽出し、座標をもとに画像から領域を切り出す。
- (ii) 切り出した領域の特徴量を計算する。
- (iii) 最後に、特徴量とラベルを紐づけて画像分類器に学習させる。ここで、画像分類器とは、3.1節の画像分類を行うものである。よって、検出対象の座標については学習の対象外となる。

(3) 評価時

- (i) 評価データの画像から、検出対象の物体が存在する可能性がある座標を切り出す。
- (ii) 切り出した物体候補領域の特徴量を計算する。
- (iii) 学習済みの画像分類器に、(ii) で計算した特徴量を入力し、ラベルを予測させる。
- (iv) (iii) の予測結果のうち、そのラベルである可能性が最も高いと予測された座標を予測結果とする。

評価データの画像から予測された物体の座標およびラベルと、評価データの座標およびラベルを比較することで、物体検出アルゴリズムの性能を評価する。

4 マーカ検出アルゴリズム

物体検出をしたい対象と一定の位置関係があるマーカがあれば、それをもとに物体検出の対象とする範囲を限定できる。限定した範囲以外に対する物体検出を行わないことで、誤検出の可能性を低下させることができ、処理の高速化も期待できる。

マーカを用いた補助手法について、**図 7** に例を示す。文字領域の検出アルゴリズムを用いて、**図 7** から青枠(あいうえお)の文字領域を検出したいとする。このとき、文字に類似した模様がある赤枠内の部分も文字領域として検出されてしまう可能性がある。これに対し、青枠と一定の位置関係があるマーカ(星印)を用いて、文字領域の検出対象とする範囲を、緑枠内に限定する。これにより、文字領域の検出対象とする範囲から、赤枠内の部分を除外することができる。

図 2 に示した管理シールに対して項目領域の検出を行う場合、QR コード^{注2)}をマーカとして、項目領域の検出対象とする範囲を限定できる。



図 7 マーカ検出による物体検出の補助

4.1 QR コードの位置検出アルゴリズム

機械学習の手法をもとにした、QR コードの位置検出アルゴリズムについて説明する。

一般的に、QR コード読み取りソフトウェアは、QR コードの3個所の角に存在するファインダパターンを画像処理の手法で検出することで、QR コードを認識している。具体的には、ファインダパターンをどの方向に走査しても、黒1:白1:黒3:白1:黒1の白黒比を満たすことを利用している(**図 8**)。

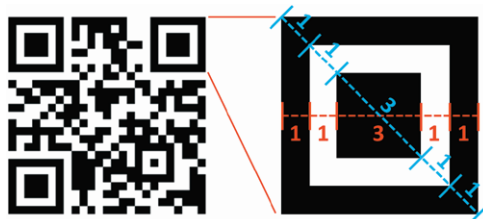


図 8 QR コードとファインダパターン

このようなソフトウェアは、通常、カメラを有する機器にて使用する。よって、ピントなどを調整しながら繰り返しQRコードを認識することが可能である。

一方、既存の写真に対して検出を行う場合を考えると、ブレやピントのズレによって前述の白黒比の条件が満たされなくなり、この手法ではQRコードの認識に失敗する可能性が高い。これに対し、ファインダパターンを機械学習により検出し、さらに、ファインダパターン間の位置関係が直角二等辺三角形をなしているという条件をもとに、QRコードの位置を検出する手法がある⁽²⁾。

図 9 にQRコードの検出例を示す。このように、機械学習を用いることで、写真が多少不鮮明であっても、QRコードの位置が検出可能である。ここで、赤丸が検出されたファインダパターン、黄線がファインダパターン間を結んだ直角二等辺三角形を示している。

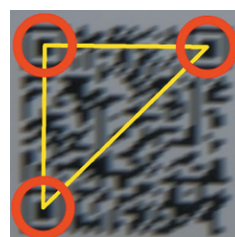


図 9 QR コードの検出例

5 項目領域の検出

今回適用した手法および、その結果を以下に示す。

5.1 手法

電力量計の写真(**図 1**)から、物体検出アルゴリズムを用いて、項目領域を検出した。また、管理用シールの

写真(図2)から、マーカによる補助手法を用いて、項目領域を検出した。

(1) 銘板の項目領域検出手法

3.2節の物体検出アルゴリズムを適用し、項目領域を検出した。

まず、機械学習を用いる際の手法について説明する。機械学習による物体検出では、物体候補領域の切り出しにスライディング・ウィンドウ法^{注3)}を使用し、特徴量の計算にHOG^{注4)}、画像分類器にSVM^{注5)}を用いる手法がある。しかし、スライディング・ウィンドウ法では、画像全体を走査するため、処理に時間を要する。そこで、項目領域が四角形で囲われていることに着目し、画像から四角形の輪郭を持つ領域を切り出し、認識を行った。四角形検出には、画像処理のオープンソース・ソフトウェアであるOpenCV^{注3)}を使用した。

次に、深層学習を用いる際の手法について説明する。深層学習を用いた比較的処理が高速な物体検出の代表例として、SSD^{注4)}やYOLO^{注5)}がある。YOLOとSSDを比較すると、YOLOのほうが処理は高速ではあるが、SSDのほうが検出したい物体の大きさの大小の差に影響を受けにくいと、今回はSSDを用いた。また、始めに写真全体(図1)から銘板のみを切り出し、次の中から各項目領域の検出を行った。これは、SSDやYOLOが、画像全体に対して小さすぎる領域の検出を苦手とするためである。

(2) QRコードをマーカとした項目領域検出手法

4.1節の機械学習によるQRコードの位置検出アルゴリズムを用いて、管理シールの写真(図2)から、QRコードの位置を検出する。次に、その位置をもとに、項目領域の検出を行った。処理のステップを(i)～(iii)に示す。なお、項目領域を検出するために、文字の領域を検出する深層学習のアルゴリズム^{注6)}を別途用いた。

- (i) QRコードの位置検出を行う。
- (ii) (i)をもとに項目領域の検出を行う領域を定める。
- (iii) (ii)の領域に対して項目領域の検出を行う。

5.2 結果

(1) 銘板の項目領域検出結果

機械学習を用いた結果、検出された項目領域の誤分類が生じた。この原因としては、画像処理による四角形検出が、写真の撮影条件の変化に対応できなかったことや、特徴量計算のアルゴリズムおよび画像分類器の性能の限界が考えられる。

次に、深層学習を用いた結果を図10に示す。各枠の付近にあるパーセンテージが検出の根拠となる信頼度(その項目領域である可能性)で、90%以上の高い信頼

度で検出できた。このように、項目領域ごとに検出を行い、それに対して文字認識を行うことで、銘板の両脇に3本ずつある線や、銘板の下部にある矢の形のマークなどを文字として誤認識をすることがなくなった。

深層学習では、画像の特徴量計算をするアルゴリズムおよび、その特徴量を用いた認識アルゴリズムが相互に最適化するように学習されるため、良い結果が得られたと考えられる。

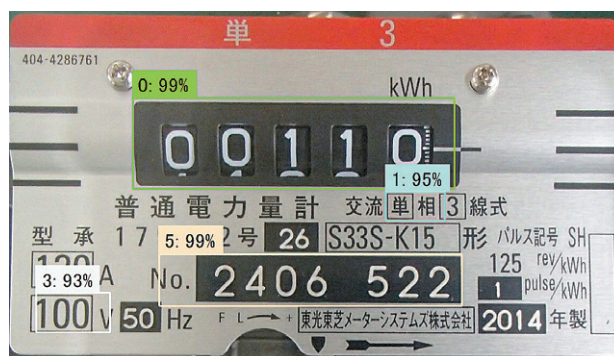


図10 深層学習による項目領域の検出例

(2) QRコードをマーカとした項目領域検出結果

図11に、検出したQRコードの位置、および、それをもとにした項目領域の検出例を示す。赤丸がファインダパターン、黄線がファインダパターンを結んだ三角形を示す。また、青枠がQRコードの位置をもとに項目領域の検出対象とした領域、赤枠内が項目領域として検出された部分である。機械学習によるQRコードの位置検出(赤丸および黄色三角形)は、十分な精度で行うことができた。

一方、深層学習による項目領域の検出(赤枠)に関しては、一連の文字列で意味を持つ領域が分かれて検出されるなどの課題が生じた。この課題については、教師データをさらに作成して学習を行うことで、検出精度を高められると考えられる。

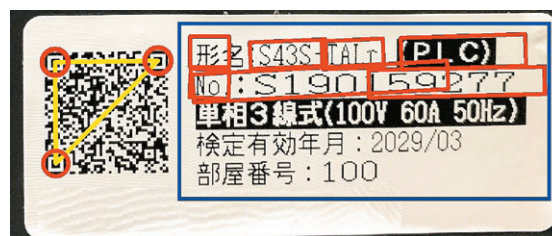


図11 QRコードおよび項目領域の検出例

6 おわりに

6.1 まとめ

電力量計の銘板内の項目領域を90%以上の高い信頼

度で検出すること（図 10），および，QR コードの位置検出をもとに項目領域を検出することができた（図 11）。

銘板内の項目領域の検出では機械学習による手法は良い結果を得られなかったが，深層学習では良い結果が得られた。一方で，QR コードの位置検出では機械学習による手法で十分な精度が得られた。

このように，AI を適用する際には，適用したい対象と認識の難易度，目標とする認識精度，および入手可能なデータ量などによって機械学習と深層学習を適宜使い分けの必要がある。

6.2 今後の取り組み

検出した項目領域の内容を認識する上で，文字認識が必要となる。電力量計などの銘板の場合，光の反射や，項目領域を囲んでいる四角形の枠，特殊な記号などが存在する。これに対し，一般的な文字認識アルゴリズムは，紙の上に印刷・手書きされた通常の文字を対象としており，銘板の内容をすべて文字認識できるわけではない。よって今後は，銘板の内容を認識するための，専用の文字認識アルゴリズムの構築を含めた検討を進めていく。

今回は，一例として電力量計および管理シールに対して物体検出アルゴリズムによる項目領域の検出を行った。十分な量の教師データを確保する必要があるが，画像認識における物体検出アルゴリズム自体は，ほぼ任意の物体の検出が可能であり，さまざまな応用が考えられる。また，QR コードなどのマーカを検出しその位置を得ることは，AR マーカ^{注6)}などの検出にも応用できる。

今後とも，画像認識に限らず，AI 技術全般への理解をさらに深めていくとともに，AI を活用した課題解決に積極的に取り組んでいく。

■参考文献

(1) Li Liu, *et al.* : “Deep Learning for Generic Object

Detection: A Survey”, IJCV (2018)

(2) L. F. F. Belussi and N. S. T. Hirata : “Fast component-based QR code detection in arbitrarily acquired images”, Journal of Mathematical Imaging and Vision (2012)

(3) <https://opencv.org>

(4) Wei Liu, *et al.* : “SSD: Single Shot MultiBox Detector”, ECCV (2016)

(5) Joseph Redmon, *et al.* : “You Only Look Once: Unified, Real-Time Object Detection”, CVPR (2016)

(6) Xinyu Zhou, *et al.* : “EAST: An Efficient and Accurate Scene Text Detector”, CVPR (2017)

■語句説明

注1) 特徴量：データがどのような特徴を有しているかを抽出し数値化したもの。例えば，画像を特徴量に変換することで，写真の手ぶれやボケ，物体への光の当たり方などの影響を軽減することができる。

注2) QR コードは（株）デンソーウェーブの登録商標。

注3) スライディング・ウィンドウ法：入力画像の全体から，さまざまな大ききで，総当たりでいくつもの画像を切り出していく手法。

注4) HOG (Histograms of Oriented Gradients)：隣接画素間の輝度の差（輝度勾配）をもとにした特徴量。

注5) SVM (Support Vector Machine)：データの特徴量をもとに，そのデータがどの答えに属するかをサポート・ベクトルとよばれる数理的手法を用いて決定する。

注6) AR マーカ：AR（拡張現実）の実装のために用いられる。QR コードに類似したマーカであり，これをコンピュータに認識させ，現実の映像に仮想の物体を重ねて描画したり，情報を表示したりする。

高田 晃平

技術研究所 ICT技術グループ 所属

高野 裕基

技術研究所 ICT技術グループ 所属