

## 技術資料

## Technical Data

# ディープラーニングを用いた丸棒鋼本数カウント技術の開発

木全謙太\*, 吉田たける\*

## Development of Bar Counting Inspection Technology for Bundled Round Steel Bars by Using Deep Learning

Kenta KIMATA and Takeru YOSHIDA

### Synopsis

Specialty steels are used in parts for automobiles, aircraft, and other industrial machinery. Since each steel grade has different characteristics, identification control in special steel production lines is important. One example of identification control is counting the number of bundled round steel bars. To count the number, it is important to count without excess or deficiency. Conventionally, operators count visually, but this method may cause errors due to human error. To solve this problem, we developed an image processing method for counting the number of bundled round steel bars using a camera, bar lighting, and AI technology.

## 1. 緒言

大同特殊鋼(株) (以下、当社という) では、構造用鋼、快削鋼、軸受鋼、ばね鋼、ステンレス鋼などの多岐に渡る特殊鋼を製造している。特殊鋼は鋼種ごとに有する特性が大きく異なり、自動車、航空機をはじめ、各種産業での重要部品として使用されるため、その製造過程で鋼種を特定する識別管理は極めて重要である<sup>1)</sup>。これら工場内で管理される特殊鋼は、一見するとほとんどが同じ色や形をしており、その違いを外観から見分けることは難しい。そこで、当社では各製造過程で、製品を個体ごとに識別記号を打刻する方法や、2次元バーコードを付与する方法、もしくは製造ロットごとに製造予定の数量と実際の製造数量が合致するかを確認する方法などを適用し、識別管理を実施している。近年では、識別記号の

打刻は一部を除いて2次元バーコードへの切り替えが進み、自動で識別管理が可能な体制が整ってきている。一方で、製造数量の確認は作業者の目視によるカウントで対応しているケースが多く、自動化が求められている。製造数量の確認の例としては、丸棒鋼の製造ラインにおける本数カウントがある。本数カウントは、丸棒鋼を前工程から受け入れる際や、次工程へ払い出す際に、上位から指示される正解本数と照合する作業である。本作業は、整列された状態もしくは結束された状態(以下、結束状態という)でコンベアによる搬送中に実施され、整列された状態の場合は、ステレオカメラ法による3D形状測定を用いた手法<sup>2)</sup>により、本数カウントの自動化が進んできている。しかし、結束状態ではいまだ作業者の目視によるカウントで対応しているため、本数を実際よりも過剰に検出する場合(以下、過検出という)や、

2023年9月28日 受付

\* 大同特殊鋼(株)技術開発研究所(Corporate Research &amp; Development Center, Daido Steel Co., Ltd.)

少なく検出する場合（以下、見逃しという）が懸念される。そこで本稿では、結束状態の丸棒鋼における自動本数カウント方法について報告する。

結束状態の丸棒鋼を自動でカウントする手法としては、計測重量から本数を予測する方法や、鋼材の端面をカメラで撮像して解析する画像処理法が挙げられる<sup>3)</sup>。しかし、計測重量による方法は寸法公差による重量ばらつきの影響を受けやすく、結束される丸棒鋼の本数が多い状況では、高い予測精度を得るのが困難である。一方で、画像処理法では寸法公差の影響を受けにくい、丸棒鋼の端面状態により発生する輝度のバラつきが予測精度に影響するため、その発生抑制もしくは影響を受けにくい画像処理方法の開発が課題である。

そこで、本稿では輝度のバラつき影響を受けにくい本数カウント手法について、下記3点の検討結果を述べる。また、開発仕様を Table 1 に示す。

- (1) 画像処理による検討
- (2) Deep Learning による検討
- (3) Deep Learning の実用化に向けた取り組み

Table 1. Development Specification.

Target	Bundled round steel bars
Number of pieces per bundle	2 - 100
Material diameter	$25 \leq \phi \leq 85$ mm
Condition of steel edge surface affecting brightness variation	Surface roughness, With or without a label, Painted or not, Sloping end face
Target inspection time per bundle	Less than 0.5 sec
Target counting accuracy	100 %

## 2. 画像処理による検討

まず、画像処理によるカウント方法を検討した。一般に、金属表面を対象とした画像処理システムは、カメラ、照明、画像処理部の三要素から構成される。

### 2. 1 カメラ

結束状態の丸棒鋼の撮像は、鋼材端面の中心部が撮像中心となるように配置したエリアカメラ1台で実施する。エリアカメラの仕様は、丸棒鋼との距離関係や、撮像に必要な視野サイズ（幅 950 mm × 高さ 710 mm）から 2592 × 1944 画素とした。

### 2. 2 照明

製造ラインの制約上、照明から鋼材端面までの距離が 1 m 以上ある状態で均一に鋼材端面を照射するため、挟角配光形のバー照明をカメラの周囲に 90° ごと 4 つの面に計 6 台配置した。エリアカメラと照明の位置関係を Fig. 1 に示す。

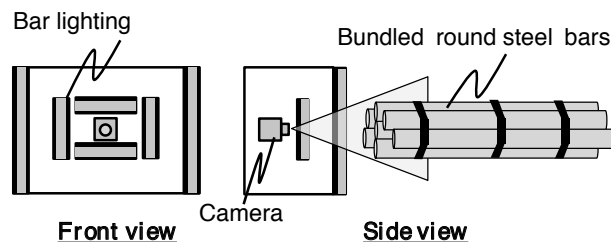


Fig. 1. Image-acquisition method for counting bundled round steel bars.

### 2. 3 画像処理による本数カウント方法

Fig. 1 に示す構成で撮像した画像に対して、以下のステップで本数カウントを実施することとした。次項に主な処理内容の説明と実施例を Fig. 2 に示す。

- [Step1] 検査範囲の画像切出し
- [Step2] 大津の二値化処理<sup>4)</sup>による鋼材領域の検出
- [Step3] 検出した鋼材領域に対し距離変換を適用
- [Step4] 閾値処理によるノイズ除去
- [Step5] ラベリング処理による鋼材本数の算出

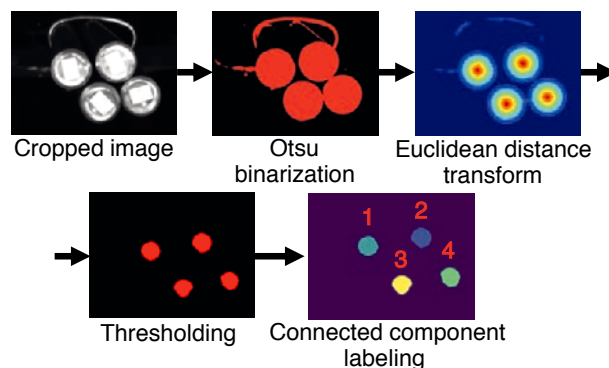


Fig. 2. Conventional image processing procedures.

Step2 で検出した鋼材領域が正解の鋼材領域か否かの判断には、鋼材領域が円形状であることに着目し、Step3 にて Fig. 3 と式 (1) に示すユークリッド距離による距離変換<sup>5)</sup>を実施した。距離変換は、入力として与えた 2 値 (0, 1) 画像の各画素について、そこから最

も近い0画素への距離を求める処理である。これにより、検出した鋼材領域の形状的な特徴が各領域の最大距離として出力される。そしてStep4にて、Step3で求めた各領域の最大距離が、通材予定の丸棒鋼の出力する距離と合致するかを確認することで、検出した鋼材領域が正解の鋼材領域かを判断している。

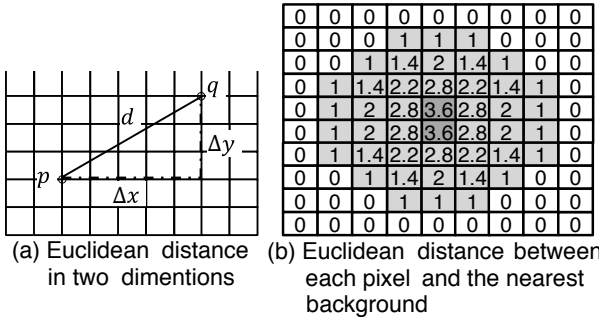


Fig. 3. Euclidean distance transform method.

$$d = \sqrt{\Delta x^2 + \Delta y^2} \tag{1}$$

## 2. 4 本数カウント結果

画像処理法による正解本数と予測本数の一致率（以下、カウント精度という）の算出結果をTable 2に示す。4208枚の画像に対する正解数は4023枚となり、画像処理によるカウント精度は95.6%であった。また、画像処理法による本数カウント例をTable 3に示す。正解本数と予測本数が不一致となった画像を確認した結果、Table 1で示したように鋼材端面の状態が、表面粗度、ラベルの有無、塗装の有無、面取りによる傾斜の有無など種々あるため、輝度のバラつきが想定よりも大きく、大津の二値化で検出する鋼材領域が一部欠落し、距離変換後の閾値処理でノイズとして除去されることで、見逃しが発生していた。大津の二値化は、輝度値のヒストグラムが双峰性をなすことを前提とした統計的な手法のため、画像全体の輝度変化には対応できるが、同じ画像内で局所的に輝度変化するものには対応できない。よって、従来の画像処理ではこれらすべてに対応して鋼材領域を検出することは難しい。正解本数と予測本数が不一致の場合は、作業者の目視によるカウントで対応する必要がある。そこで、さらなるカウント精度の向上に向け、Deep Learningによる本数カウントを試みた。

Table 2. Counting results by Image processing.

	OK (Number matching)	NG (Number not matching)	Total
Number of images	4023	185	4208
Counting accuracy = 4023 ÷ 4208 = 95.6 %			

Table 3. Example of image processing results.

Input image	Otsu binarization	Distance transform	Result
			OK (20 = 20)
Correct number 20		Predicted number 20	
			NG (54 ≠ 44)
Correct number 54	Lack of detection	Predicted number 44	Overlook

## 3. Deep Learningによる検討

近年、高い物体認識力を持つDeep Learningを、複雑な鉄鋼製造プロセスや、材料開発の評価プロセスに適用するための検討が進められている<sup>6)~8)</sup>。例えば原田ら<sup>7)</sup>は、U-Net<sup>9)</sup>によるセマンティックセグメンテーションを適用し、鉄鋼の組織写真中の輝度や形状、テクスチャなど、複数の特徴を考慮した、畳み込み層での特徴マップから、二相の組織の違いを画素単位で識別している。また、Chenら<sup>8)</sup>は、Mask R-CNN<sup>10)</sup>によるインスタンスセグメンテーションを適用し、鉄鋼の組織写真中の複数ある相の違いを画素単位で分類し、さらに同じ種類でも個体ごとに分割して識別している。インスタンスセグメンテーションは、画像内の物体を種類ごとに分割するセマンティックセグメンテーションと異なり、同じ種類でも個体ごとに分割するため、画像中の隣接した丸棒鋼に対して、物体間の境界を区別して各鋼材領域の検出が可能な手法である。そこで、本稿では画像処理法の課題であった輝度のバラつきによる鋼材領域の欠落に対し、Mask R-CNNによるインスタンスセグメンテーションを、Step2の二値化処理に適用することで、カウント精度の向上を試みた。使用したモデルの構造はFig. 4のとおりである。

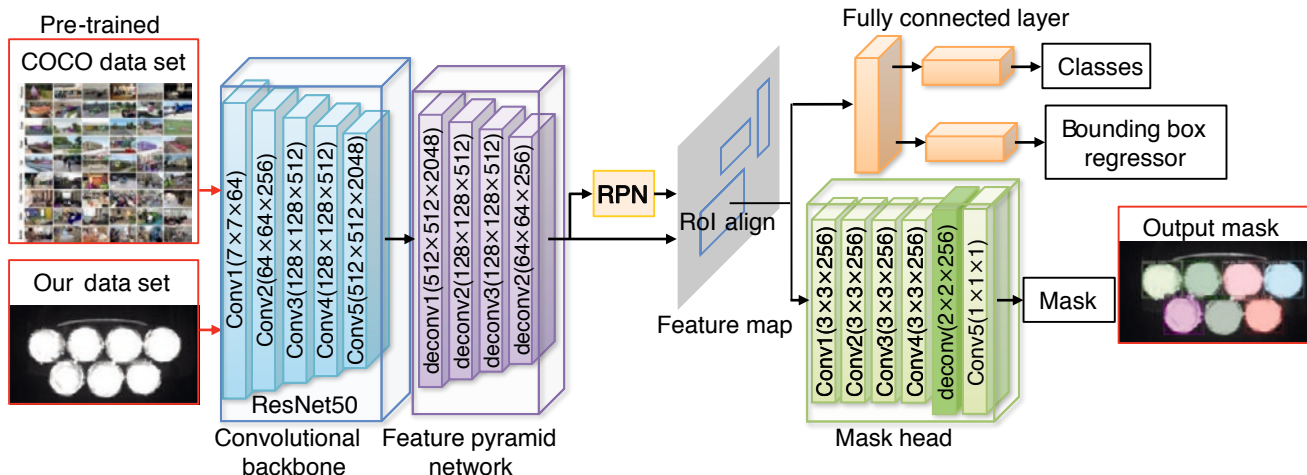


Fig. 4. Mask R-CNN architecture.

### 3. 1 Deep Learningの学習

Mask R-CNN の学習には、Table 4 に示すように画像データを学習用、評価用、テスト用データセットに分割して用いた。Microsoft COCO<sup>11)</sup> による事前学習済みの重みを初期値として利用し、独自の学習データによってネットワーク全体を再度学習する転移学習により、重みを更新した。機械学習アルゴリズムの挙動を制御するハイパーパラメータ（学習率、バッチサイズなど）は、初期値を変化させながら評価用データセットにおける、鋼材領域の検出誤差が十分に小さくなる値に設定した。

Table 4. Calculation condition.

Model	Mask R-CNN
Backbone	ResNet50-FPN
Number of training data	120 Training: Validation = 100 : 20
Number of test data	4208
Optimizer	Stochastic Gradient Descent (SGD)

### 3. 2 鋼材領域の検出精度評価

Mask R-CNN における鋼材領域の検出精度を、人間が教示した正解画像と出力画像の類似度から評価した。定量評価には2つの集合 A と B の類似度を表す、Jaccard index<sup>12)</sup>（以下、Ji という）と Dice index<sup>13)</sup>（以下、Di という）を用いた。本評価では、Mask R-CNN もしくは大津の二値化で検出した鋼材領域と、正解の鋼材領域で類似度を算出している。

$$\text{Jaccard index} = \frac{|A \cap B|}{|A \cup B|} \tag{2}$$

$$\text{Dice index} = \frac{2|A \cap B|}{|A| + |B|} \tag{3}$$

Mask R-CNN における鋼材領域の検出精度を、前述の画像処理と比較した結果を Fig. 5 に示す。なお、左から順に入力画像、正解画像、Mask R-CNN による出力画像、画像処理による出力画像である。この例における検出精度は Mask R-CNN で Ji = 0.935、Di = 0.966、画像処理で Ji = 0.901、Di = 0.948 となり、どちらも Mask R-CNN

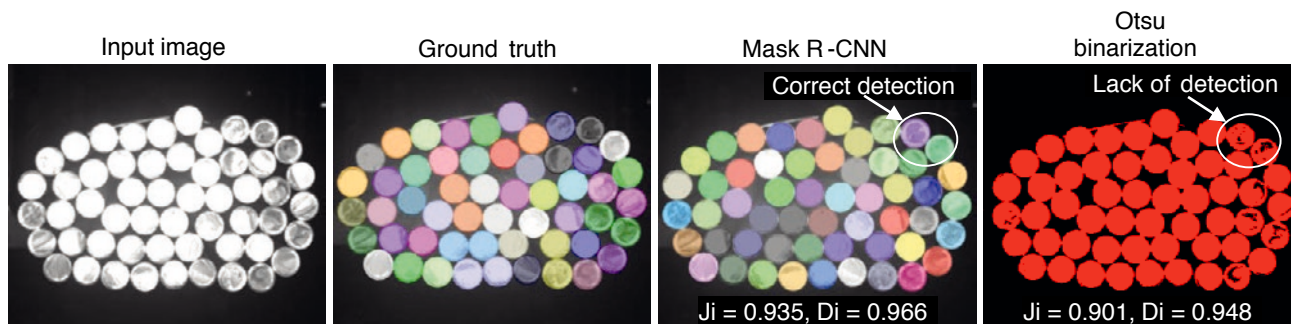


Fig. 5. Comparison of detection accuracy in Mask R-CNN and Otsu binarization.



が高い検出精度を示した。また、出力画像を比較すると、鋼材領域の一部が欠損していた画像処理に対して、Mask R-CNN は欠損なく、輝度のバラつきがある鋼材に対しても正確に領域を検出可能なこと示した。

### 3. 3 本数カウント結果

テスト用データセットに対するカウント精度結果を Table 5 に示す。その他の処理条件は変更せず、Step2 の処理のみを大津の二値化から Mask R-CNN に切り替えた結果、カウント精度は 99.7 % であった。前述したとおり、画像処理によるカウント精度が 95.6 % であったことから、Mask R-CNN を用いた本数カウントがより高精度であることを示した。

Table 5. Counting results by Mask R-CNN.

	OK (Number matching)	NG (Number not matching)	Total
Number of images	4196	12	4208

Counting accuracy =  $4023 \div 4208 = 95.6 \%$

Mask R-CNN による本数カウント例を Table 6 に示す。本表では Mask R-CNN による過検出例と、丸棒鋼の一部が隠れたことによる見逃し例を示す。過検出例では、Mask R-CNN が検出した鋼材領域と重複して、端面に貼られたラベルも鋼材領域として検出していた。これは、訓練データに同形状で同サイズのラベルが貼られた丸棒鋼の画像が存在し、実質的にラベルの特徴も学習しているためであると考えられる。

Mask R-CNN で検出した鋼材領域は、距離変換を実施

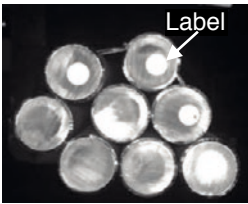
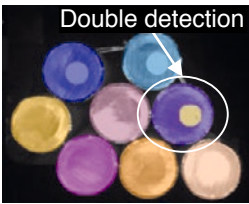
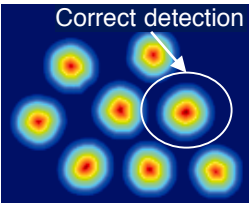
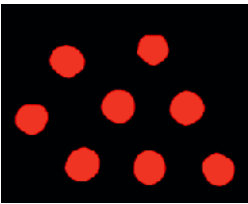


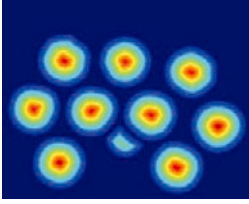
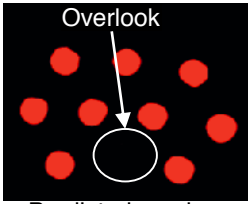
する直前で二値画像に変換される。このとき、過検出したラベルの領域は鋼材領域との区別がなくなり、かつ鋼材領域の方が常に大きいため、実質的にラベルの領域は鋼材領域に取り込まれて除去される。そのため、予測本数はラベルの過検出を除外した本数が出力される。このように、本手法では Mask R-CNN の検出した鋼材領域に対し、形状や径、領域の重複などの識別ルールを適用して検証することで、本数カウントのロバスト性を向上させている。

また見逃し例では、画像奥側へ約 200 mm 凹んだ丸棒鋼が検出されていなかった。この原因は、凹んだ丸棒鋼が周囲に隠れて領域全体が検出されず、距離変換で出力される距離が小さく、閾値の範囲外となり、ノイズとして除去されるためであった。この見逃しは、距離変換後の閾値範囲を緩和することで解消されるが、同時にノイズによる過検出が発生するため、画像処理上での対応が困難である。今後はこの見逃し発生を抑制するため、製造ラインにて撮像直前の端面整列を実施する予定である。

## 4. Deep Learningの実用化に向けた取り組み

これまで、Mask R-CNN を用いて高精度に本数をカウントする手法を提案してきた。しかし、本数カウントを従来の画像処理から Mask R-CNN で代替し、現場導入するためには、安定した動作、かつ高性能な GPU (Graphics Processing Unit) を搭載した PC が必要である。しかし、複数ラインへの導入を見込む場合は導入コスト

Table 6. Example of Mask R-CNN results.

Sample	Input image	Mask R-CNN	Distance transform	Thresholding	Result
Over-detection by deep learning					OK (8 = 8)
Overlook due to hidden bars					NG (10 ≠ 9)

が高額となる。そのため、安価で高性能なPCの選定とMask R-CNNの演算負荷を低減する工夫が必要である。そこで、Mask R-CNNによる本数カウント装置の実用化に向けた取り組みとして、Jetson AGX Xavier<sup>®</sup>による高速化を検討した。

Jetson AGX Xavier<sup>®</sup>は、NVIDIA<sup>®</sup>製のGPUを搭載したエッジコンピュータであり、高性能で低価格、かつ低消費電力なため、本数カウント装置の本番環境用として選定した。本番環境の仕様をTable 7に示す。

Table 7. Production environment specifications.

PC	Jetson AGX Xavier <sup>®</sup>
OS	Ubuntu 18.04
Power consumption	30 W
Software	Python 3.6.9 Pytorch 1.7.1 CUDA 10.2

Jetson AGX Xavier<sup>®</sup>によるMask R-CNNを用いた本数カウントを現場導入するためには、画像1枚あたりのMask R-CNNによる処理時間を0.5秒以下にする必要がある。要求処理時間内に収めるためには、処理速度の高速化が不可欠である。Mask R-CNNは、入力する画像サイズに応じて大量のパラメータを調整するため、膨大な浮動小数点演算が必要である。よって、高速化を実現するには、調整するパラメータ量を減らすべく、入力画像サイズを縮小する、もしくは、演算で使用する浮動小数点の桁数を減らすことが有効である。しかし、これらの方法では扱う情報量の低下に伴う、Mask R-CNNによる鋼材領域の検出精度の低下が懸念される。そこで、Mask R-CNNの演算に対し、従来の単精度浮動小数点数から、半精度浮動小数点数を用いるAutomatic Mixed Precision<sup>14)</sup>（以下、AMPという）への変更と、Miralievら<sup>15)</sup>のように入力画像サイズの縮小を試みながら、処理時間と $J_i$ による検出精度の関係を調査した。

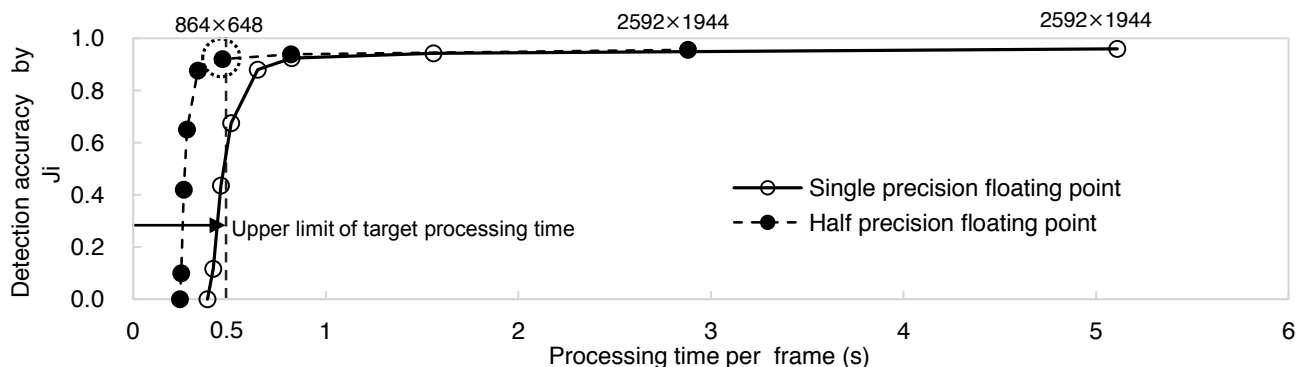


Fig. 6. Detection accuracy and processing time depending on AMP and input image size.

入力画像サイズは、2592 × 1944画素から4:3のアスペクト比を保ったまま、1296 × 972、864 × 648、648 × 486・・・のように、元画像サイズに対して整数倍で縮小した。なお、入力画像には画像縮小による影響が最も大きい、通材する中で最小径の丸棒鋼（入力画像サイズが2592 × 1944画素の場合で、直径が140画素/本）に対して評価した。その結果をFig. 6に示す。検出精度は、両者の浮動小数点数で入力画像サイズが縮小しても、864 × 648画素まではほぼ一定で推移しているが、次の648 × 486画素以下で急激に低下していることが分かる。これは、入力画像サイズを従来の2592 × 1944画素に対して648 × 486画素へと極端に縮小したことにより、丸棒鋼一本当たりの解像度が直径140画素から35画素まで低下し、検出する鋼材領域が過剰に小さくなったことが原因だと考えられる。

また、処理時間はAMPの適用および入力画像サイズの縮小に伴い低下し、目標処理時間0.5秒以下を達成する条件の中で、AMP適用後の864 × 648画素で検出精度が $J_i = 0.92$ と最も高く、高速化前の $J_i = 0.935$ とほぼ同等である。これにより、検出精度を維持しつつ、処理時間を5.11秒から0.46秒と、約11分の1まで低減することが可能となった。

## 5. 結 言

本稿では、結束状態の丸棒鋼における本数カウントに対し、画像処理法およびDeep Learningによりカウントする手法と、実用化に向けた取り組みについて紹介した。

- (1) 機械学習の1つであるDeep Learning (Mask R-CNN)を用いた鋼材領域の検出手法を提案し、画像処理法と比較してカウント精度が向上した。
- (2) Mask R-CNNによる本数カウント装置の実用化に向け、Jetson AGX Xavier<sup>®</sup>による演算を高速化する手

法を提案し、AMPと入力画像サイズの縮小を適用することで、鋼材領域の検出精度を維持しつつ、処理時間を約11分の1まで低減した。

(文献)

- 1) 山田龍三: 電気製鋼, 77(2006), 225.
- 2) 高橋信幸, 渡邊光, 森山稔: 電気製鋼, 93(2022), 29.
- 3) 株みどり精密: 特許第 5479385号(2014).
- 4) 大津展之: 電子情報通信学会論文誌 D, J63-D(1980), 349.
- 5) P. E. Danielsson: Computer Graphics and Image Processing, 14(1980), 227.
- 6) 森大輔, 湯藤隆夫, 岡本有史: 電気製鋼, 90(2019), 53.
- 7) 原田絵美, 森大輔: 電気製鋼, 93(2022), 85.
- 8) D. Chen, D. Guo, S. Liu and F. Liu: Symmetry, 12(2020), 639.
- 9) O. Ronneberger, P. Fischer and T. Brox: U-Net: Convolutional Networks for Biomedical Image Segmentation, 2015, (arXiv:1505.04597).
- 10) K. He, G. Gkioxari, P. Dollár and R. Girshick: Mask R-CNN, 2017, (arXiv:1703.06870).
- 11) T.-Y. Lin, M. Maire, S. Belongie, L. Bourdev, R. Girshick, J. Hays, P. Perona, D. Ramanan, C. L. Zitnick and P. Dollár: Microsoft COCO: Common Objects in Context, 2014, (arXiv:1405.0312).
- 12) P. Jaccard: Bulletin de la Société Vaudoise des Sciences Naturelles, 37(1901), 241.
- 13) L. R. Dice: Ecology, 26(1945), 297.
- 14) P. Micikevicius, S. Narang, J. Alben, G. Diamos, E. Elsen, D. Garcia, B. Ginsburg, M. Houston, O. Kuchaiev, G. Venkatesh and H. Wu: Mixed precision training, 2017, (arXiv:1710.03740).
- 15) S. Miraliev, S. Abdigapporov, J. Alikhanov, V. Kakani and H. Kim: Edge Device Deployment of Multi-Tasking Network for Self-Driving Operations, 2022, (arXiv:2210.04735).



木全謙太



吉田たける