

# 物体認識アルゴリズム YOLO を利用した鉗子検出 AI の構築

北里大学医療衛生学部医療工学科臨床工学専攻

藤本 浩徳

## 要旨

医療器材の管理業務は、安全な医療を提供する上で重要な役割を担っている。本研究では物体認識による医療器材管理システムの開発を目指し、医療現場で使用されている鉗子を検出する AI を構築し、評価を行った。鉗子検出 AI は、500 枚の鉗子画像と対応するアノテーションデータから学習し、AUC 0.981、感度 0.988、特異度 0.980 となり高い鉗子検出能力を示した。物体認識技術が、医療器材に応用可能であることが示唆された。

## 背景・目的

臨床工学技士が行う業務の 1 つに医療器材の管理業務が存在する。例えば人工呼吸器や血液浄化装置、人工心肺装置といった生命維持管理装置の保守点検業務が挙げられる。医療器材管理業務は現代の高度な医療を安全かつ円滑に提供する上で重要な役割を担っている。さらに、医療現場では大小合わせて数多くの医療器材を使用している。特に手術時に使用する器材は、種類および数の管理が不十分であると円滑な手術ができないだけでなく、異物残存などの医療事故となる危険性がある。このような事故を防ぐために、術前、術中、術後に種類とその数を目視確認している。しかしながら目視カウントミスによるインシデントは約 1 年間で 7 回程度発生し、紛失は 6 回発生したという報告<sup>1)</sup>がある。ガーゼの異物残存事故は 2016 年から 2018 年で年間 20 件程度報告されている<sup>2)</sup>。このように医療現場では器材カウントを徹底する必要性が求められている<sup>1)</sup>。

器材管理は、非接触 RFID タグによるシステムが開発されるなど機械化も進んでいる。RFID タグシステムは、器材の種類によって効率的なタグの設置方法等を実用化に向けて検討されている段階である。このような背景をふまえ、病院内にある監視カメラ等の映像から物体を認識する技術を利用した医療器材の管理システムとして、手術や処置で共通して用いられる代表的な医療器材である鉗子を画像から検出し数量を認識する AI を構築し評価を行った。

## 方法

### 1. 物体認識アルゴリズム YOLO<sup>3)</sup>

物体認識とは、1 枚の画像に写っている多種または複数の物体の一般名と位置を特定することである。一般名の特定はクラス分類であり、検出した物体がそのクラスに属する確率が求められる。位置の特定は、検出した物体が実際に特定した位置に存在する確率が求められる。この二つを掛け合わせて得られた値を信頼度と呼ぶ。信頼度は、一般名と位置の特定がどの程度うまく行っているかの指標である。本研究における鉗子の認識は、物体認識アルゴリズム YOLOv2 608×608 (以下 YOLO<sub>608</sub>) を基に新たに鉗子認識 AI を構築した。YOLO<sub>608</sub> は、物体認識アルゴリズムの中で最速の物体検出速度<sup>3)</sup>を実現している。公開されている学習済み AI である YOLO<sub>608</sub> は約 80 種類の物体を識別可能であるが、医療器材を検出することはできない。しかし、独自に学習用のデータセットを作成し、新規に学習を行うことで新たな医療器材認識 AI を構築することが可能である。

### 2. 学習用データセット作成

今回検出する医療器材は、手術、処置および透析室など多くの医療現場で共通して用いられ、数の管理が煩雑

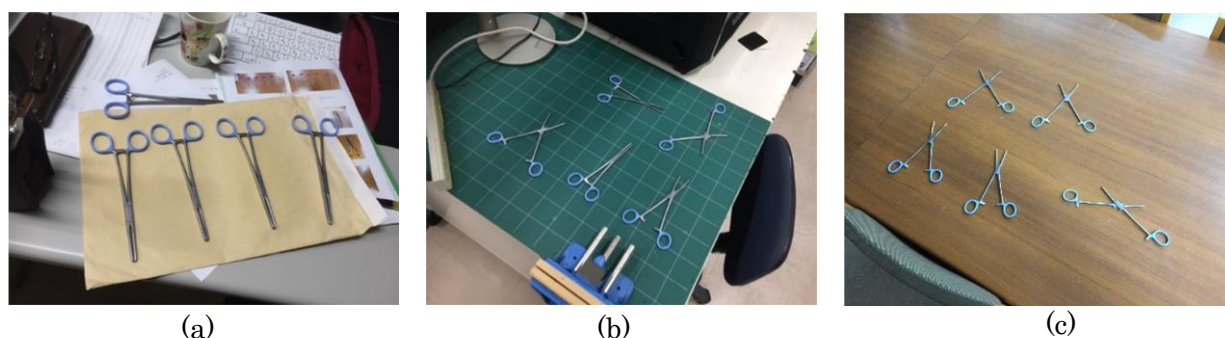


図 1 鉗子学習用データセット例

鉗子を検出する YOLO<sub>鉗子</sub> の学習に使用した 500 枚の鉗子画像例。背景、鉗子の本数および状態等多様な画像を学習させるため考慮した。(a: 茶背景、鉗子 5 本、全て閉じた状態。b: 緑背景、鉗子 5 本、一部の鉗子が開いた状態。c: 木目背景、鉗子 5 本、鉗子が全て開いた状態。)

となる鉗子とした。学習用データセットは、様々な環境、撮影距離、撮影角度から複数の鉗子を密着もしくは重なりあわないように撮影した 500 枚の画像 (図 1) を用いた。500 枚の学習用画像 1 枚 1 枚に、画像中に写っている鉗子の位置と大きさを記述したテキストデータであるアノテーションの付与を行った。未学習の YOLO<sub>608</sub> が実装されたライブラリ darknet<sup>4)</sup>を使って、学習用画像と対応するアノテーションからなる学習用データセットを学習させ、鉗子検出 AI である YOLO<sub>鉗子</sub>を新たに構築した。画像中の鉗子の数量に注目した場合、YOLO<sub>鉗子</sub>は画像に鉗子があるか、ないかの 2 群を判別する 2 値のクラス分類器とみなすことができる。

### 3. YOLO<sub>鉗子</sub>の識別能力評価方法

ROC 曲線 (Receiver Operating Characteristic curve) は、医学では診断や検査など罹患有無のような 2 群判別の識別能力の評価に使用される。ROC 曲線は縦軸に真陽性率、横軸に偽陽性率をプロットする。AUC (Area Under the Curve) と呼ばれる ROC 曲線の下側の面積が大きいほど 2 群の識別能力が高いことを意味する。真陽性率、偽陽性率を軸とした曲線下の面積なので、AUC の最大値は 1.0 となる。一方、AI 分野ではクラスに属する (鉗子であるか) か、属さないか (鉗子でないか) といった 2 群を識別する AI の評価に用いられる。今回構築する AI は手術器材の数量をカウントするために使用する。そのため鉗子有無の 2 群を判別する AI である YOLO<sub>鉗子</sub>は、AUC や混同行列を用いた評価を行うことになる。YOLO<sub>鉗子</sub>の識別能力の評価のために、学習用とは別に新たに 200 枚の評価用画像を用意した。評価用データセットの内訳は、「鉗子が写っている鉗子あり画像」と「鉗子が写っていない鉗子なし画像」を 100 枚ずつ計 200 枚用意した。鉗子あり画像は、1~5 個の鉗子を規則正しく並べた画像と乱雑に 5 個鉗子を並べた画像を 50 枚ずつ用意した。鉗子なし画像は、部屋の背景をランダムに撮影した画像とハサミやスパナなどの形や材質が比較的似ているものを撮影した画像 50 枚を用意した。ROC 曲線作成にあたり鉗子を正しく鉗子であると判断できているものを真陽性 (True Positive、以下 TP)、鉗子ではないものを正しく鉗子ではないと判断したものを真陰性 (True Negative、以下 TN)、ハサミ等の鉗子ではないものを鉗子であると誤って判断したものを偽陽性 (False Positive、以下 FP)、鉗子があるにもかかわらず鉗子と判断できなかったものを偽陰性 (False Negative、以下 FN) と定義した (図 2)。200 枚の評価用データセットから YOLO<sub>鉗子</sub>で鉗子を検出し、検出された物体の信頼度がある閾値以上の場合を「鉗子である」、未満の場合を「鉗子ではない」とした。なお TN、FN の場合は信頼度を 0 とした。0~1 の範囲で閾値を変化させ、閾値ごとの混同行列から真陽性率、偽陽性率を算出し、ROC 曲線を作成した。ROC 曲線から Youden index を用いて最適閾値を求め、その時の混同行列から感度、特異度を評価指標として求めた。

### 4. 撮影条件検証方法

YOLO<sub>鉗子</sub>は、手術室の天井などに設置したカメラから複数の手術器材を検出し数量を認識するシステムでの使用を想定しているため、「鉗子とカメラの距離」、「画像の画素数」、「鉗子の向き」、「鉗子に対するカメラの角度」で検出結果に違いが生じるのか検証を行った。

#### 4.1 鉗子とカメラの距離についての検証

画素数を約 800 万ピクセルで一定にした。机に配置した鉗子から 22.5 cm~230 cm までの距離でカメラを設置し、それぞれの距離から鉗子が画面中央に写るよう真上から撮影を行った。YOLO<sub>鉗子</sub>が撮影画像から鉗子を検出したときの信頼度を比較した。

#### 4.2 画像の画素数についての検証

机に配置した鉗子から 50 cm の距離でカメラを設置し、約 800 万ピクセルの画素数で鉗子が画面中央に写るよう真上から撮影を行った。この 1 枚の画像の画素数を徐々に少なくした画像を用意した。YOLO<sub>鉗子</sub>がこれら

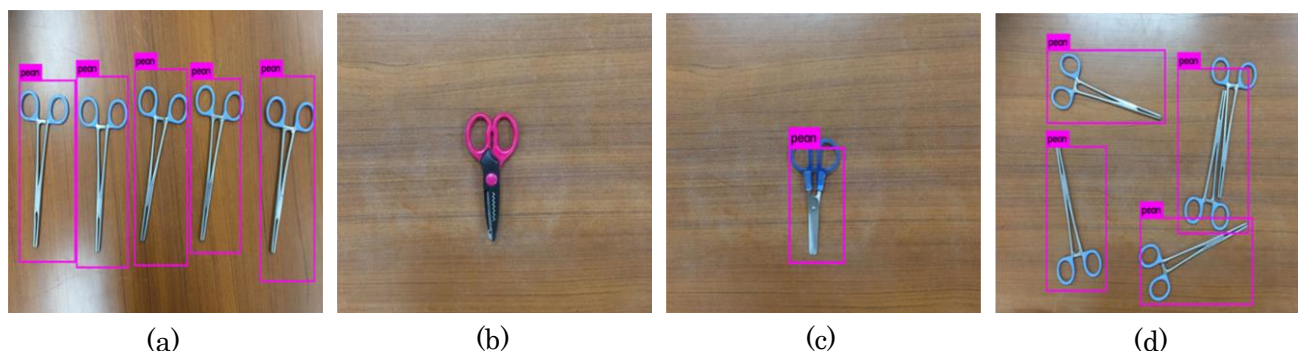


図 2 YOLO<sub>鉗子</sub>による検出結果の例

(a) : TP の例。5 つの鉗子をすべて正しく鉗子(pean)と識別した。(b) : TN の例。はさみを鉗子ではないと正しく識別した。(c) : FP の例。はさみを鉗子であると誤認した。(d) : FN の例。密着した 2 つの鉗子のうち 1 つを鉗子ではないと誤認識した。

の画像から鉗子を検出したときの信頼度を比較した。

#### 4.3 画像中の鉗子の向きを検証

机に配置した鉗子から 50 cm の距離でカメラを固定し、約 800 万ピクセルの画素数で鉗子が画面中央に写るよう真上から撮影を行った。鉗子を 0~330 度まで約 30 度ずつ回転させ撮影した 12 種類の画像から YOLO<sub>鉗子</sub> が鉗子を検出したときの信頼度を比較した。

#### 4.4 鉗子とカメラの角度についての検証

鉗子の長軸を Y 軸、短軸を X 軸として、机に固定した鉗子を約 800 万ピクセルの画素数で画面中央に写るよう撮影を行った。カメラを鉗子 Y 軸 X 軸に対して 0~150 度まで約 30 度ずつ変え、12 種類の画像を撮影した。この時どの角度においても鉗子とカメラの距離が 50 cm となるように撮影を行った。これらの画像から YOLO<sub>鉗子</sub> が鉗子を検出したときの信頼度を比較した。

## 結果

### 1. YOLO<sub>鉗子</sub> の識別能力評価結果

ROC 曲線は真陽性率 1、偽陽性率 0 となるグラフ左上の点に曲線が近づくほど識別能力が高いことを示し、同時に AUC は 1 に近づくことが知られている。YOLO<sub>鉗子</sub> の ROC 曲線はその曲線の形状と AUC 0.981 であり非常に高い識別能力を持つことが示唆された (図 3)。YOLO<sub>鉗子</sub> は、信頼度が最適閾値の 0.88 のときの混同行列 (表 1) から感度 0.988、特異度 0.980 で鉗子の有無を判断できた。

### 2. 撮影条件検証結果

鉗子が撮影された画像に対して、最適閾値以上の信頼度で鉗子を検出できれば TP (鉗子を鉗子と正しく検出する) となる。一方、最適閾値未満の信頼度で鉗子を検出した場合は FN (鉗子があるにもかかわらず検出できない) となる。従って、信頼度が 0.88 以上となる撮影条件を検証する必要がある。

#### 2.1 鉗子とカメラの距離についての検証

鉗子とカメラの距離を変化させた時の鉗子検出の信頼度は、至近距離 22.5 cm で最も信頼度が低く、50 cm 付近をピークに鉗子とカメラの距離が離れるに従い信頼度も徐々に低下した (図 4)。200 cm 以上で信頼度 0.88 未満であった。42.5 cm~190 cm であれば鉗子を正しく検出できることが分かった。

#### 2.2 画像の画素数についての検証

画素数を変化させた時の鉗子検出の信頼度は、約 7000 ピクセル以下の画素数で信頼度が 0.88 未満となり、6500 ピクセルで検出不能であった (図 5)。約 8000 ピクセル~800 万ピクセルの画像であれば、鉗子を正しく検出できることが分かった。

#### 2.3 画像中の鉗子の向きを検証

撮影されている鉗子の向きを変化させた時の鉗子検出の信頼度は、0.92~0.95 となり 0.88 未満となることはなかった。よって検証を行ったすべての鉗子の向きについて、問題なく鉗子を検出できることが分かった。

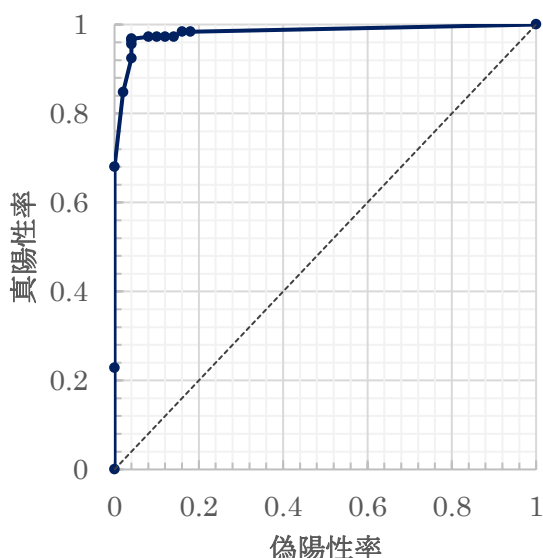


図 3 YOLO<sub>鉗子</sub> の ROC 曲線

200 枚の評価用データセットにおける鉗子識別結果から信頼度の閾値ごとに真陽性率と偽陽性率を求めプロットした。

AUC は 0.981。Youden index による最適閾値は 0.88。

表 1 最適閾値時の混同行列

		Actual class	
		鉗子あり	鉗子なし
Predicted class	鉗子あり	395	2
	鉗子なし	5	100

Actual class は実際に撮影画像に鉗子が写っているか、いないかを示す。一方 Predicted class は撮影画像から鉗子検出 AI が鉗子を検出し、鉗子があるかないかを判断した結果。TP は 395、FP は 2、FN は 5、TN は 100 となった。FP ははさみを誤検出し、FN は鉗子の密着または重なりによる未検出が主な原因。

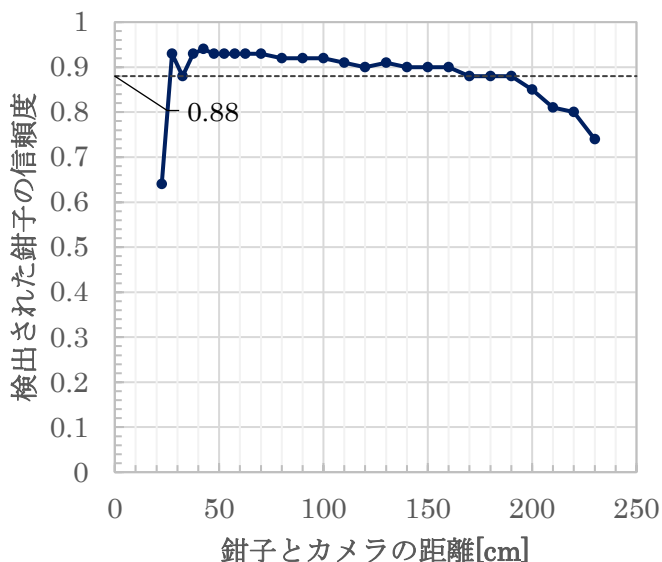


図4 鉗子からのカメラの距離と信頼度

机に配置した鉗子とカメラの距離を変化させ、約 800 万ピクセルで撮影した時の鉗子検出信頼度の変化。

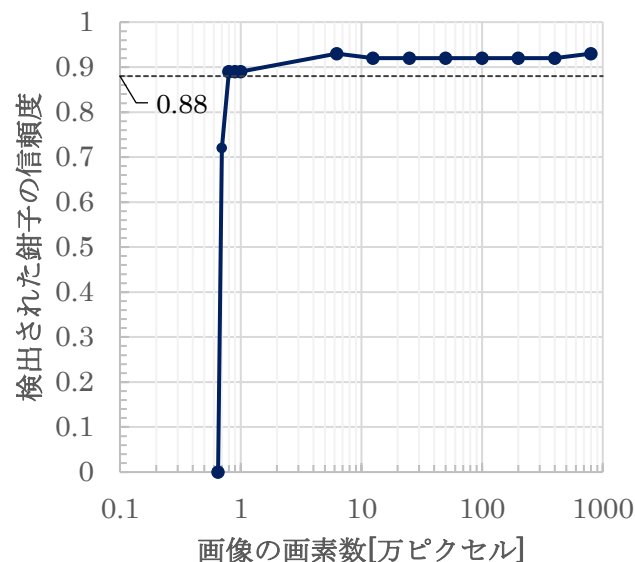


図5 撮影画素数と信頼度

50 cm の位置から 800 万ピクセルの画像を撮影し、画像処理で画素数を徐々に低下させたときの鉗子検出信頼度の変化。

#### 2.4 鉗子とカメラの角度についての検証

どの角度においても信頼度は 0.90~0.95 となり 0.88 未満となることはなかった。よって検証を行ったすべての鉗子とカメラの角度について、問題なく鉗子を検出できることが分かった。

### 考察

#### 1. YOLO 鉗子の識別能力評価

YOLO 鉗子は非常に高い識別能力があるといえる。検出ミスの多くは、鉗子同士が密着している、鉗子同士が重なっている場合であった。学習用データセットで鉗子の密着や重なりを除いた影響を受けていると考えられた。そのため手術器材の個数を検出する際には器材同士を密着させすぎず、ある程度の距離を保つといった条件が必要である。識別能力評価では学習用と評価用にデータセットを分割したホールドアウト検証であり k 分割の交差検証ではないため、データセットを変更した場合に同等の性能が得られるかは検討が必要である。

#### 2. 撮影条件検証

鉗子の向きと鉗子に対するカメラの角度は鉗子検出にほとんど影響がないと考えられる。すなわち必ずしも手術器材の真上にカメラを設置する必要はなく状況に応じてカメラを自由に設置することが可能である。鉗子からの距離と撮影画像のピクセルでは、それぞれ正しく検出できる範囲が判明した。これらの条件から、撮影に使用するカメラのレンズ画角や画素数といった基本性能を決定できるようになった。

#### 3. 医療機器管理システム

医療機器のなかには手術器材のように RFID タグシステムで用いられるタグなどを装着できないものがある。物体認識を利用した手術器材の検出による管理システムはこのような欠点を補うことができる可能性がある。また検出したい機器の学習用データセットを用意すれば、追加して検出可能である。医療器材である鉗子を高い識別能力で検出できたため、他の手術器材や複数の医療機器の検出も十分に可能であることが示唆された。

### まとめ

独自の学習用データセットを用いて高い識別能力で鉗子を検出できる AI を構築した。他の医療器材を含めて医療機器に応用可能なシステムに組み込むことで AI を用いた医療機器管理システムの構築が可能であることが示唆された。

### 参考文献

- 1) 平山陽子, ほか: A 大学病院中央手術部におけるインシデントレポートの分析, 手術医学, 39(3), pp.139-142.
- 2) 公益社団法人日本医療機能評価機構. 事故情報収集等事業第 54 回報告書. pp.74-77. 2018. [http://www.med-safe.jp/pdf/report\\_54.pdf](http://www.med-safe.jp/pdf/report_54.pdf)(参照 2020-03-01).
- 3) J. Redmon, et al. You Only Look Once: Unified, Real-Time Object Detection. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.779-788, 2016.
- 4) J. Redmon. Darknet: Open Source Neural Networks in C. 2013-2016. <http://pjreddie.com/darknet/>(cited 2020-03-09).