

画像解析による頸椎可動域角度計測補助システム

松尾 佳奈^{*1} 藤田 浩二^{*2} 小山 恭史^{*2}

森下 真伍^{*2} 杉浦 裕太^{*1}

Cervical Spine Range of Motion Angle Measurement Assistance System by Image Analysis

Kana Matsuo^{*1}, Koji Fujita^{*2}, Shingo Morishita^{*2}, Takafumi Koyama^{*2}, and Yuta Sugiura^{*1}

Abstract — Diseases of the cervical spine often cause more serious disabilities in activities of daily living than diseases of other parts of the body, and thus require prompt and accurate diagnosis. One of the indicators to diagnose cervical spine diseases is to measure the range of motion angle. The main measurement method is manual, which is a burden for doctors. In this paper, we investigate the possibility of measuring the range of motion angle of the cervical spine from cervical X-ray images using Mask R-CNN and image processing. As a result of measuring the range of motion angle with the proposed cervical spine motion angle measurement aid system, the mean error from the true value was 3.1 degrees and the standard deviation was 2.5 degrees. The standard deviation of the specialist measurements was 2.9 degrees, while that of the proposed system was 0 degrees, indicating that there was no variation in the measurements of the proposed system.

Keywords : measurement system, deep learning and cervical spine

1. はじめに

医療において、レントゲン画像など多数の画像検査が行われているが、医師が多量の画像を読影するには時間を要し診療での時間の負担となっている。さらには夜間など、少人数の当直医で診察を行う時間帯では、専門外の領域の画像診断を求められることがある。そのため、近年AIによる自動画像診断の研究が多く行われている。

頸椎は整形外科を中心に多くの画像検査が行われる部位の一つである。加齢などで頸椎の変性が進行すると、頸部痛の他、脊髄症や神経根症といった神経疾患を引き起こす^[1]。特に脊髄症が悪化した場合、手指の巧緻運動障害や歩行障害を引き起こし、日常生活に重大な支障を来すため、迅速かつ正確な診断が要求される。

頸部を動かすとC1-C7(第一頸椎-第七頸椎)間のそれぞれで動きが発生する。特に前後屈の動作に関しては、頸椎レントゲン画像から椎体間の可動域を計測し、頸椎の不安定性を評価することがある。この評価にはC1-C7間の各可動域角度を計測する必要がある。医師が行う計測の1つであるBakke法では、頸部を前屈と後屈させたときのレントゲン画像を撮り、各椎体間の上縁と下縁に線を引く。次にそれらの線がなす角度

を求め、前屈のなす角と後屈のなす角の差を取る。ただし、C1・C2は形状が特殊であるため、C1は前弓後縁に、C2は椎体後縁に線を引く。それらの線がなす角度を求める(図1)。この方法では前屈・後屈のレントゲン画像上に合計24本の線を引くため医師の負担となっている。また、C1・C2頸椎は形状が特殊なため正確な計測が困難であるという課題がある。以上から本研究では、頸椎レントゲン画像からの可動域角度計測を自動化するシステムを実装することで医師の可動域角度計測の支援を目指す。

2. 関連研究

2.1 医療画像セグメンテーション

画像処理を用いる方法として、Huaifeiらはコントラストのヒストグラムによって大まかな頸椎領域を抽出し、カーブフィッティングを用いて領域を推定する方法を提案した^[2]。また、Lecronらは頸椎のセグメンテーションのためにエッジ検出とハフ変換などを用いて各椎骨の角を検出する手法を提案した^[3]。このような画像処理技術を用いたセグメンテーションでは、多くのパラメータを設定し閾値を決めねばならず実装が複雑になりやすい。また、ヒストグラムに注目する場合は頭部が影響しないようにするため、元のレントゲン画像から頸椎部位だけを切り出す処理が必要になる。

近年では深層学習を用いたセグメンテーション手法も提案されている。セグメンテーション技術にはFully Convolutional Networks (FCN)^[4]、U-Net^[5]、Mask R-CNN^[6]などがある。ArifらはFCNベースのSP-

^{*1}: 慶應義塾大学 理工学研究科

^{*2}: 東京医科歯科大学 大学院医歯学総合研究科

^{*1}: Faculty of Science and Technology, Keio University

^{*2}: Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University

Net という Convolutional Neural Network (CNN) を提案し頸椎をセグメンテーションした^[7]. U-Net などの既存手法と比べて、頸椎部分に異物が写っている画像についても頑健に頸椎領域を検出できる. 魚住らは Mask R-CNN を用いて胸部レントゲン画像から肺領域を高精度に抽出した^[8]. また、彼らは U-Net を利用した場合と比較しており、抽出精度に関しては U-Net の場合のが高いが、Mask R-CNN の場合は疾患によるレントゲン画像の陰影の変化や肺の位置の変化に伴う多様性に対するロバスト性が高いと述べた. 医療画像におけるセグメンテーションは一般的な物体のセグメンテーションに比べて疾患によって検出対象の形状が多様に変化することがある. また、Mask R-CNN は他のセグメンテーションに特化したネットワークと異なりインスタンスセグメンテーションを行うため、頸椎の各椎体を区別して処理できる利点もある. このため本研究では Mask R-CNN を用いて頸椎レントゲン画像から頸椎領域を抽出する.

2.2 診断支援システム

セグメンテーション技術や画像処理、深層学習を用いて医療画像の分析・診断を行い、医師の診断を支援する研究が行われている. Choi らはモアレ画像と腰椎レントゲン画像のセットを CNN によって学習させることでモアレ画像から脊柱配列を推定し、脊柱のカーブの大きさを表すコブ角を計測するシステムを開発した^[9]. 真値との平均誤差は 3.8 度であり専門医による計測と同等の精度を実現した. Alomari らは画像の輝度と勾配ベクトル流を用いて腰椎の椎間板をセグメンテーションし、椎間板の高さを計測し椎間板の異常を検知するシステムを開発した^[10]. Young らは深層学習と画像処理を用いて頸椎レントゲン画像から椎体前軟部組織の厚さを計測し、医師によるむくみ診断の流れを参考にしてむくみを診断する手法を提案した^[11]. 本研究では椎体間の可動域角度計測に注目し、医師の診断の流れを参考にして実装したシステムを用いて頸椎の可動域角度を計測した.

3. 提案システム

3.1 システムの概要

本研究で提案するシステムの流れを図 2 に示す. まず Mask R-CNN によって頸椎レントゲン画像から頸椎領域を推定する. 推定した領域を 3 または 4 の頂点数の単純な多角形に近似する. 頸椎 n と頸椎 $n+1$ の多角形の辺の中心が近い辺同士を椎体間にある辺とする. その辺の傾きによって前屈画像と後屈画像における椎体間のなす角をそれぞれ計測し、差を求めることで可動域角度を求める.

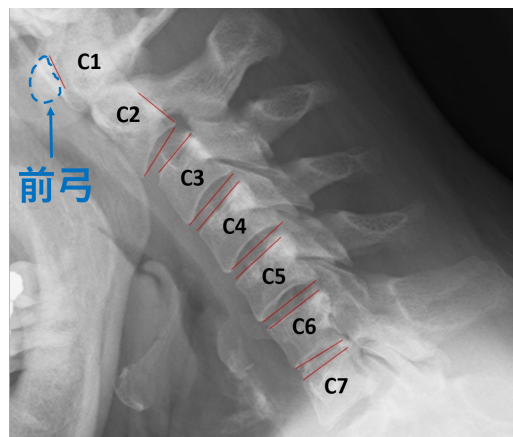


図 1 計測のためのアノテーション例
Fig.1 Example of annotation for measurement

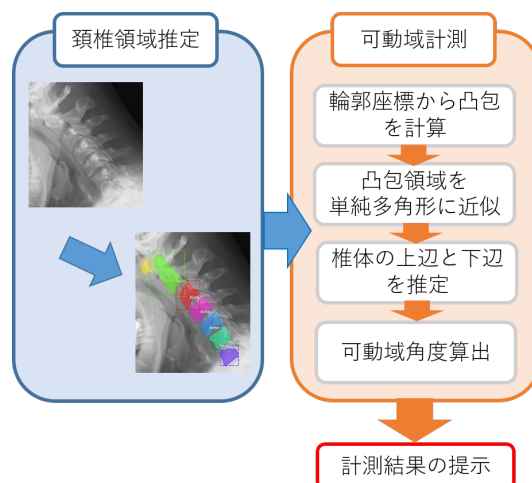


図 2 システム構成
Fig.2 System configuration

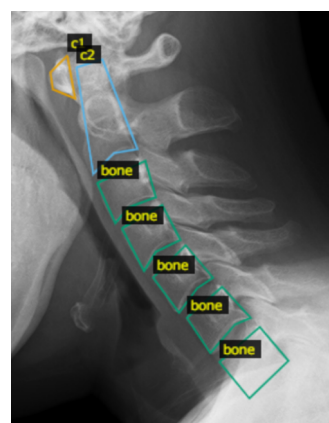


図 3 アノテーション領域
Fig.3 Annotation area

3.2 Mask R-CNN による頸椎領域の推定

学習データは頸椎を前屈・後屈させた際のレントゲン画像である. 頸椎専門の放射線科医の指導に基づき頸椎マスク情報を作成し、原画像とのペアを学習デー

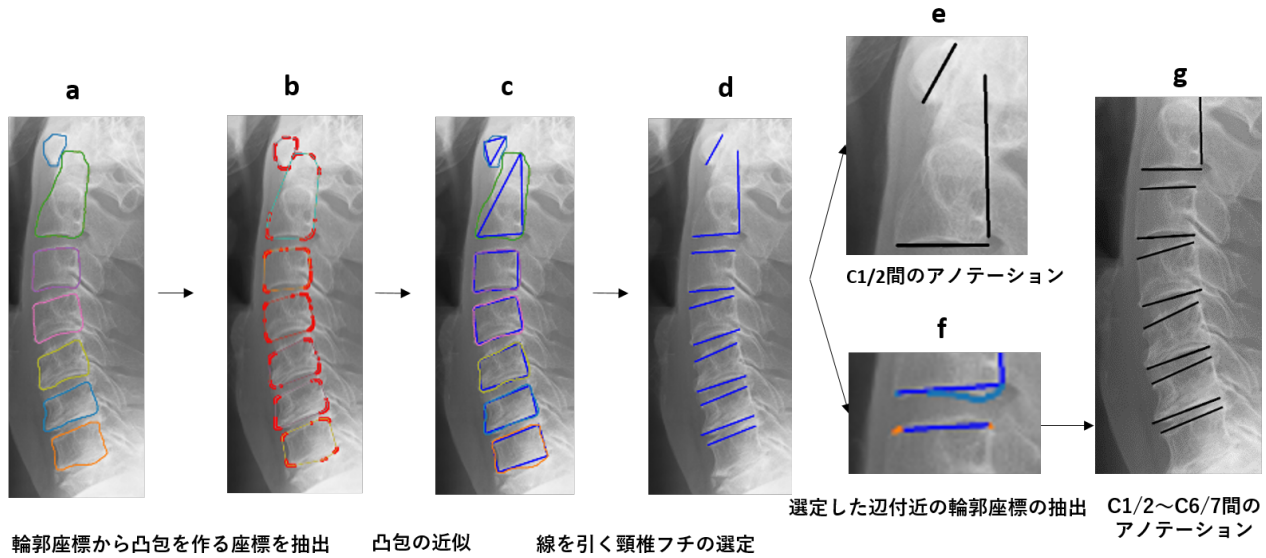


図4 自動アノテーションまでの流れ
Fig.4 Flow of automatic annotation

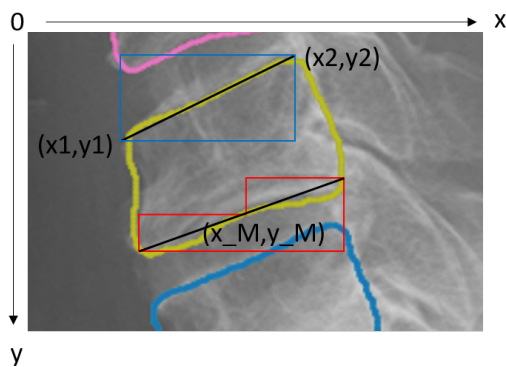


図5 線が通る座標の抽出範囲
Fig.5 Extraction range of coordinates
through which the line passes

タとした。マスクする頸椎領域は、C1とC2は形状が特殊なためC3以下とは異なり、C1領域は前弓とし、C2領域は椎体から歯突起まで含んだ部位とした。領域のラベルは、C1、C2、その他の頸椎とした(図3)。

3.3 推定頸椎領域からの可動域角度計測

Mask R-CNNによって推定された領域の矩形座標と面積情報は配列に格納されているがC1-C7の順に並んでいない。このため、両画像の推定領域の頸椎位置を対応させる必要がある。計測準備として、矩形のy座標を昇順にし、推定領域をC1-C7の順に並び変える。しかし画像中の一部の頸椎領域を推定できない場合がある。この場合は頸椎nと頸椎n+1の矩形の左上と右下のy座標の差分を取り、頸椎nの矩形の高さの半分以下であれば、頸椎nと頸椎n+1は連続しているとみなした。この条件に外れたものは誤

推定領域または、推定不可の頸椎領域があると判定できる。次に推定頸椎領域から領域の輪郭座標を取得する(図4a)。輪郭座標から凸包を作る座標を選択し、凸包領域を3または4の頂点数の多角形に近似する(図4b c)。このあと近似した多角形の各辺の中点を算出する。頸椎nと頸椎n+1の間で算出した中点同士が近くなる辺を計測に用いる頸椎の縁とみなす(図4d)。ただし、C3以下の近似多角形は凸包領域に内接するように近似されるため、近似多角形の辺が推定領域の輪郭座標から離れている場合がある。そのため、近似多角形における選択された辺付近にある輪郭座標を抽出する(図4f)。近似多角形における選択された辺の左端の座標をx1, y1とし、右端をx2, y2、中点をx_M, y_Mとする(図5)。輪郭座標のx座標群をverts_x, y座標群をverts_yと表すと、注目する椎体の上辺は(1)式の条件を満たす輪郭座標であり、図5の青枠内にある輪郭座標である。下辺は(2)式の条件を満たす輪郭座標であり、図5の赤枠内にある輪郭座標である。それらの座標に最小二乗法を用いて、なす角度を求めるための線を引き(図4g)。引いた2線の傾きをa, bとすると、なす角度はタンジェントの加法定理から(3)式によって算出される。

$$(x1 \leq \text{verts_x} \leq x2) \cap (y2 \leq \text{verts_y} \leq y1) \quad (1)$$

$$\begin{cases} (x1 \leq \text{verts_x} \leq x_M) \cap (y_M \leq \text{verts_y} \leq y1) \\ (x_M \leq \text{verts_x} \leq x2) \cap (y2 \leq \text{verts_y}) \end{cases} \quad (2)$$

$$\tan\Theta = \frac{a-b}{1+ab} \quad (3)$$

4. 評価

4.1 頸椎領域の学習モデルの条件

本研究で用いたデータセットは、東京医科歯科大学医学部附属病院の放射線科に通う 20 歳以上 100 歳未満の患者 484 人における頸椎を前屈・後屈させた際のレントゲン画像の 968 枚である。この 968 枚のうち 922 枚（患者 461 人）を学習データ、残りの 46 枚（患者 23 人）をテストデータとした。

初期重みには Microsoft COCO により学習された重みを用い、作成した学習データによってネットワーク全体を再度学習することで重みを更新した。学習パラメータを表 1 に示す。学習データ 922 枚はトレーニングデータとバリデーションデータの比率が 8 対 2 になるように分けた。本実験の学習のパラメータは以下の通りである。

表 1 ハイパーパラメータ
Table 1 Hyperparameters

トレーニングデータ	756 枚
バリデーションデータ	166 枚
クラス分類	3 クラス分類
学習回数	100 epoch
画像サイズ	512× 512 px
バッチサイズ	1
学習係数	0.001
最適化手法	確率的勾配降下法

4.2 頸椎領域の学習モデルの学習結果と評価

テストデータ 46 枚を推定した。46 枚に写っている合計 322 個の頸椎のうち、推定できた頸椎は 313 個であった。このことから、全テストデータの 97%は頸椎を検出できたといえる。推定頸椎領域の評価には 2 つの集合の類似度を表す Intersection over Union (IoU) を用いる。IoU は正解領域と推定領域の共通部分を 2 つの領域の和集合で割った値であり、最大値は 1 で表される。IoU の値が大きいほど高精度に物体検出できている。各椎体の IoU と全体平均 IoU の値を表 2 に示す。テストデータ 46 枚における IoU 平均値は 0.85 であった。

4.3 研修医計測と提案システム計測の精度比較

ここでは専門医の計測を真値とし、真値と自動計測値、真値と研修医計測値の平均誤差を比べることで自動計測手法の計測精度を評価した。

計測対象のデータは頸椎領域の学習モデルの検証に用いたテストデータ 46 枚（患者 23 人）である。真値の計測は、専門医 2 名に 23 人（テストデータ）× 3 回（計測回数）の 69 回計測してもらった。計測頻度は 1 日 1 回までとし、連続で計測しないこととした。

表 2 各椎体の IoU
Table 2 IoU of each cervical spine

部位	IoU
C1	0.74
C2	0.83
C3	0.88
C4	0.88
C5	0.87
C6	0.86
C7	0.86
全体平均	0.85

真値は専門医 2 名における各 3 回分の計測値の平均とした。研修医の計測は、専門医から頸椎可動域角度計測の指導を受けた研修医 2 名に、専門医と同様の条件でテストデータを計測してもらった。テストデータにおける可動域角度計測において、自動計測できた箇所は 23 人 × 6 箇所（可動域角度の計測箇所）の 138 箇所のうち 133 箇所であった。計測できなかった残り 5 箇所の原因は、頸椎領域が推定できなかったためである。

真値と自動計測値の誤差は各椎体間において真値と比較する測値の差分を絶対値にして平均化し、計測誤差を算出した。真値と研修医の誤差は研修医計測の各 3 回において各椎体間の真値と研修医計測値の差分を絶対値にしたあと、それらを平均化し算出した。しかし本研究の手法では頸椎領域が正確に推定されなかった場合、可動域角度を正しく計測できない。自動計測をする際に画像に引かれた線が明らかに医師が線を引くであろう箇所と離れている場合はその箇所の値を外れ値とし、評価対象から外した（図 6）。外れ値は 10 個あると判定し、138 箇所のうち 123 箇所を誤差算出の評価対象とした。自動計測と研修医計測の各椎体間の可動域角度の真値との誤差平均は表 3 のようになった。また、専門医と研修医による各椎体間の計測値において標準偏差を算出した結果を表 4 に示す。自動計測においては 3 回計測を実行しても同じ値が出るため標準偏差は 0 度であった。

真値と研修医の誤差、真値と自動計測値の誤差を比較した。図 7 に研修医計測と自動計測の平均誤差の比較結果を示す。研修医計測の全体平均誤差と比べて、自動計測は 0.8 度小さい。この差が統計的に有意かを確かめるために、有意水準 5 % で両側検定の t 検定を行った。変数 1 は研修医の誤差値であり、サンプル数は研修医（2 名）× 計測数（3 回）× 頸椎間 123/138 箇所の 738 個である。変数 2 は自動計測の誤差値であり、システムの計測数（6 回）× 頸椎間 123/138 箇所の 738 個である。p=0.0002(<0.05) であり研修医計測と自動計測の平均誤差の差に有意差が見られた。

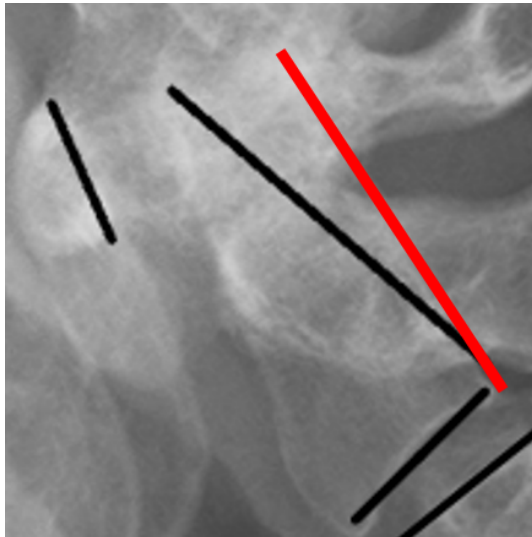


図6 外れ値とみなす基準 (赤: 医師の引く線, 黒: 自動で引いた線)

Fig. 6 Criteria for considering an outlier (Red: physician line, Black: automatic line)

表3 真値との平均誤差

Table 3 Average error compared to the specialist measurements

椎体間	自動計測 (度)		研修医計測 (度)	
	平均誤差	標準偏差	平均誤差	標準偏差
C1/2	5.0	4.0	6.6	11.0
C2/3	2.8	1.9	3.1	2.5
C3/4	2.5	2.4	3.3	2.6
C4/5	2.6	2.1	3.8	3.1
C5/6	3.6	2.3	3.3	2.4
C6/7	2.5	2.2	3.3	2.6
全体平均	3.1	2.5	3.9	4.0

表4 真値と研修医計測値のばらつき

Table 4 Standard deviation of true value and resident measured value

椎体間	真値 (度)	研修医計測値 (度)
	標準偏差	標準偏差
C1/2	4.4	5.2
C2/3	2.4	2.6
C3/4	2.9	2.9
C4/5	2.5	3.0
C5/6	2.7	2.7
C6/7	2.4	2.7
全体平均	2.9	3.2

5. 議論

図8に高精度に頸椎領域を推定できた例をあげる。緑色が正解領域を表し、赤色が推定領域を表す。この例はC1-C7まで推定でき、C1-C7のIoUを平均した値が0.85を超えている。しかしながら、図9のように一部の頸椎を誤って推定することもある。高精度に推定できなかったケースを図9に示す。ケースa,c,dが起る原因は学習データが少ないことが考えられ、ケース



図7 研修医計測と自動計測の平均誤差の比較

Fig. 7 Comparison of mean error between resident measurements and automatic measurements

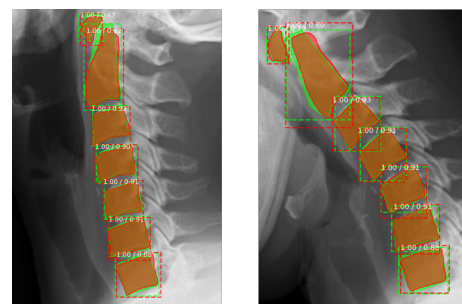


図8 推定を高精度にできた例

Fig. 8 Example of correct estimation

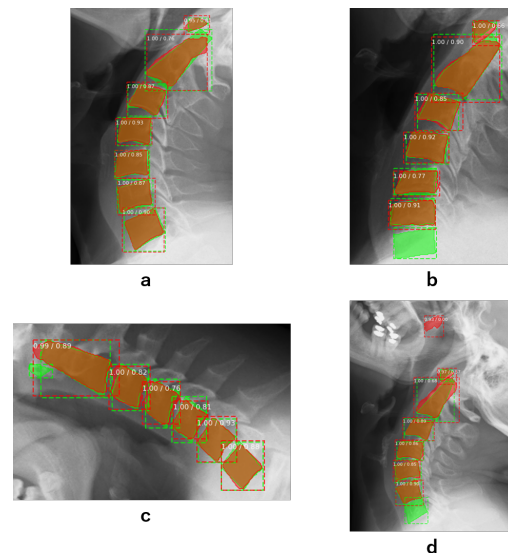


図9 推定を誤った例; 推定頸椎領域が大きく欠損 (a); C6やC7が患者の胸部に隠れて推定不可 (b); C1が推定不可 (c); 頸椎以外の領域が頸椎と推定 (d)

Fig. 9 Example of incorrect estimation ; Large missing estimation area (a); C6 and C7 are hidden in the patient's chest and cannot be estimated (b); C1 not estimable (c); Non-cervical region estimated to be cervical spine (d)

bはレントゲンの写りの問題であると考えた。しかしながら患者の写り方を変えるのは難しいため、医師が計測した過去の可動域角度の値と画像をセットで学習させ、回帰分析を用いることで本手法では計測できなかった箇所の可動域角度を計測できるか議論したい。

4.2節の結果より、各頸椎部位のIoU値を比較すると、特にC1のIoU値はほかの頸椎部位より値が小さい。また、C2のIoU値もC3-C7と比べると小さい傾向が見られた。このことから、C1・C2の頸椎領域の推定精度が低いことがわかった。加えて、提案システムの可動域角度計測精度はC1/2-C6/7間においてC1/2が特に悪かった。以上から、計測精度を下げている要因としてC1・C2の頸椎領域の推定精度が低いことがあげられる。推定精度が低い原因として、C1とC2はそれぞれ別のラベルとしており1画像に領域は1つしかないが、C3-C7は同じラベルとしたので1画像に4つの領域が写っている。このため、C1とC2のIoU値が小さい要因はC1とC2の学習データがC3-C7に比べて少ないことと考えた。このため学習データを増やしC1・C2の頸椎領域の推定精度をあげていきたい。

6. おわりに

本研究では、画像分析を用いて頸椎可動域角度を計測する頸椎可動域角度計測補助システムを提案した。Mask R-CNNによる頸椎領域の推定は、テストデータ全体の97%を推定でき、全体のIoUは0.85となった。計測値の標準偏差を調べたところ、専門医は2.9度、研修医は3.2度であったのに対し、提案システムによる計測結果は何度計測しても変わらないため0度であった。このため、計測値がばらつく問題に対しては提案システムが有利であることがわかった。提案システムの計測精度を研修医と比較した結果、真値と提案システムの平均計測誤差は研修医との平均計測誤差より0.8度小さかった。この差に対して有意水準5%のt検定をしたところ、p値は0.05以下となり有意差があったため、研修医より良い計測精度であることがわかった。

頸椎レントゲン画像の分析では可動域角度だけでなく、頸椎のアラインメントが正常に整列されているか注目する。今後は、学習データを増やし頸椎領域推定精度の向上や頸椎のアラインメントのズレの推定を行い、不具合のある頸椎部位を推定できるか検討する。

謝辞

本研究は、JSPS 科研費（課題番号 JP21H03485）、および JST AIP-PRISM（課題番号 JPMJCR18Y2）の助成を受けたものである。計測に協力してくださった東京医科歯科大学の鍋木先生、本研究に助言いた

いた家永特任助教に感謝申し上げます。

参考文献

- [1] 日本大学医学部整形外科系整形外科学分野脊椎班: 脊椎脊髄ハンドブック; 三輪書店 (2018).
- [2] Huaifei, H., Liu, H., Chen, L., Hung, C.: Image segmentation of cervical vertebra in x-ray radiographs using the curve fitting strategy; In Proceedings of the 2011 ACM Symposium on Applied Computing, SAC '11, pp.853–858, New York, NY, USA(2011).
- [3] Lecron, F., Benjelloun, M., Mahmoudi, S.: Points of interest detection in cervical spine radiographs by polygonal approximation; In International Conference on Image Processing Theory, Tools and Applications, IEEE(2010).
- [4] Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation; In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp.3431–3440(2015).
- [5] Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional networks for biomedical image segmentation; In International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 234–241(2015).
- [6] He, K., Gkioxari, G., Dollar, P., Girshick, R.: Mask R-CNN; In Proceedings of the IEEE International Conference on Computer Vision, pp.2961–2969(2017).
- [7] Arif, SMMR Al., Knapp, K., Slabaugh, G.: SP-Net: Shape prediction using a fully convolutional neural network; In International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 430–439(2018).
- [8] 魚住, 松原, 他: Mask R-CNN を用いた小児胸部 X 線画像における肺領域の自動抽出; Medical Imaging Technology, Vol. 38, No.3, pp.126–131(2020).
- [9] Choi, R., Watanabe, K., et al.: CNN-based spine and cobb angle estimator using moire images; IEEE Transactions on Image Electronics and Visual Computing, Vol.5, No.2, pp.135–144(2017).
- [10] Alomari, R., Chaudhary, V., Dhillon, G.: Computer aided diagnosis system for lumbar spine; ISABEL '11, pp.1–5, New York, NY, USA(2011).
- [11] Young, H., Sewon, K., Suh, J., Hwang, D.: Learning radiologist's step-by-step skill for cervical spinal injury examination: Line drawing, prevertebral soft tissue thickness measurement, and swelling detection; IEEE Access, Vol.6, pp.55492–55500(2018).