

学位論文の全文に代えてその内容を要約したもの

愛知学院大学

乙 第 号	論文提出者 村田 淳
論文題目 パノラマ X 線画像における上顎洞炎診断のための深層 学習システム	

I. 緒言

上顎洞炎はしばしば根尖性歯周炎や辺縁性歯周炎に関連して引き起こされるが 1, 2)、片側性の症状や画像所見を有するものは歯性上顎洞炎とみなされ、患者は歯科医院を受診することも多い。したがって、歯科医師は初診時に撮影されることの多いパノラマ X 線画像を用いて、上顎洞炎の有無を判定する必要に迫られることもしばしばである。上顎洞の画像所見で 4 mm 以下の粘膜肥厚は健常とされ、上顎洞全体の 3 分の 1 を超える粘膜肥厚や液体貯留、粘液貯留嚢胞様の所見が見られる場合に上顎洞炎と診断される 3, 4)。しかしながら、臨床研修歯科医師などの読影経験の浅い観察者にとっては、パノラマ X 線画像で上顎洞内やその周囲に発生した病変を正しく診断することは困難なことも多い 5)。なぜなら硬口蓋や下鼻甲介などの構造が画像上で複雑に重なり合っていることや、それらの所見がパノラマ X 線装置の断層域との位置関係によって変化するからである 6-8)。さらに、上顎洞後壁部に発生した小さな病変はパノラマ X 線画像で常に検出できるとは限らないことも読影を困難にする原因となる 9)。

このような状況に対して、申請者の研究室では以前の研究で、経験の少ない観察者の支援を目的として、片側性の上顎洞炎を検出するコンピュータ支援検出(computer-aided detection: CAD) システムを開発した 5)。この CAD システムを用いることによって、経験の浅い観察者の診断能力は十分な経験を有する歯科放射線科医に匹敵するほどに改善された 5)。しかしながら、このシステムでは診断は左右の上顎洞の透過性を比較することによってなされるため、適用できるのは片側性の上顎洞炎のみであった。

近年、人工知能や深層学習の方法が注目されるようになった。畳み込みニューラルネットワーク(convolution neural network: CNN) は深層学習のためのアルゴリズムであり 10)、畳み込み層やプーリング層などの幾つかの特徴的な機能を持った層を積み上げることで構成される。畳み込み層では画像からのエッジ抽出等の特徴を抽出し、プーリング層では抽出された特徴が平行移動などでも影響を受けないようにロバスト性(強靱性)を与える 11)。深層学習システムの医用画像への応用は進行中であり、いくつかの研究が病変検出 12)、病変分類 13)、領域分割 14)、画像再構築 15, 16) や自然言語処理 17) などの分野において、すでに発表されている。歯科領域でも下顎智歯の発育段階の分類などの試みが深層学習による人工知能を用いて行われている 18)。申請者の研究室でも、口腔癌患者における頸部リンパ節転移の CT 診断 19) やパノラマ X 線画像による下顎第一大臼歯の遠心過剰根の判定に応用し 20)、満足できる結果を得ている。しかしながら、未だ歯科領域への応用は十分とは言えない状況にある。

本研究の目的は、深層学習システムをパノラマ X 線画像における上顎洞炎の診断に適用し、その診断能を明らかにすることである。

II. 対象と方法

本研究は愛知学院大学歯学部倫理委員会の承認を得て (No. 496)、ヘルシンキ宣言に則って行われた。

1. 症例と検査法

愛知学院歯学部附属病院の画像データベースに保管された画像から、パノラマ X 線画像検査とコンピュータ断層撮影 (CT) あるいは歯科用コーンビーム CT (CBCT) 検査を行った症例で、上

顎洞に炎症性変化を有する症例と有しない症例を2007年4月から2018年5月の間で連続的に抽出した。上顎洞炎の有無はCTあるいはCBCTの所見によることとし、粘膜肥厚、液体貯留あるいは粘液嚢胞様の所見が上顎洞の3分の1以上にみられるものを上顎洞炎有りとし、以下上顎洞炎群とする(3-5)。また粘液貯留嚢胞も大きなものでは上顎洞炎と同様に不透過性の病変となるので、今回上顎洞炎群に含めた。上顎洞炎と診断された症例はすべて頬部の腫脹や疼痛あるいは上顎歯周囲からの排膿などの症状のために検査を行ったものであった。埋伏歯、顎変形症、顎関節疾患あるいは顎骨に限局した嚢胞などで検査を行った症例で、CTあるいはCBCT画像で上顎洞全域が十分に含まれ、粘膜肥厚が4mm以下のものを健常な上顎洞とした。パノラマX線画像検査はCTあるいはCBCTに先行し、両者の間隔は3週以内のものを用いたが、抗菌薬の投与後でCTあるいはCBCTで上顎洞炎の所見が明らかでないものは除外した。線維性異形成症や口腔癌などの腫瘍性病変も除外した。

訓練データとするために、それぞれ100例(男性50例、女性50例)からなる4グループの症例が2018年5月末日を起点として、検査日を遡って、平均年齢に差が出ないように抽出された(表1)。4グループとは両側上顎洞が健常なもの、左側上顎洞に炎症を有するもの、右側上顎洞に炎症を有するもの、両側上顎洞に炎症を有するものである。テストデータのためには、両側上顎洞が健常な30例(男性15例、女性15例)と左側上顎洞に炎症を有する60例(男性30例、女性30例)の2グループが抽出された。これらは訓練データ用の症例との重複を避けるために2010年4月末日を起点として、検査日を遡って抽出された。選択された症例数を表1に示す。グループ間で平均年齢に差はなかった。

CT検査はAsterion TSX(キャノンメディカルシステム)を使用して、管電圧120kVp、管電流50mA、スライス厚0.5mm、スライスピッチ0.3mm、撮影領域20cmで行われた。CBCT画像はAlphard Vega(朝日レントゲン)を使用して得られた。0.2mmのボクセルサイズを有するIモード(102x102mm)を使用し、撮影条件は管電圧80kVp、管電流8mA、360度回転で照射時間17秒とした。パノラマX線画像はVeraview epocs(モリタ製作所)を用いて、管電圧75kVp、管電流9mA、照射時間16秒で取得した。

2. 訓練データ用画像の準備とデータ拡張

訓練用の画像パッチを作成するために、片側の顎洞を含むように200x200ピクセルの矩形の関心領域をパノラマX線画像上に設定した(図1)。画像処理ソフトウェア(Adobe Photoshop ver. 13, Adobe System Co. Ltd.)のマクロ機能を用いて、全症例のパノラマX線画像において両側の顎洞に設定された関心領域が半自動的に切り取られた。結果的に、健常な顎洞と炎症性の顎洞それぞれについて400の画像パッチが作成された(表2)。

学習モデルの信頼性を高めるために、データ拡張がソフトウェア(IrfanView, <https://www.irfanview.com/>)を使用して行われ、訓練データの数が増強された。画像の輝度、コントラスト、鮮鋭度などの変更がなされた結果、健常および炎症性の顎洞それぞれについて6000の画像パッチが作成された。

3. テストデータの準備

テストデータ用の画像パッチが訓練用の画像パッチの作成と同様の手法で行われた(表2)。健

常群として、30 例の両側に健常な上顎洞を有する患者のパノラマ X 線画像から左右側それぞれの上顎洞を切り取り 60 の画像パッチが、また上顎洞炎群として、片側に炎症性の上顎洞を有する患者のパノラマ X 線画像から炎症側の上顎洞を切り取り 60 の画像パッチが作成された。

4. 深層学習の過程

深層学習システムは Ubuntu ver. 16.04.2 の動作システム(OS: operating system)、11GB の NVIDIA GeForce GTX 1080Ti の GPU (Graphics processing unit) 上に構築された。深層学習トレーニングシステムは DIGITS Library ver.5.0 (NVIDIA; <https://developer.nvidia.com/digits>)を使用した。DIGITS には、フレームワークとして Café と Torch が、ネットワークとして LeNet、AlexNet、GoogLeNet が標準で組込まれている。今回は、フレームワークとして Café を、ネットワークとして AlexNet を採用した。AlexNet は 5 層の畳み込み層と 3 層の全結合層を有するネットワークである。DIGITS の「分類 (Classification)」を使用して学習過程を遂行した。

訓練用の画像パッチが深層学習システムに適用され 200 エポックの学習が行われた。テスト用の画像パッチを作成された学習モデルに適用して、正診率、感度、特異度、陽性的中率、陰性的中率を求めた。さらに ROC カーブ(receiver-operating characteristic curve)を描いてカーブ下の面積(AUC)を求めた。

5. 観察者の診断能との比較

20 年以上の経験を有する 2 名の歯科放射線専門医と卒後 1 年未満の臨床研修歯科医師 2 名が別々に上顎洞炎の有無をパノラマ X 線画像上で判定した。数例のパノラマ X 線画像で練習を行った後に、実際の判定を行った。深層学習の評価に使用したテストデータ用画像と同じ上顎洞(120 上顎洞)をモニター(RadiForce G20; EizoNanao Corp., Ishikawa)上で判定した。観察者は左右側それぞれの上顎洞に対して上顎洞炎の有無を 4 段階で評価した。すなわち、1: 上顎洞炎が存在しない。2: 上顎洞炎がおそらく存在しない。3: 上顎洞炎がおそらく存在する。4: 上顎洞炎が存在する。の 4 段階である。判定結果に基づいて、正診率、感度、特異度、陽性的中率、陰性的中率および AUC を決定し、深層学習の診断能と比較した。

6. 統計的解析

グループ間の年齢の違いは t 検定によって検討した。AUC の比較は χ^2 検定によった。観察者間の判定の一致度は Cohen の κ 値を、歯科放射線専門医と研修歯科医師の診断能の差は Mann-Whitney の U テストを使用した。危険率 5%未満を統計的に有意とした。

III. 結果

学習過程に要した時間は、作成されたデータセットを DIGITS にインポートするのに 26 秒、12000 の訓練用画像パッチで 200 エポックの学習過程を遂行し学習モデルが作成されるまでに 28 分、120 のテスト用画像パッチを学習モデルに適用して病変の有無をテストするのに 9 秒であった。

図 3 に 200 エポックの学習過程を示す。Loss (損失) は学習による予測と実際の診断との違いを

示すもので、学習過程がすすむと 0 になった。Accuracy (正診率) は学習過程の精度を示しており、早期の段階で 85% を超えてその後緩やかに安定していた。

深層学習システムによる正診率、感度、特異度はそれぞれ 87.5%、86.7% および 88.3% であった (表 3)。歯科放射線専門医による診断ではそれぞれ、89.6%、90.0% および 89.2% であったが、臨床研修歯科医師では 76.7%、78.3% および 75.0% であった。深層学習による診断能は歯科放射線専門医と同等であった。

AUC は深層学習システムで 0.875、歯科放射線専門医で 0.896、臨床研修歯科医師で 0.767 であった。深層学習システムの AUC は歯科放射線専門医と同等で、臨床研修歯科医師よりも高く有意差を認めた (深層学習システム vs 歯科放射線専門医: $p=0.4670$ 、深層学習システム vs 臨床研修歯科医師: $p=0.0018$ 、歯科放射線専門医 vs 臨床研修歯科医師: $p=0.0001$ 、 χ^2 検定)。

2 名の歯科放射線専門医における判定は κ 値で 0.733 と十分な一致度であったが、臨床研修歯科医師における一致度は 0.508 と中等度の値を示した。

IV. 考察

本研究で作成した深層学習システムでは、正診率 87.5%、感度 86.7%、特異度 88.3%、ROC 解析における AUC は 0.875 と歯科放射線専門医と同等で十分に高い診断能を示した。一方、臨床研修歯科医師は深層学習システムおよび歯科放射線専門医に比較して低い診断能であった。これらの結果に基づくと、本研究で作成された人工知能を用いたシステムは、パノラマ X 線画像による上顎洞炎の診断において、特に経験の少ない観察者に対して、診断支援が可能となることが示された。

多層畳み込みニューラルネットワークに画像パッチを入力すると、深層学習システムが学習過程を繰り返し、学習モデルを作成する (11)。このようにして作成された学習モデルは病変の検出 (12) や分類 (13)、領域分割 (14) や自然言語処理 (17) など様々な領域で利用されている。本研究ではパノラマ X 線画像における上顎洞炎の診断について、5 層の畳み込み層と 3 層の全結合層を有する AlexNet を使用して (21)、病変分類の学習過程が遂行された。GoogleNet (12) などのさらに進化したニューラルネットワークの応用も可能かもしれないが、本研究で使用した 12000 の画像パッチの学習では、AlexNet は十分に有効であり、短時間で問題なく学習モデルを作成できた。深層学習では、臨床データなどデータが少ないときにはデータ拡張の手法を用いることが多く (21-24)、輝度、コントラストやノイズなどを変化させることによって、訓練データを数倍に拡張できる。十分に有効な学習モデルを作成するためには訓練データの数をどれほどにすればよいかを事前に確定する方法はない方法はないが、本研究で 87.5% の高い診断能が得るには、使用した訓練データ数は適切であったと考えられた。

申請者の研究室では以前の研究で、片側性の上顎洞炎を診断する CAD システムを開発した (5)。このシステムでは、サブトラクションの技術を利用して、左右の上顎洞の透過性を比較することによって、片側性の上顎洞炎を検出することが可能となった。このシステムを利用した診断支援によって、臨床研修歯科医師など経験の少ない観察者の診断能を経験豊富な歯科放射線専門医の診断能に匹敵するまでに向上することができた。しかしながら、経験豊富な歯科放射線専門医では CAD システムを使用しなくても高い診断能を示し、その利用によって診断能が向上することはなかった。この CAD システムの結果と本研究のそれを直接的に比較することはできないが、両研

究において歯科放射線専門医は本研究の人工知能システムと同等の高い診断能を示した。以前の CAD システムは、読影者の診断支援を目的として作成されたが、本研究におけるシステムはそれ自身で上顎洞炎を分類でき、高い診断能を示したことを考慮すると、パノラマ X 線画像の自動診断を可能とするシステムと位置付けることができる。

パノラマ X 線画像で上顎洞疾患を診断する際には、歯科放射線科医は上顎洞の形態や透過性などの変化に注意を払うが、同時にそれらを反対側と比較する。このような考え方や以前の CAD システムにおける方法を考慮すると、今回の深層学習システムでも、片側性上顎洞炎の場合、訓練データ用画像パッチは両側の上顎洞を含むように切り出したほうが良いかもしれない。しかし、臨床においては、両側に炎症を持つ患者や両側ともに健常な患者が存在するし、人工知能による診断過程は歯科放射線科医のものと異なるかもしれない。本研究では片側の上顎洞について学習モデルを作成したが、今後両側の上顎洞を含む画像パッチ用の ROI の設定も必要かもしれない。

本研究はパノラマ X 線画像の完全自動診断システム確立を目指す研究の一部と位置付けることができる。この目的を達成するためには多くの解決しなければならない問題がある。本研究では画像パッチは市販のソフトウェアを用いて半自動的に行ったが、最終ゴールに達するには、完全自動化された画像パッチの切り出しシステムの開発が必要不可欠である。さらに、臨床応用を目指した研究も企画されなければならない。つまり、他の施設および他のパノラマ装置で撮影された画像を収集して訓練データを増加することによって、さらに高精度の学習モデルを作成することが可能となる。

V. 結論

パノラマ X 線画像による上顎洞炎の診断に深層学習システムを適用したところ、正診率 87.5%、感度 86.7%、特異度 88.3%、ROC 解析における AUC は 0.875 と高い診断能を示した。この診断能は臨床研修歯科医師よりも高く歯科放射線専門医と同等であった。パノラマ X 線画像による上顎洞炎の診断には、深層学習による診断支援が十分に可能であることが示された。