



## Diagnosis of Developmental Dysplasia of the Hip by Ultrasound Imaging Using Deep Learning

衣笠, 真紀

(Degree)

博士 (医学)

(Date of Degree)

2024-03-25

(Resource Type)

doctoral thesis

(Report Number)

甲第8763号

(URL)

<https://hdl.handle.net/20.500.14094/0100489988>

※ 当コンテンツは神戸大学の学術成果です。無断複製・不正使用等を禁じます。著作権法で認められている範囲内で、適切にご利用ください。



(博士課程関係)

## 学位論文の内容要旨

# Diagnosis of Developmental Dysplasia of the Hip by Ultrasound Imaging Using Deep Learning

ディープラーニングによる発育性股関節形成不全の超音波画像診断

神戸大学大学院医学研究科医科学専攻

小児高度専門外科学

指導教員：大島義博 教授（神戸大学連携大学院客員教授）

衣笠 真紀

# Diagnosis of developmental dysplasia of the hip by ultrasound imaging using deep learning

## ディープラーニングによる発育性股関節形成不全の超音波画像診断

### 【背景および目的】

発育性股関節形成不全（以下 DDH、developmental dysplasia of the hip）の診断は、適切な時期に正確に行うことが良好な治療成績につながる。DDH の診断に超音波画像検査（以下、エコー検査）は有用であるが、一方で乳児に対するエコー検査は、一定以上の研鑽を積んだ者でなければ適切な画像の描出や診断が困難である。そこで、我々は人工知能 AI (artificial intelligence) を用いた診断が可能となれば乳児股関節エコー検査による画像診断の普及の一助となると考えた。AI の学習モデルとして様々なアルゴリズムが提唱されているが、その中でも脳の神経細胞を模したニューラルネットワークを多数組み合わせて、自動的に特徴量を学習させる手法はディープラーニングと呼ばれている。医療画像診断の分野では、画像の特徴を学習するディープラーニングが広くおこなわれており、本研究でも DDH のエコー診断にディープラーニングを使用し、その診断精度を 5 名の整形外科医と比較検証した。

### 【対象および方法】

対象は 2016 年から 2021 年の期間に、生後 6 か月未満で DDH を疑われ受診した患者とし、診察と単純レントゲン検査、Graf 法によるエコー検査を行い、10 年以上的小児整形外科医の臨床経験を持つ 2 名の医師が DDH と診断した 60 名 64 股、正常と診断した 131 名 262 股とした。エコー動画から Graf 法のスタンダードプレーン（腸骨外縁がプローブと平行で、腸骨下端と関節唇が明瞭に描出されている）の静止画像を切り出し 2,400 枚の訓練データおよび 600 枚の検証データを用い、最終的に 214 枚のテストデータで AI の精度を調べた。エコー画像において Graf 分類では 1, 2a, 2b を正常、2c, D, 3, 4 を DDH とした。解析ソフトは MatLab の Deep learning Toolbox (MathWorks, Natick, MA, US) を用いた。複数の事前学習済みモデルを比較検討する目的で畠み込み層が 53 層の MobileNet, 82 層の Efficient Net, 18 層の SqueezeNet を使用し学習を行った。モデルの性能評価指標としては、混合行列に基づいて機械学習の分野で広く使用される正解率、適合率、再現率、F 値を用いた。さらにモデルが判断に重要とした領域をヒートマップにて表示する‘判断根拠の可視化’を occlusion sensitivity, image LIME, gradient-weighted class activation mapping 以下 Grad-CAM という手法を使用して行った。また、214 枚のテストデータのエコー画像診断を診察時に携った小児整形外科医と異なる 5 名の整形外科医が各々行い、整形外科医の診断の正解率を調べた。

### 【結果】

AI の正解率、適合率、再現率、F 値は SqueezeNet ではそれぞれ、0.9907, 0.9878, 0.9878, 0.9878、MobileNet でそれぞれ、0.9533, 0.8780, 1.0000, 0.9351 であり、EfficientNet では全て 1.0000 であった。一方、5 名の整形外科医によるエコー画像診断の正解率は 0.930, 0.953, 0.958, 0.963, 0.977 であった。

診断根拠の可視化では骨性臼蓋や腸骨下端、関節唇、骨頭が関心領域として示された。

### 【考察】

AI による DDH の画像診断が可能かどうかという研究は近年いくつか報告されている。特に単純レントゲン写真における DDH の AI による診断は、高い感度と特異度で可能とされている。しかしエコー画像における DDH の AI による診断の研究は、これまでのいずれの報告も Graf 分類の 3 や 4 にあたる完全脱臼例のデータ数が少なく、これらの診断の精度については言及されていない。本報告では、60 名 64 股の Graf 3 または 4 にあたる完全脱臼例のデータを含んでおり、DDH のエコー画像という分野において比較的多いと言える。

また AI による診断において、機械が判断に重要とした領域、つまり関心領域の可視化を行ったところ、骨性臼蓋や腸骨下端、関節唇、骨頭が関心領域として示された。

従来の AI 研究では機械の判断根拠はブラックボックスとされているが、医療用の AI 実用化にむけては、機械が疾患であると判断した根拠が人間に理解可能である必要があると考えられる。本研究の結果は、日常臨床の感覚と合致するものであった。特に乳児健診では正常例が大多数であり診断を即時にしていく必要性を考えると、すべてのエコー画像で  $\alpha$  角や  $\beta$  角を計測することは現実的ではなく、角度測定せずに正常と DDH を判別する場合が多い。本研究でも AI は角度測定をすることなく、5 名の整形外科医の診断に劣らない正確さで DDH と正常を判別することが出来ることが示された。

本研究の課題としては 3 つ挙げられる。1 つめは診断に不適切な画像（スタンダードプレーンでない画像）を除外している点である。今後は AI が適切なスタンダードプレーンを抽出し画像診断を完結できるようにすることが望まれる。2 つめはデータ数をさらに増やすことである。AI の深層学習においてデータ数は多いほど診断精度は増すと考えられる。3 つめは Graf 分類の 2a や 2b は正常に含んでいる点である。実際にこれらは完全な正常とは断言できず、慎重に経過観察をする必要がある。Graf 分類の 2a や 2b を識別するためには検査時の月齢や  $\alpha$  角、 $\beta$  角も判断のために必要な情報でありこれらを含めたモデルの構築が必要である。

### 【結論】

エコー画像の AI による DDH の診断能力は整形外科医と比較して遜色なく、臨床現場に応用できる可能性がある。

神戸大学大学院医学研究科（博士課程）

論文審査の結果の要旨			
受付番号	甲 第 3334 号	氏名	衣笠 真紀
論文題目 Title of Dissertation	Diagnosis of Developmental Dysplasia of the Hip by Ultrasound Imaging Using Deep Learning ディープラーニングによる発育性股関節形成不全の超音波診断		
審査委員 Examiner	主査 Chief Examiner 佐藤 純子 副査 Vice-examiner 野津 実二 副査 Vice-examiner 酒井 訓一		

(要旨は1,000字～2,000字程度)

発育性股関節形成不全（以下 DDH、developmental dysplasia of the hip）の診断は、適切な時期に正確に行うことが良好な治療成績につながる。DDH の診断に超音波画像検査（以下、エコー検査）は有用であるが、一方で乳児に対するエコー検査は、一定以上の研鑽を積んだ者でなければ適切な画像の描出や診断が困難である。今回、研究者らは人工知能 AI (artificial intelligence) を用いた診断が可能となれば乳児股関節エコー検査による画像診断の普及の一助となると考え、DDH のエコー検査による診断にディープラーニングを使用し、その診断精度を 5 名の整形外科医と比較検証した。

## 方法

対象は生後 6 か月未満で DDH を疑われ受診した患者とし、診察と単純レントゲン検査、Graf 法によるエコー検査を行い、10 年以上的小児整形外科医の臨床経験を持つ 2 名の医師が DDH と診断した 60 名 64 股、正常と診断した 131 名 262 股とした。エコー動画から Graf 法のスタンダードプレーン（腸骨外縁がプローブと平行で、腸骨下端と関節唇が明瞭に描出されている）の静止画像を切り出し 2,400 枚の訓練データおよび 600 枚の検証データを用い、最終的に 214 枚のテストデータで AI の精度を調べた。

エコー画像において Graf 分類では 1, 2a, 2b を正常、2c, D, 3, 4 を DDH とした。解析ソフトは MatLab の Deep learning Toolbox (MathWorks, Natick, MA, US) を用いた。複数の事前学習済みモデルを比較検討する目的で MobileNet, Efficient Net, SqueezeNet を使用し学習を行った。モデルの性能評価指標としては、混合行列に基づいて機械学習の分野で広く使用される正解率、適合率、再現率、F 値を用いた。さらにモデルが判断に重要とした領域をヒートマップにて表示する‘判断根拠の可視化’を occlusion sensitivity, image LIME, gradient-weighted class activation mapping (Grad-CAM) という手法を使用して行った。また、214 枚のテストデータのエコー画像診断を診察時に携った小児整形外科医と異なる 5 名の整形外科医が各々行い、整形外科医の診断の正解率を調べた。

## 結果

AI の正解率、適合率、再現率、F 値は SqueezeNet ではそれぞれ、0.9907, 0.9878, 0.9878, 0.9878、MobileNet でそれぞれ、0.9533, 0.8780, 1.0000, 0.9351 であり、EfficientNet では全て 1.0000 であった。一方、5 名の整形外科医によるエコー画像診断の正解率は 0.930, 0.953, 0.958, 0.963, 0.977 であった。

診断根拠の可視化では骨性臼蓋や腸骨下端、関節唇、骨頭が関心領域として示された。

## 考察および結論

エコー画像における DDH の AI による診断のこれまでの報告では Graf 分類の 3 や 4 にあたる完全脱臼例のデータ数が少なく、これらの診断の精度については言及されていなかった。今回の研究は、60 名 64 股の Graf 分類 3, 4 にあたる完全脱臼例のデータを含んでおり、AI による DDH の診断能力は整形外科医と比較して遜色ないという結果であった。また従来 AI 診断において判断根拠は解明されていないが、関心領域の可視化を行ったところ、骨性臼蓋や腸骨下端、関節唇、骨頭が関心領域として示された。

本研究は、DDH の超音波画像診断について、ディープラーニングを用いた AI による診断精度を研究したものであるが、従来は解明されていなかった Graf 分類 3, 4 を含む完全脱臼例に対する AI の診断精度の評価と、AI の関心領域の可視化を行った世界で初めての論文である。よって AI による DDH の超音波画像診断について、重要な知見を得たものとして価値のある業績であると認める。

よって、本研究者は、博士（医学）の学位を得る資格があると認める。