



Deep learning model for predicting gestational age after the first trimester using fetal MRI

小路田, 泰之

(Degree)

博士 (医学)

(Date of Degree)

2022-03-25

(Resource Type)

doctoral thesis

(Report Number)

甲第8299号

(URL)

<https://hdl.handle.net/20.500.14094/D1008299>

※ 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。



(課程博士関係)

学位論文の内容要旨

Deep learning model for predicting gestational age after the first trimester using fetal MRI

胎児 MRI にて妊娠第 2 期以降の在胎週数を予測する深層学習モデルの構築

小路田 泰之, 松尾 秀俊, 神田 知紀, 西尾 瑞穂, 祖父江 慶太郎,
野上 宗伸, 河野 淳, 堀 雅敏, 村上 卓道

神戸大学大学院医学研究科医科学専攻
放射線診断学分野
(指導教員: 村上 卓道 教授)

小路田 泰之

【目的】

在胎週数を正確に予測することは、妊娠の適切な治療や管理を決定するために不可欠である。在胎週数の予測には通常最終月経の日付と超音波検査の所見が用いられる。超音波検査による妊娠年齢の判定には、頭囲、腹囲、児頭大横径 (BPD: Biparietal Diameter)、大腿骨長などに基づいて行われる。超音波による在胎週数の予測は妊娠第 1 期では誤差が 1 週間以内と比較的正確だが、第 2 期では 1~2 週間、第 3 期では 3 週間~1 ヶ月ほどの誤差が生じる可能性がある。以前の研究では、妊娠第 1 期に定期的な超音波検査を受けた女性は、妊娠第 2 期のみ超音波検査を受けた女性に比べて、妊娠中期以降の分娩誘発の必要性が有意に少なかったとされる。このことから、妊娠第 2 期以降の在胎週数を正確に予測する方法が求められる。

深層学習とは複数の処理層で構成された計算モデルを用いて、データセット内の項目を自動的に分類することができる機械学習の一種である。畳み込みニューラルネットワーク (CNN: Convolutional Neural Network) は、画像解析を行うための一般的な深層学習手法であり、医療用画像の認識タスクに適している。

胎児の脳の形状は在胎週数に応じてダイナミックに変化するため、CNN ベースの深層学習モデルを用いて妊娠第 2 期以降に取得された脳 MRI から在胎週数を予測できるのではないかと仮説を立てた。比較対象として、深層学習モデルに用いた画像と同じ画像から BPD を測定して週数を予測した。

【方法】

2014 年 1 月から 2019 年 6 月の間に当院で画像診断がなされた 225 件の胎児 MRI が対象となった。臨床データの不足や妊娠第 1 期の胎児の画像、出生前後に先天奇形のある胎児の画像は除外した。胎児が複数回の MRI 検査を受けていた場合は、最初の MRI 検査で得られた画像のみを使用した。参照基準となる在胎週数は、最終月経日と第 1 期の超音波検査の測定値から産科医が推定した週数を用いた。

胎児 MRI のデータセットのうち 3 分の 2 は訓練データセット、6 分の 1 は検証データセット、6 分の 1 はテストデータセットの 3 つのサブセットに分けた。

研究には胎児脳における視床と透明中隔のレベルの single-shot fast spin echo 法の T2 強調画像軸位断を使用した。画像が斜めに傾いていたり、胎動により解剖学的構造が不明瞭となっていたりしていた場合は、側脳室の形状を参考にして視床と透明中隔のレベルに最も近い画像を選択した。学習に用いるために取得画像から胎児の頭部を中心として四角形の領域を切り取り、切り取った画像を最近傍補間法で 150 × 150 ピクセルに統一した。

今回比較として、テストデータセットの MRI から BPD を測定し、在胎週数を算出した。

本研究では胎児 MRI を入力データとし、在胎週数を予測する深層学習回帰モデルを

開発した。一般性を高めるために Large-Scale Visual Recognition Challenge 2014 で使用された画像で事前に学習された識別モデルである VGG16 を回帰用に改良し、転移学習に使用した。プログラミング言語には Python (3.7.6) (<http://www.python.org>)、深層学習フレームワークには Keras 2.2.4 (<https://github.com/fchollet/keras>) と TensorFlow 1.15.4 (<http://tensorflow.org/>) を使用した。データ値を正規化するために、入力画像の画素値をプログラム内部で 0~256 の整数値として扱い、256 で割って 0~1 の連続値に変換した。同様に出力 (在胎週数) も 40 で割って 0~1 の連続値に変換した。学習済みの VGG16 の入力層を胎児 MR 画像のサイズに合わせて 150×150 ピクセルにリサイズし、出力層直前のシグモイド層を線形活性化層に変更し、出力次元を 1 に変更した。畳み込み層を含む VGG16 ネットワークの 13 層において、学習可能なパラメータを解放して転移学習を行った。回帰損失関数には Mean absolute error を用い、最適化器には stochastic gradient descent を用いた (学習率=0.0001)。修正した VGG16 ネットワークは、バッチサイズを 1 とし、最大 50 のエポックを用いて学習した。検証データセットでの Loss が改善されなかった時点で学習を終了し、モデルを修正した後テストデータで予測を行った。各エポックにおいて 126 枚の画像が学習に使用され、これらの画像は過学習を避けるために、回転 (-180°~180°)、垂直または水平方向の反転、0.8~1.2 倍の拡大・縮小によって画像の数を増やした。最後に学習したモデルを用いてテストデータセットに含まれる胎児の在胎週数を予測し、その結果を非正規化した。

モデルまたは BPD によって予測された在胎週数は、Lin の一致相関係数 (ρc) を用いて参照基準と比較し、McBride の定義に従って評価した (0.9 未満は poor、0.90 以上 0.95 以下は moderate、0.95 より大きく 0.99 以下は substantial、0.99 より大きい場合は almost perfect)。系統的なバイアスを評価するために、Bland-Altman プロットを用いてモデルまたは BPD によって予測された在胎週数を参照基準と比較した。

【結果】

225 件の胎児画像のうち基準を満たした 184 件が対象となった。胎児の平均在胎週数は 29.4 週 (14.0~41.4 週) で、妊娠第 2 期は 25% (n=46)、妊娠第 3 期は 75% (n=138) であった。すべての胎児に明らかな異常は見られなかった。

184 件の画像のうち 100 件 (54.3%) が斜めになっており、31 件 (16.8%) がアーチファクトを含んでいた。また 23 件 (12.5%) が斜めになっていると同時にアーチファクトを含んでおり、全体では 108 件 (58.7%) が斜めになっているかあるいはアーチファクトを含んでいた。

Lin の一致相関係数 (ρc) ではモデルの予測値と参照基準との一致率は substantial であり ($\rho c=0.964$, 95%信頼区間 [CI]: 0.928-0.982)、BPD の予測値と基準値との一致率は moderate ($\rho c=0.920$, 95%CI: 0.845-0.959) であった。モデルまたは BPD

の予測値と基準値との差を示す Bland-Altman プロットでは、モデル予測値は基準値よりも 0.43 ± 1.47 週 (平均±標準偏差, 95%CI: 0.13-0.99) 短く、BPD 予測値は 0.97 ± 2.37 週 (95%CI: 0.07-1.87) 長かった。モデル予測と比較して BPD 予測は妊娠期間が長くなるほど誤差が大きくなる傾向にあった。モデル予測値の誤差の下限 (-3.31, 95%CI: -4.27~-2.34) は、BPD 予測値の下限 (-3.67, 95%CI: -5.23~-2.11) とほぼ同じであった。モデル予測の誤差の上限 (2.45, 95%CI: 1.48-3.41) は、BPD 予測の上限 (5.61, 95%CI: 4.06-7.18) よりも小さかった。

【結論】

深層学習は妊娠第 2 期以降の胎児の脳 MRI から比較的高い精度で在胎週数を予測することができた。深層学習による在胎週数の予測は、妊娠第 1 期にスクリーニング検査を受けていない妊娠に対する出生前医療に良い影響を与える可能性がある。

論文審査の結果の要旨			
受付番号	甲 第3150号	氏名	小路田 泰之
論文題目 Title of Dissertation	Deep learning model for predicting gestational age after the first trimester using fetal MRI 胎児MRIにて妊娠第2期以降の在胎週数を予測する深層学習モデルの構築		
審査委員 Examiner	主査 寺井 義人 Chief Examiner 副査 野津 寛之 Vice-examiner 副査 福本 巧 Vice-examiner		

(要旨は1,000字~2,000字程度)

妊娠週数を正確に予測することは、妊娠の適切な治療や管理を決定するために不可欠である。妊娠週数の予測には通常最終月経の日付と超音波検査の所見が用いられる。超音波検査による妊娠年齢の判定には、頭囲、腹囲、児頭大横径 (BPD: Biparietal Diameter)、大腿骨長などに基づいて行われる。超音波による妊娠週数の予測は妊娠第1期では誤差が1週間以内と比較的正確だが、第2期では1~2週間、第3期では3週間~1ヶ月ほどの誤差が生じる可能性がある。胎児の脳の形状は妊娠週数に応じてダイナミックに変化するため、畳み込みニューラルネットワーク (CNN: Convolutional Neural Network) ベースの深層学習モデルを用いて妊娠第2期以降に取得された脳MRIから妊娠週数を予測できるのではないかという仮説を立てた。比較対象として、深層学習モデルに用いた画像と同じ画像からBPDを測定して週数を予測した。

【方法】

2014年1月から2019年6月の間に当院で画像診断がなされた225件の胎児MRIが対象となった。胎児MRIのデータセットのうち3分の2は訓練データセット、6分の1は検証データセット、6分の1はテストデータセットの3つのサブセットに分けた。

研究には胎児脳における視床と透明中隔のレベルのsingle-shot fast spin echo法のT2強調画像軸位断を使用した。画像が斜めに傾いていたり、胎動により解剖学的構造が不明瞭となっていたりしていた場合は、側脳室の形状を参考にして視床と透明中隔のレベルに最も近い画像を選択した。学習に用いるために取得画像から胎児の頭部を中心として四角形の領域を切り取り、切り取った画像を最近傍補間法で150 × 150ピクセルに統一した。本研究では胎児MRIを入力データとし、妊娠週数を予測する深層学習回帰モデルを開発した。一般性を高めるためにLarge-Scale Visual Recognition Challenge 2014で使用された画像で事前に学習された識別モデルであるVGG16を回帰用に改良し、転移学習に使用した。プログラミング言語にはPython (3.7.6) (<http://www.python.org>)、深層学習フレームワークにはKeras 2.2.4 (<https://github.com/fchollet/keras>)とTensorFlow 1.15.4 (<http://tensorflow.org/>)を使用した。最後に学習したモデルを用いてテストデータセットに含まれる胎児の妊娠週数を予測し、その結果を非正規化した。

モデルまたはBPDによって予測された妊娠週数は、Linの一致相関係数(ρ_c)を用いて参照基準と比較し、McBrideの定義に従って評価した(0.9未満はpoor、0.90以上0.95以下はmoderate、0.95より大きく0.99以下はsubstantial、0.99より大きい場合はalmost perfect)。系統的なバイアスを評価するために、Bland-Altmanプロットを用いてモデルまたはBPDによって予測された妊娠週数を参照基準と比較した。

【結果】

225件の胎児画像のうち基準を満たした184件が対象となった。胎児の平均妊娠週数は

29.4 週 (14.0~41.4 週) で、妊娠第 2 期は 25% (n=46)、妊娠第 3 期は 75% (n=138) であった。すべての胎児に明らかな異常は見られなかった。

184 件の画像のうち 100 件 (54.3%) が斜めになっており、31 件 (16.8%) がアーチファクトを含んでいた。また 23 件 (12.5%) が斜めになっていると同時にアーチファクトを含んでおり、全体では 108 件 (58.7%) が斜めになっているかあるいはアーチファクトを含んでいた。

Lin の一致相関係数 (ρ_c) ではモデルの予測値と参照基準との一致率は substantial であり ($\rho_c=0.964$, 95%信頼区間 [CI]: 0.928-0.982)、BPD の予測値と基準値との一致率は moderate ($\rho_c=0.920$, 95%CI: 0.845-0.959) であった。モデルまたは BPD の予測値と基準値との差を示す Bland-Altman プロットでは、モデル予測値は基準値よりも 0.43 \pm 1.47 週 (平均 \pm 標準偏差, 95%CI: 0.13-0.99) 短く、BPD 予測値は 0.97 \pm 2.37 週 (95%CI: 0.07-1.87) 長かった。モデル予測と比較して BPD 予測は妊娠期間が長くなるほど誤差が大きくなる傾向にあった。モデル予測値の誤差の下限 (-3.31, 95%CI: -4.27~-2.34) は、BPD 予測値の下限 (-3.67, 95%CI: -5.23~-2.11) とほぼ同じであった。モデル予測の誤差の上限 (2.45, 95%CI: 1.48-3.41) は、BPD 予測の上限 (5.61, 95%CI: 4.06-7.18) よりも小さかった。

【結論】

深層学習は妊娠第 2 期以降の胎児の脳 MRI から比較的高い精度で妊娠週数を予測することができた。深層学習による妊娠週数の予測は、妊娠第 1 期にスクリーニング検査を受けていない妊娠に対する出生前医療に良い影響を与える可能性がある。

本研究は、深層学習回帰モデルにより妊娠第 2 期以降の胎児の脳 MRI から比較的高い精度で妊娠週数を予測することを研究したものであるが、従来検査間の誤差が生じやすい妊娠第 2 期以降の妊娠週数を正確に予測するために、人工知能を用いた新しい手法により正確に診断することが出来るツールを開発し得た点で価値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。