

MetaFormer を用いた胎児心拍数陣痛図によるハイリスク出産の予測 Prediction of high-risk births by cardiotocography data using MetaFormer

辺見 一成¹⁾ Mohannad Alkanan²⁾ 柴田 千尋¹⁾
Kazunari Hemmi Chihiro Shibata

宮田 康平³⁾ 宮本 新吾³⁾ 今村 利朗³⁾ 福西 広晃²⁾
Kohei Miyata Shingo Miyamoto Toshiro Imamura Hiroaki Hukunishi

1 はじめに

近年、多くの分野で深層学習の活用が進んでおり、特に医療の分野では一部の疾患については既に実用化されている。周産期医療では、子宮内の胎児は様々な要因によりリスクに直面するため、胎児心拍数陣痛図 (CTG) と呼ばれるデータを用いて医師が危険性を目視により判断している。そのため、CTG に対して畳み込みニューラルネットワーク (CNN) を用いてハイリスク出産の予測を行う研究が行われており、既に一定程度の精度が得られることが知られている。本稿では MetaFormer を用いてさらなる精度の向上が可能かを検討する。また、適したデータ拡張を行うことで精度が向上するか検討する。

2 先行研究

Alkanan ら [1] は、妊娠週数や出産経験回数などのメタデータおよび CTG 画像そのものを入力とし、出産時の胎児の状態を危険かどうかの 2 値で予測するモデルを、EfficientNet をベースに提案している。また、国内の複数の医療機関から得た約 3 万件のデータから学習を行い、交差検定法を用いて、予測精度が ROC-AUC で 0.9 以上となることを示している。一方で、病院ごとにメタデータを含むデータの性質が異なる事がわかっており、同じ医療機関から得たデータが、テストデータと訓練データの両方にまたがるため、交差検定法は、新規の医療機関における予測値を必ずしも正しく反映しているとは言えない。そのため、本稿では、MetaFormer を用いた手法の提案の他に、上記 3 万件のデータとは別に、異なる医療機関から新たに得た約 750 件データを、新規のテスト用データセットとして、再度予測精度の検証を行った。

3 胎児心拍数陣痛図 (CTG)

胎児心拍数陣痛図 (CTG) とは、分娩監視装置による胎児心拍数と陣痛の連続記録で子宮収縮に対する胎児の心拍数変化により胎児の健全性を推測するものである。本来、胎児心拍数と子宮収縮の波形は別のグラフとして記録されており、横軸の時間軸のみ共通している。縦軸の単位はそれぞれ bpm と mmHg である。それらを 1 つの画像にまとめ、上段に心拍数、下段に子宮収縮の波形を記録する。また、学習時にはこの画像をリサイズしてから使用する。別のグラフに記録されていた CTG と学習時に使用する CTG を図 1 に示す。

4 データセット

本稿で使用するデータセットのラベルは、分娩後の胎児の状態を「良い(負例)」「悪い(正例)」の 2 値で表現したものであり、その閾値は、先行研究 [1] と同様に、分娩 1 分後および 5 分後のアプガースコアが 6 未満かどうか

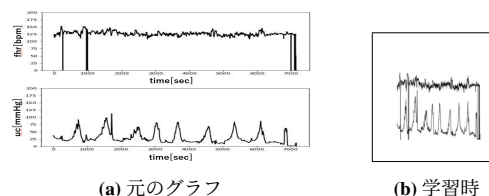


図 1: 胎児心拍数陣痛図 (CTG)

かである。また、データの性質上、負例がほとんどであるため、負例をランダムに除くことにより、正例と負例との比率を調整する。使用するデータセットのラベルの調整後の比率を表 1 に示す。

表 1: データセットのラベル比率

データセット	正例の数	負例の数
訓練用	670	3756
テスト用	38	711

5 提案手法

5.1 MetaFormer による精度の検証

MetaFormer を用いて精度の検証を行う。MetaFormer とは特定の性質を持ったモデル構造のグループを指し、Embedding, Token Mixer, Channel MLP の 3 要素から構成される。本稿で使用する VisionTransformer(ViT)[2] と PoolFormer[3] は Token Mixer にそれぞれ Attention と Pooling を用いたモデルである。本稿で使用するモデルの構造を図 2 に示す。

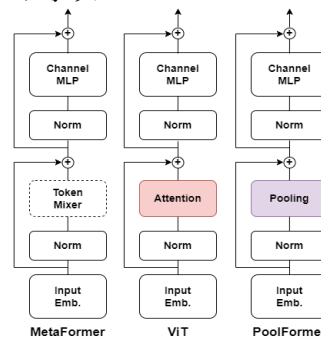


図 2: MetaFormer の構造図 [3]

5.2 データ拡張による精度の検証

画像データに対してデータ拡張を行い、その有無で精度の検証を行う。データ拡張を行う理由は、本稿で使用するデータセットの正例のデータ数が負例のデータ数に比べて少ないからである。本稿では Terrance DeVries ら [4] が提案した Cutout と Sangdoon Yun ら [5] が提案した CutMix を行う。Cutout は、モデルの正則化を目的としたデータ拡張で、画像データの一部をランダムにマスクする手法である。また、本稿では使用する画像データに対して新たな Cutout(スリット型 Cutout) を提案する。本稿

1) 法政大学
2) 東京工科大学
3) 福岡大学

で扱うデータは時系列データであり、同時刻の胎児心拍数と子宮収縮の波形を同時にマスクすることが重要である。スリット型 Cutout では、マスクする矩形の縦幅を画像データの縦幅と同じ縦幅で固定し、垂直な矩形で一部がマスクされるようにする。元画像とスリット型 Cutout が行われた画像を図 3 に示す。以下の図では、視認しやすくするために灰色の矩形でマスクしている。学習時に使用する画像データには白色の矩形でマスクする。

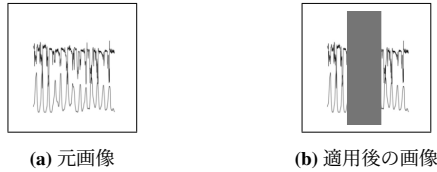


図 3: スリット型 Cutout

CutMix とは Cutout で画像に対してマスクする矩形を他の画像のバッチに置き換えたデータ拡張である。また、バッチの面積に比例してラベルを混合させる。CutMix についても同様にスリット型 CutMix を提案する(図 4)。

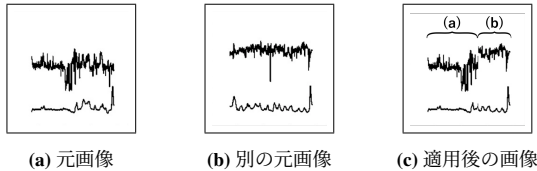


図 4: スリット型 CutMix

6 実験結果

6.1 MetaFormer による精度の検証結果

MetaFormer の 5 つの異なるモデルとベースラインのモデルで精度の検証を行った。PoolFormer からは Patch Embedding の埋め込みサイズが小さく (s), PoolFormer のブロック数がそれぞれ 12,24,36 であるモデルを使用した。ViT からは Base, Large, Huge の 3 つのタイプのうち、Base でありパッチサイズがそれぞれ 16,32 であるモデルを使用した。さらにベースラインとして CNN である EfficientNet の b2 を使用した。以上のモデルで精度を検証した結果を表 2 に示す。表 2 より、PoolFormer の 3 つのモデルは CNN と同程度の精度を示した。さらに PoolFormer の s24 と s36 は、新規のデータセットの ROC-AUC において EfficientNet-b2 よりも高い精度を示した。また、ViT の 2 つのモデルはそれ以外の 4 つのモデルと比べて精度が低かった。

表 2: MetaFormer を用いた結果

model	params	10-fold 交差検定 (mean \pm std)	ROC-AUC
EfficientNet-b2	7.7M	0.8276 \pm 0.0277	0.6624
PoolFormer_s12	11.4M	0.8159 \pm 0.0374	0.6591
PoolFormer_s24	20.8M	0.8190 \pm 0.0435	0.7050
PoolFormer_s36	30.3M	0.8241 \pm 0.0351	0.7159
ViT_Base_patch16	85.8M	0.7746 \pm 0.0426	0.6350
ViT_Base_patch32	87.4M	0.7611 \pm 0.0450	0.6401

6.2 データ拡張による精度の検証結果

データ拡張による精度の検証を行った。データ拡張を行った結果を図 5, 図 6 に示す。図 5 より、データ拡張であるスリット型 Cutout やスリット型 CutMix を行うことで 10-fold 交差検定の精度が向上する。また図 6 より、スリット型 Cutout については新規のデータセットに対しても精度が向上した。以上のことから本稿で使用するデータに対してスリット型 Cutout とスリット型 CutMix を行うことでモデルの精度が向上することが分かる。特にスリット型 Cutout については新規のデータセットに対しても精度が向上した。

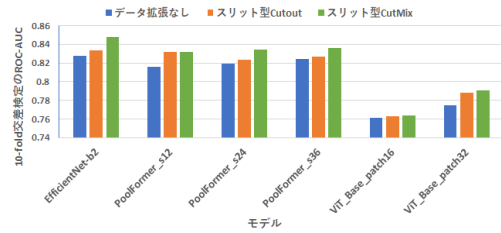


図 5: データ拡張を行った 10-fold 交差検定の ROC-AUC

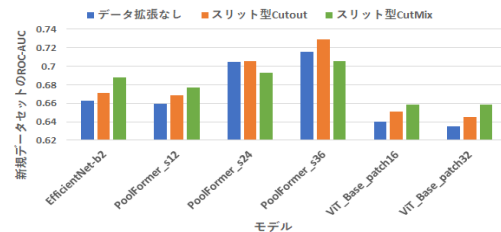


図 6: データ拡張を行った新規のデータセットの ROC-AUC

7 まとめ

本稿では、MetaFormer を用いてさらなる精度の向上が可能かを検討した。また、適したデータ拡張を行うことで精度が向上するか検討した。結果としては PoolFormer は CNN と同程度の精度を示した。データ拡張についてはスリット型 Cutout とスリット型 CutMix 共は精度の向上を示した。またスリット型 Cutout は新規のデータセットに対しても精度の向上を示したことから新規の医療機関のデータに対しても高い精度が期待できる。

参考文献

- [1] Mohannad, A. Shibata, C. Miyata, K. Imamura, T. Miyamoto, S. Fukunishi, H. and Kameda, H. :Predicting High Risk Birth From Real Large-scale Cardiotocographic Data Using Multi-input Convolutional Neural Networks(Mar.2021)
- [2] Dosovitskiy, A. Beyer, L. Kolesnikov, A. Weissenborn, D. Zhai, X. Unterthiner, T. Dehghani, M. Minderer, M. Heigold, G. Gelly, S. Uszkoreit, J. Houshy, N. equal technical contribution, equal advising, Google Research and Brain Team :An Image Is Worth 16x16 Words:Transformers For Image Recognition At Scale(Sep.2021)
- [3] Yu, W. Luo, M. Zhou, P. Si, C. Zhou, Y. Wang, X. Feng, J. Yan, S. Sea AI Lab and National University of Singapore :MetaFormer Is Actually What You Need For Vision(Nov.2021)
- [4] DeVries, T. Taylor, G. University of Guelph, Canadian Institute for Advanced Research and Vector Institute :Improved Regularization Of Convolutional Neural Networks With Cutout(Aug.2017)
- [5] Yun, S. Han, D. Oh, S. Chun, S. Choe, J. Yoo, Y. Clova AI Research, NAVER Corp. Clova AI Research, LINE Plus Corp. and Yonsei University :CutMix: Regularization Strategy To Train Strong Classifiers With Localizable Features(May.2019)