

深層学習による精子グレード分布推定と EMD を用いた精度評価

中川 勇人[†] 濱上 知樹[†]

湯村 寧[‡] 竹村 徹平[‡] 齋藤 智樹[‡] 上野 寛枝[‡] 山本 みずき[‡]

横浜国立大学[†] 横浜市立大学[‡]

1 はじめに

生殖医療による出生児が増加していることや、2022 年 4 月に政府が不妊治療の保険適用範囲を拡充する予定であることから、生殖医療の高度化への期待が高まっている。その中でも体外受精において顕微授精が注目されている。顕微授精は精子を直接卵子に注入する不妊治療法であり、その作業は胚培養士が担当する。胚培養士は顕微鏡観察によって精子を選別し、着床確度が高いと推測される精子を受精に使用する。この作業は正常な着床の成功率に影響するため、高度な専門知識・技術が求められる。

そこで本研究では、精子の選別作業における、熟練した胚培養士の技術共有や負担軽減のために生殖医療支援 AI クラウドシステムの開発を進めている。生殖医療支援 AI クラウドシステムはクラウドを活用した胚培養士の精子の評価結果の収集と、集めたデータに対して AI を活用した精子の評価結果の分布の推定を行うものである。

本研究の目的は収集した胚培養士による精子の評価結果の推定である。胚培養士は精子グレードクラスを選択することにより評価を行う。精子グレードクラスは胚培養士による A, B, C, D, E の 5 段階評価である (A が最も着床確度が高く、E が最も低い)。胚培養士の間でも異なる精子グレードを選択することがあり、複数の胚培養士の評価結果は精子グレード分布として得られる。精子グレード分布は各精子グレードクラスへの帰属度 (胚培養士の選択率) の分布である。グレード分布には、主に精子の着床確度の高さ、専門家の評価のばらつきの二つの情報が含まれる。これらの情報ができるだけ失われないように、深層学習により分布を直接推定する。

2 グレードクラス間の距離指標

分布間の違いを表す指標として、回帰問題に広く用いられる Mean Squared Error (MSE) や 2 つの確率分布間の差異を表す Kullback-Leibler divergence (KL divergence) が挙げられる。しかし、これらの指標では各クラスを独立に見たときの分布の違いを表すに過ぎず、グレードクラス間のずれ方の違いを考慮できない。

そこで本研究ではグレードクラス間の違いを考慮することができる Earth Mover's Distance (EMD) [1] を用いる。EMD は輸送問題によって定義される距離指標である。輸送問題は、各供給地から需要地への輸送コストが与えられたときに、最小の輸送コストで需要地の需要を満たすような輸送方策を考える問題である。供給地から需要地への輸送コストにより、グレードクラス間の違いを定義できる。したがって、EMD ではグレードクラス間の違いを考慮できることがわかる。

3 精子グレード分布推定モデル

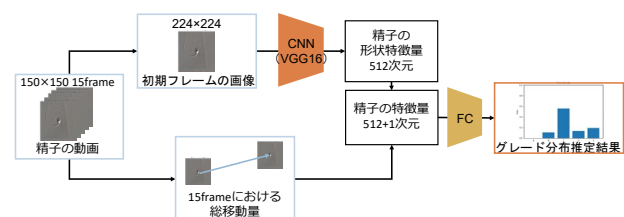


図 1 精子グレード分布推定モデル

複数人の胚培養士の評価結果の推定には図 1 に示す精子のグレード分布推定モデルを使用する。このモデルでは 1 秒間 (15frame) における精子の動画から複数人の胚培養士の評価結果である精子グレード分布を推定する。

精子の形状特徴には初期フレームの画像に VGG16 [2] の特徴抽出部分を通して得られる 512 次元の特徴量を使用する。精子の移動特徴には 1 秒間の精子の総移動量を使用する。得られた精子の形状特徴と移動特徴を統合し、全結合層を通して精子グレード分布が得られる。また、このモデルの学習には損失関数 MSE を使用する。

Deep learning predicting sperm grade distribution and evaluation using EMD

[†] (Hayato Nakagawa, Tomoki Hamagami), Yokohama National University

[‡] (Yumura Yasushi, Teppei Takemura, Tomoki Saito, Hiroe Ueno, Mizuki Yamamoto), Yokohama City University

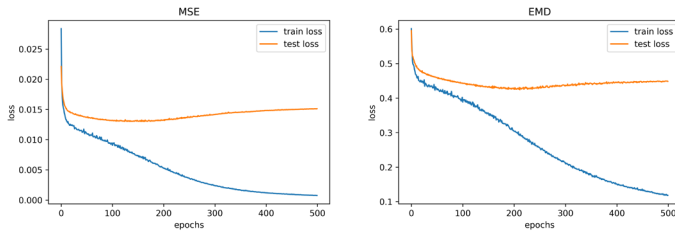


図2 学習曲線(MSE) 図3 EMDのepochごとの変化

表1 EMDのepochごとの変化

		MSE Best Model (161epoch)	EMD Best Model (221 epoch)
Train	MSE	0.00712	0.00442
	EMD	0.348	0.281
Test	MSE	0.0130	0.0137
	EMD	0.434	0.424

4 実験

4.1 実験設定

3章で述べた精子グレード分布推定モデルの学習において、MSEで評価したときの最高精度のモデル(MSE Best Model)とEMDで評価したときの最高精度のモデル(EMD Best Model)における推定結果の違いを比較する。複数の観点から評価するため、推定結果の違いを表す指標には複数の指標を用いた。用いた指標は精子の着床確度の高さに関わるグレード分布における重心の差(Ave Grade Distance)、精子の評価結果のばらつきを表すグレード分布におけるエントロピーの差(Entropy Distance)、分散の差(Variance Distance)、大多数の胚培養士の評価結果の推定精度を表す最大選択率のグレードクラス的一致率(Max Grade Precision)である。使用するデータは動画数: 615, 評価者数: 42, 総評価数: 24533である。学習用データとテストデータに8対2に分割し、損失関数にMSE、最適化手法にSGDを用いて学習、推定を行った。

4.2 実験結果

学習曲線(MSE)を図2, EMDのepochごとの変化を図3にそれぞれ示す。また、図2, 図3からMSEで評価したときの最高精度のモデルとEMDで評価したときの最高精度のモデルが得られる(表1)。これら2つのモデルの比較結果を表2に示す。

表2 MSEとEMDのベストモデルの比較

	MSE Best Model (161epoch)	EMD Best Model (221 epoch)
Ave Grade Distance	0.398	0.391
Entropy Distance	0.294	0.260
Variance Distance	0.337	0.295
Max Grade Precision(@1)	64.2%	61.7%
Max Grade Precision(@2)	86.1%	88.6%

4.3 考察

図2と図3からMSEによるepochごとの変化とEMDによるepochごとの変化が異なることが確認できた。表2を見ると、ほとんどの評価指標においてEMD Best Modelが上回った。これはEMDのほうがグレード分布間の違いとして適切であるためだと考えられる。

最大グレードの一致率(Max Grade Precision)では、EMD Best Modelがわずかに劣ったが、2位まで許容したMax Grade Precision(@2)ではEMD Best Modelが上回った。推定対象の分布は最大グレードと2位のグレードの確信度の差が小さいこともあるため、2位まで許容したときの一致率の方が重要である。

図4はMSE Best Modelに比べ、EMD Best Modelのほうが良い推定を行えたときの例である。

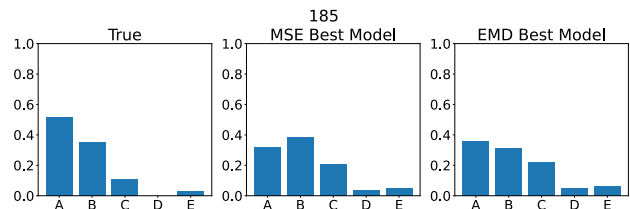


図4 EMD Best Modelのほうが良い推定を行えた例

5 おわりに

本研究では、グレード分布間の距離指標としてEMDを用い、評価分布を推定する方法を提案した。実際のグレード推定モデルの学習でMSE Best ModelよりEMD Best Modelのほうが多くの評価指標において精度が良いことが確認できたため、EMDの優位性が確認できた。今後はグレード分布間の距離指標として適切であることがわかったEMDを損失関数に用いた学習を行う。また、今回実験に使用した精子グレード分布推定モデルを活用し、オンラインによる精子グレード分布推定に応用の予定である。

参考文献

- [1] Rubner, Yossi, Carlo Tomasi, and Leonidas J. Guibas. "The earth mover's distance as a metric for image retrieval." International journal of computer vision, vol.40, No.2, p99-121, 2000.
- [2] K.Simonyan, A.Zisserman:"Very Deep Convolutional Networks for Large-Scale Image Recognition", ICLR, 2015.