

## 生殖医療支援を目的とした顕微鏡動画像における精子検出 Detection of sperms in microscopic videos for assisted reproductive technology

佐々木 勇人<sup>†</sup> 足立 一樹<sup>†</sup> 濱上 知樹<sup>†</sup>  
 山本 みずき<sup>‡</sup> 竹島 徹平<sup>‡</sup> 湯村 寧<sup>‡</sup>  
 Hayato SASAKI Kazuki ADACHI Tomoki HAMAGAMI  
 Mizuki YAMAMOTO Teppei TAKEISHIMA Yasushi YUMURA

### 1. はじめに

男性不妊治療における顕微授精では、培養士の肉眼による精子選別作業が受精率に大きな影響を及ぼす。しかしながら精子選別作業は経験に基づく高度な知識と技術を要すだけでなく、限られた視野・時間のもとで有望な精子を見つけ出さなければならない。特に、この選別作業による受精確率は50~70%程度であり、そのうち着床にまで至る確率はわずかに5%である。加えて一回の顕微授精に要する費用は80万円程度であるため、培養士の負担だけでなく、受診者への負担も極めて大きい。

以上のような問題に対して、本研究では培養士の精子選別の支援を知的処理により実現する。培養士および受診者への負担を軽減することができれば、社会的な問題である少子化問題の解決に大きく貢献することが期待される。現在計画している支援システムは図1のような構成となる。

本稿は支援システムの一つ目の要素にあたる精子検出について扱う。精子検出は後に続く形態・運動特徴の抽出や精子のスコアリングの前処理であり、高い Recall を維持する必要がある。すなわち、選別される可能性のある精子は確実に残す一方で、多少の偽陰性(精子以外の細胞を精子であると誤判定すること)は許容する。しかしながら、偽陰性が増加して Precision が低下してしまうと、精子検出移行の段階で、計算コストが無駄に増加してしまう。したがって、Recall と Precision のバランスを評価することが重要となる。また、精子検出は培養士が精子を見つけ出す際の補助としての役割もある。したがってユーザビリティの向上のためには、精子の検出中に動的に Recall を変更できることが望ましい。

以上のことから、本稿では精子検出における Recall を動的に調整できる手法を構築すると同時に、Recall と Precision のバランスの定量評価を行う。

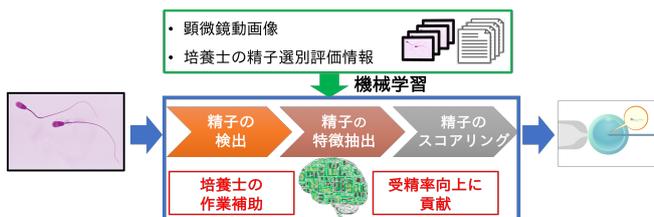


図 1: 精子選別支援システムの全体像

### 2. 精子検出

図2に示すように顕微鏡動画像中には精子以外にも白血球や精母細胞など、精子以外の細胞も多数存在する。精子検出のためにはノイズとなる精子以外の細胞を取除くことが求められる。精子以外を取り除くという目的のもと、精子検出は次の三段階で行う。

1. 候補点(キーポイント)の検出
2. 候補点の判定(精子の識別)
3. 精子の追尾

1段階目では、十分に高い Recall のもとで精子もしくは精子に類似した細胞を対象として候補点を検出する。この操作は、精子の検出を画像全体ではなく局所領域に絞り込むことで、後の工程の計算コストを削減する狙いがある。

2段階目では、1段階目で検出した候補点が精子であるかどうかを詳しく判定する。

3段階目では、前段で精子と判定された点に関して時間・空間的な追尾を行う。時間方向の情報を加えることで、1段階目や2段階目では取りこぼしてしまった点を、前時刻の情報により補うことができ、結果として精子検出の Recall が上昇する。

この3段階の手順のうち特に重要なのが2段階目の精子の識別である。精子の構造は大まかに頭部、尾部に分けられる。顕微鏡で精子を観察する際、精子は画像のX、Y方向だけでなくZ方向(奥行方向)にも移動するため、必ずしも精子に対してピントが合うとは限らない。ピントが合わない場合、精子の尾部は画像中で確認できない場合も多いため、安定した精子検出を実現するためには頭部を検出することが求められる。しかしながら、顕微鏡画像中には精子の頭部に似た円形・楕円形の細胞が多数存在するため、精子の頭部と他の細胞との識別率が精子検出率に大きな影響を与える。



図 2: 顕微鏡画像の例(黄色の枠内が精子)

<sup>†</sup>横浜国立大学

<sup>‡</sup>横浜市立大学附属市民総合医療センター

本研究の精子検出では SIFT[1] や SURF[2] といった局所画像特徴抽出手法により候補点検出および特徴抽出を行い (1 段階目)、それらの特徴を機械学習により精子とそれ以外のクラスに分類する (2 段階目)。そして精子と判定された候補点を Kanade-Lucas Tracking (KLT)[3] により追尾する。この 3 段階をフレームごとに繰り返すことで、顕微鏡動画中の精子の検出を行う。

本稿では、2 段階目の精子の識別を中心的に扱う。

## 2.1. 局所画像特徴

SIFT と SURF はともに勾配ベースの特徴点 (キーポイント) 検出器・特徴記述子であり、照明変化やスケール変化、回転に対するロバスト性がある。SURF は SIFT の高速化を図った手法であり、SIFT のキーポイント検出を近似計算に置き換えることで高速化を達成している。

また、SIFT と SURF は特徴ベクトル (特徴記述子) の構成も異なる。まず SIFT と SURF で共通しているのが、スケールを  $s$  としたときにキーポイントを中心とした  $20s \times 20s$  の領域を  $4 \times 4$  個の領域に分割するという処理である。両手法とも、それらの 16 ブロックに関してベクトルを割り当てる。

SIFT の場合は各ブロック内の勾配方向を 8 方向に分割し、各方向の勾配強度と頻度に基づいて 8 つのビンのヒストグラムを作る。これを 8 次元のベクトルとすると SIFT の次元数は  $4 \times 4 \times 8 = 128$  次元となる。一方 SURF の場合は X 方向と Y 方向の勾配を符号付きで計算する。ブロックの特定の方向に対する勾配は  $2s \times 2s$  の haar-wavelet の応答  $d_x, d_y$  により算出する。具体的には

$$\mathbf{v} = \left( \sum d_x, \sum d_y, \sum |d_x|, \sum |d_y| \right) \quad (1)$$

を各ブロックに割り当てる。したがって SURF の次元数は  $4 \times 4 \times 4 = 64$  次元となる。

SURF は SIFT と比べて半分の次元数である一方で、SURF には 128 次元に拡張したものが存在する (SURF-128)。SURF-128 では次のように  $\mathbf{v}$  を  $\mathbf{v}'$  に拡張する。

$$\mathbf{v}' = \left( \sum_{d_y > 0} d_x, \sum_{d_x > 0} d_y, \sum_{d_y > 0} |d_x|, \sum_{d_x > 0} |d_y|, \sum_{d_y < 0} d_x, \sum_{d_x < 0} d_y, \sum_{d_y < 0} |d_x|, \sum_{d_x < 0} |d_y| \right) \quad (2)$$

ただし  $\sum_{d_y > 0} d_x$  は  $d_y > 0$  の場合の  $d_x$  の総和を表す。

## 2.2. 識別器

本研究では候補点として検出された局所画像特徴 (実数ベクトル) が精子に由来するものか否かを学習により識別する。実数ベクトルの識別器として高い性能を有するのが Support Vector Machine (SVM)[4] である。

### 2.2.1. Class Imbalance

精子識別は精子とそれ以外の 2 クラス分類と捉えることができるが、ここで問題となるのが Class Imbalance

である。Class Imbalance とは、クラスごとのデータ数の偏りのことを意味する。Class Imbalance が生じている場合、正解率は高くなりやすいものの、少数派クラスの Recall が小さくなってしまいう問題が発生する。精子とそれ以外の細胞の個数を比べると顕微鏡動画に映っている精子の個数は後者よりも遥かに少ない。Class Imbalance への対処としては

- リサンプリング
- 識別器損失関数への重み付け

が挙げられる。

今回の精子検出においては、識別器損失関数への重み付けで対処する。本研究で扱う顕微鏡動画中の精子は不妊治療の過程で取得された検体であるため、動きが活発ではないものがほとんどである。したがって、動画の各フレームから抽出した局所画像特徴はすでに重複を多く含んでいるためリサンプリング手法は採用しない。

SVM ではマージン内に入ってしまった訓練データに対する罰則を、ハイパーパラメータ  $C$  により調整する。したがって、SVM で Class Imbalance に対処する場合には、式 (3) に示す訓練データのラベルに応じた重み  $w$  を  $C$  に乗ずる。

$$w = \begin{cases} \frac{N_+ + N_-}{2N_+} & (\text{正例}) \\ \frac{N_+ + N_-}{2N_-} & (\text{負例}) \end{cases} \quad (3)$$

ただし  $N_+$  および  $N_-$  は正例・負例の訓練データ数を表す。このように誤識別時の罰則に関してクラスごとに重みを変更することにより、少数派のクラスが軽視されにくくなる。

### 2.2.2. Recall の調整

精子検出時には Recall を動的に変更することが求められる。すなわち、識別器の学習後に決定関数  $f$  による判定を動的に変更しなければならない。SVM の識別関数の変更は、テストデータ  $\mathbf{x}$  に対して、SVM の decision value を  $f(\mathbf{x})$ 、識別関数を  $\text{sgn}(f(\cdot))$  としたとき、式 (4) 中のしきい値  $b_{th}$  を調整することで実現する。

$$y_{pred}^{b_{th}} = \text{sgn}(f(\mathbf{x}) - b_{th}) \quad (4)$$

ただし  $y_{pred}^{b_{th}}$  はしきい値  $b_{th}$  の元での予測結果である。

## 3. 実験

精子検出における Recall と Precision を定量的に評価する方法として、Precision-Recall 曲線 (以降 PR 曲線とする) の Area Under the Curve (AUC) と F 値を算出した。AUC は PR 曲線下部領域の面積を表しており、1 に近い値を取るほど識別器の性能は高い。F 値は Precision と Recall の調和平均で式 (5) で表現される。

$$F\text{-measure} = \frac{2\text{Recall} \cdot \text{Precision}}{\text{Recall} + \text{Precision}} \quad (5)$$

Recall や Precision の算出は、局所画像特徴ベクトルの識別結果に基づいて行なう。具体的な実験手順は以下のとおりである。

1. 訓練データを用いて候補点の識別器を訓練する
2. テストデータを用いて SVM の Decision Value を算出する
3. Decision Value のしきい値を変化させ、各しきい値で Recall と Precision を算出する
4. Recall-Precision 曲線を引き、AUC を計算する

データセットは次の手順で作成した。

1. 顕微鏡動画を前後半で分割し、前半を識別機の訓練用、後半をテスト用とする
2. 顕微鏡動画からパッチ画像を切り出す
3. 各パッチ画像から局所画像特徴を抽出する

パッチ画像の切り出しは手作業で行った。このとき、精子の頭部を中心として切り出したパッチ画像を正例パッチ、精子以外の細胞を中心としたパッチ画像を負例パッチと設定している。正例パッチに関しては精子の頭部以外からキーポイントが検出されないようにパッチサイズを  $30\text{px} \times 30\text{px}$  に調整している。

正例パッチから抽出された局所画像特徴を正例、負例パッチから抽出された特徴を負例としてラベル付を行った。正例パッチからは 0~2 個の局所画像特徴が抽出された。一方、負例パッチからはより多くの特徴が抽出された。なお、キーポイント検出の際にはすべて SURF を用いている (特徴ベクトルの記述子は別)。

### 3.1. データ諸元

局所画像特徴抽出は SIFT(128 次元) および SURF(64 次元), SURF-128(128 次元) で行った。顕微鏡動画から切り出されるフレームサイズは  $960\text{px} \times 720\text{px}$  である。実験に用いたデータの諸元を表 3.1 に示す。

表 1: 顕微鏡動画から抽出された各種データ

項目	訓練	テスト
フレーム数	4000	3996
局所画像特徴数 (正例)	4889	3607
局所画像特徴数 (負例)	48409	50814

### 3.2. 識別器 (SVM) のハイパーパラメータ

SVM のハイパーパラメータ  $C$  と  $\gamma$  を grid search により決定した。

グリッドサーチは  $C$  を  $(10^{-3}, 10^{-2}, \dots, 10^3, 10^4)$  から選択、 $\gamma$  を  $(10^{-2}, 10^{-1}, \dots, 10^2, 10^3)$  から選択して、訓練データに対して最も F 値が高い組み合わせを探索する。グリッドサーチの結果、 $C = 100.0$ ,  $\gamma = 10.0$  と設定した。

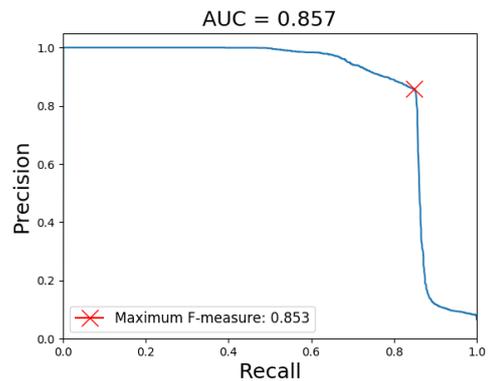


図 3: SIFT に関する PR 曲線

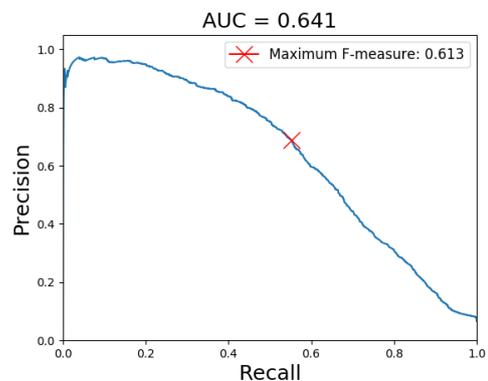


図 4: SURF に関する PR 曲線

### 3.3. Recall-Precision 曲線

Decision value に関するしきい値  $b_{th}$  を変化させることで、訓練済みの SVM に関して PR 曲線をプロットした。しきい値  $b_{th}$  を変化させる範囲は  $\mathbf{x}_{tr}$  を訓練データとして  $[\min_{\mathbf{x}_{tr}} f(\mathbf{x}), \max_{\mathbf{x}_{tr}} f(\mathbf{x})]$  と定めた。PR 曲線はテストデータを元にプロットする。プロット数はテストデータ数のおよそ  $1/10$  である 5,000 点とした。その結果が図 3、図 4 および図 5 である。

図が示すように、SIFT と SURF とでは AUC と F 値に明確な差が生じていることがわかる。

### 3.4. 局所画像特徴の可視化

Isomap[5] は多様体学習の一つであり、高次元空間のデータを低次元空間へ写像することでデータ同士の位相関係を可視化できる。テストに用いた局所画像特徴ベクトルを Isomap で可視化することで、SIFT と SURF とで AUC や F 値に差が生じた原因を考察する。

Isomap により局所画像特徴を可視化した結果が図 6(SIFT)、図 7(SURF) および図 8(SURF-128) である。高次元データを強制的に二次元平面に写像しているため、当然ながら正例 (赤) と負例 (青) との重なりが生じやすくなる。注目すべきはクラスごとのデータのまとまり具合である。定性的ではあるものの、SURF および SURF-128 の可視化結果では正例が複数箇所に散らばっている一方、SIFT の可視化結果では正例が小さな分散のもとまとまっていることがわかる。

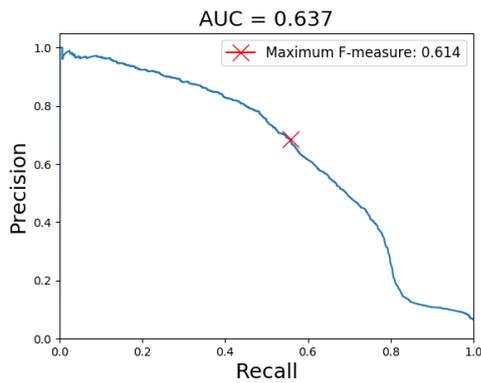


図 5: SURF-128 に関する PR 曲線

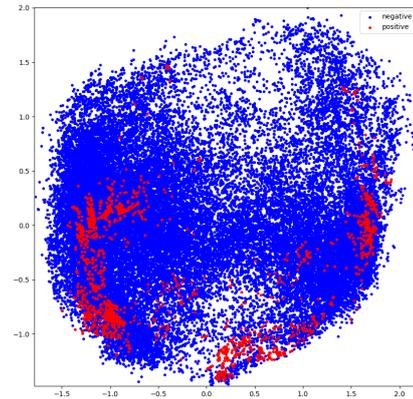


図 7: Isomap による SURF の可視化

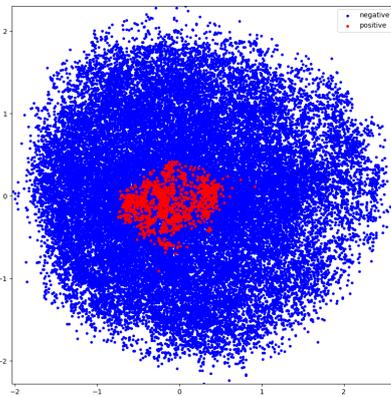


図 6: Isomap による SIFT の可視化

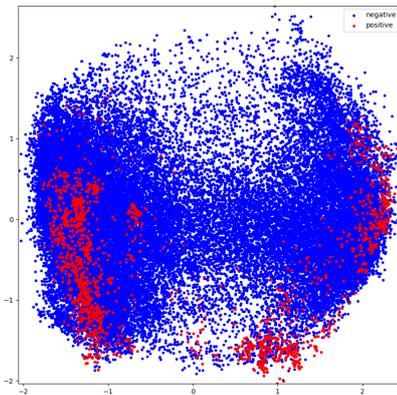


図 8: Isomap による SURF-128 の可視化

#### 4. おわりに

顕微授精における精子選別作業の支援を目的とした精子検出について、局所画像特徴と機械学習を利用した精子識別の評価を行った。

SVM の Decision value に関して、しきい値の変更により Recall を調整するモデルの評価を行った結果、SIFT と SURF とで AUC および F 値に明確な差が現れた。

この評価結果から局所画像特徴として SIFT のほうが高い識別性能を持つことが明らかになった。一方、SIFT は計算コストの高いため、後に続く精子の形態・運動特徴の抽出や精子の評価の段階で足かせになることが予想される。今後は今回の結果をベースラインとして、Deep Learning のような特徴抽出方法も同時に学習する手法を活用し、精子画像に特化した特徴抽出器、識別器の検討を進める。

#### 参考文献

[1] David G Lowe “Distinctive image features from scale-invariant keypoints”. *International journal of computer vision*, Vol. 60, No. 2, pp. 91–110, 2004.

[2] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. “Surf: Speeded up robust features”. In Ale Leonardis, Horst Bischof, and Axel Pinz, editors, *European conference on computer vision*, pp. 404–417. Graz, Austria, Springer, 2006.

[3] Jean-Yves Bouguet “Pyramidal implementation of the affine lucas kanade feature tracker description of the algorithm”. *Intel Corporation*, Vol. 5, No. 1-10, p. 4, 2001.

[4] Corinna Cortes and Vladimir Vapnik “Support vector machine”. *Machine learning*, Vol. 20, No. 3, pp. 273–297, 1995.

[5] Joshua B Tenenbaum, Vin De Silva, and John C Langford “A global geometric framework for non-linear dimensionality reduction”. *science*, Vol. 290, No. 5500, pp. 2319–2323, 2000.