

# 細胞数カウントに向けた CNN を用いた尿中有形成成分の分類

浅倉健太<sup>1</sup> 越村三幸<sup>2</sup> 池田大輔<sup>2</sup> 藤田博<sup>2</sup>

**概要:** 尿沈渣検査では、尿中成分である赤血球、白血球、上皮細胞、細菌などを正確に分類・カウントする必要がある。本研究では、顕微鏡画像において尿中の各有形成成分の分類を行うことを目的とする。深層学習の一手法である畳み込みニューラルネットワークを用い、100 検体の顕微鏡画像から切り出した有形成成分の画像約 1 万枚で分類学習を行ったところ、正答率 97.26% のモデルを得た。学習画像の取得に使用した機器の後継機として、解像度の向上等を行った別モデルにおいて追加調査をした所、同等の性能とならないことを確認したのでこれらを合わせて報告する。

**キーワード:** 細胞数カウント, 尿沈渣, 顕微鏡画像, 深層学習, Fully Convolutional Networks, LeNet

## Toward Cell Counting in Urinalysis: Making a Naive Neural Network Model for Early Study

KENTA ASAKURA<sup>†1</sup> MIYUKI KOSHIMURA<sup>†2</sup>  
DAISUKE IKEDA<sup>†2</sup> HIROSHI FUJITA<sup>†2</sup>

**Abstract:** In urinalysis, we need to classify and count urine sediment particle like Red Blood Cells(RBCs), White Blood Cells(WBCs), epithelial cells or bacteria in a microscopic examination. This study was designed to assess the classification performance of Convolutional Neural Network(CNN) in urine microscopic images. We collected 100 urine samples and clipped 10,128 images with three kinds of labels; RBCs, WBCs and others. Then, we obtained a model with 97.26% accuracy in classifying. Also, we have applied the model to the successor machine so as to examine its performance in another modality.

**Keywords:** cell counting, urinary sediment, urinalysis, microscopic image, deep learning, Fully Convolutional Networks, LeNet

### 1. はじめに

細胞数のカウントはそれ自体が定量的な指標であり、これを用いて作られた半定性的な指標と合わせて様々な用途で利用されている。細胞数カウントを用いた検査はその需要の大きさから業務効率化のため自動化が望まれる一方で、製品として考えると精度をいかに保障するかが問題となる。自動運転や医療機器の分野においては、AI の説明可能性によって動作の妥当性を示す試みもある。

我々は現在、尿中有形成成分分析装置 USCANNER (東洋紡製) における深層学習応用に取り組んでいる。USCANNER は、撮像、物体検出、分類、カウントを自動で行う。画像を用いた自動化では、画像自体を保存しておくことで同じ画像を人が確認することが出来るため、説明可能性確保の前段階として検証可能性が確保される。USCANNER において、機械学習により行われている撮像後の作業を深層学習で置き換えるのが我々の目標である。USCANNER は現在、撮像部の変更が予定されている。そこでまず初期検討として、現世代の USCANNER (E) で取得された画像群を用いて畳み込みニューラルネットワーク (Convolutional Neural Network, CNN) による分類実験を行い、深層学習が

適用可能であるか調べた。その後、次世代機である USCANNER premio で取得された画像群における性能実験を行い、モデルの違いによる影響を調査した。

### 2. 手法

本研究では、USCANNER の機能のうち分類部分を置き換えることを考える。USCANNER が認識する有形成成分は成分ごとに大きさが異なるため、そのうちの赤血球および白血球を対象として {赤血球, 白血球, その他} の 3 種類に分類する。

#### 2.1 データの取得

本研究における対象成分はそれぞれ異なる病気の疑いを示唆するものであり、学習に十分なデータを得るためにはこれらが含まれる検体を選ぶ必要がある。そこで本研究では、USCANNER によって取得されたデータのうち、専門家が有形成成分に関して主観評価を行い、有形成成分同士の重なりが少ないデータを選定した。

以下、USCANNER (E) により取得されたデータを学習データ、USCANNER premio で取得されたデータを検証データと呼ぶ。学習データは専門家により選定された 100 件、それぞれ 25 枚の画像 (1600×1200[pixel]) と、USCANNER

<sup>1</sup> 九州大学大学院システム情報科学府  
Graduate School of Information Science and Electrical Engineering, Kyushu University  
<sup>2</sup> 九州大学大学院システム情報科学研究院  
Faculty of Information Science and Electrical Engineering, Kyushu University

が検出した有形成分の位置情報、分類、および分類毎のカウントを含む。また、同データから、専門家が対象成分に関して画像を確認しながら USCANNER の分析結果を編集し、Ground Truth (以下、GT) を作成した。検証データは赤血球を多く含む 4 件および追加取得された白血球を多く含む 4 件の計 8 件であり、DPI の変更に伴い画像のサイズが 2048×1536[pixel]である以外は学習データと同様である。

## 2.2 CNN の構成および学習

本検討では、入力画像のサイズ変更に対応できるよう全層畳み込みニューラルネットワーク (FCN: Fully Convolutional Network) を用いた。モデルの構造は LeNet1 を参考に全結合層を畳み込み層に置き換えてモデルを作成し、深層学習を適用した場合のベースラインを定めることを目的とした。

CNN の入力は 40×40[pixel] の RGB 画像であり、各ピクセルの値を 1/256 として 0-1 の範囲に収まるよう線形変換を行う scale 層を用いた。また、提案手法におけるニューラル層は、LeNet が畳み込み 2 層、全結合 2 層の計 4 層であったため、これに合わせて畳み込み 4 層とした。LeNet 等で見られる畳み込みの pooling 層であるが、本検討では使用しない。全結合層を排した影響として、ニューラル層の出力が入力に応じた広さを持つ形状になるため、ニューラル層の次に global pooling 層において出力の広さを 1×1 に変換した後、Softmax による活性化を行う。

表 1 FCN-1 の構造

層	入力	カーネル	活性化関数	出力
Data	3@40x40	-	-	3@40x40
Scale	3@40x40	-	-	3@40x40
Conv1	3@40x40	5x5	-	50@40x40
Conv2	50@40x40	9x9	-	125@40x40
Conv3	125@40x40	1x1	ReLU	700@40x40
Conv4	700@40x40	1x1	-	3@40x40
global pooling	3@40x40	-	Softmax	3@1x1

表 2 FCN-2 の構造

層	入力	カーネル	活性化関数	出力
Data	3@40x40	-	-	3@40x40
Scale	3@40x40	-	-	3@40x40
Conv1	3@40x40	5x5	-	50@20x20
Conv2	50@20x20	5x5	-	125@10x10
Conv3	125@10x10	11x11	ReLU	700@10x10
Conv4	700@10x10	1x1	-	3@10x10
global pooling	3@10x10	-	Softmax	3@1x1

## 2.3 分類実験用データの作成

USCANNER により取得される画像には複数の有形成分が含まれる。そのためまず事前準備として、元画像から USCANNER が出力した有形成分の位置情報を基に部分画像を切り出すプログラムを作成した。これにより、分類のための部分画像生成を自動化できる。部分画像のサイズは、学習データにおいて白血球の直径がおおよそ 37[pixel]であったため、USCANNER が出力する位置情報を中心に 40×40[pixel]とし、切り出した画像に GT を基としたラベル付けを行った。その後、再度対象成分のラベルが付いている切り抜き画像を確認し、修正を行った。結果として、学習データに 3 種合計で 14033 枚の画像を得た。詳細を表 3 に示す。

表 3 学習データの構成

種別	学習用	評価用	合計[枚]
赤血球	1162	387	1549
白血球	1767	589	2356
その他	7596	2532	10128

## 2.4 追加検討

モデルの学習は学習データのモーダルに最適化される。本研究において、学習データの作成には専門家の確認を行って GT を作成する必要があるため、現行機のデータを用いて学習したモデルが次世代機のデータに対応可能か検討を行った。解像度の変更により DPI は 1.28 倍になっているが、丸めるため 1.3 倍として 52×52[pixel]で切り出し、linear に 40×40[pixel]へ変形した。追加検討においては赤血球および白血球のみに着目する。検討用に初め 4 件 219 枚の画像を得たが、そのうち 3 枚の白血球画像に関して正答率 100% であったため追加で 4 件 323 枚の画像 (うち白血球は 291 枚) を得たのでこれらを合わせて計 542 枚の画像として評価する。

## 3. 実験 1

### 3.1 FCN の作成

モデルの説明可能性を下げる要因として最適化によりハイパーパラメータ探索を行うことが挙げられる。本研究ではハイパーパラメータ探索を行わずにモデルを策定する試みとして以下のように定めた。まず、神経科学における受容野を参考に、モデルを理解しやすいよう全層の出力サイズを 40x40 とした。また、受容野の考えに基づき、Conv2 のカーネルサイズを出力サイズが 2 倍になるのに合わせて 5x5 から 9x9 とした。次に入力が mnist のようなグレースケールではなく RGB の画像を使用するため次元数を増加させた。次に、全結合層の出力はベクトルであり、ベクトルは広さ 1x1 のテンソルであるから、全結合層から畳み込み層に変更した Conv3 および 4 のカーネルサイズを 1x1 としてモデル FCN-1 を作成した。

表 4 FCN-1 の混同行列

		予測			合計
		赤血球	白血球	その他	
正解	赤血球	347	34	6	387
	白血球	28	521	40	589
	その他	6	54	2472	2532
	合計	381	609	2518	3508

実験結果の混同行列を表 4 に示す. 表 4 より, 全体の正答率は 95.21% である一方, 赤血球および白血球に関する性能指標は低く, 赤血球に関する感度は 89.66%, 白血球では 88.46% となった.

### 3.2 考察

対象成分の分類感度が低い原因を考える. 感度は正答率より低いもののチャンスレベルよりは高いことから, ある程度は対象成分に関する学習が出来ており, FCN-1 の構造に性能を低下させる積極的な要因があると考えられる.

そこで Conv4 の出力 1x1 に対応する入力を考えてみると, Conv4 および Conv3 ではカーネルサイズが 1x1 であるため padding が 0 となり 1x1 であるが, Conv2 および Conv1 ではそれぞれ 4 および 2 の padding が必要となるため 1+4x2+2x2 で 13x13[pixel] の広さが受容野となる. 白血球の大きさがおよそ 37x37[pixel] であるから不適切な受容野が設定されていることが分かる.

以上を受け, 次節では受容野に関して改善を行った.

## 4. 実験 2

### 4.1 FCN の改良

実験 1 より, 受容野の改善を考えた. まず LeNet と同様に出力サイズに関して Conv1 および Conv2 のストライド (stride) を 2 とすることで半分になるようにし, 同時に conv2 のカーネルサイズを 5x5 にすると 13x13 の受容野となり, 実験 1 と等しい. 次に全結合層から畳み込み層に変更した Conv3 および Conv4 について考える. LeNet の第 4 層はベクトルからベクトルへの変換であるから Conv4 のカーネルサイズは 1x1 で正しいが, 第 3 層は前層の出力全てからベクトルへの変換であるから Conv3 のカーネルサイズが不適切であると言える. 以上から, Conv3 のカーネルサイズを Conv2 の出力サイズである 10x10 以上となるように大きくすればよいと考えられる.

以上より, Conv3 のカーネルサイズを 11x11 (カーネルサイズは奇数にする必要がある) としたモデル FCN-2 を作成し, 実験 1 と同様に学習・評価を行った. FCN-2 の混同行列を表 5 示す. また, 2 値分類としてみた場合の性能指標の値を表 6 に示す. 全体の正答率は, 97.26% であり, 多クラス分類における性能指標 [2] を計算すると, マイクロ平均は全体の正答率と同じく 97.26%, マクロ平均では正答率 98.18%, 適合率 95.09%, 再現率 95.86% であった.

表 5 FCN-2 の混同行列

		予測			合計
		赤血球	白血球	その他	
正解	赤血球	347	34	6	387
	白血球	28	521	40	589
	その他	6	54	2472	2532
	合計	381	609	2518	3508

表 6 各性能指標の値

	赤血球	白血球	その他	平均
正答率	.9875	.9778	.9800	.9818
適合率	.9409	.9223	.9897	.9509
再現率	.9457	.9474	<u>.9826</u>	.9586
特異度	<u>.9926</u>	<u>.9839</u>	.9734	.9833

### 4.2 考察

医療の分野で一般に使われる評価指標として感度と特異度があり, このうち感度は再現率と等しい. 特異度のマクロ平均を計算すると 98.33% となった. 表 6 より, それぞれの分類で見ると, 赤血球および白血球の分類では感度が特異度より低く, その他の分類では逆に感度の方が特異度より高くなった. FCN-2 における, 受容野を考慮したカーネルサイズの設定は有効であると考えられる.

## 5. 追加検討

学習時と異なるモーダルでは期待される性能が発揮されないことが予想される. 追加検討として, 学習時と異なるモーダルにおける性能調査を行った. 2.4 節で述べた検証データにおいて FCN-2 を評価した結果を混同行列として表 7 に示す. また, 実験 2 と同様に各性能指標の値を計算した結果を表 8 に示す.

検証データにその他が含まれないため分母が該当ラベルの総数となる再現率で比較を行うと, 実験 2 では赤血球白血球それぞれ 94.57%, 94.74% (表 6) に対し, 学習時と異なるモーダルではそれぞれ 71.37% および 71.09% (表 8) と明らかに低い. 一方で, 統計的情報のみ利用し, 全て赤血球白血球と答えるなどした場合はチャンスレベルとして考えると, 今回は全て白血球と答えることで正答率 54.24% を得られるが, このチャンスレベルよりは高い.

以上から, 以下のことが言える. まず, モーダル変更においては, モーダル変更後の学習データを学習に使用することが望ましいと考えられる. 次に, チャンスレベルとの比較で, モーダル変更前に獲得された特徴量の一部がモーダル変更後においても効果的に働いていると思われる. モーダル変更前のデータに希少な検体がある場合などは, 一般にマルチモーダルにおいて行われているように, 両モーダルのデータを使用することや, 転移学習を行うことも有用であると考えられる.

表 7 マルチモーダルにおける FCN-2 の混同行列

		予測			合計
		赤血球	白血球	その他	
正解	赤血球	177	36	3	248
	白血球	10	209	75	294
	その他	0	0	0	0
合計		187	245	10	542

表 8 マルチモーダルにおける標の値

	赤血球	白血球	その他	平均
正答率	.8506	.7768	-	.8137
適合率	.9465	.8531	-	.8998
再現率	.7137	.7109	-	.7123
特異度	<u>.9660</u>	<u>.8548</u>	-	.9104

## 6. おわりに

本研究では、顕微鏡画像診断における深層学習の導入検討として実験を行い、正答率 97.26% のモデルを得た。また、追加検討により、異なるモーダルにおいてモデルの性能が低下することを確認した。

本研究のように、診断機器において深層学習を導入する場合、導入機器における学習データの収集が望ましいが、診断機器ではどこに設置されている製品でも同じ結果が求められるため、発売後の再学習は禁止されると予想される（専門部会の議事録が公開されている[3]）。また、本結果が示すように、モデル変更の影響は少なくないと考えられるため、ベースラインとして変更前のモデルを生成することは有用であるが、同じモーダルで学習したモデルよりは性能が低下すると予想できる。この点に関して、モデル変更後の機器と深層学習のモジュールのリリース時期をずらす等、深層学習の導入を促進するには何らかの打開策が必要であると考えられる。

また、今後の課題として、本研究では性能を向上させるための手法を適用しなかったが、ハイパーパラメーター探索や一般画像認識で高性能を記録しているモデルの利用などが挙げられる。また、展望として、FCN-1 の Conv4 のように、出力サイズを入力サイズと同じにすると、分類に関するヒートマップ（図 1）が得られる。ヒートマップからカウントを回帰する研究として[4]が挙げられる。また、このヒートマップをさらに発展したものとして、画像のセグメンテーション[5][6]が挙げられる。このセグメンテーションは、物体検出において画像の 2 値化等が用いられているが、これに代わる技術である。

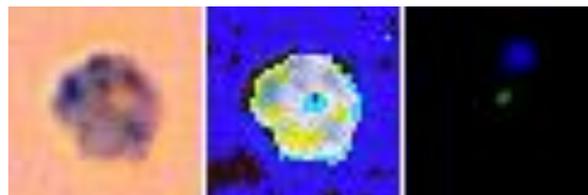


図 1 FCN により得られるヒートマップの例  
(元画像（左，白血球），処理後の画像（中央），Conv4 の出力を加工したもの（右）。右図では緑が白血球，青がその他を表し，値は緑が一番大きいため max pooling により最終的な出力は白血球となり一致する。)

**謝辞** 本研究は東洋紡株式会社との共同研究によるものである。ご討論いただいた同社の奥田洋平氏に感謝する。

## 参考文献

- [1] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” Proc. IEEE, vol.86, no. 11, pp.2278–2324, 1998.
- [2] R. Sebastian and M. Vahid, “多クラス分類のための性能指標,” Python 機械学習プログラミング 達人データサイエンティストによる理論と実践, 2nd ed., 株式会社クイープ訳, 福島真太郎監訳. 株式会社インプレス, 東京, 2018, p.207.
- [3] “AI 専門部会 | 独立行政法人 医薬品医療機器総合機構”. <https://www.pmda.go.jp/rs-std-jp/subcomm-3rd/0007.html>, (参照 2018-11-08).
- [4] D. Onoro-Rubio and R. J. Lopez-Sastre, “Towards Perspective-Free Object Counting with Deep Learning,” ECCV 2016 14th Eur. Conf. Amsterdam, Netherlands, pp.615–629, 2016.
- [5] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” Medical Image Computing and Computer-Assisted Intervention -- MICCAI 2015, pp.234–241, 2015.
- [6] B. Romera-Paredes and P. H. S. Torr, “Recurrent instance segmentation,” Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol.9910 LNCS, pp.312–329, 2016.