

## 【原著】

# 畳み込みニューラルネットワーク(CNN)を用いた末梢血白血球分類スクリーニング技術の検討

佐々木亜実\*<sup>1</sup> 小田未来\*<sup>1</sup> 野坂大喜\*<sup>2</sup> 中野学\*<sup>2</sup> 藤岡美幸\*<sup>2</sup> 高見秀樹\*<sup>2</sup>

(2020年1月15日受付, 2020年2月27日受理)

**要旨:** 医療用 AI は, 熟練者の経験に基づいて診断された患者症例を教師データとして大量に学習することで, ユーザーの経験値に左右されず, かつ EBM に即した診断を実現する次世代の医療技術である。本研究では AI による末梢血白血球分類スクリーニング技術の有用性を検討した。対象は MG 染色を行った健常人末梢血塗抹標本 57 例とし, 典型的白血球形態画像を用いた転移学習と Fine-tuning によって得た CNN モデルを用いて白血球画像分類精度を比較評価した。その結果, 背景なし教師画像による教師学習では 5 分類で 0.990, 6 分類において 0.822 の Accuracy を示した。一方, 背景あり教師画像による学習では 5 分類で 0.992, 6 分類において 0.879 の Accuracy を示した。CNN による末梢血白血球分類スクリーニング技術は Accuracy が高く有用であると考えられるものの, その臨床応用化においてはカットオフ値など境界領域細胞の判定保留条件などを検討する必要がある。

**キーワード:** 人工知能, 畳み込みニューラルネットワーク, 血液形態検査, 白血球分類

## I. はじめに

近年, 人間の思考回路を反映させた判断を可能とする自動化技術として人工知能(Artificial Intelligence, AI)が急速に発展している。AI は言語の理解や推論, 問題解決などの知的行動を人間に代わってコンピュータに行わせる技術である。AI は既に様々な商品・サービスに組み込まれ活用が始まっている。社会実装の成功例としては, 画像認識と音声認識から得られた情報を基に, 車両の運行情報・地図情報・位置情報などの他の情報を加えて, 車両がおかれた状況を識別する自動運転技術がある。本技術により衝突の可能性の予測, 安全を保つために最適な運転, 目的地に到達するための最適経路の検索が可能となっている<sup>1)</sup>。農業分野においては, 画像からウメの果実の熟度を判断する技術が研究されている。AI が果実の最適収穫期を判定することで, 品質の向上とバラツキの低減が期待されている<sup>2)</sup>。同様に医療分野においても, CT や MRI などの放射線画像診断, 病理画像診断, 内視鏡画像診断等において診断専門医支援のための AI 開発が進められている<sup>3)</sup>。放射線画像診断では, マンモグラフィ画像検査において深層学習技術によるコンピュータ支援型検出システムの有用性が検討されているほか, 軟部組織と石灰化の所見を組み合わせ, 疑わしい領域の所見を判断し, 癌が存在する疑いのレベルを表す AI 診断補助システムが検討されている。Alejandro Rodríguez-Ruiz らの研究報告では, 放射線科医は追加の読

影時間を必要とせずに, マンモグラフィでの乳がんの検出における診断性能を改善したとの結果が示されている<sup>4)</sup>。また病理診断分野においては, Andrew A. Borkowski らは, Google と Apple が提供する AI の機械学習を用いて, 肺扁平上皮癌, 肺癌, 結腸腺癌の区別能力の比較を行ったところ, どちらも非常に良好に機能し, 組織病理学的画像を高精度で識別および分類可能であったと報告している<sup>5)</sup>。以上のことから AI 技術は医療分野においても画像診断を中心としてその臨床応用化の取り組みが開始されている。医療用画像診断において多くの取り組みが開始されている一方, 同様に形態学的解析を必要とする血液形態学診断においては AI 技術を用いた研究報告は少ない。従来, 血液形態学検査においては末梢血液標本や骨髓塗抹標本について顕微鏡下での白血球分類を行うことで, 異常白血球や幼若細胞を検出し, 白血病やリンパ腫, 骨髓異形成症候群などの診断を行ってきた。しかし骨髓塗抹標本での白血球細胞分類については技術の修得において熟練までに長いトレーニング期間が必要とされ, 血液検査専門技師や血液専門内科医の不足とも重なって自動診断技術の実用化が強く望まれている。加えて, 千葉らはフローサイトメトリー法による血球分類技術の評価検討を行い, フローサイトメトリー方式では正常血液細胞の場合には分析精密度は高いものの, 化学療法による血球形態の変化や血液疾患患者検体の分析では異常なスキッタグラムパターンを示し正確性が劣ることを報告している<sup>6)</sup>。そのためこれらの課題に対応し, EBM に即した分析技術開発が必要とされており, AI を用いた新たな形態解析技術が期待されている。そこで本研究では, 新たな形態解析技術として畳み込みニューラルネットワーク(CNN)による末梢血白血球分類スクリーニング技術の検討を行った。

\*1 弘前大学医学部保健学科  
Hirosaki University Graduate School of Health sciences  
〒036-8564 青森県弘前市本町 66-1 TEL:0172-39-5918  
66-1, Honcho, Hirosaki-shi, Aomori, 036-8564, Japan  
\*2 弘前大学大学院保健学研究科  
Hirosaki University Graduate School of Health sciences  
〒036-8564 青森県弘前市本町 66-1 TEL:0172-39-5918  
66-1, Honcho, Hirosaki-shi, Aomori, 036-8564, Japan  
Correspondence Author h16m2317@hirosaki-u.ac.jp

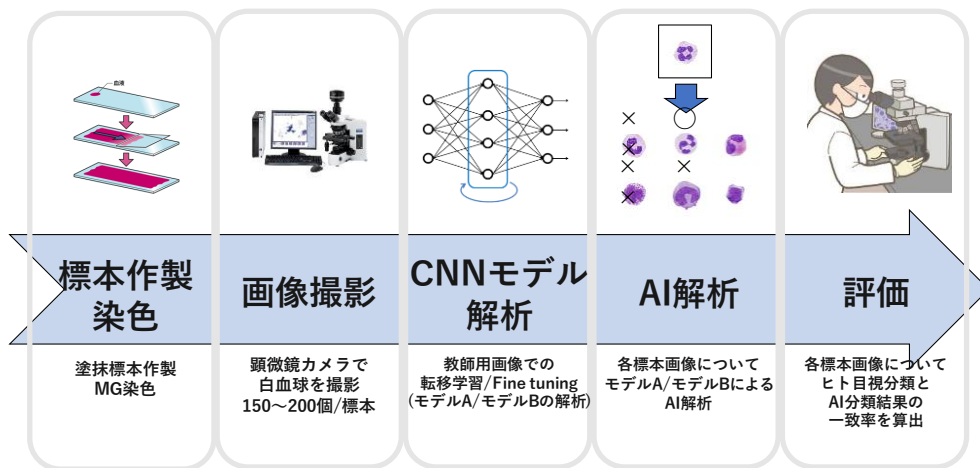


図1 末梢血白血球分類のためのCNNモデル作成から標本解析までのフローチャート

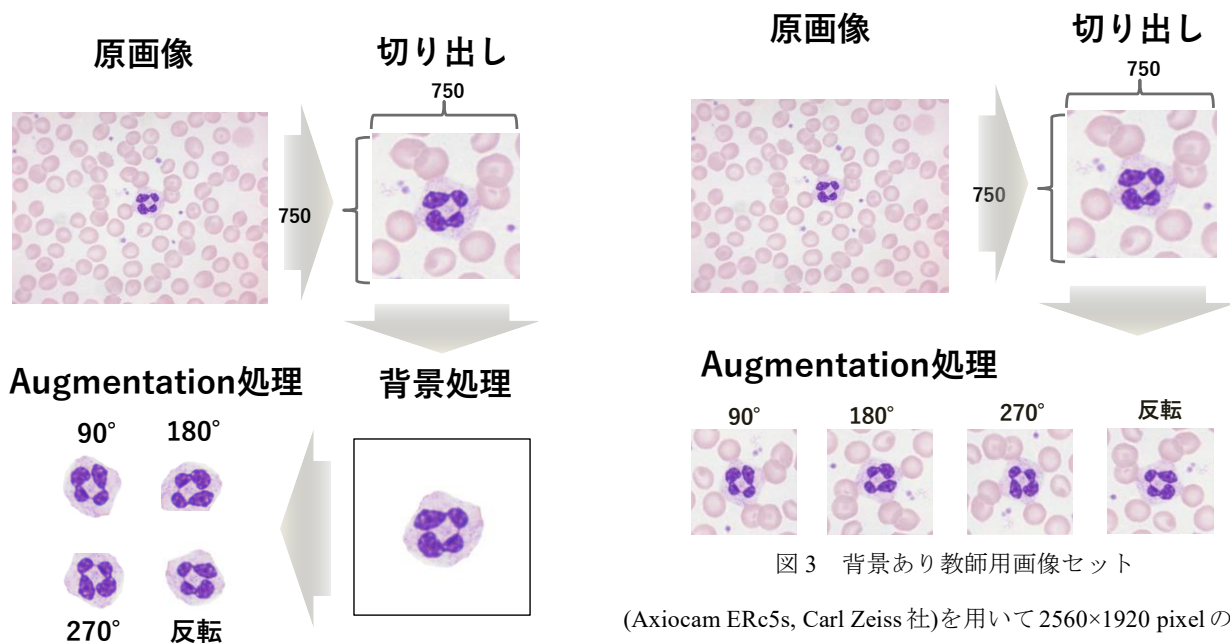


図2 背景なし教師用画像セット

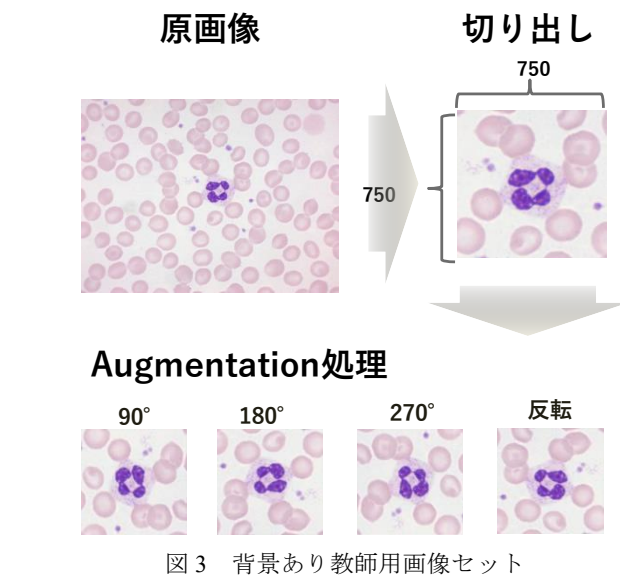


図3 背景あり教師用画像セット

## II. 研究対象と方法

図1に末梢血白血球分類のためのCNNモデル作成からAIによる標本解析までのフローチャートを示す。

### 1. 対象

#### (1) 教師学習用画像

May-Grünwald・Giemsa (MG)染色を行った健常人薄層血液塗抹標本40例とした。MG染色標本は対物100倍油浸レンズを用いて顕微鏡観察し、顕微鏡カメラ(Axiocam ERc5s, Carl Zeiss社)を用いて典型的細胞像を示す正常白血球細胞を2560×1920 pixelの解像度にて撮影、合計1335枚の正常白血球画像を得た。

#### (2) 推論評価用画像

MG染色を行った健常人薄層血液塗抹標本57例とした。MG染色標本は対物100倍油浸レンズを用いて顕微鏡観察し、1標本につき150~200個の白血球細胞を顕微鏡カメラ

(Axiocam ERc5s, Carl Zeiss社)を用いて2560×1920 pixelの解像度にて撮影した。

教師学習および推論評価に供した薄層血液塗抹標本はともに教育用標本であり、その利用においては承諾を得るとともに、匿名化を行った。

### 2. 深層学習システムの構成

深層学習ライブラリにはNnabla (SONY社)を、開発環境にはAnaconda3.0, Python3.5を使用した。また解析用ハードウェアにはCPUにIntel(R) Core(TM) i7-8700 3.2GHz, GPUにNVIDIA GeForce GTX 1070 8GB, OSにMicrosoft Windows 10 professionalを使用した。

### 3. 画像データセットの作成

#### (1) 教師用画像データセットの作成

撮影した教師学習用白血球画像は目視により桿状核好中球(Band), 分節核好中球(Segment), 好酸球(Eosinophil), 好塩基球(Basophil), 単球(Mono), リンパ球(Lymph)に6分類した。判定基準は日本検査血液学会が示す方法に従い<sup>7,8)</sup>, 好中球はBandとSegmentとの鑑別基準として、Bandは直

径 12~15 $\mu$ m, 核の長径と短径の比率が 3:1 以上, かつ核の最小幅部分が最大幅部分の 1/3 以上で長い曲がった核を持つこととした。一方, Segment は分葉した核の間は核糸でつながるが, 核の最小幅部分が十分に狭小化した場合は核糸形成が進行したとみて Segment と判定した。6 細胞群に分類した画像データを対象として①背景なし白血球画像データセットと②背景あり白血球画像データセットの 2 つを教師用画像学習データとして作成した。図 2, 3 に各教師用画像データセット作成手順を示す。

#### ① 背景なし白血球画像データセット

白血球細胞を中心として周辺部を 750×750 pixel でトリミング後, 赤血球や血小板など背景情報を消去し, 白血球細胞のみの情報を抽出した。得られた画像はさらに学習データに対して変形やノイズを加えることで学習データ数を増加させる画像処理(Augmentation 処理)を施した。各画像に対し Augmentation 処理として 90 度, 180 度, 270 度回転, 反転処理等を加えることで画像枚数を水増しさせ, Band, Segment, Eosinophil, Basophil, Mono, Lymph を同等の割合で含む合計 2385 枚の教師用画像セットを生成した。その後 CNN での学習用画像解像度として 480×480 pixel に解像度を変更した。

#### ② 背景あり白血球画像データセット

白血球細胞を中心として周辺部を 750×750 pixel でトリミング後, Augmentation 処理を施した。各画像に対し 90 度, 180 度, 270 度回転, 反転処理等を加えることで画像枚数を水増しさせ, Band, Segment, Eosinophil, Basophil, Mono, Lymph を同等の割合で含む合計 3284 枚の教師用画像セットを生成した。その後 CNN での学習用画像解像度として 480×480 pixel に解像度を変更した。

#### (2) 推論評価画像データセットの作成

白血球画像撮影後に目視により Band, Segment, Eosinophil, Basophil, Mono, Lymph に 6 分類した。鑑別基準は教師用画像データセットの作成と同様である。各標本において撮像した白血球細胞画像について画像サイズを 750×750 pixel でトリミング処理し, その後さらに CNN での推論評価用解像度として 320×320 pixel を生成した。

### 4. CNN モデルでの画像解析方法

#### (1) 背景なし画像学習による CNN モデル作成と推論評価

背景なし画像による教師学習には CNN モデルとして Kaiming He らが開発した ResNet-18 について分類項目数を 6 分類に改変して使用した<sup>9)</sup>。改変後の CNN 条件は図 4 に示した。改変 ResNet-18 を用い, 背景なし教師用画像データセット 2385 枚についてパラメータとして収束最適化法(Optimizer)は AdaBound 法, 学習回数は 300 Epoch を設定し転移学習を行い, 白血球細胞分類用に重み付けがなされた最適化 CNN モデル(WBC モデル A)を得た。その後, 本 WBC モデル A を用いて各標本画像に対して推論評価を行い白血球細胞の 5 分類と 6 分類での予測値を計算処理した。

Layer Name	Output Size	ResNet-18
conv1	112×112×64	7×7, 64, stride2
		3×3 max pool, stride2
conv2_x	56×56×64	3×3, 64
		3×3, 64
conv3_x	28×28×128	3×3, 128
		3×3, 128
conv4_x	14×14×256	3×3, 256
		3×3, 256
conv5_x	7×7×512	3×3, 512
		3×3, 512
average pool	1×1×512	7×7 average pool
fully connected	6	512×6 fully connections
softmax	6	

図 4 白血球分類用 ResNet-18 モデル

Training と Validation では Input Size は 480×480 pixel とし, Augmentation 処理により 1 学習毎に Contrast, Brightness, Magnification, Angle を変更した。推論評価においては Input Size は 320×320 pixel とし, Augmentation 処理は行わなかった。

#### (2) 背景あり画像学習による CNN モデル作成と推論評価

前述の背景なし最適化 CNN モデルである WBC モデル A を予備転移学習(Pre-training)モデルとして, さらに背景あり教師用画像データセット 3284 枚についてパラメータとして Optimizer は AdaBound 法, 学習回数は 300 Epoch を設定し追加転移学習(Fine tuning)を行い白血球細胞分類用に重み付けがなされた最適化 CNN モデル(WBC モデル B)を得た。その後, 本 WBC モデル B を用いて各標本画像に対して推論評価を行い白血球細胞の 5 分類と 6 分類での予測値を計算処理した。

#### (3) 教師用画像の背景処理の有無による分類精度比較

WBC モデル A と WBC モデル B の白血球分類認識精度について Wilcoxon の順位和検定により統計学的に検証した。有意水準  $p < .05$  を有意差有りとした。

## III. 結果

### 1. 背景なし画像学習による推論評価

図 5 に CNN による学習結果を示す。TRAINING ERROR 曲線と VALIDATION ERROR 曲線には乖離は認められず, VALIDATION ERROR 曲線, COST 関数曲線はいずれも最終的に 0.1 未満まで収束した。図 6 に 2 例の推論評価結果を示す。A-1 は 57 例中 Accuracy が最も高値を示した標本であり, A-2 は 57 例中最も低値を示した標本である。Accuracy は 5 分類/6 分類では A-1 において 0.985/0.949, A-2 において 0.985/0.646 であり 0.303 の差が認められた。表 1 (A-1, A-2)に全 57 標本の推論評価結果を示す。Accuracy は 5 分類で  $0.990 \pm 0.011$  (Mean  $\pm$  SD), 6 分類で  $0.822 \pm 0.062$  (Mean  $\pm$  SD)を示した。

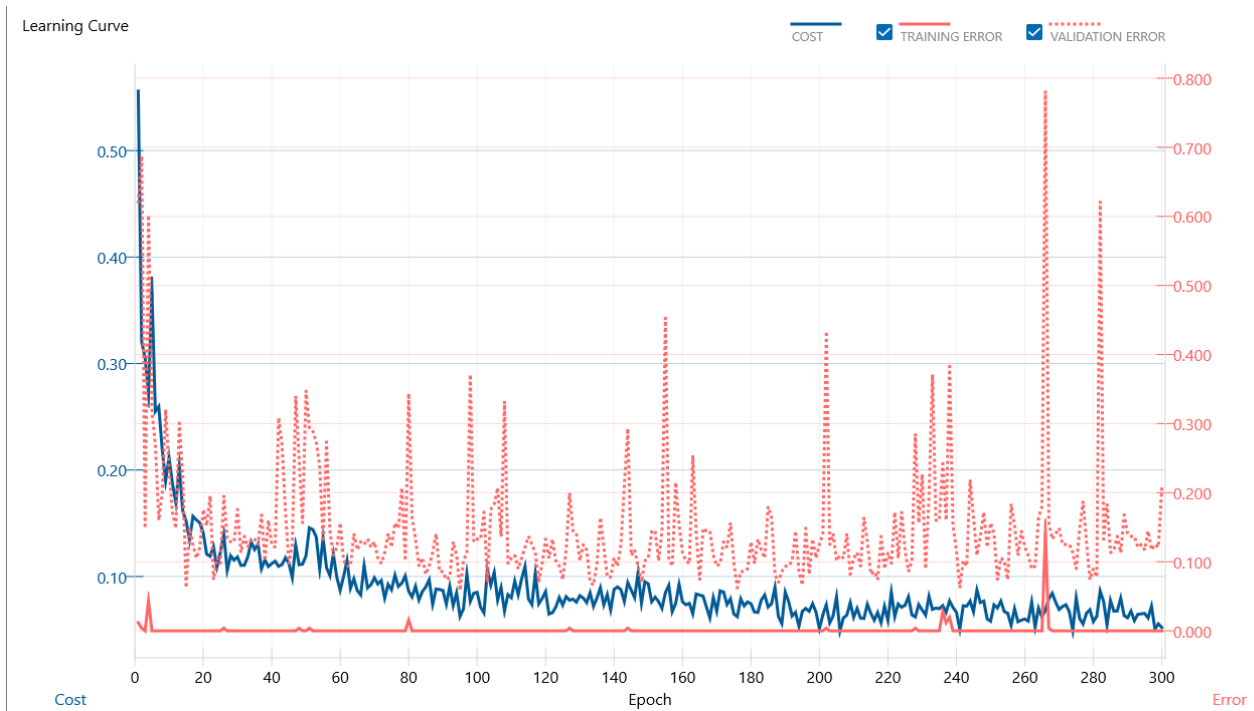


図5 背景なし画像での学習結果

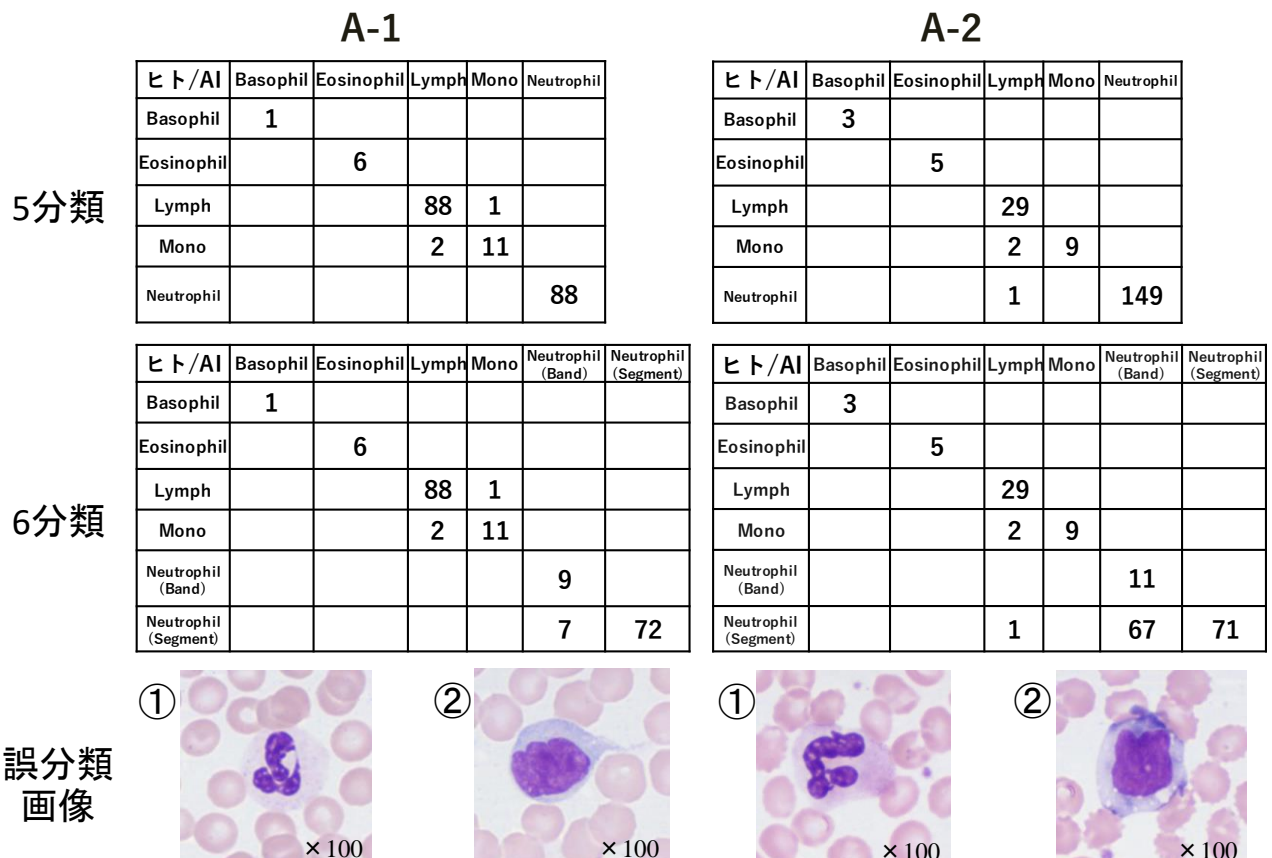


図6 WBCモデルAによる健常人標本2例の推論評価結果

白血球細胞分類でのヒト分類結果との一致率は、5分類/6分類においてA-1で0.985/0.949、A-2で0.985/0.646であった。両症例とも①はSegmentとBand間での誤分類細胞、②はMonoとLymph間での誤分類細胞である。

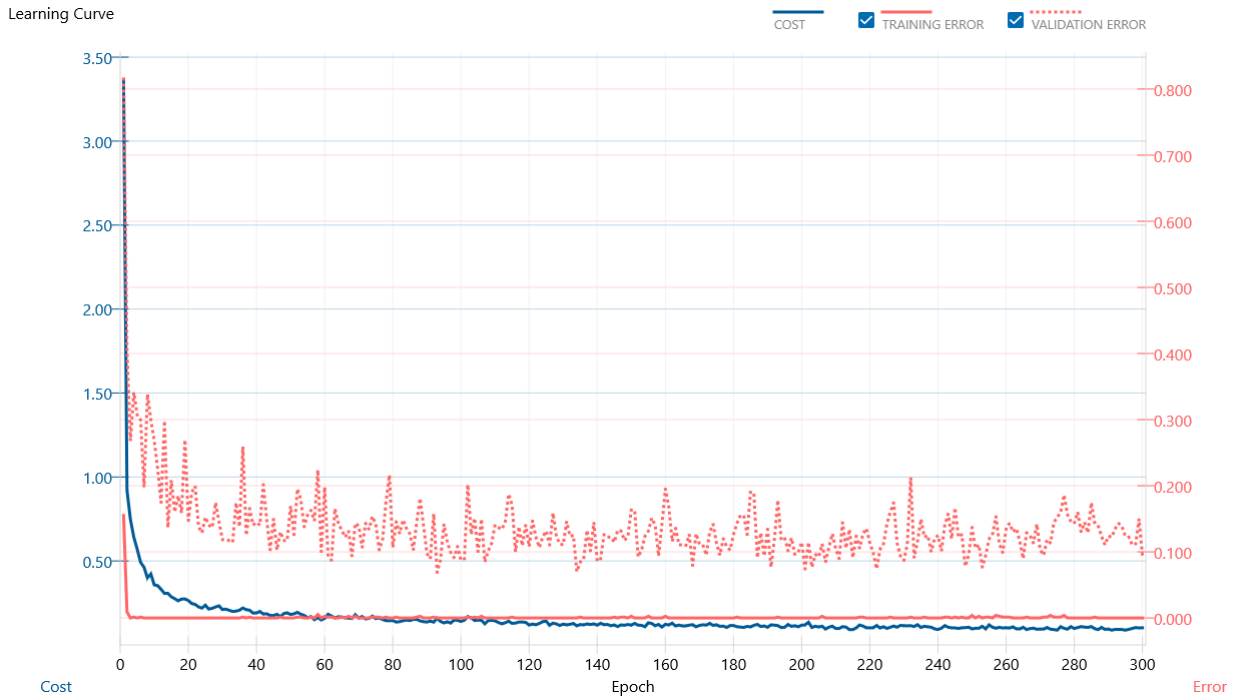


図7 背景あり画像での学習結果

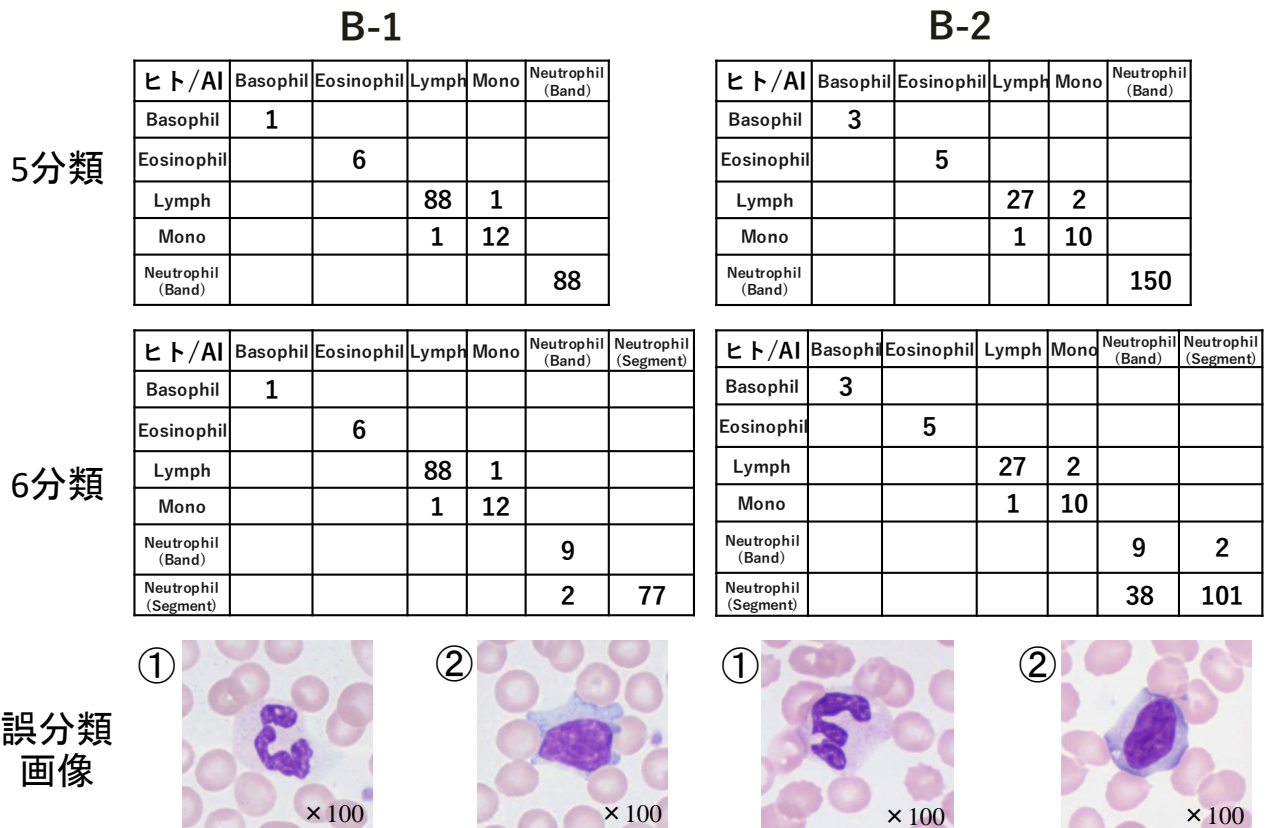


図8 WBCモデルBによる健常人標本2例の推論評価結果

白血球細胞分類でのヒト分類結果との一致率は、5分類/6分類においてB-1で0.990/0.980, B-2で0.985/0.783であった。両症例とも①はSegmentとBand間での誤分類, ②はMonoとLymph間での誤分類細胞である。

表1 全症例での推論評価結果

A1: WBCモデルAによる5分類

ヒト/AI	Basophil	Eosinophil	Lymph	Mono	Neutrophil	Accuracy
Basophil	116					1.000
Eosinophil		246				1.000
Lymph			3179	34		0.989
Mono			61	596		0.907
Neutrophil	14		3		6917	0.998
合計						0.990

A2: WBCモデルAによる6分類

ヒト/AI	Basophil	Eosinophil	Lymph	Mono	Neutrophil (Band)	Neutrophil (Segment)	Accuracy
Basophil	116						1.000
Eosinophil		246					1.000
Lymph			3179	34			0.989
Mono			61	579			0.905
Neutrophil (Band)	1		1		849	80	0.912
Neutrophil (Segment)	13		2		1812	4176	0.696
合計							0.822

B1: WBCモデルBによる5分類

ヒト/AI	Basophil	Eosinophil	Lymph	Mono	Neutrophil	Accuracy
Basophil	116					1.000
Eosinophil		246				1.000
Lymph			3184	29		0.991
Mono		1	43	596		0.931
Neutrophil	9	1		1	6923	0.998
合計						0.992

B2: WBCモデルBによる6分類

ヒト/AI	Basophil	Eosinophil	Lymph	Mono	Neutrophil (Band)	Neutrophil (Segment)	Accuracy
Basophil	116						1.000
Eosinophil		246					1.000
Lymph			3184	29			0.991
Mono		1	43	596			0.931
Neutrophil (Band)	3				810	118	0.870
Neutrophil (Segment)	6	1		1	1152	4843	0.807
合計							0.879

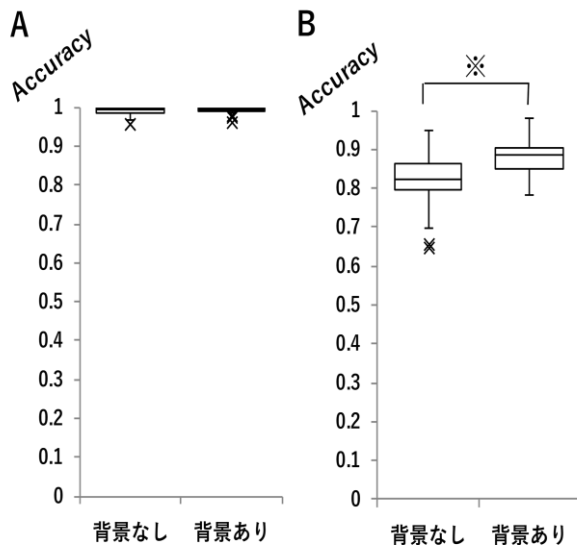


図9 背景処理の有無による分類精度比較

A: 5分類精度比較 B: 6分類精度比較

5分類では両CNNモデルで0.990以上のAccuracyが得られ有意差は認められない。一方、6分類においては背景あり画像学習において分類精度の有意な向上(p<0.001)が認められた。

## 2. 背景あり画像学習による推論評価

図7にCNNによる学習結果を示す。TRAINING ERROR 曲線と VALIDATION ERROR 曲線には乖離は認められず、VALIDATION ERROR 曲線、COST 関数曲線はいずれも最終的に0.1未満まで収束した。図8に健常人標本2例の推論評価結果を示す。B-1は57例中Accuracyが最も高値を示した標本であり、B-2は57例中最も低値を示した標本で

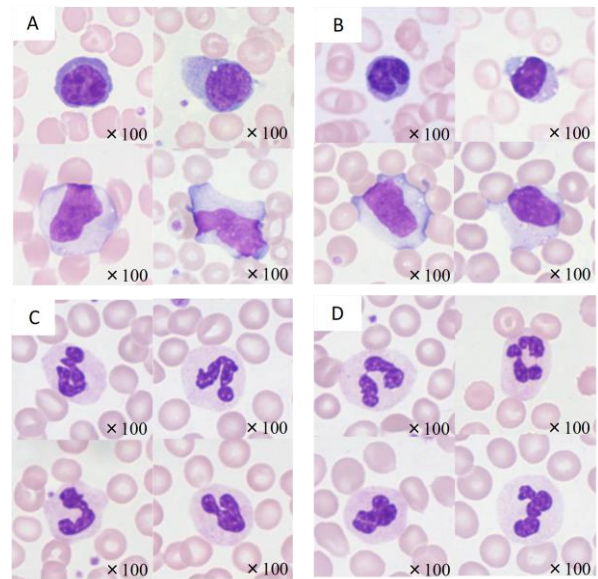


図10 ヒト目視分類-AI分類間での不一致細胞例

- A: Monoに分類されたLymph
- B: Lymphに分類されたMono
- C: Segmentに分類されたBand
- D: Bandに分類されたSegment

ある。Accuracyは5分類/6分類ではB-1において0.990/0.980、B-2において0.985/0.783であり最大0.197の差が認められた。表1(B-1, B-2)に全57標本の推論評価結果を示す。Accuracyは5分類で0.992±0.008(Mean±SD)、6分類で0.879±0.039(Mean±SD)を示した。

## 3. 教師用画像の背景処理の有無による分類精度比較

図9にWBCモデルAとWBCモデルBとの分類精度の

比較結果を示す。WBC モデル A と WBC モデル B との間で、全 57 例での推論評価結果の Accuracy は 5 分類で 0.002, 6 分類で 0.057 の上昇を示した。Wilcoxon の順位和検定は 5 分類で有意差は認められず ( $p=0.473$ ), 6 分類では有意差が認められた ( $p<.001$ )。最も Accuracy が上昇した細胞群は Segment であり, 0.11 の上昇が認められた。図 10 に, Lymph と Mono, Band と Segment の誤認識した細胞の例を示した。Lymph と Mono, Band と Segment の鑑別において, 誤認識された細胞に細胞質形状, 細胞質の濃淡, 核形状など顕微鏡観察上の共通の形態学的な傾向は認められなかった。

#### IV. 考察

本研究では健康成人末梢血塗抹標本を用いて CNN による末梢血白血球分類スクリーニング技術としての有用性を検討した。その結果 5 分類での Accuracy は 0.992 に達し, スクリーニング技術として臨床的要求を充分満たす精度が得られた。一方, 6 分類では Accuracy は 0.879 にとどまり, 白血球分類スクリーニング技術としての臨床的要求を満たすためには改良の余地が認められた。5 分類と 6 分類において精度が大きく乖離した原因として, 好中球である Band と Segment 間での誤分類が大きな原因となっており, 6 分類での精度向上においては好中球の分類精度向上が課題となる。分類精度の向上においては, 本研究において背景あり画像による追加学習において有意差が認められたことから, 教師学習の段階において解析対象となる白血球細胞の形態学的特徴を CNN に与えるだけでなく, 除外すべき情報を加えることが精度向上に有効であると考えられる。しかしこのことは, 津田らによる AI を用いた物体認識精度の報告<sup>10)</sup>とは反している。津田らの研究によると, AI はあるカテゴリーの画像がカテゴリーとは無関係な特徴を持った画像を多く含んでいる状況においては学習用データのバイアスによる誤認識が発生することが示されている。本研究においては, WBC モデル B の導出において対象物のみによる WBC モデル A に対して Fine tuning を行ったことで学習データバイアスが回避された可能性があり, このことは齋藤による AI を用いた病理組織類似画像検索による診断支援ツールの開発の試みにおいて, ImageNet による Pre-training 後の CNN モデルに対して胃生検病理組織 HE 画像を追加学習させることで, 高分類精度な CNN モデルが導出されたとの報告によって支持される<sup>11)</sup>。しかしながら CNN が画像のどの特徴箇所を見て判断しているかは本法のみでは不明である。この課題に対しては, 学習済み CNN モデルが学習用データのバイアスの影響を受けているかどうかを判定する技術として Grad-CAM による解析が有効であることが Ramprasaath らの研究により明らかとなっている<sup>12)</sup>。白血球画像データ解析において本手法が効果を示した要因については本研究に使用したデータに対して Grad-CAM を用いた解析, また他の医療用形態解析 AI にお

ける研究を通じて明らかにできると考える。また CNN では分析の階層を増加させることで分類精度の向上が見込めることが Karen らの研究により報告<sup>13)</sup>されているが, 階層の深さは推論評価モデルの作出解析に要する時間とのトレードオフとなることに留意する必要がある。また Qingshan She らの報告によると, ラベル付けされた教師用画像だけでなく, ラベル付け無しでの教師画像学習も分類精度の向上に寄与することが示されている<sup>14)</sup>。これは半教師用画像学習と呼ばれる手法であり, 臨床検査データではラベル付けされていない多数のデータを基に学習できる本法は症例数を増加させ, 多様な形態解析に対応させる上で有用であるものと考えられることから, 分類精度の向上においてはこれらの手法の導入を検討する必要があると考えられる。

このような AI 分析技術とのハイブリッド化で分析精度向上は見込まれるものの, 一方で臨床検査データはヒトと AI との相関性はもちろんのこと, 従来技術との相関性を必要とすることから, データ補正についても検証が必要となる。機械学習は原理として, あらかじめ装置に設定したパラメータに基づき, 白血球の特徴情報を抽出し, データベースと照合の上で白血球分類を行う。機械学習法では標本が一定の理想的条件下で作製されたことを前提条件としており, 個々の医療者の標本作製に伴う細胞形態変化誤差までは考慮することなく一律に判断を行っている<sup>15)</sup>。それに対して, AI による教師あり学習では与えられた教師用画像から特徴量を自動的に抽出する点において相違があり, 教師用画像の準備過程において検査技術者が作製した多様な条件下での標本を学習させることで, 標本作製時に生じる手技的な誤差を含めた最適なパラメータを得ることが可能となる。本研究においても前処理過程で Augmentation 処理を加えることで画像を AI に自己作成させ拡張を図っている。よって両者は手技的な誤差の判断において基準が異なっており, このような基準値の相違を明確化する必要がある。このような相違点を考慮した例として, 機械学習方式とヒトによる目視分類での好中球判定の相違基準の策定がある。機械学習方式では, 白血球画像に対して RGB 分解を行い, 個々の 2 値化した分解画像に対して解析を行うことでパラメータ毎の実測値を得, 核の最低幅が一定基準 ( $2 \mu\text{m}$  以下) であるかどうかで Segment と Band 判別している。一方, ヒトによる目視分類では同基準での判定は困難であることから, 今回我々が判定基準として使用した渡邊らの方法に従っている<sup>8)</sup>。本基準では好中球は Band と Segment との鑑別基準として, Band は直径  $12\sim 15 \mu\text{m}$ , 核の長径と短径の比率が 3:1 以上, かつ核の最小幅部分が最大幅部分の 1/3 以上で長い曲がった核を持つこととしており, 実測値とは基準が異なっている。CNN モデルによる解析ではヒトの目視基準に準じた解析が行える一方, 機械学習の基準とは異なった判断となることから, これら 2 法間でのデータ解釈の相違が生じやすい形態学的な境界領域での細胞

についての判定について、明らかにする必要があると考えられる。特に、Band と Segment の分類においては目視分類でヒトは2次元画像を基に3次元構造を想像して判断することも無意識に行っていることから、フローサイトメーターによる血球分析データを参考とするなどといったことも分類精度の向上において有用であると考えられる。特に Viktor H. Koelzer らは、AI 解析においてはデータが慎重に前処理されず、他の情報が除去されていない場合、AI がアーチファクトを誤認する可能性があるため、トレーニングデータセットは、高度に標準化された方法で生成されている必要がある<sup>16)</sup>と述べており、AI に三次元の情報を提供し、基準が明確となった細胞を学習させることは Band と Segment の分類だけでなく、Mono と Lymph など他の細胞の分類精度の向上にもつながると考えられる。

また判定基準に起因する問題点としては、Lymph は形態学的に Small type と Large type とが大きく形態学的特徴を異にするにもかかわらず、分類上は1つのカテゴリーにまとめられていることがあげられる。あまりにもかけ離れた2種の形態学的特徴を1つにまとめていることにAIが少数ながら誤分類判定をした可能性がある。ヒトが Lymph と Mono とを誤分類するのは、Small type と Mono の分類時ではなく Large type と Mono の分類時であり、これらは極めて類似した形態学的特徴を有している。Small type+Large type 対 Mono という大きなカテゴリー間で分類を行っていることでAIの判定基準が曖昧になり、上手く分類が機能していないことが推測される。このことから、Lymph と Mono との分類においては改めて Small type—Large type—Mono の3カテゴリーに再分類を行った上でCNN学習させることで、大リンパ球—単球間での特徴量抽出をさらに厳密化させ得るか否かを検討する必要がある。仮に本検討により分類精度が向上した場合には、AIによる血球解析ではAI分析用にサブカテゴリーを定義し直した上で、ヒトの目視分類用カテゴリーへと転換することで、精度向上が図れる可能性がある。

以上の検討を行うことで、従来技術やヒトによる目視観察に合致したデータ解析が行われ、精度向上が期待できると考えられる。しかしながら細胞形態解析においてAIは人では可視化できない周波数情報などを特徴量として計測し、因子としても使用している。そのため、AIによる血球分類精度向上においては、染色法自体の変更によるさらなる精度向上の可能性も否定できない。そのため、今後の臨床応用化研究においては、AIスクリーニング技術用の染色法についても再検討することでAIによる形態分析に最適化させ、異常細胞の鑑別にも応用することも必要であると考えられる。

## V. 結語

白血球形態分類においてCNNによるスクリーニング技術はAccuracyが高く有用であるものの、その臨床応用化においてはカットオフ値など境界領域細胞の判定保留条件などを検討する必要があり、さらなる精度向上や異型細胞などにも対応可能な技術として研究開発を進める必要がある。

**謝辞** 本研究は総務省戦略的研究開発推進事業 SCOPE ならびに JSPS 科研費(19K21737)の助成を受けております。

## 引用文献

- 1) 平成28年版情報通信白書. 総務省: 232-241, 2016. <https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h28/pdf/index.html> (アクセス日:2020-01-11)
- 2) 建本聡: 深層学習を利用したウメ「露苗」の画像による熟度分類. 農業情報研究, 28(3): 108-114, 2019.
- 3) 木田智士: 医療画像分野におけるAIの活用 医用画像情報学会雑誌. 35(2): 25-29, 2018.
- 4) Alejandro Rodríguez-Ruiz, Elizabeth Krupinski, et al: Detection of Breast Cancer with Mammography: Effect of an Artificial Intelligence Support System. Radiology, 00: 1-10, 2019.
- 5) Andrew A. Borkowski, Catherine P. Wilson: Comparing Artificial Intelligence Platforms for Histopathologic Cancer Diagnosis. Fed Pract, 36(10): 456-463, 2019.
- 6) 千葉直子: 症例から学ぶ 検査データと血液像の読み方 UniCel DxH800 (ベックマン・コールター). Medical Technology, 43(3): 242-248, 2015.
- 7) 渡邊真一郎, 坂場幸治, 他: 好中球系細胞標準化の経緯と血球形態標準化合同ワーキンググループについて. 医学検査, 64(6): 3-6, 2015.
- 8) 坂場孝治: 好中球桿状核球と分葉核球の識別基準および白血球目視分類の共用基準範囲. 臨床検査, 60(12): 1418-1427, 2016.
- 9) Kaiming He, Xiangyu Zhang, et al: Deep Residual Learning for Image Recognition. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 90: 2016.
- 10) 津田諒太: 人工知能学会資料 SIG-Challenge, 44(6): 30-31, 2016.
- 11) 齋藤勝彦: AIを用いた病理組織類似画像検索による診断支援ツールの開発の試み. 臨床病理, 66(補): 226. 2018.
- 12) Ramprasaath R. Selvaraju, Michael Cogswell: Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. Int J Comput Vis, doi:10.1007/s11263-019-01228-7, 2019.
- 13) Karen Simonyan, Andrew Zisserman: VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE SCALE IMAGE RECOGNITION. ICLR 2015 proceedings, 2015.
- 14) Qingshan She, Bo Hu: A hierarchical semi-supervised extreme learning machine method for EEG recognition. Medical & Biological Engineering & Computing, 57: 147-157, 2019.
- 15) 下前哲也: 特集最近の医用検査診断血液増自動分類. 精密工学会誌, 53(4): 530-533, 1987.
- 16) Viktor H. Koelzer, Korsuk Sirinukun Wattana.: Precision immunoprofiling by image analysis and artificial intelligence. Virchows Archiv, 474: 511-522, 2019.



## 【Original Article】

# Artificial intelligence analysis with convolutional neural network for peripheral blood smear screening in leukocyte classification

AMI SASAKI\*<sup>1</sup> MIKU ODA\*<sup>1</sup> HIROYUKI NOZAKA\*<sup>2</sup>  
MANABU NAKANO\*<sup>2</sup> MIYUKI FUJIOKA\*<sup>2</sup> HIDEKI TAKAMI\*<sup>2</sup>

(Received January 15, 2020; Accepted February 27, 2020)

**Abstract:** Medical AI is a next-generation medical technology that presents a diagnosis based on EBM regardless of the experience of clinical laboratory technologists. The technology is characterized by learning large amounts of patient data diagnosed by experts based on years of experience. In this study, we examined the clinical usefulness of screening technology with AI for peripheral leukocyte classification. The subjects were 57 healthy person's peripheral blood smears performed MG staining. The CNN model learned with teacher images of mature leukocyte cells that show typical morphology, and parameter tuning for optimization was performed. We performed additional learning and fine-tuning on this CNN model with various leukocyte images. The accuracy of 5 classifications showed 0.990 and 6 classifications showed 0.822 respectively in the additional learning with background-less images. Contrast, the accuracy of 5 classifications showed 0.992 and 6 classifications showed 0.879 respectively in the additional learning with images included background. It was cleared that the mature leukocyte cell morphology screening with CNN was highly accurate and useful. However, it is necessary to examine the cutoff value and the judgement pending condition for the boundary area cells in the clinical application.

**Keywords:** *Artificial intelligence, Convolutional neural network, Hematological morphology test, Leukocyte classification*

---

\*1 Hirosaki University School of Health Sciences, 66-1, Honcho, Hirosaki-shi, Aomori 036-8564, Japan

\*2 Hirosaki University Graduate School of Health Sciences, 66-1, Honcho, Hirosaki-shi, Aomori 036-8564, Japan  
E-mail: h16m2317@hirosaki-u.ac.jp