

畳み込みニューラルネットワークを用いた蚊幼生の識別

Sanchez Ortizy Alejandra¹ Nakano Miyake Mariko¹ Tuennermann Henrik² 寺本 陶冶²
庄野 逸^{2,a)}

概要：メキシコでは多くの病気が蚊によって媒介される。特に *Aedes* と呼ばれる蚊の幼生が黄熱やデング熱といった熱病の媒介原因になっている。この媒介原因を早期に除去するには蚊の幼生 (Larva) が発生している箇所を処理することが要求される。蚊のライフサイクルは非常に短時間で世代交代が行われるため、手軽なシステムで識別を行う必要がある。本研究では、スマートフォンを撮像デバイスとして用い、畳み込みニューラルネットワーク (CNN) を用いたシステムを提案する。本研究では比較的小さい 10^3 程度の画像データセットしか用意できなかったため、CNN の学習には転移学習を用いた。結果として、このような比較的小数の蚊幼生画像データセットの識別問題においても十分な精度の性能を示すことが確認できた。

Mosquito Larva Classification based on a Convolution Neural Network

SANCHEZ ORTIZY ALEJANDRA¹ NAKANO MIYAKE MARIKO¹ TUNNERMANN HENRIK² TERAMOTO TOYA² HAYARU SHOONO^{2,a)}

1. はじめに

デング熱 (Dengue Fever: DF) や、デング出血熱 (Dengue Hemorrhagic Fever: DHF)、チクングニア (CHIKV)、ジカ熱といった熱病は人間の健康に重大な問題を起こすことが知られている。これらの病気は、*Aedes Aegypti* (図 1) に代表される *Aedes* 属の蚊を介して拡散されることが知られている。表 1 に示すように近年のメキシコではこれらの熱病が大きな問題となっている。[1], [2], [3]

表 1 メキシコにおける患者数 2015-2017

	DF and DHF	CHIKV	Zika
Number of patients	14,138	13,163	11,851
Death	55	4	2

病気の発生をくい止めるための戦略の一つとして、蚊の幼生の繁殖を防ぐことが挙げられる。*Aedes* 蚊は、卵、幼

生、蛹、成虫という 4 つのライフサイクルを介して増殖することが知られており最初の 3 つの段階は水辺で行われる。*Aedes* 蚊のライフサイクルは短時間に行われることが知られており、このライフサイクルは 1~2 週間で 1 サイクルを構成する。したがって、*Aedes* 蚊の幼生の繁殖を防ぐには非常に短時間の期間に蚊の幼生所在を突き止め、燻蒸する必要がある。このため、比較的ポータブルな蚊の識別システムが必要となってくる。そこで、蚊のデータ収集のためにスマートフォンを利用し、スマートフォンから送られてきた画像を、解析することを考える。本研究ではこの画像解析に対して機械学習を適用し、判別することを考えている。

目的となる蚊の幼生第 8 節部に縞状の紋様があり、これ



図 1 *Aedes Aegypti* 蚊. 胴体部に白スポットがあるのが特徴

¹ National Polytechnic Institute,
Av. Santa Ana no. 1000, Col. San Francisco Culhuacan, Mexico D.F.,
04430, MEXICO

² 電気通信大学情報理工学研究所
〒182-8585 東京都調布市調布ヶ丘 1-5-1

a) shouno@uec.ac.jp

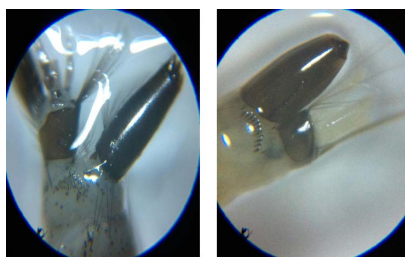


図2 蚊幼生と第8節.

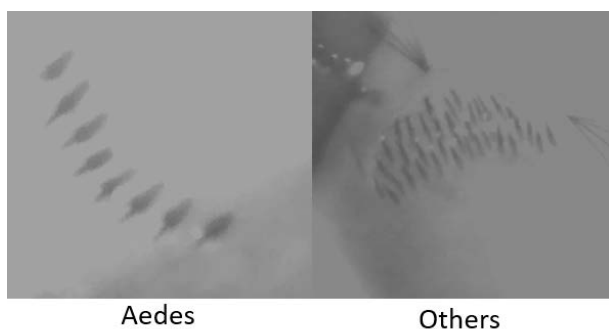


図3 前処理後の入力画像例

が識別のための鍵となる。著者らのグループは、この模様部分を画像として入力し、画像判別を行うことでシステムを構築してきた。本研究では、この判別システムに VGG16 と呼ばれる深層学習機械を用いて識別システムを構築することを試みている。

2. データ諸元

データの収集にはスマートフォンを用いて行った。データの画像は、専門家の指導のもと得られた 1,118 枚の蚊の幼生第8節画像を収集している。これらの画像には“Aedes”および“その他”のラベルがつけられている。ラベルの判断は Mexico Hidalgo 州の Public Health Laboratory によってなされたものを用いている。図2に蚊幼生とその第8節の典型的な画像例を示す。o 得られた画像は、スポットの映り込みなどがあるため、画像認識に悪影響を与えるので Total Variation 法を用いたノイズ除去を行った。この前処理を施した画像例を図3に示す

3. VGG16 モデルによる識別システムの構築

本研究では、VGG16 と呼ばれる畳み込みニューラルネットワーク (CNN) を用いて画像判別を行う。VGG16 は図5に示すようなニューラルネットワークモデルでリファレンスとして使われるモデルである [4]。一般に深層学習においては大量のデータを学習に必要とする。しかしながら本研究で適用するようなデータ数が十分でない場合は、なんらかの工夫が必要となる。本研究では、転移学習を用いて学習を補助することを考えた。転移学習は異なるドメインの知識をターゲットとなるドメインの知識としてもちいる

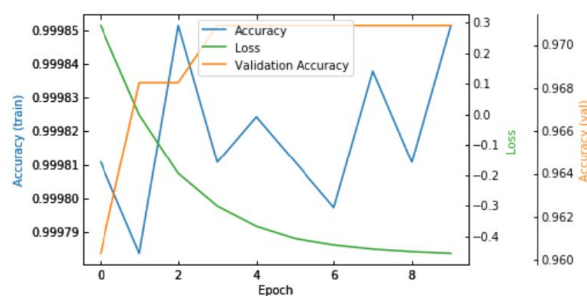


図4 convergence of accuracy performance in epochs

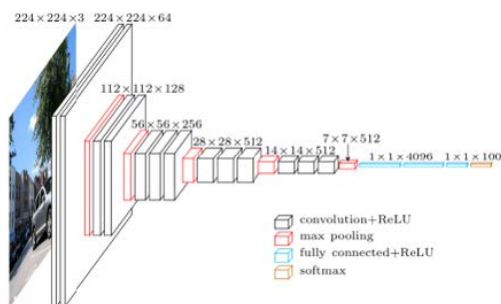


図5 The VGG16 architecture.

機械学習のテクニックである。ここでは、Suzuki らの手法に基づき [5]、一度自然画像を学習させて成熟させたネットワークを蚊の幼生判別に用いるネットワークの初期状態として適用した。

学習には、正例、負例あわせて 750 枚の画像を用い、識別性能を測るために 378 枚の画像を用いた。

4. 結果とまとめ

図4に識別能力の結果を示す。図4の横軸は学習エポック数を表し、縦軸はそれぞれ、精度 (Accuracy) と学習検証用精度 (Validation-Accuracy) および、学習におけるロス関数の推移を表している。転移学習をもちいると、非常に短時間に収束していることが見て取れる。結果として、我々の用いた方法を用いると 97% 近い精度がテスト画像に対して得られることが判明した。これは従来の k-means 方を用いた識別精度 (60%) や、SVM を用いた識別精度 (79%) よりも向上しており、十分実用に耐えうるシステムを構築できていることが示唆される。

今後の課題としては、どのような特徴解析が識別に効いているのかなどの解析を行っていく予定である。

謝辞

本研究の一部は科学研究費補助金 基盤研究 (C) 16K00328 および新学術領域研究 16H01542 から補助を受けている。

参考文献

- [1] “Panorama Epidemiologico de Fiebre por Dengue y Fiebre Hemorragica por Dengue”, Secretaria de Salud de Mexico,

- Direccion General de Epidemiologia, July 2016.
- [2] T. Uribarren-Berrueta, “Dengue y otras infecciones no hemorragicas: Fiebre Chikungunya, Zika Fiebre del nilo occidental y otros Abrovirus”, Departamento de Microbiologia y Parasitologia, Facultad de Medicina, UNAM. 2016.
 - [3] S. B. Halstead, “Pathogenesis of dengue: challenges to molecular biology”, Science, vol. 239, pp. 476-481, 1988.
 - [4] K. Simonyan, A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition”, Visual Geometry Group, Department of Engineering Science, University of Oxford, In ICLR, 2015.
 - [5] A. Suzuki, S. Suzuki, S. Kido, H. Shouno, A 2-staged Transfer Learning Method with Deep Convolutional Neural Network for Diffuse Lung Disease Analysis, IFMIA2017, P1-17, 2017