

研究論文

機械学習による食品から分離された微生物のコロニーの 検出と菌種の推定

長谷川 撰^{*1}、日渡 美世^{*1}、安田(吉野) 庄子^{*2}

Detection of Microbial Colonies Isolated from Foods and Estimation of Species by Machine Learning

Osamu HASEGAWA^{*1}, Miyo HIWATASHI^{*1} and Shoko YOSHINO-YASUDA^{*2}Food Research Center ^{*1 *2}

食品から分離した 5 種類の微生物(細菌及び酵母)について、標準寒天培地プレート上のコロニーを撮影し、その画像から種類を判別することを試みた。各菌を培養したプレートを回転させながら照明の状態とカメラの露出を変化させ、1 枚のプレートから撮影条件の異なる画像を効率的に生成した。この画像を教師データとして Single Shot Multibox Detector による機械学習を行ったところ、検証用画像の撮影条件が適切であった場合、各菌の平均適合率は 0.95 以上となった。

1. はじめに

加工食品の微生物管理は、製品の生菌数の検査を中心に行われている。また、製造機器、原材料、仕掛品などについても生菌数の検査を行うことがある。検査によっては特定の微生物をターゲットに行うが、日常的な菌数の測定において菌の種類まで特定することはコストや時間の面から困難である。しかしながら、汚染源の特定や変敗のリスク評価を効果的に行うには、微生物の菌叢まで把握することが望まれる。

一方、菌数測定の際に培養して得られるコロニーは、菌の種類に応じて形状や色、光沢などの情報を有している。これらの情報は数値化や言語化が困難であることから、熟練者が経験をもとに菌叢の判断の参考に活用する程度にとどまっている。通常生菌数の検査で得られるコロニーの外観から菌叢を予測できれば、遺伝子解析や質量分析などによる菌種の特定と比べてコストと時間を節減できると期待される。そこで本研究では、機械学習によりこれらの情報から微生物の種類を推定する技術を開発することとした。

2. 実験方法

2.1 食品からの微生物の分離と同定

愛知県内で購入した惣菜、和菓子及び洋菓子を微生物の分離源として用いた。試料中の微生物の分離は、標準寒天培地を用いた混釈法で行い、35℃で約 48 時間培養後に検出したコロニーを新しい同プレートに継代し、冷

蔵保管して用時使用した。

取得した微生物の一部について、MALDI-TOF MS による微生物同定システム (AXIMA 微生物同定システム、(株)島津製作所)及び 16S rDNA 配列による同定を行った。MALDI-TOF MS による同定は、エタノール-ギ酸抽出法により試料調製を行い、解析はソフトウェア SARAMIS Premium (データベースは Version 4.13) (バイオメリュウ、(株)島津製作所)を使用した。16S rDNA 配列による同定は、第十七改正日本薬局方の「遺伝子解析による微生物の迅速同定法」に準じて行った。

2.2 機械学習に用いる写真撮影用プレートの作製

乳酸菌は MRS 培地、乳酸菌以外の細菌は NBRC802 培地、酵母は YM 培地で 1 晩培養したものを適当な菌数となるように希釈し、直径 9cm のシャーレに標準寒天培地を固化させたプレートに塗抹して 35℃で約 48 時間培養した。これを機械学習のためのコロニー画像の撮影に用いた。

2.3 コロニーの写真撮影

プレート上のコロニーの写真撮影は、Raspberry Pi 3 Model B+ にカメラモジュール Raspberry Pi Camera V2 を接続して行った。撮影時の光源には昼白色の LED を使用した。カメラモジュールはプレートの真上となる位置に設置した。光源はプレートを挟むように 2 か所設置(照明 1)するとともに、それよりやや高い位置 1 か所に 3 列並べて設置(照明 2)した。照明の点灯パターンは照明 1 の 2 か所を点灯させたものを L0、照明 2

のうちの1列、2列、3列を点灯させたものをそれぞれL1、L2、L3とした。さらに、カメラの露出補正を最も暗いE0から最も明るいE4まで5つのパターンを用意することで、1枚のプレートから照明と露出補正の20通りの組み合わせを用意した。プレートを少しずつ回転させながら、照明と露出補正の各組み合わせについて10枚ずつ、合計で200枚の写真を撮影した。

2.4 機械学習

コロニーの検出は Single Shot Multibox Detector¹⁾ (SSD)で行った。SSDはPyTorchを用いた実装^{2),3)}を一部改変したものを使用した。今回の学習ではバッチサイズを12とし、10エポックの学習を行った。

学習後、教師用の画像データとは別に撮影した画像データを用いて推論を行った。

3. 実験結果及び考察

3.1 食品からの微生物の分離と同定

2.1の方法で食品から微生物を分離し、同定した結果を表1に示す。MALDI-TOF MSによる同定ではSARAMIS Premiumのクラスター解析機能を用いて分類し、同一クラスターに属する菌株を同一種とした。一部の株については16S rDNA配列を解析した。その結果、113株について表1のように菌種を推定した。

機械学習に用いる菌株の選定には、コロニーの検出頻度が高いもの、複数の食品から検出されたものを優先し、

4種類の細菌 *Microbacterium arborescens*、*Staphylococcus warneri*、*Leuconostoc citreum*、及び *Kocuria kristinae* を選定した。また、生菌数の検査では細菌とともに酵母が検出されることがあるため、酵母で検出頻度の高かった *Candida krusei* を追加し、合計5種類の菌株を機械学習用の菌株として選定した。

3.2 コロニーの写真撮影と教師データの作成

選定した菌株を用いて2.2の方法でプレートを作成した。これらを写真撮影し、図1に示した緑色の方眼1マス単位で縦横それぞれ8分割し、300×300ピクセルの画像を切り出した。これらのうち、コロニーの存在しない画像を除いて教師用のデータセットを作成した。切り出される画像の枚数はコロニーの生育状況により異なっていたが、写真1枚から約25枚の画像を得ることができた。

*S. warneri*の写真から切り出した画像の例を図2に示す。(b)は(a)と比べて照明に近い場所にあったコロニーのため、(a)よりも明るく、照明の映り込みが顕著であった。また、カメラとコロニーの位置関係により、シャーレ底面に映り込むコロニーの影の位置も(a)と(b)では異なっていた。機械学習では、学習に変化の乏しい画像ばかりを用いると汎化性に劣るが、今回の撮影条件では1枚のプレートからも変化に富んだ画像が得られることが確認できた。

3.3 SSDによる学習と推論の実施

表1 食品から分離した微生物

種名	分離された株数			
	惣菜	和菓子	洋菓子	合計
<i>Acinetobacter johnsonii</i>	1			1
<i>Bacillus</i> sp.	2			2
<i>Candida krusei</i>			8	8
<i>Candida pelliculosa</i>	2			2
<i>Curtobacterium</i> sp.	1			1
<i>Kocuria kristinae</i>	1	3	1	5
<i>Lactobacillus</i> sp.			5	5
<i>Lactococcus lactis</i>	1			1
<i>Leuconostoc citreum</i>	13			13
<i>Microbacterium arborescens</i>	22			22
<i>Sphingomonas</i> sp.	2			2
<i>Staphylococcus</i> sp.		4		4
<i>Staphylococcus epidermidis</i>		8		8
<i>Staphylococcus sciuri</i>	1			1
<i>Staphylococcus warneri</i>		21	16	37
<i>Stenotrophomonas maltophilia</i>	1			1
合計	47	36	30	113

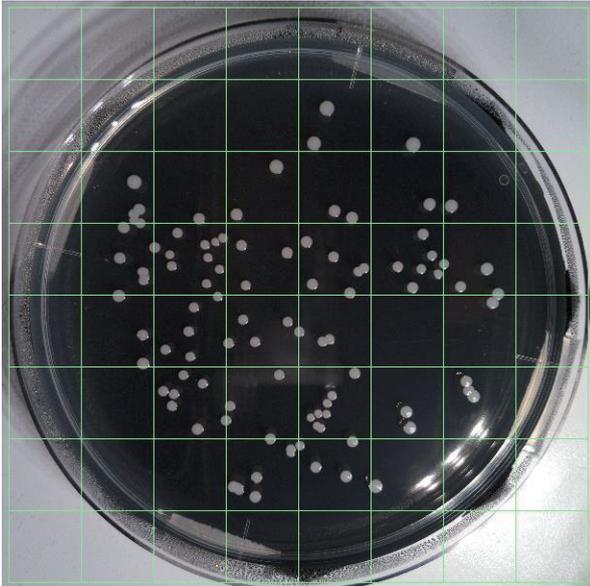


図1 プレートを撮影した写真からの機械学習に用いる画像の切り出し

5 菌株の画像 109,285 枚を教師データとして用いて SSD による学習を行った。学習後、教師データとは別に撮影した 11,992 枚の画像(図3)を用いて推論を行った。菌株ごとにそれぞれの照明条件、露出補正条件での平均適合率を求めた結果を表2に示す。

S. warneri 及び *C. krusei* は L0~L3 の照明条件、E0~E4 の露出補正条件全てにおいて平均適合率が 0.94

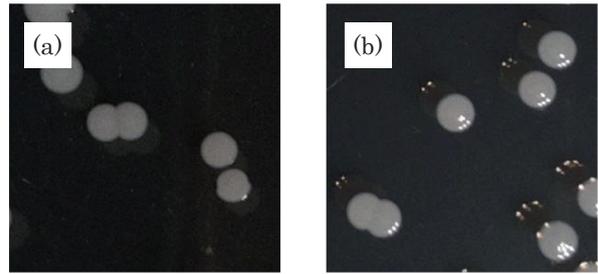


図2 *S. warneri* の写真から切り出した画像の例

以上であった。一方、*M. arborescens* や *L. citreum* の平均適合率は L0~L3 及び E0~E4 の条件によって大きく異なり、画像が暗くなる照明 L1 や L2、照明 L0 で露出補正 E0 の場合に平均適合率が低かった。これらの菌株のコロニーは他と比べて小さいことから、小さいコロニーでは画像が暗い場合にうまく検出できなくなる可能性が示唆された。また、*M. arborescens* は他の菌株と比べて透明なコロニーであったことから、検出が困難であったと推察された。

全ての菌株に対する平均適合率は、照明 L0 で露出補正 E2 の条件が最も高かった。この条件では *M. arborescens* や *L. citreum* の平均適合率も高かった。今回、光源や露出補正の条件を変化させて撮影した画像を教師データとして用いたが、推論を行う場合にはさまざまな条件で撮影する必要はなく、高い平均適合率となっ



(1) *M. arborescens* (2) *S. warneri* (3) *K. kristinae* (4) *L. citreum* (5) *C. krusei*

図3 推論に用いた画像の例

表2 推論結果の平均適合率

(1) *M. arborescens*

	E0	E1	E2	E3	E4
L0	0.80	0.93	0.97	0.96	0.89
L1	0.75	0.72	0.76	0.62	0.64
L2	0.85	0.84	0.75	0.74	0.91
L3	0.85	0.94	0.92	0.84	0.84

(2) *S. warneri*

	E0	E1	E2	E3	E4
L0	0.97	0.97	0.97	0.97	0.96
L1	0.94	0.95	0.94	0.94	0.95
L2	0.96	0.96	0.95	0.95	0.98
L3	0.96	0.97	0.97	0.96	0.96

(3) *K. kristinae*

	E0	E1	E2	E3	E4
L0	0.86	0.96	0.95	0.96	0.92
L1	0.93	0.94	0.91	0.92	0.94
L2	0.85	0.88	0.91	0.89	0.93
L3	0.96	0.95	0.93	0.96	0.96

(4) *L. citreum*

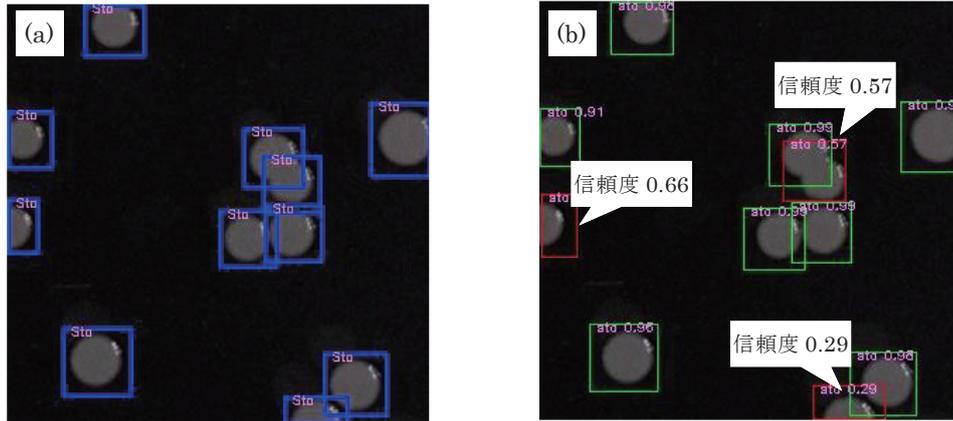
	E0	E1	E2	E3	E4
L0	0.66	0.97	0.97	0.95	0.88
L1	0.86	0.83	0.86	0.73	0.83
L2	0.88	0.90	0.85	0.77	0.84
L3	0.91	0.95	0.95	0.92	0.89

(5) *C. krusei*

	E0	E1	E2	E3	E4
L0	0.95	0.96	0.96	0.96	0.97
L1	0.96	0.95	0.97	0.97	0.95
L2	0.95	0.95	0.97	0.95	0.97
L3	0.97	0.97	0.95	0.97	0.96

(6) 全ての菌株の平均

	E0	E1	E2	E3	E4
L0	0.85	0.96	0.97	0.96	0.92
L1	0.89	0.88	0.89	0.83	0.86
L2	0.90	0.91	0.88	0.86	0.92
L3	0.93	0.95	0.94	0.93	0.92



(a) 青色の枠は正解として与えたデータ。
 (b) 緑色の枠は信頼度 0.9 以上、赤色は信頼度 0.9 未満として検出されたコロニー。

図 4 *S. warneri* の画像における推論の例

た条件で撮影すればよいと考えられた。

3.4 推論における信頼度の閾値の設定

推論により検出されたコロニーの例を図 4 に示す。推論では、検出されたコロニーの分類がどの程度信頼できるかが 0 から 1 の範囲の値で示され、この値が 1 に近いほど信頼できると判定している。単独で存在しているコロニーの信頼度が 0.9 を超えているのに対し、画像の端で一部が切り取られていたり、他のコロニーと重なっていたりするコロニーの信頼度は 0.7 を下回っていた。検出の漏れ(偽陰性)を減らすためには信頼度の閾値を小さくする必要があるが、誤検出(偽陽性)の可能性も高くなると推察されることから、信頼度の閾値を変えることで偽陽性や偽陰性がどのように変化するかを調べた。図 5 に照明 L0、露出補正 E2 でのコロニーの検出結果を示す。信頼度の閾値が 0.9 の場合、偽陰性となるコロニーは約 1/3 を占めていたが、閾値を小さくすることで急激に減少し、閾値が 0.5 の場合には 1/10 が偽陰性となった。一方、擬陽性は閾値を小さくしてもあまり増加しなかった。このことから、今回の学習結果では、信頼度の閾値を小さくすることが有効であると考えられた。

4. 結び

本研究の結果は、以下のとおりである。

- (1) 学習時には光源やカメラの露出補正の条件を変化させて撮影した教師データを用い、推論時には最も適切な撮影条件のデータを用いた場合、全体の平均適合率は 0.97 となった。
- (2) 推論において信頼度の閾値を 0.9 から 0.5 とすることで偽陰性となるコロニー数は大きく減少する一方、偽陽性はあまり増加しなかった。

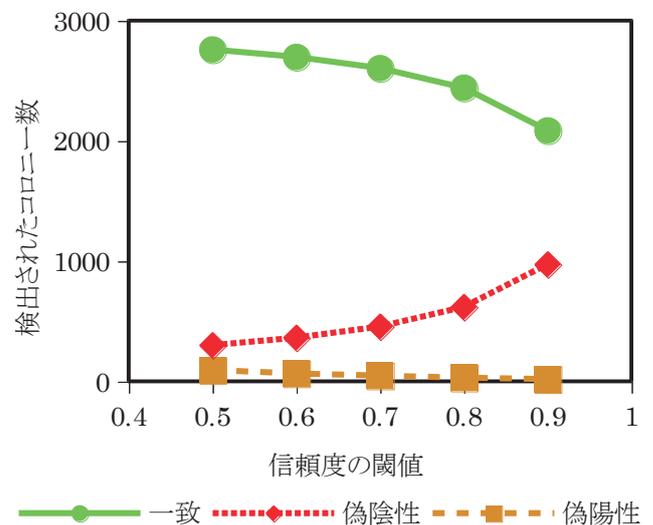


図 5 照明 L0、露出補正 E2 の撮影条件におけるコロニーの検出結果

付記

本研究は、公益財団法人人工知能研究振興財団の助成を受けて実施した。

文献

- 1) W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu and A. C. Berg: *European Conference on Computer Vision*, 21(2016)
- 2) SSD: Single Shot MultiBox Detector, in PyTorch, <https://github.com/amdegroot/ssd.pytorch>, (2020/10/26)
- 3) 宮本圭一郎, 大川洋平, 毛利拓也: *PyTorch ニューラルネットワーク実装ハンドブック*, 181(2018), 秀和システム