

研究論文

機械学習による微生物のコロニーの検出と菌種の推定における 食品成分の影響

長谷川 撰*1、日渡美世*1、安田(吉野) 庄子*2

Effect of Food Composition on the Detection of Microbial Colonies and Estimation of Species by Machine Learning

Osamu HASEGAWA*1, Miyo HIWATASH*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。食品を含まない教師データと含んだ教師データで学習を行い、推論に与える影響を評価した結果、食品を含んだ教師データで学習した重みを用いることで、食品を含んだ画像の推論結果がより正しくなった。偽陽性や偽陰性の生じる割合は菌株によって大きく異なる結果が得られたが、食品ごとの偽陽性や偽陰性の生じる割合は、菌株の違いと比べて差は小さかった。

1. はじめに

前報¹⁾では、DNA 解析や MALDI-TOFMS などの時間やコストのかかる方法を用いなくても、機械学習を応用することで微生物のコロニーの画像から、5 種類の菌種の推定が可能であることを示した。この画像は食品を含まない菌液を塗抹し、培養したものから得たものであるが、実際の食品の検査では、菌液中に食品由来の固形物や栄養素が含まれており、これが機械学習による菌種の推定に影響を与える可能性がある。そこで食品成分の存在が推論結果に与える影響を評価し、食品成分が存在する場合でもよりよい推論結果が得られる方法を検討した。

2. 実験方法

2.1 試料

食品は当センターで調製したスポンジ、パウンドケーキ及び市販の小倉あん、スパゲティサラダを使用した。

微生物は昨年度食品より分離した *Microbacterium arborescens*、*Staphylococcus warneri*、*Leuconostoc citreum* 及び *Kocuria kristinae* の 4 種類の細菌と *Candida krusei* の 1 種類の酵母を使用した。

2.2 希釈菌液の調製

微生物は前報¹⁾に従い液体培地で培養した後、菌数がおおよそ 10^4 cfu/mL から 10^5 cfu/mL となるように希釈し、希釈菌液とした。

2.3 塗抹用菌液の調製

食品をオートクレーブで滅菌した後、10 倍量となるように滅菌水と混合してストマッカー処理し、食品 10 倍希釈液とした。これをさらに滅菌水で 10 倍に希釈し、食品 100 倍希釈液とした。これらを 2mL とり、希釈菌液を 0.02mL 添加して混合したものを、標準寒天培地を固化させたプレートに 0.1mL 塗抹し、35°C で約 48 時間培養した。また、食品を含まない菌液についても同様に培養した。これらを機械学習のためのコロニー画像の撮影に用いた。

2.4 コロニーの写真撮影

プレート上のコロニーは、前報¹⁾と同様の撮影装置を用いて写真撮影を行った。光源はプレートを挟むように撮影装置の 2 か所に設置した昼白色の LED を使用した。撮影時のカメラの露出補正は、教師用のデータについては 5 段階に変化させて撮影し、推論用のデータについては 5 段階のうちの中間にあたる条件で撮影した。

2.5 機械学習

コロニーの検出は Single Shot Multibox Detector²⁾ (SSD)で行った。SSD は PyTorch を用いた実装^{3),4)}を一部改変したものを使用した。

教師データは、食品を含まない菌液を培養して得られた画像のデータセットと、食品 10 倍希釈液または 100 倍希釈液を含んだ菌液及び食品を含まない菌液を培養して得られた画像のデータセットの 2 つのパターンを用意した。これらのデータセットを用いて機械学習を行った後、推論を実施した。

*1 食品工業技術センター 分析加工技術室 *2 食品工業技術センター 保蔵包装技術室

3. 実験結果及び考察

3.1 食品の希釈液及び培養後のプレートの外観

食品 10 倍希釈液の外観を **図 1** に示す。小倉あんは固形分が速やかに沈殿したが、他の食品については全体が白く濁った状態であった。



図 1 食品 10 倍希釈液の外観



図 2 食品 10 倍希釈液を含んだ菌液を塗抹し、培養したプレートの外観

食品 10 倍希釈液を含んだ菌液を塗抹し、培養したプレートの外観を **図 2** に示す。食品を含んだ菌液を塗抹した場合、プレート表面には食品由来の固形分や油脂が映り込んでいることが確認できた。

3.2 食品の有無が学習と推論に与える影響

教師データ作成時の食品の有無が推論に与える影響を評価するため、食品を含んだ教師データと含まない教師データを用いて学習を行い、これらの重みを用いて推論を行った結果を **表 1** に示す。教師データとして食品を含まない画像で学習した場合、スポンジの 10 倍希釈液を含んだ *Candida* の平均適合率とパウンドケーキの 10 倍希釈液を含んだ *Microbacterium*、*Staphylococcus*、*Leuconostoc*、*Candida* の平均適合率は 0.9 未満であったが、食品を含んだ画像で学習を行った場合はパウンドケーキの 10 倍希釈液を含んだ *Leuconostoc* については

表 1 教師データにおける食品の有無が推論結果に与える影響

		食品なしの教師データでの学習			食品ありの教師データでの学習		
		10倍希釈	100倍希釈	食品なし	10倍希釈	100倍希釈	食品なし
スポンジ	<i>Microbacterium</i>	0.924	0.977		0.975	0.977	
	<i>Staphylococcus</i>	0.943	0.987		0.987	0.991	
	<i>Kocuria</i>	0.974	0.989		0.980	0.989	
	<i>Leuconostoc</i>	0.993	0.998		0.993	0.996	
	<i>Candida</i>	0.822	0.973		0.970	0.969	
パウンドケーキ	<i>Microbacterium</i>	0.818	0.935		0.954	0.950	
	<i>Staphylococcus</i>	0.889	0.985		0.991	0.989	
	<i>Kocuria</i>	0.970	0.986		0.978	0.989	
	<i>Leuconostoc</i>	0.874	0.984		0.894	0.984	
	<i>Candida</i>	0.875	0.957		0.978	0.974	
小倉あん	<i>Microbacterium</i>	0.900	0.998		0.969	0.996	
	<i>Staphylococcus</i>	0.968	0.984		0.988	0.989	
	<i>Kocuria</i>	0.944	0.986		0.977	0.990	
	<i>Leuconostoc</i>	0.980	0.994		0.980	0.997	
	<i>Candida</i>	0.963	0.956		0.979	0.981	
スパゲティサラダ	<i>Microbacterium</i>	0.904	0.943		0.974	0.989	
	<i>Staphylococcus</i>	0.996	0.991		0.997	0.993	
	<i>Kocuria</i>	0.984	0.998		0.991	0.998	
	<i>Leuconostoc</i>	0.940	0.985		0.975	0.991	
	<i>Candida</i>	0.981	0.969		0.984	0.974	
食品なし	<i>Microbacterium</i>			0.915			0.902
	<i>Staphylococcus</i>			0.984			0.985
	<i>Kocuria</i>			0.992			0.994
	<i>Leuconostoc</i>			0.970			0.984
	<i>Candida</i>			0.958			0.973

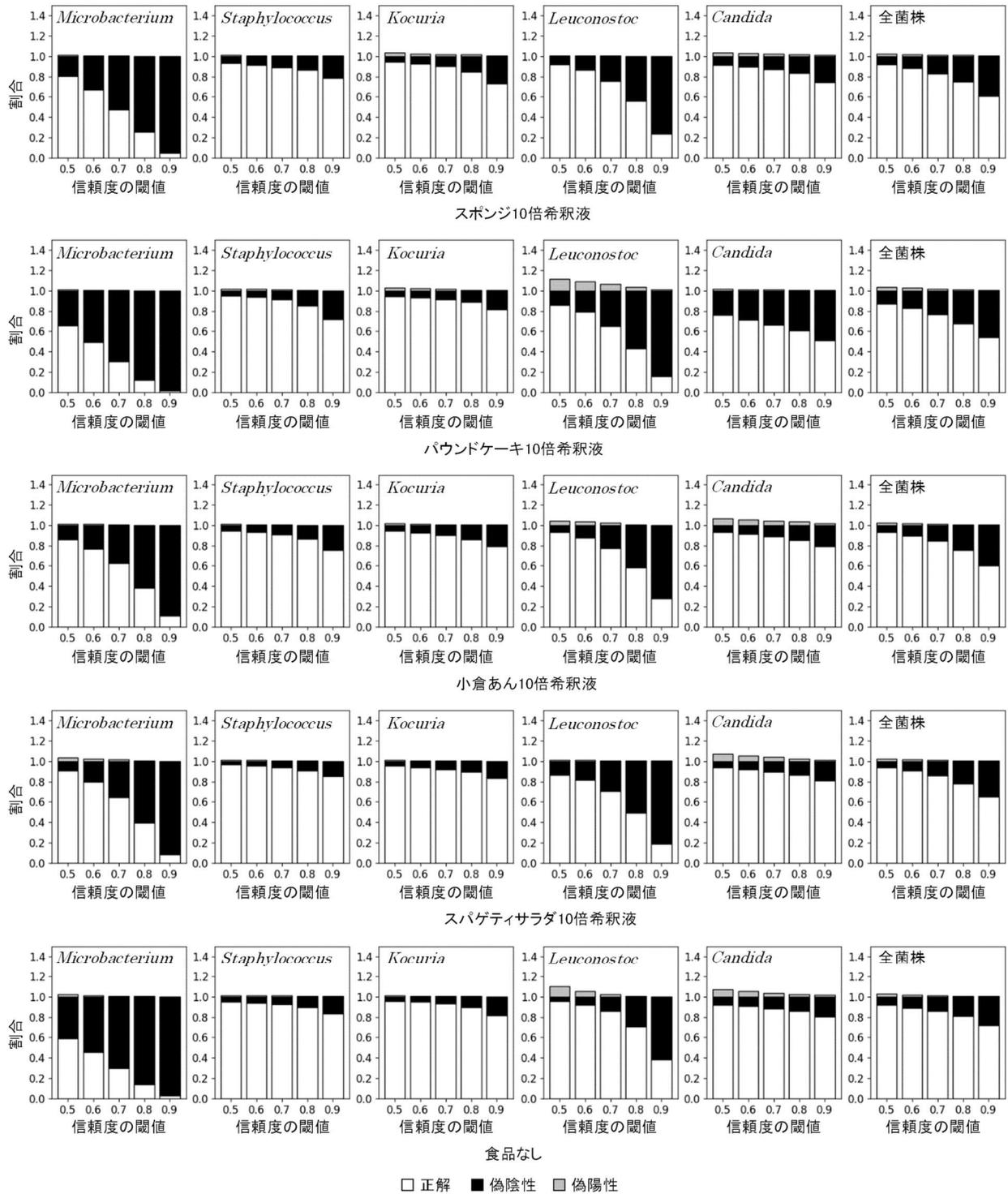


図3 信頼度の閾値が菌株ごとの推論結果に与える影響

0.874 から 0.894 に上昇し、それ以外のものについては 0.9 以上に上昇した。

一方、スポンジの 100 倍希釈液を含んだ *Leuconostoc*、*Candida*、小倉あんの 100 倍希釈液を含んだ *Microbacterium*、食品を含まない *Microbacterium* については平均適合率が低下していた。特に、食品を含まない *Microbacterium* は他のものと比べて低下の幅がやや

大きかった。しかしながら、食品を含まない教師データを用いて学習を行った場合に平均適合率が低かったものが、食品を含んだ教師データで学習を行うことで高くなる傾向にあり、平均適合率が高かったものはあまり低下していなかったことから、教師データとして食品を含んだものを用いて学習した重みで推論を行うことは、食品を含んだ試料に対してより良い推論結果を得る手段とし

て適切であると考えられた。

3.3 信頼度の閾値の設定

機械学習による推論では、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。しかしながら、推論結果には存在する菌株を検出できない偽陰性や、正解とは異なるコロニーや物体を検出してしまう偽陽性が含まれ、信頼度の閾値をいくつに設定するかによってその割合は変わってくる。そこで、閾値を 0.5 からとしたときに、偽陰性や偽陽性がどのように変化するか確認することとした。

図 3 に食品を含んだ教師データを用いて学習を行った重みで食品の 10 倍希釈液を含んだものと食品なしのものについて推論を行い、正解と偽陰性の合計を 1 とした場合の正解、偽陰性、偽陽性の割合を示す。食品の種類によらず、*Microbacterium* や *Leuconostoc* は偽陰性が多くなる傾向にあった。閾値を小さくすることで偽陰性は減少するが、閾値を 0.5 とした場合の偽陰性の割合は、*Staphylococcus* では 0.07 以下となったのに対し、*Microbacterium* ではスパゲティサラダを含んだ場合には 0.09、パウンドケーキを含んだ場合 0.34 であり、食品の種類によって偽陰性の割合が大きく異なっていた。

閾値を小さくすることで偽陽性は増加するが、増加の傾向は菌株や食品により異なっていた。*Leuconostoc* ではパウンドケーキを含む場合や食品なしの場合に偽陽性の増加の割合が大きくなる傾向があったのに対し、*Candida* では、小倉あんやスパゲティサラダを含む場合や食品なしの場合に大きくなる傾向があった。一方、偽陰性の割合が大きかった *Microbacterium* については、偽陽性の割合はあまり増加しなかった。

全菌株についての偽陰性や偽陽性の割合を見た場合、

偽陰性となる割合は食品によって多少の差があるものの、同じ食品で菌株間について比べた場合のような大きな差はなかったことから、菌株の違いに比べて食品の違いは偽陰性や偽陽性の割合に与える影響は少ないと考えられた。このような結果から、信頼度の閾値は検出しようとする菌の種類によって適切な閾値は異なるものの、おおむね 0.5 から 0.7 程度に設定するのがよいと思われた。

4. 結び

本研究の結果は、以下のとおりである。

- (1) 食品を含んだ教師データで学習した重みを用いることで、食品を含まない教師データで学習した重みを用いた場合と比べて推論における平均適合率を高くすることができた。
- (2) 偽陽性や偽陰性の生じる割合は菌株によって大きく異なった。一方、食品ごとの偽陽性や偽陰性の生じる割合は、菌株の違いと比べてあまり差がなかった。

文献

- 1) 長谷川撰, 日渡美世, 安田(吉野)庄子: あいち産業科学技術総合センター研究報告, **10**, 74(2021)
- 2) W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu and A. C. Berg: European Conference on Computer Vision, **21**(2016)
- 3) SSD: Single Shot MultiBox Detector, in PyTorch, <https://github.com/amdegroot/ssd.pytorch>, (2020/10/26)
- 4) 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