

대면적 산술현미경과 머신러닝을 이용한 박테리아 검출

† 안 장 환

Utilizing Fourier Ptychographic Microscopy and Convolutional Neural Network for Detecting Bacteria

JangWhan Ahn

Key Words : *Bacteria(박테리아), Fourier ptychographic microscopy(대면적 산술 현미경), Convolutional neural network, Deep learning, Image classification(이미지 분류)*

ABSTRACT

Establishing early warning systems for virus or bacteria infection is crucial in biodefense. This study focuses on developing the rapid and accurate detection and classification system with five different bacteria. Using Fourier ptychographic microscopy(FPM), which iteratively stitches together a number of low-resolution intensity images in Fourier space to produce a wide-field of view and high-resolution image, two different types of image data, which are absorption and phase, were produced per each bacterium which later be utilized for training neural network. Three convolutional neural network models were developed: the first model was trained only with absorption image data while another was trained with phase image data and the last one as the multi model where two different informative images were put together as an input. These models were experimented upon various standards such as optimizer, regularization, epoch and hyperparameters. As a result, two specific models were selected for each three cases(absorption, phase and multi). Finally, the best model selected was the one trained with absorption and phase image simultaneously which could classify given images with 91.4% test accuracy. Overall, through utilizing FPM and machine learning, classification of bacteria in rapid speed and high accuracy was achieved.

1. 서 론

최근 COVID-19 바이러스를 비롯하여 여러 바이러스나 세균에 대해 방어체계를 갖추는 Bio-Defense에 대한 중요성이 증가하게 되었다. 이에 따라 실내에서 많은 인원이 활동할 시에 생화학적 이슈에 대해 실시간으로 조기 대응 체계를 갖추는 것이 필요한 상황이다.

이를 위해서는 공기 중의 박테리아를 다량으로 포집하고, 검출한 다량의 샘플 박테리아에 대해 빠르고 정확한 검출 및 분류확인 시스템을 구축하는 것이 중요하다. 따라서 이 시스템을 구축하기 위해서는 첫 번째 다량의 박테리아를 검출하고, 두 번째 빠르고 정확한 검출 및 분류를 한다는 것이 보장되어야 한다.

광학계에서 이미징 플랫폼의 처리율은 광학장치의 Space Bandwidth Product(SBP)¹에 의해 제한된다. 이 때 SBP는 광학계 장치요소들의 geometric aberration에 의해 제한되어 임의로 증가시킬 수 없다¹. 이는 제한된 SBP안에서 이미지의 해상도(resolution, number of pixels)와 field of view(FOV) 두 가지 사이의 타협을 가져온다².

그러나 본 연구에서 개발하고자 하는 시스템에서는 고화질 이미지를 통해 다량의 데이터셋을 마련해야하기 때문에 큰 SBP를 가지는 이미지가 바람직하다.

현재 사용하는 이미징 기술로는 명시야 현미경과 lensless holography로 나뉜다. 명시야 현미경에는 기계적인 방법론 자체에 내재하는 한계 SBP가 존재한다. 그리고 기계적인 방법으로 SBP를 향상시키는 것은 정교한 제어, optical alignment, motion tracking이 필요하기 때문에 차선택이 될 수밖에 없다². Lensless holography중 digital in-line holography³⁻⁴는 인접한 샘플들에 대해 잘 작동하지 않는다. Contact-imaging microscopy⁵⁻⁶는 샘플이 센서에 근접해야 잘 작동한다는 조건이 필요하다. 따라서 두 가지 모두 큰 SBP를 얻는데 효과적인 방법이 아니다.

박테리아를 검출하는 방법 중 culture-based methods는 수 일이 소요된다. Polymerase Chain Reaction(PCR)기반 방법은 종 간의 구분이 가능할 정도(species-level specificity)의 정확성을 제공하지만 수 시간이 소요되고, 전문적인 장비와 인력을 필요로 한다. Photoluminescence and

surface-enhanced Raman scattering methods는 수 분 밖에 소요가 되지 않지만, 외부인자들과 구별하는 labeling 과정이 필요하다⁷. 대부분의 검출 방법은 아주 작은 이미지 사이즈가 필요할 정도로 민감하고 이는 실시간 검출을 위한 방법으로 이용하기에 적절하지 않다.

위 한계점을 극복하고 목표하는 시스템을 구축하기 위해 필요한 두 가지 조건, 다량의 박테리아 검출과 빠르고 정확한 검출 및 분류를 만족시키기 위해 Fourier ptychographic microscopy(FPM)을 사용해 높은 SBP와 FOV를 얻고, 머신러닝 기법을 활용해 다량의 대상에 대해 정확하고 빠르게 박테리아를 검출해낼 수 있다.

FPM 장비로는 박테리아 세포들의 광흡수(absorption)와 위상(phase)이미지를 얻을 수 있다. 머신러닝 기법을 적용한 부분은 딥러닝에서 이미지 처리 기술 중 눈에 띄는 발전을 보인 convolutional neural network(CNN)를 이용해 모델을 설계했다. CNN모델은 FPM장비로 얻은 광흡수 이미지만을 이용하여 학습시키는 모델, 위상 이미지만을 이용한 모델 그리고 두 가지 정보를 모두 활용하는 모델을 설계하였다. 그리고 optimizer, regularization, hyperparameter에 따라 달라지는 성능을 비교해 가장 빠르고 정확하게 이미지 분류를 수행하는 모델을 개발하였다.

2. Results

데이터셋을 마련하기 위해 포집된 박테리아 5종에 대해 FPM 현미경 장비를 이용해 광흡수와 위상 이미지를 얻었다. 광흡수와 위상 이미지 데이터에 대해 개별 박테리아의 이미지를 얻을 수 있도록 threshold값을 사용한 segmentation을 진행한 뒤, 이를 학습 데이터셋(trainset)으로 활용해 Single_absorp_BactConvNet,

Single_phase_BactConvNet 그리고 두 종류의 정보를 함께 활용하는 Multi_BactConvNet이라 명명한 뉴럴 네트워크를 학습시켰다. 각 뉴럴 네트워크에서 학습가능한 매개변수들은 오차 역전파 알고리즘(error backpropagation algorithm)⁸⁻⁹을 통해 반복적으로 개선되었고, 최종적으로 선정된 모델들에 대해서는 네트워크에 입력된 적 없는 새로운 평가 데이터셋(testset)으로 성능을 측정하였다. 이를 바탕으로 입력된 5종의 박테리아 이미지에 대해 실제 정답과 얼마나 일치하는지를 평가하였다.

FPM 장비를 통해 얻은 이미지에 대해 segmentation을 진행해 5종의 박테리아 각 2000장(광흡수 1000장, 위상 1000장), 총 1만 장의 이미지 데이터셋을 만들었다. 그리고 모델 학습에 사용되는 trainset 7200장, 학습과정에서 평가에 활용되는 validation set 1800장, 그리고 학습이 완료된 모델에 대한 성능평가에 사용되는 testset 1000장으로 나누어주었다.

뉴럴 네트워크를 학습시켜 데이터를 분류하는 문제는 비선형적으로 분포해 있는 데이터 분류를 해결하는 방법이라고 할 수 있다⁸⁻⁹. 따라서 이 연구에서 모델을 개발 및 평가하는 과정은 학습에 광흡수정보만을 이용했을 때, 위상정보만을 이용했을 때, 그리고 모두 이용하는 모델을 설계하고 학습시켜 각 세 가지에 대해 새로운 데이터를 보여주었을 때, 데이터를 얼마나 비선형적으로 잘 구분하는지를 평가하는 과정이다.

BactConvNet은 크게 Single 모델과 Multi 모델 두 가지로 구분된다. Single 모델은 Fig. 1에서 볼 수 있듯이 광흡수 이미지와 위상이미지에 사용되는 모델로서, 각 세 번의 convolution layer, nonlinearity, pooling 연산을 거쳐 처리된다. 그리고 fully connected layer를 통과해 다섯 가지 중

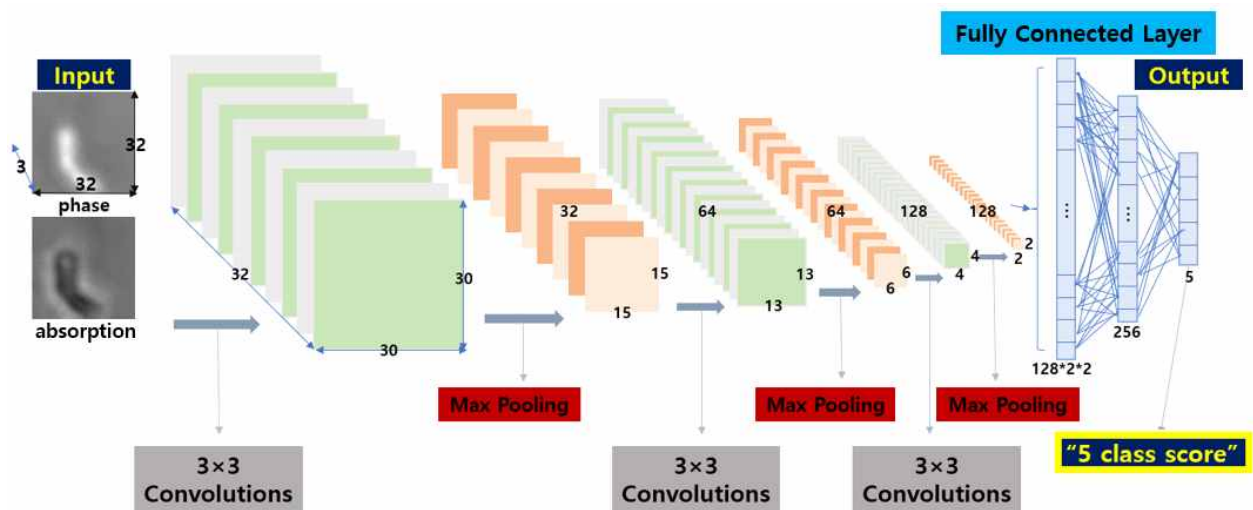


Fig. 1 Single input neural network model architecture

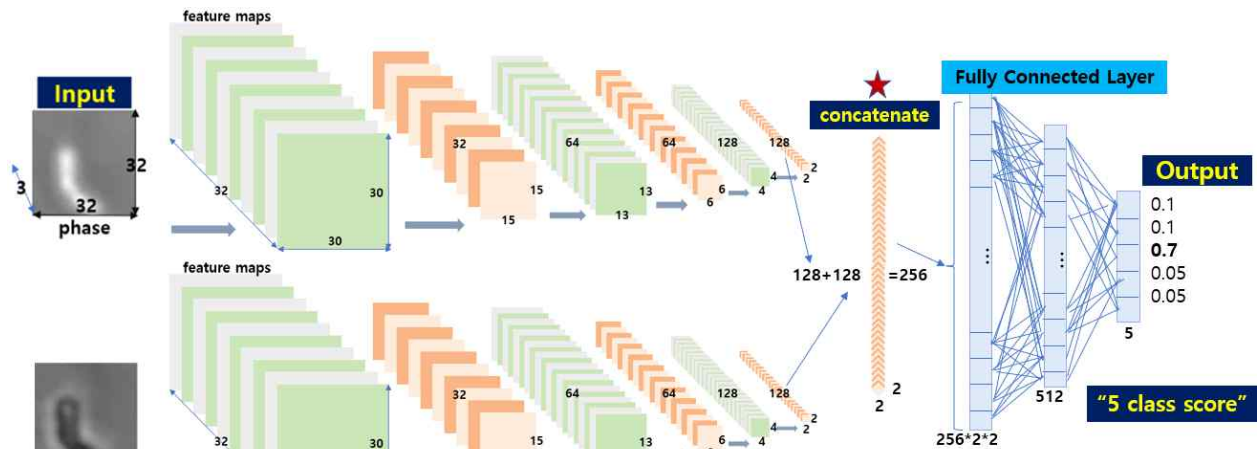


Fig. 2 Multi input neural network model architecture

Layer (type)	Output Shape	Param #
Conv2d-1	[-1, 32, 30, 30]	896
ReLU-2	[-1, 32, 30, 30]	0
MaxPool2d-3	[-1, 32, 15, 15]	0
Conv2d-4	[-1, 64, 13, 13]	18,496
ReLU-5	[-1, 64, 13, 13]	0
MaxPool2d-6	[-1, 64, 6, 6]	0
Conv2d-7	[-1, 128, 4, 4]	73,856
ReLU-8	[-1, 128, 4, 4]	0
MaxPool2d-9	[-1, 128, 2, 2]	0
Dropout-10	[-1, 512]	0
Linear-11	[-1, 256]	131,328
ReLU-12	[-1, 256]	0
Dropout-13	[-1, 256]	0
Linear-14	[-1, 5]	1,285

Total params: 225,861
Trainable params: 225,861
Non-trainable params: 0

(a)

conv2d (Conv2D)	(None, 30, 30, 32)	320	['input_2[0][0]']
conv2d_3 (Conv2D)	(None, 30, 30, 32)	320	['input_3[0][0]']
max_pooling2d (MaxPooling2D)	(None, 15, 15, 32)	0	['conv2d[0][0]']
max_pooling2d_3 (MaxPooling2D)	(None, 15, 15, 32)	0	['conv2d_3[0][0]']
conv2d_1 (Conv2D)	(None, 13, 13, 64)	18496	['max_pooling2d[0][0]']
conv2d_4 (Conv2D)	(None, 13, 13, 64)	18496	['max_pooling2d_3[0][0]']
max_pooling2d_1 (MaxPooling2D)	(None, 6, 6, 64)	0	['conv2d_1[0][0]']
max_pooling2d_4 (MaxPooling2D)	(None, 6, 6, 64)	0	['conv2d_4[0][0]']
conv2d_2 (Conv2D)	(None, 4, 4, 128)	73856	['max_pooling2d_1[0][0]']
conv2d_5 (Conv2D)	(None, 4, 4, 128)	73856	['max_pooling2d_4[0][0]']
max_pooling2d_2 (MaxPooling2D)	(None, 2, 2, 128)	0	['conv2d_2[0][0]']
max_pooling2d_5 (MaxPooling2D)	(None, 2, 2, 128)	0	['conv2d_5[0][0]']
concatenate (Concatenate)	(None, 2, 2, 256)	0	['max_pooling2d_2[0][0]', 'max_pooling2d_5[0][0]']
flatten (Flatten)	(None, 1024)	0	['concatenate[0][0]']
dense (Dense)	(None, 512)	524800	['flatten[0][0]']
dropout (Dropout)	(None, 512)	0	['dense[0][0]']
dense_1 (Dense)	(None, 256)	131328	['dropout[0][0]']
dropout_1 (Dropout)	(None, 256)	0	['dense_1[0][0]']
dense_2 (Dense)	(None, 5)	1285	['dropout_1[0][0]']

Total params: 842,757
Trainable params: 842,757
Non-trainable params: 0

(b)

Fig. 3 (a) Number of parameters used in Single input model and (b) in Multi input model

해당하는 확률점수로 계산된다. Multi 모델 Fig. 2은 광 흡수 이미지와 위상 이미지가 병렬적으로 Single 모델에서와 같이 세 번의 연산을 거친 뒤, $128 \times 2 \times 2$ 크기의 feature map 들을 이어 붙여준다. 그리고 이어 붙여진 feature map은 fully connected layer를 통과하여 다섯 가지 확률점수 결과로 출력된다.

위 모델에서의 핵심은 convolution filter를 통과한 뒤 nonlinearity 연산과정(ReLU)을 거친 입력이미지가 class label로 nonlinearly mapping되는 과정이다. 이를 위해 신경망에서는 매우 많은 변수를 통한 최적화가 필요하고 이는 학습과정을 느리고 어렵게 만들 수 있다. 그러나 CNN에서는 convolution filter의 cross correlation 연산을 통해 공간 정보를 유지하고, pooling 과정을 통해 receptive field를 빠르게 만들어내기 때문에 현저히 적은 수의 매개변수를 통해 학습할 수 있다^{8,10}. BactConvNet 모델에서 입력되는 이미지의 크기는 $3 \times 32 \times 32$ 로 크기 않기 때문에 3×3 의 convolution filter를 사용해 공간정보의 손실을 적게 하고, 비선형성을 추가한 뒤, receptive field를 효과적으로 얻기 위한 max pooling 과정을 세 번씩 거친다.

그리고 실제 입력된 박테리아의 정답 label과 모델에서 예측한 출력값을 통해 모델의 성능을 평가하는 지표인 loss function을 계산한 뒤, 오차 역전파를 통해 매개변수들을 갱신해나간다.

설계한 Single 모델의 매개변수의 개수는 225,861개였고, Fig. 3(a)를 보면 fully connected layer에서 절반 이상의 수를 차지함을 알 수 있다. Multi 모델, Fig. 3(b)는 842,757개로 이 중 657,413개가 class score를 도출해내는 fully connected layer에서 학습할 때 사용되었다.

뉴럴 네트워크에서는 매개변수를 갱신하기 위해 오차 역전파 알고리즘과 기울기 하강법(gradient descent)을

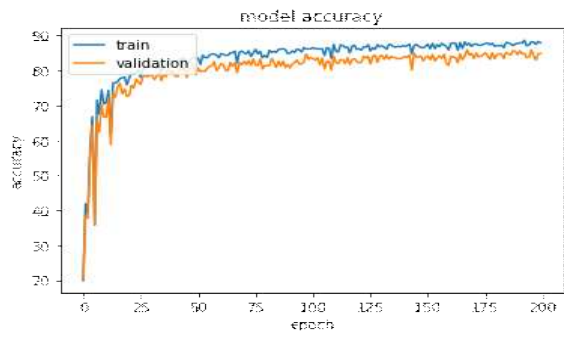


Fig. 4 (a) Single absorption model train & validation accuracy
optimizer = Adagrad, learning rate = 0.01, epochs = 200

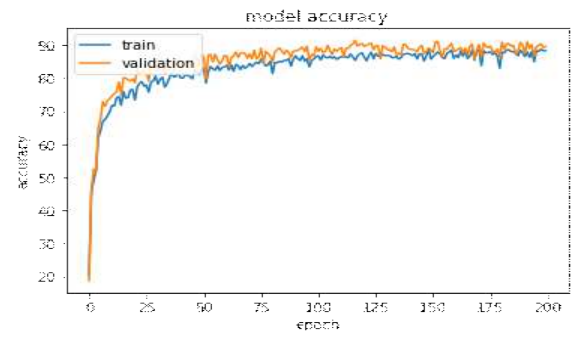


Fig. 4 (b) Single absorption model train & validation accuracy
optimizer = Adam, learning rate = 0.001, epochs = 200

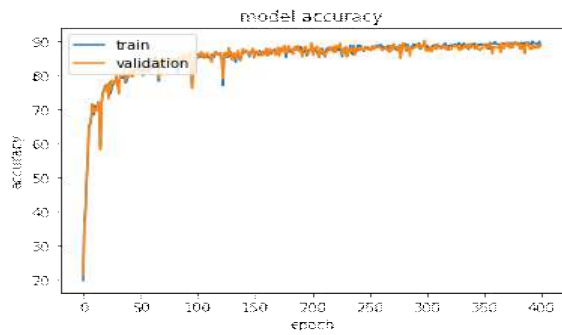


Fig. 5 (a) Single absorption model train & validation accuracy
optimizer = Adagrad, learning rate = 0.01, epochs = 400

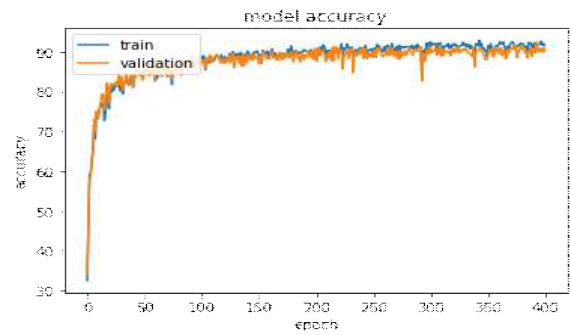


Fig. 5 (b) Single absorption model train & validation accuracy
optimizer = Adam, learning rate = 0.001, epochs = 400

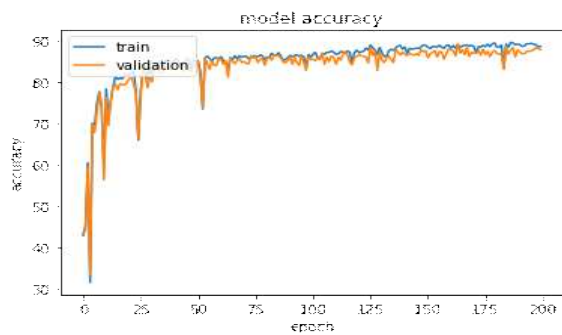


Fig. 6 (a) Single phase model train & validation accuracy
optimizer = Adagrad, learning rate = 0.01, epochs = 200

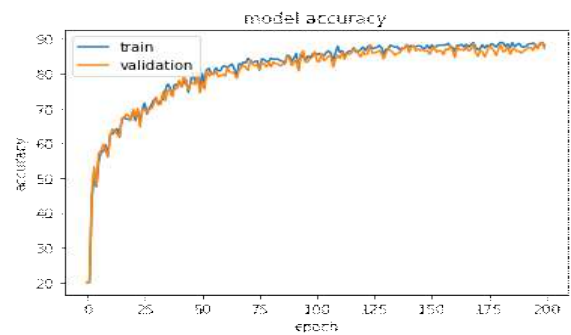


Fig. 6 (b) Single phase model train & validation accuracy
optimizer = Adam, learning rate = 0.0001, epochs = 200

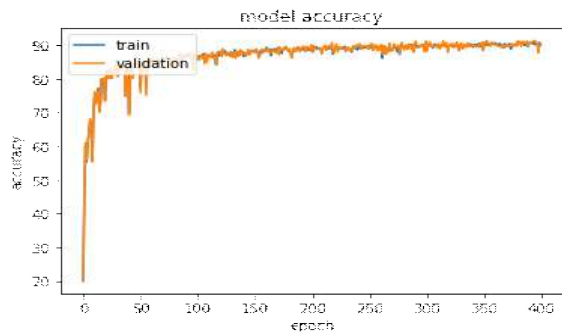


Fig. 7 (a) Single phase model train & validation accuracy
optimizer = Adagrad, learning rate = 0.01, epochs = 400

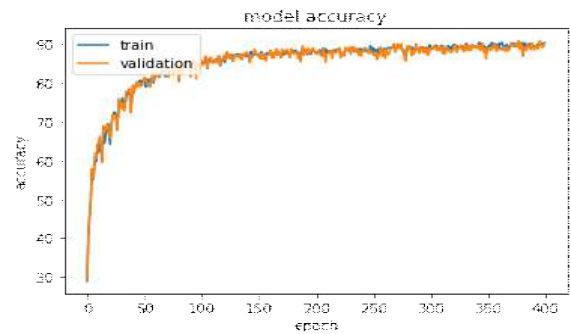


Fig. 7 (b) Single phase model train & validation accuracy
optimizer = Adam, learning rate = 0.0001, epochs = 400

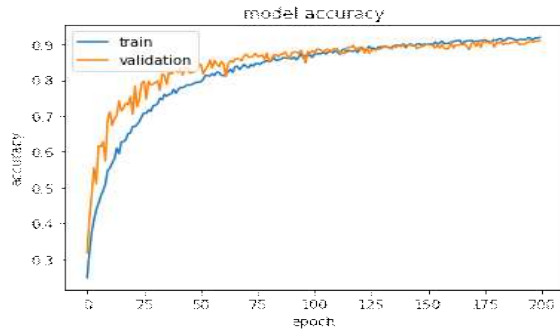


Fig. 8 (a) Multi model train & validation accuracy
optimizer = Adagrad, learning rate = 0.001, epochs = 200

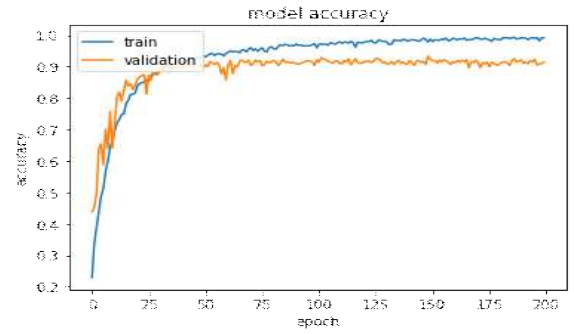


Fig. 8 (b) Multi model train & validation accuracy
optimizer = Adam, learning rate = 0.0001, epochs = 200

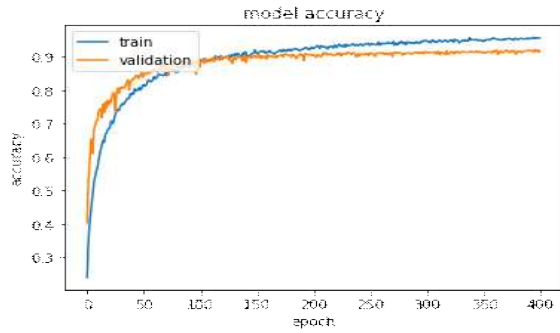


Fig. 9 (a) Multi model train & validation accuracy
optimizer = Adagrad, learning rate = 0.001, epochs = 400

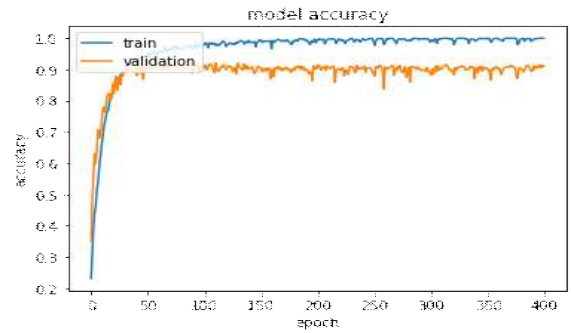


Fig. 9 (b) Multi model train & validation accuracy
optimizer = Adam, learning rate = 0.0001,
epochs = 400, L2 regularization

적용하여 매개변수를 적절한 값에 수렴시킨다. 이러한 매개변수 최적화 과정에서 다양한 optimizer를 사용할 수 있고, 본 실험에서는 Stochastic Gradient Descent(SGD)^{8,9}, Adagrad, RMSprop, Adam을 사용하였다. 각 optimizer에 대해서는 learning rate(lr)를 0.01, 0.001, 0.0001로 택해주어 총 12가지의 모델을 학습했다.

Single_absorp_BactConvNet에서 12가지를 200epochs로 학습시키는 실험에서 좋은 train accuracy와 안정적인 loss 값을 나타내는 2가지 Fig. 4(a)와 Fig. 4(b)를 선정해주었다.

그 후 epoch 수를 증가시키고 regularization을 적용했다. 각 2가지에 대해 400epochs로 학습횟수를 증가시킨 것, 이로 인해 발생할 수 있는 overfitting을 방지해주기 위해 batch normalization¹¹을 적용한 것, L2 regularization⁹을 적용한 것 중 학습이 잘 되고 정확도 성능이 좋은 2가지 Fig. 5(a)와 Fig. 5(b)를 선택했다. 결과적으로 Fig. 5(a)와 Fig. 5(b) 2가지를 광합수 이미지만을 사용하는 CNN모델에서 가장 좋은 두 가지 모델로 선정했다.

Single_phase_BactConvNet도 Single_absorp_BactConvNet에서의 실험과정과 동일한 과정을 거쳤으며 12가지 실험 모델 중 Fig. 6(a)와 Fig. 6(b) 2가지를 선정할 수 있었다. 그 후 정규화를 진행하여 선정된 모델은 Fig. 7(a)와 Fig. 7(b)이다.

Multi_BactConvNet도 Single 모델에서와 같은 과정으로 실험하였고 처음 12가지 중 Fig. 8(a)와 Fig. 8(b) 그리고 정규화 후 Fig. 9(a)와 Fig. 9(b)를 선정했다.

위 실험과정을 거쳐 Single_absorp_BactConvNet, Single_phase_BactConvNet, Multi_BactConvNet에서 각각 두 가지씩 최종 후보 모델에 포함시켰다(Fig. 5(a), Fig. 5(b), Fig. 7(a), Fig. 7(b), Fig. 9(a), Fig. 9(b)). 6가지 모델의 validation accuracy는 모두 약 90 % 정도의 정확도에 수렴하였다.

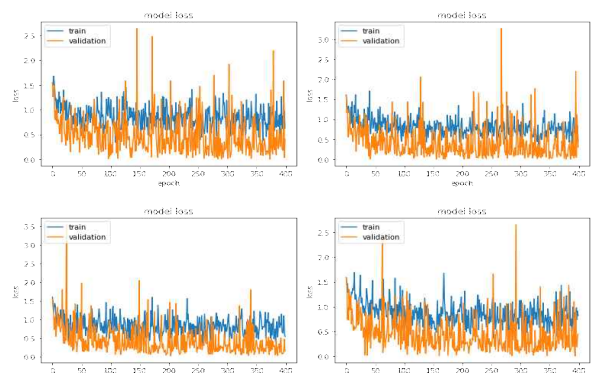


Fig. 10 (a) Single input model loss

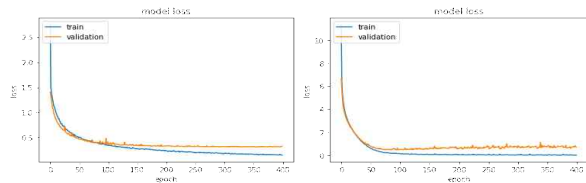


Fig. 10 (b) Multi input model loss

Fig. 10 (a) Single input model의 loss 그래프와 Fig. 10 (b) Multi input model의 loss 그래프를 비교해보면 Single와 Multi 두 모델 모두 평균적으로 loss값이 잘 수렴하는 것을 알 수 있지만, Single input model에서는 loss값의 진동이 심한 반면에 Multi input model에서는 안정적으로 수렴하는 것을 보아 Multi input model이 더 일관된 성능을 보이는 것을 알 수 있다.

선정한 모델들에 대해 train & validation accuracy와 loss를 그래프를 통해 확인한 뒤, testset으로 정량적인 평가를 시행해 최종 모델 선정과정을 진행했다. Testset 이미지는 학습과정에 전혀 포함되지 않은 이미지들로, 선정된 6가지 모델들이 각각의 학습을 완료한 뒤에 입력되었다. 결과는 Fig. 11 및 Table 1과 같았다.

Adagrad lr = 0.001을 사용한 Multi 모델이 91.4%로 가장 높은 정확도를 보였고, 다음으로 Adam lr = 0.0001을 사용한 Single phase 모델이 90.2%의 정확도를 보였다. 그리고

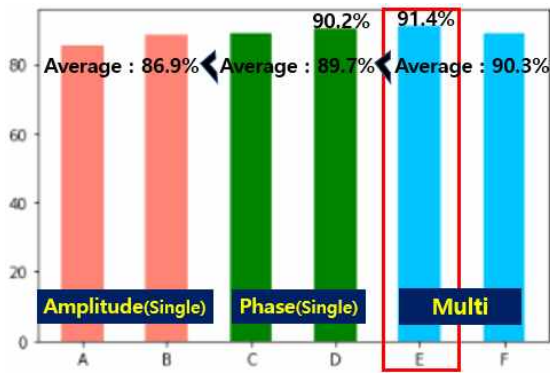


Fig. 11 Selected model testset accuracy

Table 1 Selected model testset accuracy

Absorption		A	Adagrad lr = 0.01	acc (%)	85.4
				loss	0.4311
		B	Adam lr = 0.001	acc (%)	88.4
				loss	0.2181
Phase		C	Adagrad lr = 0.01	acc (%)	89.2
				loss	0.2729
		D	Adam lr = 0.0001	acc (%)	90.2
				loss	0.7231
Multi		E	Adagrad lr = 0.001	acc (%)	91.4
				loss	0.3292
		F	Adam lr = 0.0001	acc (%)	89.2
				loss	0.7239

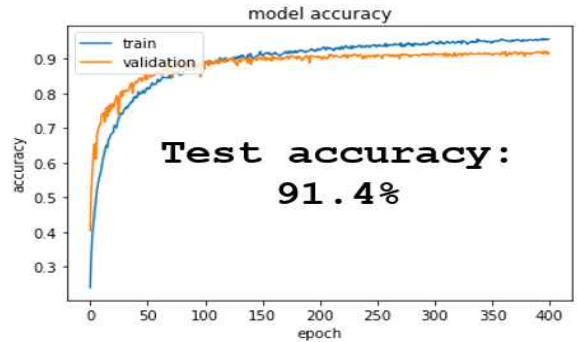


Fig. 12 (a) The best model testset accuracy
(multi adagrad lr=0.001)

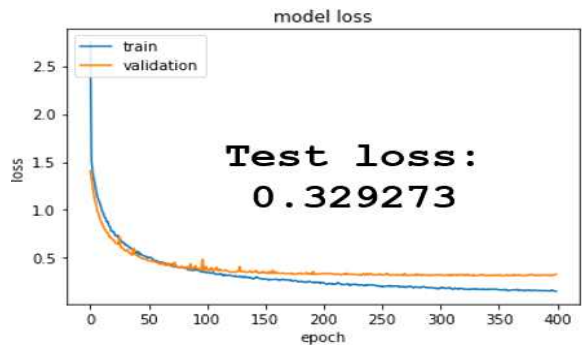


Fig. 12 (b) The best model testset loss
(multi adagrad lr=0.001)

Table 2 The best model confusion matrix

Predicted Actual	BC	BR	MI	SA	SE
BC	100	0	0	0	0
BR	0	86	7	0	7
MI	0	6	93	0	1
SA	3	5	0	85	7
SE	1	4	1	2	92

평균적으로 Multi 모델은 90.3%, Single phase 모델은 89.7%, Single absorption 모델은 86.9%를 달성했다. 이를 통해 박테리아의 두 가지 정보를 모두 활용해 분류를 수행하는 것이 가장 효과적이고, 위상정보가 광흡수 정보보다 분류 작업에 유용한 특성을 제공한다는 것을 알 수 있다.

최종적으로 가장 높은 정확도를 달성한 Adagrad lr = 0.001 Multi 모델을 본 전체 실험에서 가장 뛰어난 모델로 정했다. 그리고 testset을 통해 confusion matrix를 도식한 결과는 Table 2과 같았다. 다섯 가지 박테리아 BC, BR, MI, SA, SE 별 100개의 실제 이미지가 주어졌을 때, 선택된 모델은 100개(100%), 86개(86%), 93개(93%), 85개(85%), 92개(92%)를 실제 박테리아 종류에 맞게 예측하는 것을 보여준다.

3. Discussion

FPM 장비를 통해 얻은 박테리아의 광흡수와 위상정보를 모두 활용하여 다량의 박테리아에 대해 정확하고 빠른 종 간의 분류를 수행하는 시스템을 제안하였고, 딥러닝을 활용한 컴퓨팅 실험을 통하여 모델을 개발 및 검증하였다.

FPM을 통해 얻은 이미지 데이터셋으로 Single 두 가지와 Multi 모델을 학습시켰다. 그리고 각 경우 12가지 모델 중 2개를 선정해 정규화를 진행한 6개 모델을 만들었고, 다시 2가지 모델을 선정했다. Single absorption과 Single phase model에서 2가지 후보, Multi에서 2가지 후보 총 6가지를 평가 데이터셋(testset)으로 정량적인 평가를 진행해 최종 모델을 선정했다.

박테리아의 광흡수 정보와 위상정보는 각 박테리아의 특성과 밀접한 관련이 있어 classifier를 만들 때 유용한 데이터로 활용할 수 있다. 두 가지 정보는 박테리아의 특성을 나타내주지만 각 종과의 관계는 직접적이지 않기 때문에 지도학습(supervised learning)을 통해 이 문제를 해결할 수 있다⁷.

BactConvNet은 학습된 데이터들을 바탕으로 얻어낸 representation learning capability를 통해 광흡수와 위상정보에서 종 간의 특징을 구분해낸다. 이 방법은 label-free 상태의 새로운 박테리아에 대해 기존 지식이나 전처리과정을 거치지 않은 채로 높은 정확도로 분류해낼 수 있고, 위 실험을 통해 개발한 모델은 평균적으로 91.4% 정확도의 분류성능을 보였다.

위 시스템을 더 완전하게 하기 위한 제안할 수 있는 사항이 있다. 모델에 입력되는 이미지는 개별의 박테리아 이미지로 입력되어야 한다. 그러나 공기 중의 박테리아를 포집할 시에 하나의 박테리아만을 분리해낼 수 없다. FPM을 통해 얻은 이미지는 고해상도를 보장해주기 때문에 푸리에 공간에서 중첩되어 얻은 고해상도의 이미지를 segmentation해야 한다. 이를 수작업으로 수많은 박테리아들을 하나씩 segmentation한다면 초기 연구에서 제안한 “빠르고 정확한 분류”의 목적에 부합하지 않는다. 따라서 분류작업을 수행하기 전에 객체탐지(object detection) 알고리즘을 통해 segmentation 작업을 수행한다면 분류작업에 입력되는 이미지를 자동적으로 얻어 본 시스템에 견고함을 더할 수 있을 것이다.

4. Materials and methods

4.1. Bacteria

본 연구에서는 연세대학교 황정호 교수님 실험실에서 제공받은 다음 다섯 종류의 박테리아가 사용되었다. (i) BC: *Bacillus cereus*, (ii) BR: *Brevibacterium casei*, (iii) MI;

Microbacterium arborescens, (iv) SA: *Staphylococcus aureus*, (v) SE: *Staphylococcus epidermidis*).

4.2. Fourier ptychographic microscopy

해상도(resolution)과 시야각(field of view, FOV)은 기존의 이미징 시스템에서 오랜 두 가지의 타협점이다. 해상도가 높으면 작고 상세한 이미지를 얻을 수 있고, 해상도가 낮으면 시야각은 커지고 큰 이미지를 얻을 수 있다¹².

Fourier ptychography(FP)는 이러한 이미징 시스템에 내재되어있는 trade-off를 해결하기 위한 방법이다¹². FP는 높은 처리율과 높은 해상도에 관해 광학계에서 가지는 물리적인 한계에 대해 computation의 영역에서 해결방안을 제시한다¹². FP는 기존의 현미경 플랫폼에서 별도의 하드웨어 조작 없이 SBP를 확장하기 위해 개발되었다. 다음의 주요 내용을 통해 FP를 이해할 수 있다.

(i) FP는 합성 개구와 위상 복원 개념을 활용해 raw data를 고해상도와 넓은 시야각으로 만들어주는 computational 기법이다. 해상도와 시야각의 타협이 있는 기존의 방법과 다르게, FP는 두 가지를 동시에 달성할 수 있다¹².

(ii) FP는 intensity 기반 측정법들로 샘플의 intensity와 phase 이미지를 얻어낼 수 있다¹².

(iii) FP에서는 computational method를 통해 수차를 조정할 수 있어 물리적인 시스템의 제약에 의한 수차로부터 자유롭다¹².

(iv) FP에서는 더 넓은 영역에 refocus할 수 있어서 수차의 한 종류인 defocus 문제를 해결할 수 있다¹².

4.3. Deep learning

BactConvNet는 각 박테리아들의 absorptitude absorption과 phase image를 이용해 설계된 CNN이다. 이 연구에서 Single_absorp_BactConvNet과 Single_phase_BactConvNet은 Pytorch framework를 통해 실험을 진행하였고, 두 가지의 정보를 같이 입력해주어야 하는 Multi_BactConvNet은 Tensorflow framework로 설계했다.

최종적으로 적용할 뉴럴 네트워크 구조는 몇 가지 다른 크기의 kernel을 시험해본 후 3*3 kernel을 택하였다. 입력 이미지의 크기가 32*32로 작기 때문에 큰 kernel을 택할 경우에는 receptive field를 빠르게 형성할 수 있지만 공간정보가 과도하게 축소된다. 따라서 3*3 kernel을 사용하되 pooling 과정을 통해 매개변수의 수가 많아지는 것을 방지해주면서 receptive field를 빠르게 형성해주어 Single 모델에서는 대략 23만개, Multi 모델에서는 84만개의 매개변수가 사용되었다. 모델에 비선형성은 매 convolution과 fully connected layer를 지날 때 Rectified Linear Unit(ReLU)로 계산해주었고, 이를 통해 딥뉴럴 네트워크에서 층이 깊어짐에 따라 발생할 수 있는 기울기 소실 문제(vanishing gradient

problem)를 방지해주었다¹³.

과적합(overfitting)발생을 방지하기 위한 방법으로는 “dropout”¹⁴, “data augmentation”¹⁰, “weight decay regularization”⁹, “batch normalization”¹¹ 을 적용했다. Dropout은 네트워크의 일부를 생략하여 과적합을 방지하는 방법으로, fully connected layer에서 한 층을 지날 때 마다 ReLU nonlinearity를 적용해준 뒤에 0.5의 확률로 적용되었다. Data augmentation이란 실제 trainset을 여러 방법으로 증강시켜 실질적인 학습 데이터셋의 규모를 키우는 방법이다¹¹. Pytorch framework에서는 이미지를 수평으로 뒤집는 RandonHorizontalFlip, 지정한 각도 범위 내에서 이미지에 회전을 적용해주는 RandomRotation, 선형변환과 이동변환을 포함하는 RandomAffine을 적용했다. Weight decay regularization⁹은 weight를 활용해 loss에 반영함으로써 모델의 복잡도를 감소시키는 정규화 방법이다. 전형적인 L1, L2 regularization이 있고, 각 모델에서 두 가지의 후보군을 선택해준 뒤에 정규화를 진행할 때, L2 regularization을 적용했다. Batch normalization¹¹은 학습 과정에서 각 배치 단위 별로 데이터가 다양한 분포를 가지더라도 각 배치별 평균과 분산을 이용해 정규화하는 것을 말한다. 하지만 Single 모델과 Multi 모델에서 batch normalization을 진행했을 때 학습에 도움되는 효과를 나타내지 않았기 때문에 본 연구에서 모델을 개발할 때는 유용한 정규화 방법이 아니었다.

학습 단계에서 매개변수들은 loss function 값을 최소화시키기 위해 오차 역전파 알고리즘⁸⁻⁹을 통해 계산된 gradient descent를 통해 갱신되었다. Loss 계산에는 실제 정답(true label)과 모델에서 예측한 정답 사이의 차이를 정량화하는 cross-entropy loss방법이 사용되었다. Chain rule을 통해 loss function을 편미분해주고, 계산된 gradient는 stochastic gradient descent(SGD), RMSprop, Adagrad, Adam에 이용되었다.

컴퓨팅 자원으로는 Google Colaboratory Pro를 통해 하나의 GPU(Nvidia Tesla P100-PCIE 16GB)와 CUDA Toolkit 11.2을 사용했다.

참고문헌

- (1) Lohmann, A. W., Dorsch, R. G., Mendlovic, D., Zalevsky, Z., & Ferreira, C. (1996). Space - bandwidth product of optical signals and systems. *JOSA A*, 13(3), 470-473.
- (2) Zheng, G., Horstmeyer, R., & Yang, C. (2013). Wide-field, high-resolution Fourier ptychographic microscopy. *Nature photonics*, 7(9), 739-745.
- (3) Xu, W., Jericho, M. H., Meinertzhagen, I. A., & Kreuzer, H. J. (2001). Digital in-line holography for biological applications. *Proceedings of the National Academy of Sciences*, 98(20), 11301-11305.
- (4) Greenbaum, A., Luo, W., Khademhosseini, B., Su, T. W., Coskun, A. F., & Ozcan, A. (2013). Increased space-bandwidth product in pixel super-resolved lensfree on-chip microscopy. *Scientific reports*, 3(1), 1-8.
- (5) Zheng, G., Lee, S. A., Antebi, Y., Elowitz, M. B., & Yang, C. (2011). The ePetri dish, an on-chip cell imaging platform based on subpixel perspective sweeping microscopy (SPSM). *Proceedings of the National Academy of Sciences*, 108(41), 16889-16894.
- (6) Zheng, G., Lee, S. A., Yang, S., & Yang, C. (2010). Sub-pixel resolving optofluidic microscope for on-chip cell imaging. *Lab on a Chip*, 10(22), 3125-3129.
- (7) Jo, Y., Park, S., Jung, J., Yoon, J., Joo, H., Kim, M. H., ... & Park, Y. (2017). Holographic deep learning for rapid optical screening of anthrax spores. *Science advances*, 3(8), e1700606.
- (8) LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, 521(7553), 436-444.
- (9) Haykin, S. (2010). *Neural networks and learning machines*, 3/E. Pearson Education India.
- (10) Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25, 1097-1105.
- (11) Ioffe, S., & Szegedy, C. (2015, June). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning* (pp. 448-456). PMLR.
- (12) Zheng, G., Shen, C., Jiang, S., Song, P., & Yang, C. (2021). Concept, implementations and applications of Fourier ptychography. *Nature Reviews Physics*, 3(3), 207-223.
- (13) Hochreiter, S. (1998). The vanishing gradient problem during learning recurrent neural nets and problem solutions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 6(02), 107-116.
- (14) Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15(1), 1929-1958.