UNIVERSITY OF CALIFORNIA

Los Angeles

Computational Imaging and Sensing in Diagnostics with Deep Learning

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Electrical and Computer Engineering

by

Calvin Clifford Brown

2020

ABSTRACT OF THE DISSERTATION

Computational Imaging and Sensing

in Diagnostics with Deep Learning

by

Calvin Clifford Brown

Doctor of Philosophy in Electrical and Computer Engineering

University of California, Los Angeles, 2020

Professor Aydogan Ozcan, Chair

Computational imaging and sensing aim to redesign optical systems from the ground up, jointly considering both hardware/sensors and software/reconstruction algorithms to enable new modalities with superior capabilities, speed, cost, and/or footprint. Often systems can be optimized with targeted applications in mind, such as low-light imaging or remote sensing in a specific spectral regime. For medical diagnostics in particular, computational sensing could enable more portable, cost-effective systems and in turn improve access to care. In the last decade, the increased availability of data and cost-effective computational resources coupled with the commodification

of neural networks has accelerated and expanded the potential for these computational sensing systems.

First, I will present my work on a cost-effective system for quantifying antimicrobial resistance, which could be of particular use in resource-limited settings, where poverty, population density, and lack of healthcare infrastructure lead to the emergence of some of the most resistant strains of bacteria. The device uses optical fibers to spatially subsample all 96 wells of a standard microplate without any scanning components, and a neural network identifies bacterial growth from the optical intensity information captured by the fibers. Our accelerated antimicrobial susceptibility testing system can interface with the current laboratory workflow and, when blindly tested on patient bacteria at UCLA Health, was able to identify bacterial growth after an average of 5.72 h, as opposed to the gold standard method requiring 18–24 h. The system is completely automated, avoiding the need for a trained medical technologist to manually inspect each well of a standard 96-well microplate for growth.

Second, I will discuss a deep learning-enabled spectrometer framework using localized surface plasmon resonance. By fabricating an array of periodic nanostructures with varying geometries, we created a “spectral encoder chip” whose spatial transmission intensity depends upon the incident spectrum of light. A neural network uses the transmitted intensities captured by a CMOS image sensor to faithfully reconstruct the underlying spectrum. Unlike conventional diffraction-based spectrometers, this framework is scalable to large areas through imprint lithography, conducive to compact, lightweight designs, and, crucially, does not suffer from the resolution–signal strength tradeoff inherent to grating-based designs.

The dissertation of Calvin Clifford Brown is approved.

Dino Di Carlo

Omai B. Garner

Mona Jarrahi

Achuta Kadambi

Yair Rivenson

Aydogan Ozcan, Committee Chair

University of California, Los Angeles

2020

Contents

List of figures

List of tables

Acknowledgements

I would like to thank Professor Aydogan Ozcan for his care, effort, and boundless energy as an advisor, mentor, and collaborator. His knowledge and experience, not only of technical content, but also of research, academia, business, and life as a scientist in general, were critical to my success as a graduate student. I would also like to thank the members of my committee, who have helped shape my work and have encouraged me to embrace new perspectives.

While I have had too many colleagues to count throughout my time at UCLA, I especially want to thank Dr. Zachary Ballard, Derek Tseng, Dr. Zoltan Gorocs, Dr. Yair Rivenson, Artem Goncharov, Dr. Mustafa Ugur Daloglu, Dr. Hyou-Arm Joung, Dr. Hatice Ceylan Koydemir, Deniz Mengu, Dr. Yichen Wu, Dr. Muhammed Veli, Kevin de Haan, Dr. Hongda Wang, and Dr. Jeremie Richard. It has been a blessing to be a part of such a large, diverse lab and I am grateful for the friends and colleagues with whom I've been able to work.

I want to thank my friends and family for the much needed support over the past five years, especially my parents Catherine and David, who have given me every opportunity to succeed and have blessed me with more love than any one child can deserve. I also would like to thank my girlfriend Ashley, who should be awarded an honorary degree for the amount of support she has given me throughout my PhD.

Lastly, I would like to thank the National Science Foundation (NSF) for its support through the Graduate Research Fellowship Program. The knowledge that the NSF believed me to be worth taking a chance on gave me confidence to face and tackle research challenges well beyond my expertise and comfort zone. I would like to thank the American Chemical Society for permission to reproduce parts of published works here. I would also like to thank the PATHS-UP research

center for its support and the University of California granting me invaluable opportunities and the gift of knowledge. *Fiat Lux*

Vita

Calvin Brown earned a BS in Engineering Physics from the University of California, Berkeley in 2015, followed by an MS in Electrical Engineering from UCLA in 2017. His current research sits at the intersection of computational imaging/sensing and artificial intelligence, particularly for medical and global health applications. In addition, he has conducted research at Lawrence Livermore National Laboratory, Berkeley Orthopaedic Biomechanics Laboratory, and Helmholtz-Zentrum Berlin in various engineering disciplines. Calvin earned the 2016 NSF GRFP fellowship, 2019 UCLA Innovation Fund, and 2019 NSF PATHS-UP Innovation Seed Fund. He has also won awards for his presentations at the 2018 Emerging Researchers National (ERN) Conference in the Technology & Engineering category (third place) and the 2016 UCLA Grad Slam (semifinalist).

Selected publications

1. **Brown** et al., "Automated, Cost-Effective Optical System for Accelerated Antimicrobial Susceptibility Testing (AST) Using Deep Learning." *ACS Photonics* **2020** 7 (9), 2527-2538. https://doi.org/10.1021/acsphotonics.0c00841

2. **Brown** et al., "Neural network-based on-chip spectroscopy using a scalable plasmonic encoder." 2020; arXiv:2012.00878

3. Daloglu et al., "Low-cost and portable UV holographic microscope for high-contrast protein crystal imaging." APL Photonics **4**, 030804 (2019); https://doi.org/10.1063/1.5080158

Chapter 1 Introduction

Here I will give a brief introduction to computational imaging and sensing, as well as an overview of my work. Chapters 2 and 3 will detail two systems that I have designed, fabricated, and tested: an automated antimicrobial susceptibility testing (AST) system and an on-chip deep learning-enabled spectroscopy framework, respectively. Chapter 4 will contain a brief conclusion.

Computational imaging and sensing

Computational imaging and computational sensing, as the names imply, take advantage of modern computational capabilities to improve the results from and expand the capabilities of classical imaging and sensing modalities.[1–4] Furthermore, by redesigning hardware sensors and software algorithms in tandem, entirely new modalities can be designed.[5] For instance, this approach can enable systems with any/all of the following characteristics: more compact, lower cost, higher fidelity or signal-to-noise ratio (SNR), lower power consumption, remote operation, etc. In this thesis, we demonstrate two novel sensing and imaging frameworks, present their performance in preliminary as well as clinical experiments, and propose potential further applications of the technologies. The first framework is an optical fiber-based system that autonomously monitors bacterial growth in real time using deep learning, with clinical results detecting antimicrobial resistance (AMR) faster than the gold standard method. The second framework is a method for designing deep learning-enabled on-chip spectrometers that can reconstruct arbitrary spectra.

Overview of dissertation

The first framework (Chapter 2) uses LED illumination and bundles of plastic optical fibers to capture changes in turbidity in each well of a standard 96-well microplate. Due to the ubiquity of plate-based assays not only in diagnostics and clinical microbiology, but also research and development and pharmaceutical production among others. The primary application we have explored thus far is antimicrobial susceptibility testing (AST). Due to the global rise of antimicrobial resistance,[6] clinicians increasingly depend on AST to inform the prescription of antibiotics to patients with bacterial infections. The gold standard AST method is broth microdilution (BMD), in which patient bacteria are incubated in the presence of candidate antibiotics and growth medium. After 18–24 h, the plate containing the bacteria is removed from the incubator and visually inspected for growth by a trained medical technologist. The presence/absence of growth in wells containing different concentrations of different antibiotics indicates which antibiotics may prove effective to treat the infection and which antibiotics the bacteria is resistant. The main drawbacks of the current gold standard method are that it requires waiting 18–24 h for incubation, it requires a trained technologist to manually read the plates to obtain results, and it is subject to variability due to human readout. The system we have demonstrated addresses these issues because it is can detect turbidity in an average of just 5.72 h of incubation, with 95.03% accuracy, and it is completely automated, eliminating the need for manual reading and human readout errors.[7] The system is entirely composed of cost-effective components, namely LEDs, plastic optical fibers, singlet lenses, Raspberry Pi computers and camera modules, and 3D printed plastic housing. Unlike available commercial AST systems, it contains no expensive objective lenses or scanning components, and does not require specialized cartridges or consumables. In addition, because the system does not depend on microscopy, it is not limited to specific bacteria or specific types of samples (e.g. urine).

The LEDs illuminate a standard 96-well microplate from above and the optical fibers below the plate relay the transmitted light to the Raspberry Pi camera for imaging. A bundle of 21 fibers sits underneath each well to sample the spatial intensity distribution of the transmitted light. This spatial information is crucial for distinguishing e.g. weak growth from settling, condensation, or bubbles. To determine which wells contain growth/turbidity at what time, a neural network takes in the fiber intensities for a given well over time and outputs a probability of turbidity between 0 and 1. We use a fully-connected network, with dropout and batch normalization for regularization and accelerating the training process, respectively. The system was validated with Staphylococcus aureus infection patients in the UCLA Health system. Our automated AST system was able to correctly identify turbidity in 95.03% of wells after an average incubation time of just 5.72 h (as opposed to the gold standard time of 18–24 h).

Moving forward, we are interested in testing the system with many more species of bacteria, as each is known to shown different growth characteristics and drug resistance profiles. In addition, we would like to add a darkfield monitoring capability, such that the system can more sensitively detect scattered light, leading to detecting growth earlier during incubation. Finally, with more data from a broader distribution of growth, we can redesign the neural network to predict turbidity before it occurs. This can be accomplished by using adding fiber intensities from neighboring wells as inputs to the network. Because neighboring wells contain two-fold dilutions of the same antibiotic, the network can use knowledge that growth has/has not occurred in the neighboring concentrations to predict the probability that growth will occur at a given concentration. This system could also be adapted—or even used directly without modification—for other plate-based tasks. For instance, viral plaque assays have long been used in vaccine development to quantify growth of viruses. These typically require much longer incubation (up to

several days) and the resulting number of plaque forming units (PFU) must be manually counted. We intend to explore viral plaque assays as another application of the system. In addition, fluorescence assays are of great importance for plate-based diagnostics. The optics of the system would of course need to be somewhat modified for a fluorescence test due to the need to excite fluorophores at a particular wavelength and then block that excitation light so that it does not drown out the fluorescent signal. We envision a system where the LED excitation passes through a bandpass filter before hitting the plate, and placing the fibers off to the side of the well, or even at an angle so that the excitation light cannot couple and propagate in the fibers. A bandpass emission filter can be placed after the imaging end of the fiber bundle. Finally, the fiber-based imaging modality could be used in situations beyond plate-based assays. We are working on a wearable glucose monitoring prototype, leveraging work by our collaborators on an implantable phosphorescence lifetime sensor, which correlates with glucose concentration. Using fiber bundle-based imaging through the skin would allow for a compact imaging system that is robust to slight misalignments (e.g. on the wrist) to be used in tandem with the implantable sensor. Because the sensor is being imaged through scattering in the skin, and because intensity of the phosphorescence signal is the only value of concern, the relatively low-resolution imaging achieved by such a fiber bundle is of no concern.

The second framework we have developed is an on-chip, deep learning-enabled spectrometer using plasmonic resonance, which could be scaled arbitrarily to perform imaging spectrometry.[8] More generally, we have developed a toolset for creating chip-scale spectrometers for any desired application. The plasmonic chip is created by molding polydimethylsiloxane (PDMS) using a silicon master fabricated in a cleanroom via electron-beam lithography. Nanostructures of various periodicities and morphologies are created in the silicon. A single master

can be used to fabricate many PDMS chips, which need only undergo deposition of a thin layer of gold (10s of nm) in a cleanroom. The optical setup for the chip-based spectrometer is simply a bare CMOS imager sensor with the spectral encoder chip placed a small distance in front of it (e.g. <1 mm). When illuminated, each region of the chip demonstrates a different transmission spectrum based on the plasmonic resonance in the gold film. Thus, spectral information is encoded in the intensities captured by the image sensor for the different nanostructured regions. The intensity patterns on the chip vary over a wide dynamic range due to constructive/destructive interference, so images are captured at various exposure times to ensure that information can be obtained without over- or undersaturation. These multiple exposure images are combined into a single high dynamic range (HDR) image. Image processing is performed to extract the intensities for each nanopatterned region, which are fed into a fully-connected neural network that outputs the spectrum of the light shone on the chip. Our system was trained and tested using an eight-channel supercontinuum laser. A random number of channels between one and eight were each set to a random wavelength between 480 and 750 nm at a random power to represent arbitrary spectra. Tens of thousands of random spectra were captured, with the first ~80% being used to train/validate the neural network and the last 20% being used for blind testing.

We believe this framework can be used to design spectrometers for various custom applications. Depending on the spectral regime of interest, different nanostructures could be used that have rich spectral features in the desired regime. In addition, because scaling the sensor only requires scaling the one-time e-beam lithographic step (the PDMS molding and gold-coating steps would not be affected), the chip could be scaled up as needed for larger field of view. And if repeating patterns of nanostructures are employed, a type of hyperspectral imaging sensor could be designed, where each “pixel” contains spectral information. Some of the clearest use cases for

this sort of hyperspectral large-FOV imaging are in remote sensing and astrophysics. For instance, monitoring soil/crop health and water content, as well as the makeup of more distant celestial objects could benefit from a more compact, lightweight imaging spectrometer that can be tailored for the spectral regime of the given application.

References

1. Duarte, M. F. et al. Single-pixel imaging via compressive sampling. IEEE Signal Processing Magazine 25, 83–91 (2008).

2. Velten, A. et al. Femto-photography: capturing and visualizing the propagation of light. ACM Trans. Graph. 32, 44:1–44:8 (2013).

3. Wang, Z. et al. Single-shot on-chip spectral sensors based on photonic crystal slabs. Nature Communications 10, 1020 (2019).

4. O'Toole, M., Lindell, D. B. & Wetzstein, G. Confocal non-line-of-sight imaging based on the light-cone transform. Nature 555, 338–341 (2018).

5. Joung, H.-A. et al. Point-of-Care Serodiagnostic Test for Early-Stage Lyme Disease Using a Multiplexed Paper-Based Immunoassay and Machine Learning. ACS Nano 14, 229–240 (2020).

6. Antimicrobial resistance. https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance.

7. Brown, C. et al. Automated, Cost-Effective Optical System for Accelerated Antimicrobial Susceptibility Testing (AST) Using Deep Learning. ACS Photonics 7, 2527–2538 (2020).

8. Brown, C. et al. Neural network-based on-chip spectroscopy using a scalable plasmonic encoder. arXiv:2012.00878 [physics] (2020).

Chapter 2 Automated optical system for accelerated antimicrobial susceptibility testing

This chapter contains material that has been published as follows: Calvin Brown, Derek Tseng, Paige M. K. Larkin, Susan Realegeno, Leanne Mortimer, Arjun Subramonian, Dino Di Carlo, Omai B. Garner, and Aydogan Ozcan. *ACS Photonics* **2020** 7 (9), 2527-2538. DOI: 10.1021/acsphotonics.0c00841

Abstract

Antimicrobial susceptibility testing (AST) is a standard clinical procedure used to quantify antimicrobial resistance (AMR). Currently, the gold standard method requires incubation for 18–24 h and subsequent inspection for growth by a trained medical technologist. We demonstrate an automated, cost-effective optical system that delivers early AST results, minimizing incubation time and eliminating human errors, while remaining compatible with standard phenotypic assay workflow. The system is composed of cost-effective components and eliminates the need for optomechanical scanning. A neural network processes the captured optical intensity information from an array of fiber optic cables to determine whether bacterial growth has occurred in each well of a 96-well microplate. When the system was blindly tested on isolates from 33 patients with *Staphylococcus aureus* infections, 95.03% of all the wells containing growth were correctly identified using our neural network, with an average of 5.72 h of incubation time required to identify growth. 90% of all wells (growth and no-growth) were correctly classified after 7 h, and 95% after 10.5 h. Our deep learning-based optical system met the FDA-defined criteria for essential and categorical agreements for all 14 antibiotics tested after an average of 6.13 h and 6.98 h, respectively. Furthermore, our system met the FDA criteria for major and very major error rates

for 11 of 12 possible drugs after an average of 4.02 h, and 9 of 13 possible drugs after an average of 9.39 h, respectively. This system could enable faster, inexpensive, automated AST, especially in resource-limited settings, helping to mitigate the rise of global AMR.

Introduction

Antimicrobial resistance (AMR) is estimated to cause over 700,000 deaths annually, with 2.8 million cases and 35,000 deaths in the United States alone.[1,2] By 2050, the number of deaths due to AMR is projected to reach as many as 10 million per year.[1] A host of factors are contributing to the global rise in AMR, such as over-prescription and abuse of antibiotics (e.g. for viral infections),[3,4] use of medically important antibiotics in agriculture for e.g. promotion of growth in livestock[5] and prevention of disease in citrus trees,[6] as well as economic and regulatory barriers to the development of new drugs.[7]

One of the most crucial tools to treat patients infected with resistant bacteria, as well as to stem the tide of global AMR, is antimicrobial susceptibility testing (AST). AST is a laboratory procedure used to determine which antibiotics will work most effectively against a given patient's bacterial infection. The gold standard method is broth microdilution (BMD), in which isolated patient bacteria are inoculated in growth medium along with a candidate antibiotic and incubated for at least 18–24 h. BMD is usually performed in a 96-well microplate, with a different antibiotic/concentration combination in each well. Neighboring wells contain successive two-fold dilutions of the same drug. After incubation, each well is inspected visually by a trained medical technologist to determine whether growth has occurred, as indicated by the presence of turbidity in the well. The minimum inhibitory concentration (MIC) for a given antibiotic is defined as the lowest concentration of the drug that successfully prevents bacterial growth. The MIC is used to

determine the categorical susceptibility of the bacteria to the drug (susceptible, intermediate, or resistant) based on concentration cutoffs published by the Clinical & Laboratory Standards Institute (CLSI).[8]

The lengthy incubation time (18–24 h or more) puts patients at risk because in the interim they may be prescribed powerful broad-spectrum antibiotics or antibiotics against which the organism is resistant. The need for a trained expert to manually read the plate strains laboratory resources and inevitably introduces human error/variability. Automated AST systems such as the bioMérieux Vitek 2[9] enable readings much earlier during incubation for certain bacteria and drugs, but these systems are relatively bulky, expensive (due to e.g. optomechanical scanning components and illumination sources), and often require the use of proprietary plates and drug panels, limiting their utility especially in resource-limited settings, where AMR is expected to take the largest toll.[10]

Numerous alternative approaches have been investigated to address the shortcomings of conventional AST. The decreasing cost of whole-genome sequencing (WGS) has made it a potentially-viable option, and it has been shown to agree with BMD for certain bacteria-drug combinations.[11–13] However, the cost remains prohibitive for most labs, even in developed countries, and there is a lack of standardization for AST protocols.[14] In addition, unless all the resistance mechanisms in question for a given sample are linked to genes with well-characterized effects—such as the mecA gene in methicillin-resistant *Staphylococcus aureus* (MRSA)[15]—WGS-based AST will provide an incomplete resistance profile, limiting its applications, especially for *emerging forms of resistance*.[16–18] Due to its ability to enable rapid, low-cost diagnostics using small sample volumes, microfluidic technology has also been investigated for AST. By confining bacteria to microscale channels or droplets, the incubation time required to identify the impact of

antibiotics on bacterial growth can be shortened considerably.[19–27] MICs can be determined straight from positive cultures (without the additional overnight isolation step) in the case of urine samples, but not for more complex samples such as blood or sputum.[26,27] Additionally, these microfluidic approaches generally require new specialized consumables and a scanning microscopy system to monitor the sample during incubation, limiting their viability in resource-limited settings. Pure microscopy-based approaches have also been demonstrated for AST.[28–34] Commercially available systems such as the Pheno (Accelerate Diagnostics)[35] and the oCelloScope (BioSense Solutions)[36] have developed a more compact form factor compared to benchtop microscopes, but still require expensive objective lenses and optomechanical scanning components to read a 96-well plate. These systems also depend on knowledge of specific organism morphologies and growth characteristics, limiting their use to certain types of bacteria. As an alternative, lensfree microscopy[37–39] eliminates the need for objective lenses, thus reducing costs and mitigating the spatial/focal drift these components can cause during time lapse imaging. Lensfree microscopy has been shown to detect bacteria over a wide field of view,[40,41] but would still require mechanical scanning to image an entire well plate. In order to capture the spatial information from a well plate without any expensive objective lenses or scanning components, we have previously demonstrated an optical fiber-based smartphone reader for AST plates *after* the incubation period to determine turbidity results.[42,43] These earlier works did not capture time lapse images of the samples and therefore were aimed to provide end-point readings, after the standard incubation period (e.g., 18–24 h).

In this work, we demonstrate an automated, cost-effective optical system for the *early detection* and *quantification* of resistance in AST using deep learning. The device can be placed directly inside a standard benchtop incubator and automatically monitor growth in all 96 wells of

a standard microplate during incubation. The plate is periodically illuminated by red, green, and blue LEDs, and the transmitted light is relayed by an array of plastic optical fibers beneath the plate to two Raspberry Pi cameras for imaging. The optical fibers enable cost-effective spatial subsampling within each well to detect localized changes during the incubation period. A neural network uses the intensity information from the images to classify each well as either turbid or non-turbid over time. This system eliminates the need to wait 18–24 h or more and does *not* rely on a trained medical technologist for readings as is necessary for conventional AST, while also being compact and cost-effective compared to commercially available automated AST systems.

Our system was blindly tested on 33 unique clinical *Staphylococcus aureus* isolates, using a panel containing varying concentrations of 14 antibiotics. 95.03% of all wells containing growth were correctly identified, with an average of 5.72 h of incubation required to identify growth. 90% of all wells were correctly classified after 7 h, and 95% after 10.5 h. The system met the FDA-defined criteria[44] for essential and categorical agreement for all 14 drugs tested after an average of 6.13 h and 6.98 h, respectively. The system met FDA criteria for major and very major error rates for 11 of 12 possible drugs after an average of 4.02 h, and 9 of 13 possible drugs after an average of 9.39 h, respectively. For each one of the drugs that did not meet the FDA criteria, only a single major or very major error was made. Some of the major and very major errors may also be due to human errors in the ground truth reading. With additional training and testing samples, the FDA criteria could potentially be met for all drugs. This system could enable inexpensive, high-throughput AST in resource-limited settings, helping treat infected patients while curbing the rise of drug-resistant bacteria.

Results

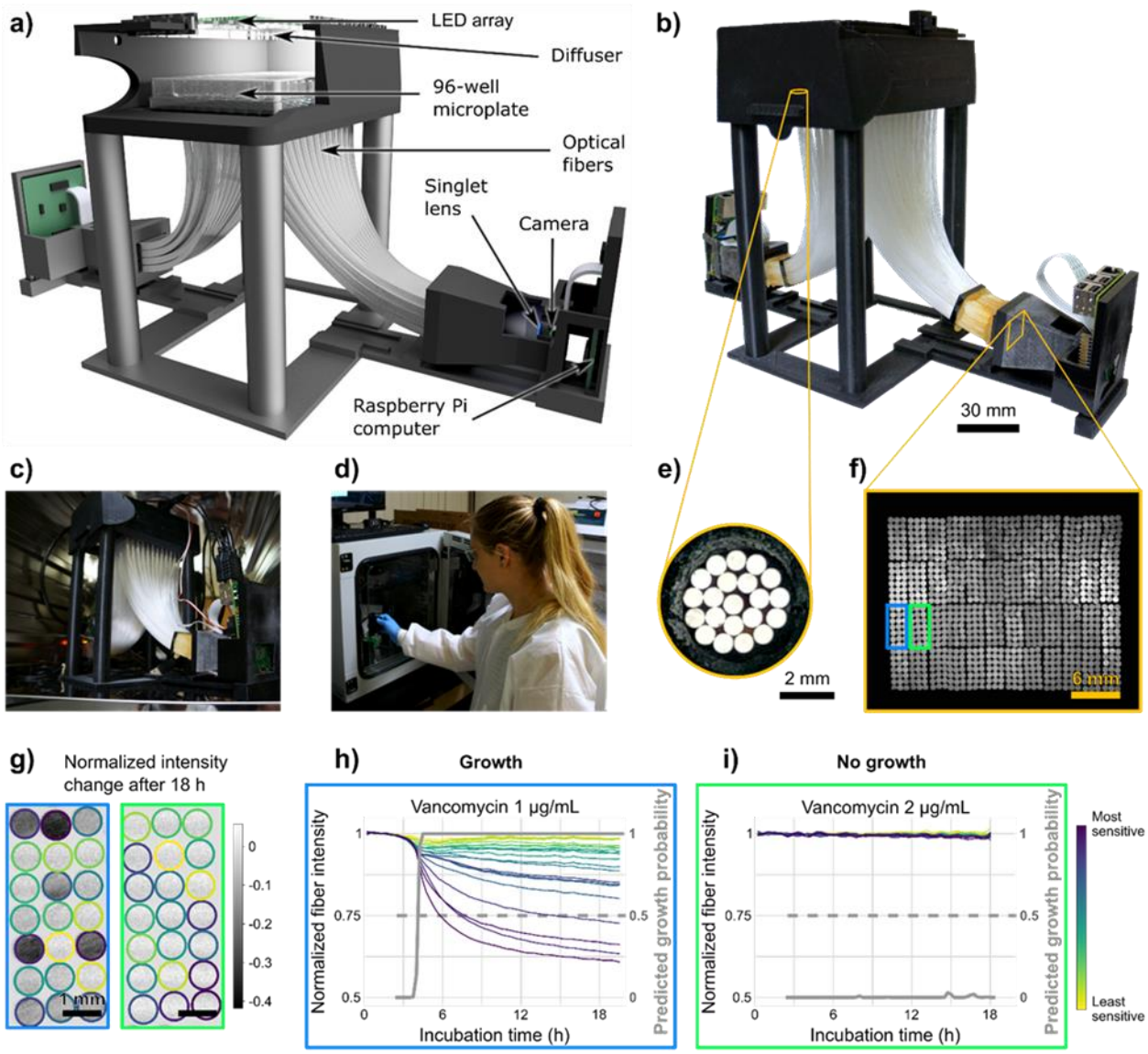


Figure 2.1: Automated antimicrobial susceptibility testing (AST) system overview. (a) Schematic and (b) photo of the device. (c) Device inside an incubator. (d) 96-well plate being loaded. (e) Close-up of 21 fibers under one well. (f) Image of fibers captured by the system. (g) Normalized fiber intensity change after 18 h incubation for two neighboring wells. Fiber intensities and neural network predicted probability of turbidity for (h) a turbid well and (i) a non-turbid well. The colormap corresponds to the random arrangement of fibers in each well. The predicted growth probability on the right axis corresponds to the gray curve in each plot, which is the output of the neural network as a function of the incubation time.

Imaging system design

The AST system (Figure 2.1a,b) is composed of cost-effective components: LEDs, plastic optical fibers, singlet lenses, Raspberry Pi computers and camera modules, and 3D printed housing. It can be placed inside any standard laboratory incubator (Figure 2.1c) and has a slot for the insertion of a standard 96-well microplate loaded with bacterial isolates, growth medium, and candidate antibiotics at various concentrations (Figure 2.1d). Two adjacent 8x8 RGB LED arrays illuminate the entire plate from above, with one LED centered over each well. A plastic diffuser beneath the LEDs ensures spatial uniformity of illumination over the wells, and the brightness of each LED is controlled by pulse width modulation to compensate for the fact that wells near the center of the plate receive more light (due to neighboring LEDs) than those at the edge. Wells containing bacterial growth scatter the incident illumination, while the wells with no growth allow the light to pass through mostly unobstructed. Below each well, a bundle of 21 plastic optical fibers (Figure 2.1e) relays the transmitted light to one of two larger bundles, which are each imaged by the combination of a singlet lens and a CMOS camera connected to a Raspberry Pi computer. A sample image is shown in Figure 2.1f. Images are periodically captured over the course of an 18 h incubation, and examples of fiber intensity changes over time for wells with and without turbidity are shown in Figure 2.1g-i.

In addition to capturing images, the two Raspberry Pis synchronously control the schedule of the illumination and image capture during incubation (Figure 2.2). Every five minutes, the LEDs are turned on and an image is captured, enabling temporal sampling of potential growth while ensuring that the bacteria are not exposed to phototoxic levels of light. The illumination cycles through the three LED colors, so that the time between images of the same color is 15 minutes.

While turbidity measurements are typically performed at a single wavelength (e.g., OD600), our system employed three illumination wavelengths to mitigate any wavelength-dependent absorbance of *S. aureus* in liquid growth medium[45]. A quality control strain of *S. aureus* was run repeatedly in the system to ensure the MICs were in the expected ranges,[46] indicating bacterial growth is not hampered by the periodic illumination (see Table 2.2). The fiber array functions to demagnify the plate area, enabling imaging of all 96 wells *without* any optomechanical scanning components, while maintaining a compact form factor.[42,43] In this case, the fibers provide a demagnification factor of ~7, while capturing spatial information within each well, which is especially important for wells showing weak or atypical growth. The number of fibers per well (21) and the focal length of the singlet lenses (50 mm) were selected to maximize the amount of information captured per well. To address future manufacturability concerns, the locations of each of the 21 fibers within the wells were not manually recorded and tracked during assembly. Instead, the rough position of the fibers within each well (and thus, the information content) was empirically determined *post hoc* from the training data (further detailed below).

The entire device measures 175 × 450 × 192 mm and the cost of the components (including all optics, electronics, and 3D printing) is under $500, which would drop significantly at higher manufacturing volume. Our system easily integrates with a typical clinical workflow, using any standard laboratory incubator and standard 96-well microplates. To operate our system, a user simply inserts a plate (Figure 2.1d), then starts the image acquisition program on one of the Raspberry Pi computers. Our system was successfully operated by five different clinical laboratory personnel.

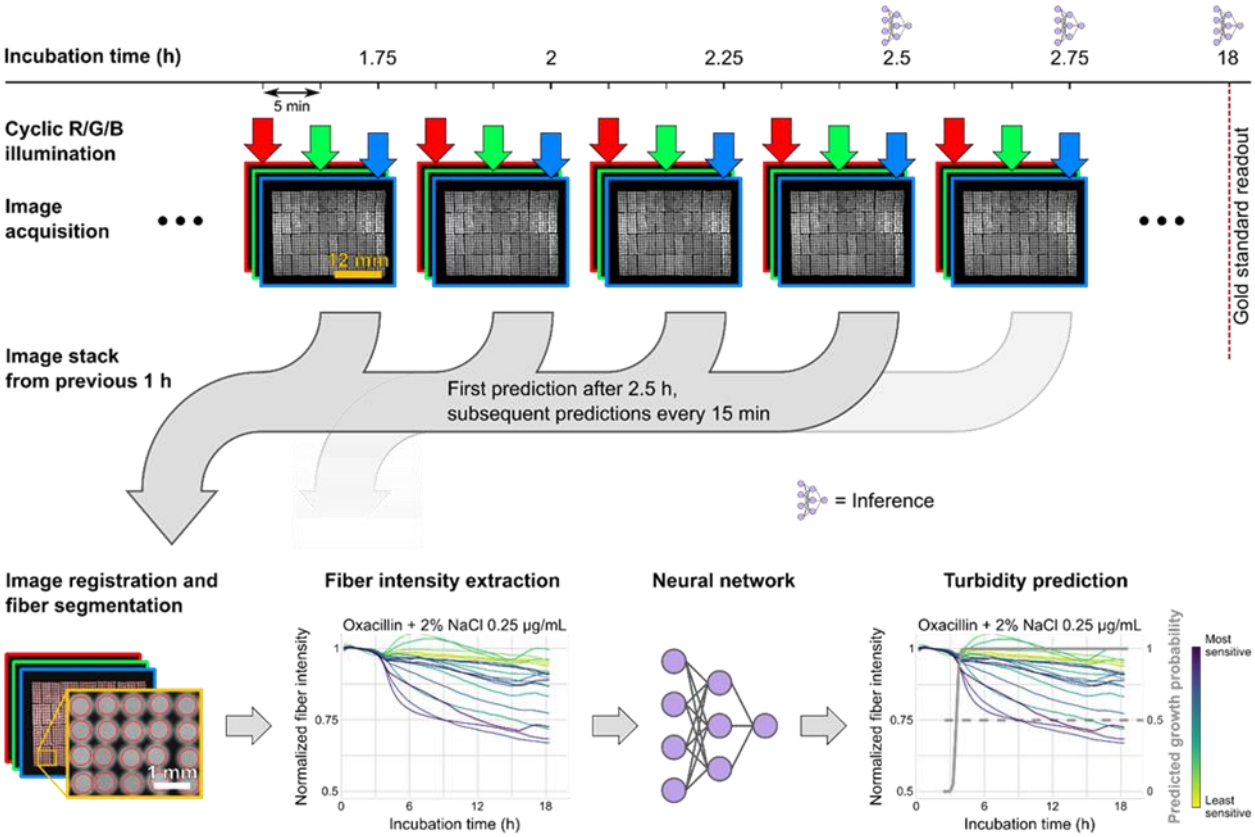


Figure 2.2: Image processing pipeline. Images are captured every 5 min under either red, green, or blue illumination. Images are registered and fiber intensities are extracted. A neural network uses the fiber intensities from the previous 1 h to predict the probability that each well is turbid at the current time point.

Image processing and neural network design

The data processing pipeline is shown in Figure 2.2. For each image, only the pixels corresponding to the illumination color are used (either red, green, or blue from the Bayer color filter array). For each plate, all subsequent images are aligned with the first image of the corresponding color using intensity-based registration, to account for any drift that may occur due to e.g. plate insertion, thermal effects, structural vibrations, etc. The first image of each color is also used to locate each fiber using the circular Hough transform, constrained by prior knowledge of the fiber grid layout. Using these fiber locations, the intensity of each fiber in each image is determined by averaging over a circular mask with a radius of 8 pixels, significantly smaller than

the radius of the fiber to avoid any edge effects due to e.g., defocus from aberrations of the singlet lens or drift over the course of the incubation period. Each temporal fiber intensity is denoised with a 30-minute moving averaging filter and normalized to its average value over the first 10 images during incubation (2.5 h), during which time detectable turbidity is not expected to develop. These preprocessing steps mitigate the effect of fiber intensity variation due to illumination, fiber polishing defects, off-axis effects, etc.

A turbidity prediction is made for each well after each image (every 5 min) starting after 2.5 h of incubation. The normalized fiber intensities for all 3 illumination colors over the previous one hour are fed into a neural network that outputs a predicted probability of turbidity in the well at the current time. This is referred to as the window slicing method in the time series classification literature.[47] *Any value above 0.5 is interpreted as turbid, while values below 0.5 are interpreted as non-turbid*. Examples of fiber intensity plots and blind testing network predictions are shown in Figure 2.3, Figure 2.7, and Figure 2.8. The turbidity classifications are then used to determine the MIC and susceptibility for each drug based on established clinical cutoffs.[8] The neural network comprises 4 fully connected hidden layers of 128 neurons each, and a binary classification output layer (Figure 2.9). Batch normalization and dropout (probability 0.5) were used after each hidden layer to accelerate training and limit overfitting, respectively. The network was trained with the Adam optimizer at a starting learning rate of 1e-3, which was decreased after the validation loss failed to improve for 20 epochs. In total, the network has 83,073 trainable parameters. Note that the network does not employ any spatial convolutional layers because the information contained in the fiber bundle images is not shift-invariant: each fiber corresponds to a fixed region of the plate. This is the reason the extracted fiber intensities—as opposed to images— are used as the input to the network. The network does not receive any prior information about the well, drug, or

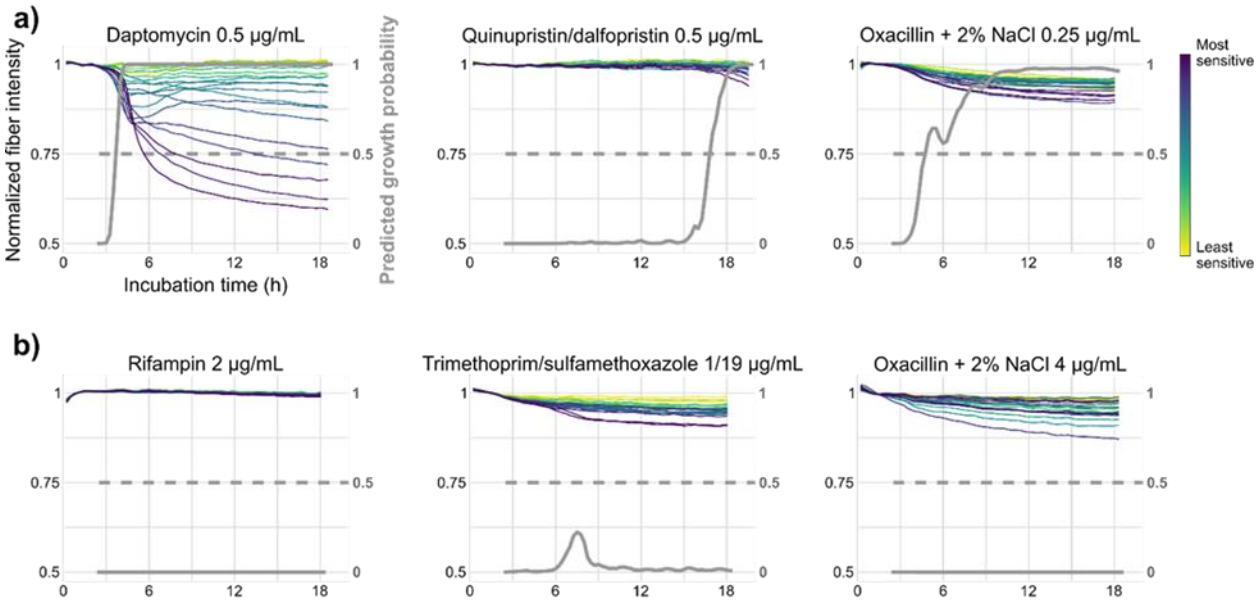


Figure 2.3: Network turbidity predictions on blind testing isolates. Fiber intensities and the panel of neural networks' predicted probability of turbidity on blind testing patient isolates of *Staphylococcus aureus* for ground truth (a) turbid and (b) non-turbid wells.

drug concentration; it makes predictions in a "well-blind" manner, which prevents it from overfitting to the specifics of the plates that were used in the experiments.

Training neural networks via supervised learning requires ground truth labels for every training sample. Because a ground truth reading can only be performed via visual inspection by the trained medical technologist *after* incubation, ground truth labels were only available for the final time point of each patient plate (~18 h). Labels for the training plates at every other image time point during incubation were created manually by inspecting the fiber intensity plots for each well, such as those in Figure 2.3, Figure 2.7, and Figure 2.8. While these labels do not constitute a ground truth, they are the best available proxy and were used to train the network to identify turbidity effectively at an earlier time point within the incubation phase. Additionally, in certain instances where the ground truth label after 18 h disagreed with the manual label, the label was changed for network training. This type of data cleaning is acceptable (and common) for

training/validation data, especially when access to ground truth is not available, but certainly must not be (and was not) employed on blind testing data as it could bias results.

As mentioned previously, the 21-fiber bundles under each well were assembled without precise control of the mapping of each fiber from the well to the image to make the device easier to potentially manufacture in large quantities. In the imaged fiber bundles (Figure 2.1f), the fibers are grouped by well, but the fibers within each well are randomly arranged. For each well, a "fiber order" was determined empirically, using the amount by which the fiber intensities dropped in the presence of turbidity, averaged over the entire training set. The fibers that showed the largest average intensity drop over the course of incubation can be assumed to be near the center of the well because the rounded bottoms of the wells eventually cause strong growth to pool in the center, although growth may *begin* anywhere in the well. Fibers that showed the smallest average intensity drop can be assumed to be near the edges of the well, where even strong growth has a limited effect on transmission at the end of incubation (due to the settling that occurs in the round-bottomed wells for Gram-positive bacteria). The resulting fiber ordering was used for training and blind testing of the network and can be seen in all fiber intensity plots. This fiber ordering ensures that the network learns a general, robust model of turbidity over all wells, instead of overfitting to the individual characteristics of the fibers of each well.

While the total number of training samples, each representing a single time point from a single well from a single patient plate, was large (263,019), the number of clinical isolates from which the training data was gathered was smaller (51). The variability among isolates accounted for much of the diversity of the dataset, both because each isolate had a unique resistance profile and because each plate was incubated on a different day by a set of rotating technicians. To ensure that the learned network model was robust to this isolate-to-isolate variability, we employed nine-

fold cross-validation by randomly splitting the 51 clinical isolate plates into nine subsets and training nine models, where each model was trained on eight of the subsets and validated on the ninth. This process was repeated 50 times and the best model for each subset was selected to form a final composite "panel" of nine neural networks. This cross-validation/composite method (a type of model bagging)[48,49] was employed to ensure that no single isolate plate exerted undue influence on the final model, as could have been the case if only a single validation set were used. Additionally, training in nine folds ensured each model was trained on a large number of dates (e.g., 45), which we found to improve performance on the validation data (Figure 2.10). Training many models was feasible because the number of layers and weights is small compared to many state-of-the-art image classification networks, which contain hundreds of millions of parameters. All training of the models was performed on a desktop computer in TensorFlow without a graphics processing unit (GPU), and in the future, due to improvements in computational power, could even be performed on the Raspberry Pi.

Clinical testing results

All experiments were performed at the UCLA Clinical Microbiology Laboratory by clinical staff, using 96-well microplates containing a Gram-positive antibiotic panel. *Staphylococcus aureus* isolates were prepared to a 0.5 McFarland standard in sterile water and then diluted in Mueller Hinton Broth. The diluted suspension was pipetted into all 96 wells and the plate was inserted into the AST system inside an incubator. 96-well microplates contained antibiotics in powder form pre-loaded into each well of the plate. Bacteria were pipetted into all 96 wells along with growth medium and inserted into the AST system inside an incubator. The plate was removed and ground truth reading was performed after 18–19 h. Initial experiments were

performed on 47 plates, each with a quality control strain of *S. aureus* with a known resistance profile to ensure the system was functioning properly and the bacteria were not experiencing phototoxicity (Table 2.2). The MICs obtained from these control runs showed that the antibiotic linezolid used in the plates did not perform as expected and failed the quality control assessment, so its wells were excluded from the study. Any wells containing antimicrobials that do not have interpretive criteria for *S. aureus* and are not routinely used for clinical management (e.g. ceftriaxone) were also excluded from the study. Next, 51 plates, each containing a *S. aureus* isolate from a unique patient, were used to generate training and validation data for the neural network. 33 additional patient plates were used for blind testing data. The blind testing plates were read after 18–19 h by *two trained technologists* and wells with discrepant readings between the technologists were not used for testing the system (Table 2.3). A single technologist was used to determine the turbidity ground truth for the training/validation data.

Examples of fiber intensities and neural network turbidity predictions for blindly tested patient plates are shown in Figure 2.3, Figure 2.7, and Figure 2.8. The average incubation time required to obtain a correct turbid prediction for each drug is shown in Figure 2.4a. 95.03% of all turbid wells were correctly identified by the network, with the average turbid well requiring just 5.72h of incubation to detect. The system detected turbidity for oxacillin in an average of 4.5 h, while it required 9 h on average for trimethoprim/sulfamethoxazole (Bactrim). Figure 2.4b shows the turbidity detection accuracy over time for all drugs. 90% of all wells were correctly classified after 7 h and 95% after 10.5 h. The MIC and susceptibility predictions for each drug were compared to the FDA-defined criteria for automated AST systems, namely essential agreement (EA), categorical agreement (CA), major error (maj) rate, and very major error (vmj) rate.[44] EA is the percentage of patients for which a drug's predicted MIC is within plus or minus one two-fold

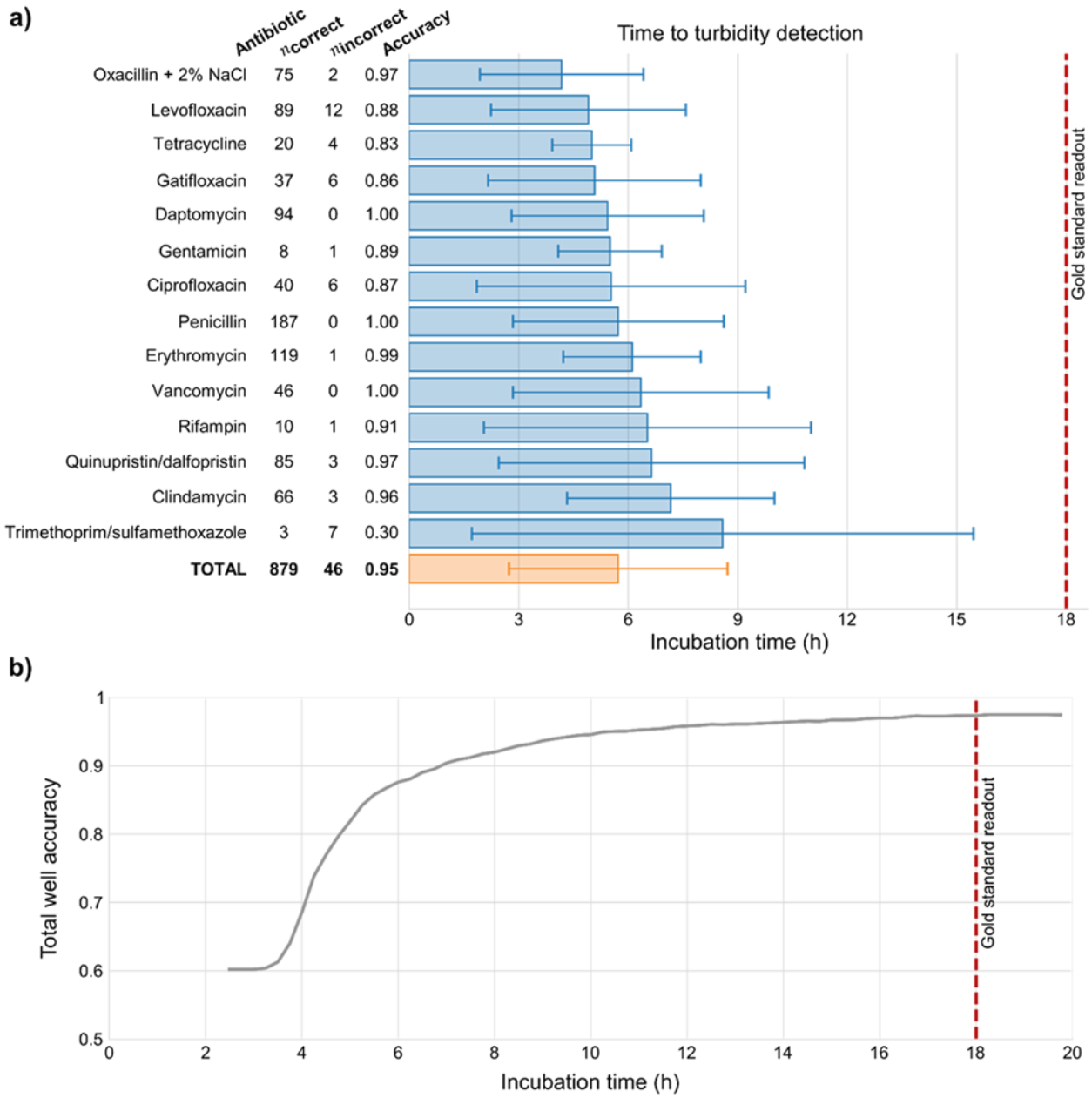


Figure 2.4: Time savings for turbidity detection. (a) Average time required for the panel of neural networks to make a correct turbidity prediction for each drug on blind testing isolates of *Staphylococcus aureus*. 95.03% of all turbid wells were correctly identified by the network, with the average turbid well requiring 5.72 h of incubation for automated detection. (b) Average well accuracy over the course of incubation. 90% of all wells were correctly classified after 7 h, and 95% after 10.5 h of incubation.

dilution of the ground truth. CA is the percentage of patients for which the predicted susceptibility category (susceptible/intermediate/resistant) matches the ground truth. Maj rate is the percentage

of all susceptible infections misclassified as resistant (i.e. false positive) and vmj rate is the percentage of all resistant infections misclassified as susceptible (i.e. false negative). The FDA requires automated AST systems to demonstrate EA and CA greater than 90%, and maj rate and vmj rate of no more than 3%.

Figure 2.5 shows the blind testing results for EA, CA, maj rate, and vmj rate for each of the 14 drugs over the course of incubation. The legend in each plot indicates the number of valid samples in the denominator of the calculation for each drug. For EA/CA this is the number of blind testing patient plates for which there was agreement between the two human readers (out of a possible 33), and for maj/vmj rate it is the number of susceptible and resistant patient infections for which the two readers agreed, respectively. EA and CA surpassed the FDA limit of 90% for all 14 drugs before the end of incubation (18–19 h), in many cases as early as 4–6 h. Note that EA/CA began near 0% at the beginning of incubation for drugs against which growth/resistance was common such as daptomycin, whereas EA/CA began higher for drugs against which growth/resistance was rare, such as rifampin.

Maj rate remained below the FDA limit of 3% for 11 of 12 possible drugs, and the drug for which it exceeded 3% (trimethoprim/sulfamethoxazole or Bactrim) was due to a single major error. It was not possible to calculate maj rate for two drugs (gatifloxacin and erythromycin) because resistance to these drugs was not observed for any clinical isolates. The maj rate for each of the other 12 drugs is plotted, but those that never move above 0% obscure one another. Vmj rate dropped below the FDA-permitted maximum of 3% before the end of incubation for 9 of 13 possible drugs. Again, the four drugs for which the system did not meet the FDA limit (levofloxacin, ciprofloxacin, trimethoprim/sulfamethoxazole, and quinupristin/dalfopristin) each

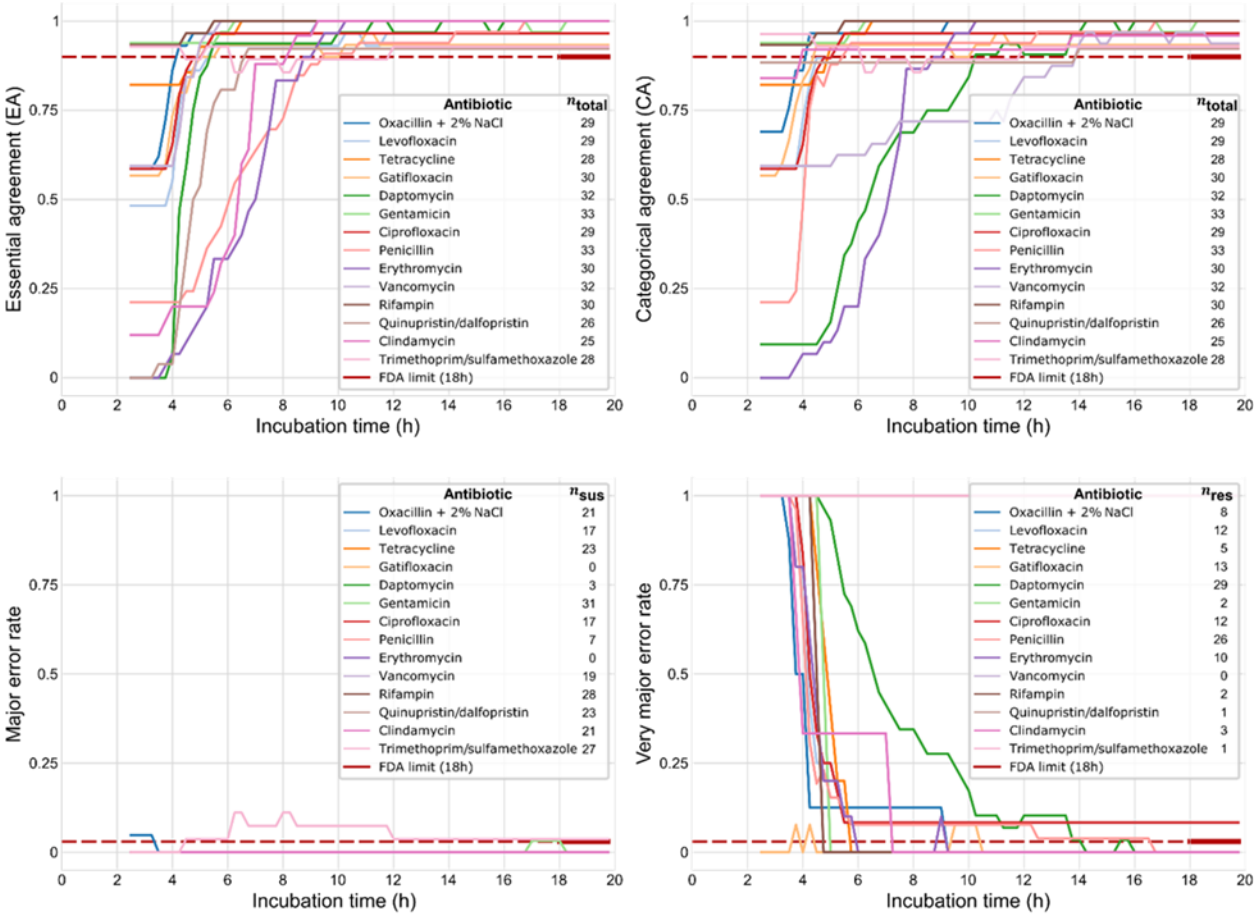


Figure 2.5: Performance metrics over time for each antimicrobial agent. Essential agreement (EA), categorical agreement (CA), major error (maj) rate, and very major error (vmj) rate as a function of the incubation time for different antibiotics on blind testing isolates of *Staphylococcus aureus*. The second column in each plot legend indicates the number of samples for the corresponding metric (total number of valid samples for EA/CA, number of susceptible samples for maj rate, and number of resistant samples for vmj rate). EA and CA surpass the FDA limit of 90% for all 14 drugs before the end of incubation. Maj rate remained below the FDA-permitted maximum of 3% for 11 of 12 possible drugs and vmj rate dropped below the FDA maximum of 3% before the end of incubation for 9 of 13 possible drugs.

experienced only a single very major error. It was not possible to calculate vmj rate for vancomycin because no clinical isolates exhibited resistance to it. Using the data from Figure 2.5, the incubation times required to meet/surpass the FDA limits for EA/CA/maj/vmj rate are listed in Table 2.1. The system met the FDA-defined criteria for EA/CA for all 14 drugs after an average of 6.13 h and 6.98 h, respectively. The system met FDA criteria for major and very major error rates for 11 of

12 possible drugs after an average of 4.02 h, and 9 of 13 possible drugs after an average of 9.39 h, respectively. These results are in line with the performance on the validation data, demonstrating that the panel of networks is not overfit (Figure 2.11, Figure 2.12, and Table 2.4).

Drug	Essential Agreement	Categorical Agreement	n_{total}	Major error rate	$n_{\text{susceptible}}$	Very major error rate	$n_{\text{resistant}}$
Oxacillin + 2% NaCl	4.25	4.25	29	3.5	21	9.25	8
Levofloxacin	5.5	5.25	29	2.5	17	N/A	12
Tetracycline	5	5	28	2.5	23	5.75	5
Gatifloxacin	5	4.5	30	-	0	10.5	13
Daptomycin	5.5	10.25	32	2.5	3	16	29
Gentamicin	2.5	2.5	33	18.25	31	5	2
Ciprofloxacin	5.25	5.25	29	2.5	17	N/A	12
Penicillin	9.5	5.5	33	2.5	7	16.75	26
Erythromycin	8.75	8.75	30	-	0	9.25	10
Vancomycin	5	14	32	2.5	19	-	0
Rifampin	2.5	2.5	30	2.5	28	4.75	2
Quinupristin/ dalfopristin	6.75	14	26	2.5	23	N/A	1
Clindamycin	8.25	4	25	2.5	21	7.25	3
Trimethoprim/ sulfamethoxazole	12	12	28	N/A	27	N/A	1
AVERAGE	**6.13**	**6.98**		**4.02**		**9.39**	

Table 2.1: Incubation time (h) required to meet FDA criteria by drug for blind testing isolates of *S. aureus*.

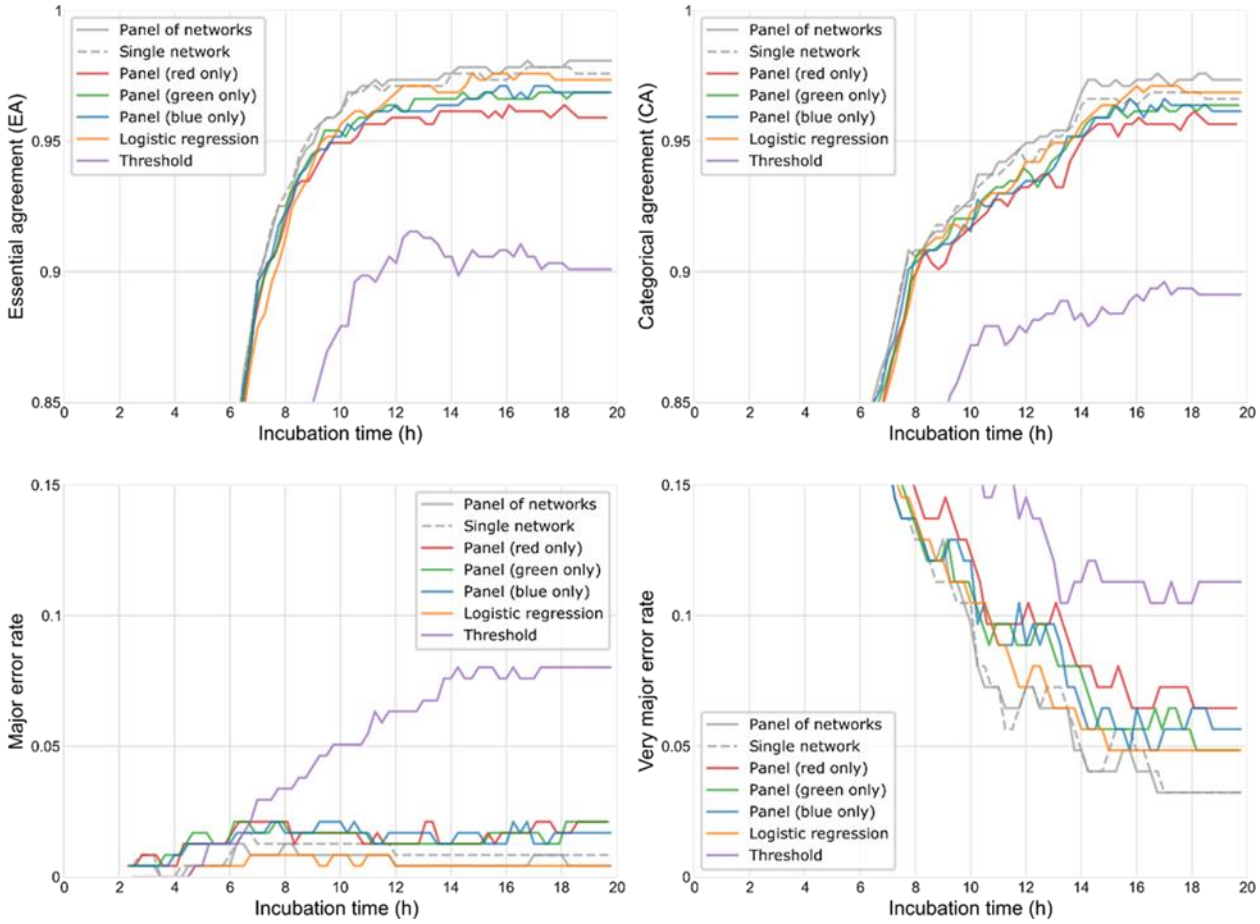


Figure 2.6: Aggregate performance metrics for different computational models. Essential agreement (EA), categorical agreement (CA), major error (maj) rate, and very major error (vmj) rate averaged over all the drugs for various models for blind testing data. "Panel of networks" refers to the panel of nine neural networks trained via cross-validation. "Single network" is only the best of the nine networks (by validation loss). "Panel (red/green/blue)" is a panel of networks that only uses images of the specified color. "Threshold" is a simple threshold-based approach in which a well is classified as turbid if fiber intensities fall below a specified threshold (see Methods section).

To investigate the performance benefits from using the composite panel of neural networks, Figure 2.6 shows the EA/CA/maj/vmj rates for the blind testing patient data averaged over all 14 drugs for several alternative approaches. The curve labeled "panel of networks" corresponds to the composite panel of neural networks, whereas "single network" refers to the best individual network (by validation loss) from the panel. Three additional panels of networks were also generated (from a sample size of 50 nine-fold cross-validations as before) using only images of a single illumination color (red, green, or blue). Finally, a logistic regression model was tested, as

well as a simple threshold-based model, which classifies the well as turbid if at least two of the three most recent images have at least one fiber intensity lower than the threshold of 0.8876. This threshold value was determined by optimizing accuracy over the training patient plates.

The panel of networks showed the best performance across all four metrics, but there was only a slight penalty in EA and maj rate by using a single network. The networks that used only images of a single color fared considerably worse than the network or panel of networks using all three colors. Among the three colors we would not expect a large difference, but green did perform the best, possibly since it used twice the number of pixels per fiber due to the Bayer filter array on the CMOS image sensor. The logistic regression performed better than the single-color networks, but not as well as the three-color network or panel of networks. The threshold-based simple approach was the worst performer, with under 90% CA and over 10% vmj rate. Figure 2.7 and Figure 2.8 show many examples of wells with predictions from the panel of networks, logistic regression, and threshold approach, demonstrating where the simpler models both failed to identify weak growth and falsely identified growth in non-turbid wells. Figure 2.13a,b show additional fiber intensities and network predictions for instances of a "skipped" well and wells with bubbles, respectively. These are well-known phenomena in AST and the network gave correct turbidity predictions in each case.

Discussion

Our system demonstrates the ability to detect turbidity and quantify resistance much sooner than the gold standard method, which requires at least 18–24 h. On the blind testing data, the system made only one major error and four very major errors across all 33 patients with 14 drugs each. The fiber intensities and network predictions for the wells corresponding to each of these

errors are shown in Figure 2.14–Figure 2.16 along with an image of the wells captured with a smartphone camera at the end of incubation. Each drug for which our system exceeded the FDA-defined limit of 3% for maj or vmj rate only experienced a single error. With additional testing samples, the maj/vmj rates may drop below 3% for all drugs. In addition, the FDA defines major and very major errors as misclassification of a susceptible/resistant organism as resistant/susceptible. However, the one major error and one of the four very major errors from the network's predictions did not include a predicted susceptibility because the predicted MIC was undefined (known as a "skipped well"). We report these instances as major/very major errors, but they can be thought of as inconclusive results, for which a human could be notified to read the MIC manually or decide to repeat the test.

The performance of the system demonstrates the potential to enable automated, cost-effective susceptibility testing with early results in resource-limited laboratories. Unlike the gold standard BMD method, our system does not require a full 18–24 h incubation or a trained technologist for plate readout and, unlike microscopy-based solutions, it requires no mechanical scanning components or bulky, costly hardware. Cost and access to trained personnel are primary factors that currently limit the reach of AST in developing regions. Our system also uses standard 96-well plates, which would allow it to more rapidly integrate with typical clinical workflow. It should be noted that other cost-effective optical monitoring systems could be imagined (e.g. a single large-core fiber under each well or an array of photodiodes underneath the entire well plate), but these would sacrifice the spatial information obtained by the fibers within each well, which is necessary to accurately identify turbidity, as evidenced by the various fiber intensity patterns shown in e.g., Figure 2.3, Figure 2.7, and Figure 2.8. While the fibers add relative complexity to the system, the manufacturing process was simplified because the approximate layout of the fiber

bundles under each well was determined empirically *post hoc* as detailed earlier. Because the system autonomously captures images during incubation without the need to remove the plate from the incubator, early results for drugs showing strong resistance can be sent to the physician as soon as they are available, while the device continues to monitor growth in the wells with the remaining drugs. Due to the phenotypic nature of our sensing mechanism, we believe it can be extended to almost any type of bacteria, or other plate-based tasks such as enzyme-linked immunosorbent assays (ELISA), virus quantification assays, and culture samples. The fiber-based subsampling of the wells could enable the streamlining of daily laboratory tasks with robust, automated readout in a compact form factor.

From Figure 2.6, it is clear that the panel of neural networks gave the best performance on the blind testing data, demonstrating an ability to discern nuanced patterns in fiber intensities. While a desktop computer was used to train the panel of networks, due to the rapidly decreasing cost of computation in embedded systems, future training could be performed on the Raspberry Pi or other compact device. In addition, a single network showed only a modest drop in performance, which could shorten computation time. Because the networks were not given knowledge of the well, drug, or concentration when making predictions, they also learned a model of turbidity that is quite general, instead of overfitting to the specifics of the plate or drugs used in the experiments. Because the second technologist who made ground truth readings for the testing data was not used for the training/validation data, the system demonstrated an ability to generalize beyond the specific patterns of an individual human reader.

Conclusion

The presented system demonstrates the ability to conduct AST much faster than the gold standard method of incubation for 18–24 h followed by visual inspection. The time savings is critical to ensuring patients receive the most effective, targeted antibiotics and to limit the global rise in antimicrobial resistance. Our system also removes the need for a trained medical technologist and integrates with the standard clinical workflow using an incubator and 96-well microplates. The system is cost-effective due to the use of off-the-shelf components and could be particularly suited to resource-limited laboratories in developing regions, where antimicrobial resistance is predicted to cause the most deaths and access to trained personnel is limited.

Methods

Imaging system

The AST system illumination is composed of two 8x8 arrays of individually addressable RGB LEDs (Adafruit Industries) whose pulse width modulation brightness is set by 2 Trinket microcontrollers (Adafruit Industries). The system contains 2016 0.75 mm plastic optical fibers (CK-30, Industrial Fiber Optics Inc.), which were epoxied and polished with a handheld polishing tool. Ferrules of appropriate dimensions were 3D printed for both the input and output bundles so that the fibers could be inserted easily and would naturally take the proper arrangement without precision handling (circular for the input end and rectangular for the output end, as shown in Figure 2.1e and f, respectively). Fiber bundles were assembled and epoxied one well at a time, then these well bundles were further combined using a larger 3D printed rectangular ferrule to create two imaging bundles, each containing fibers for half of the 96-well plate (Figure 2.1f). Each 48-well bundle is imaged by the combination of a 10.0 mm diameter × 50.0 mm focal length plano-convex lens (Edmund Optics) and a Raspberry Pi Camera Module V2 (Newark) with 1.12 µm × 1.12 µm

pixel size. The two cameras are controlled by two Raspberry Pi 3 Model B computers. Images are captured in raw 10-bit format at 8.1 MP.

Image processing and neural network

Image processing was performed in MATLAB (MathWorks) and neural network training/testing was performed in Python using TensorFlow 1.14 (Google). The logistic regression model was created in Python with the scikit-learn library (David Cournapeau), using the saga solver, the elasticnet penalty with an L1 ratio of 0.9, and an inverse regularization strength of C = 2.

Clinical testing

All experiments were performed at the UCLA Clinical Microbiology Laboratory. The AST system was placed inside a 2-cubic foot incubator (Binder) for the duration of the experiments. Initial experiments were performed using the ATCC43300 strain of MRSA in 47 plates. Confirmed clinical *S. aureus* isolates collected at the UCLA Clinical Microbiology Laboratory were tested on the platform for the remainder of the experiments. *S. aureus* isolates were prepared to a 0.5 McFarland standard in sterile water and 50 µL of this suspension was transferred into 11 mL of Mueller Hinton Broth. The dilution was inoculated into 96-well microplates (100 µL of bacterial suspension per well) containing a commercially available Gram-positive antibiotic panel (Sensititre Gram Positive MIC plates, ThermoFisher Scientific) shown in Table 2.5. Following bacterial inoculation, single plates were loaded into the incubator for 18–19 h. At the end of incubation, plates were removed and turbidity was manually assessed by trained personnel. For training/validation data (51 clinical plates), plates were read by a single reader. For testing data (33 clinical plates), plates were read by two readers to assess and mitigate interpersonal variances

among readers. MIC was determined by identification of the first well without turbidity for increasing drug concentrations. Interpretation of susceptibility was determined in accordance to Clinical & Laboratory Institute Standards 2019.[8]

Appendix

Supplementary figures

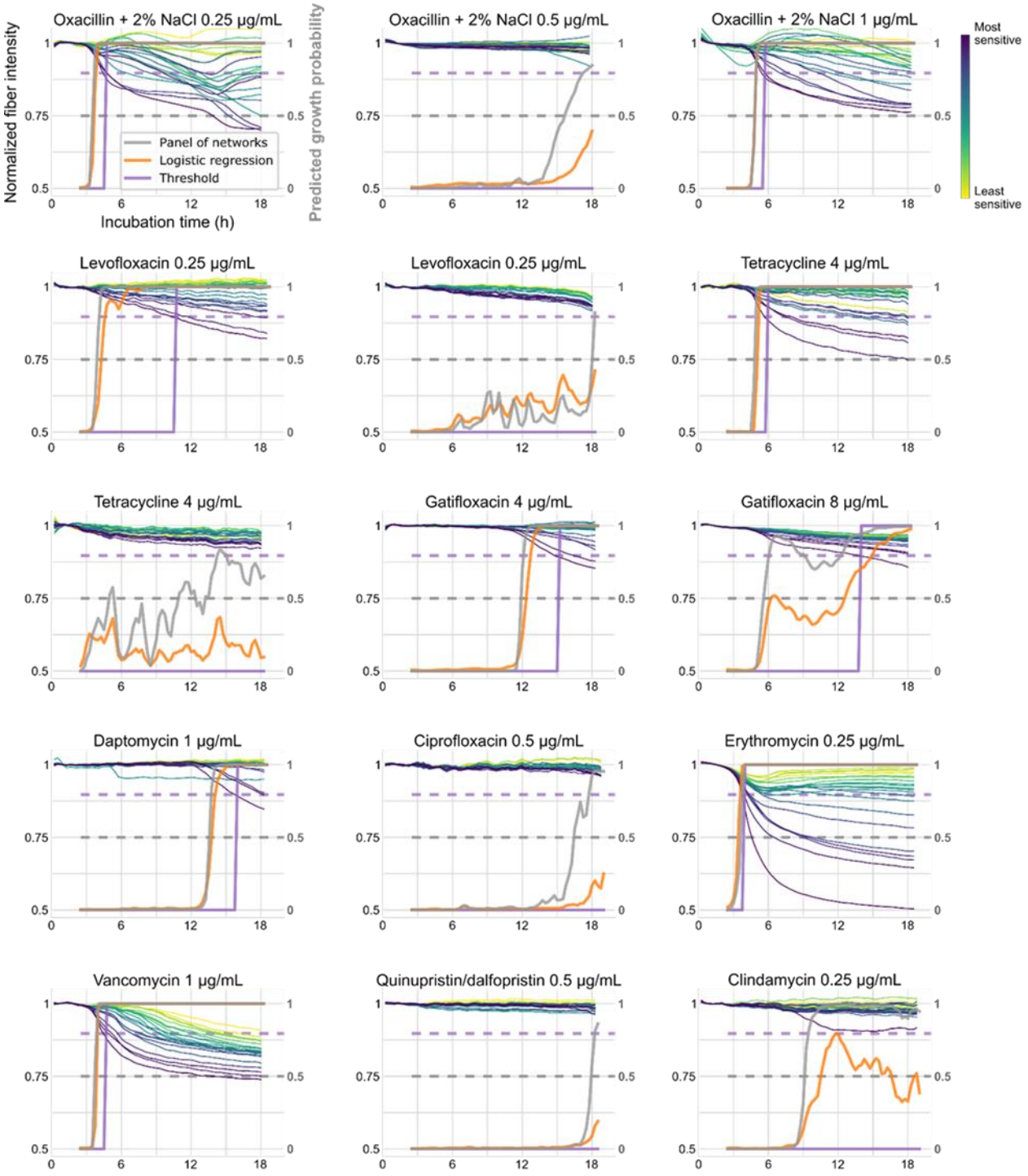

Oxacillin + 2% NaCl 0.25 µg/mL
Oxacillin + 2% NaCl 0.5 µg/mL
Oxacillin + 2% NaCl 1 µg/mL
Normalized fiber intensity
Predicted growth probability
Incubation time (h)
Panel of networks
Logistic regression
Threshold
Most sensitive
Least sensitive
Levofloxacin 0.25 µg/mL
Levofloxacin 0.25 µg/mL
Tetracycline 4 µg/mL
Tetracycline 4 µg/mL
Gatifloxacin 4 µg/mL
Gatifloxacin 8 µg/mL
Daptomycin 1 µg/mL
Ciprofloxacin 0.5 µg/mL
Erythromycin 0.25 µg/mL
Vancomycin 1 µg/mL
Quinupristin/dalfopristin 0.5 µg/mL
Clindamycin 0.25 µg/mL

Figure 2.7: Further examples of model predictions for turbid wells. Fiber intensities for *turbid* wells and our automated turbidity predictions from the panel of neural networks (gray), logistic regression (orange), and simple threshold-based approach (purple). The signature of turbidity varies considerably from well to well, often fooling the logistic regression and threshold approaches, especially for weak growth signals. The panel of neural networks generally provides the best accuracy and the most time savings by reaching more than 50% probability of turbidity (gray dashed line) after a shorter incubation time than the other models. The threshold-based model predicts a well to be turbid when at least one fiber drops below 0.8876 intensity (purple dashed line) in two of the previous three images.

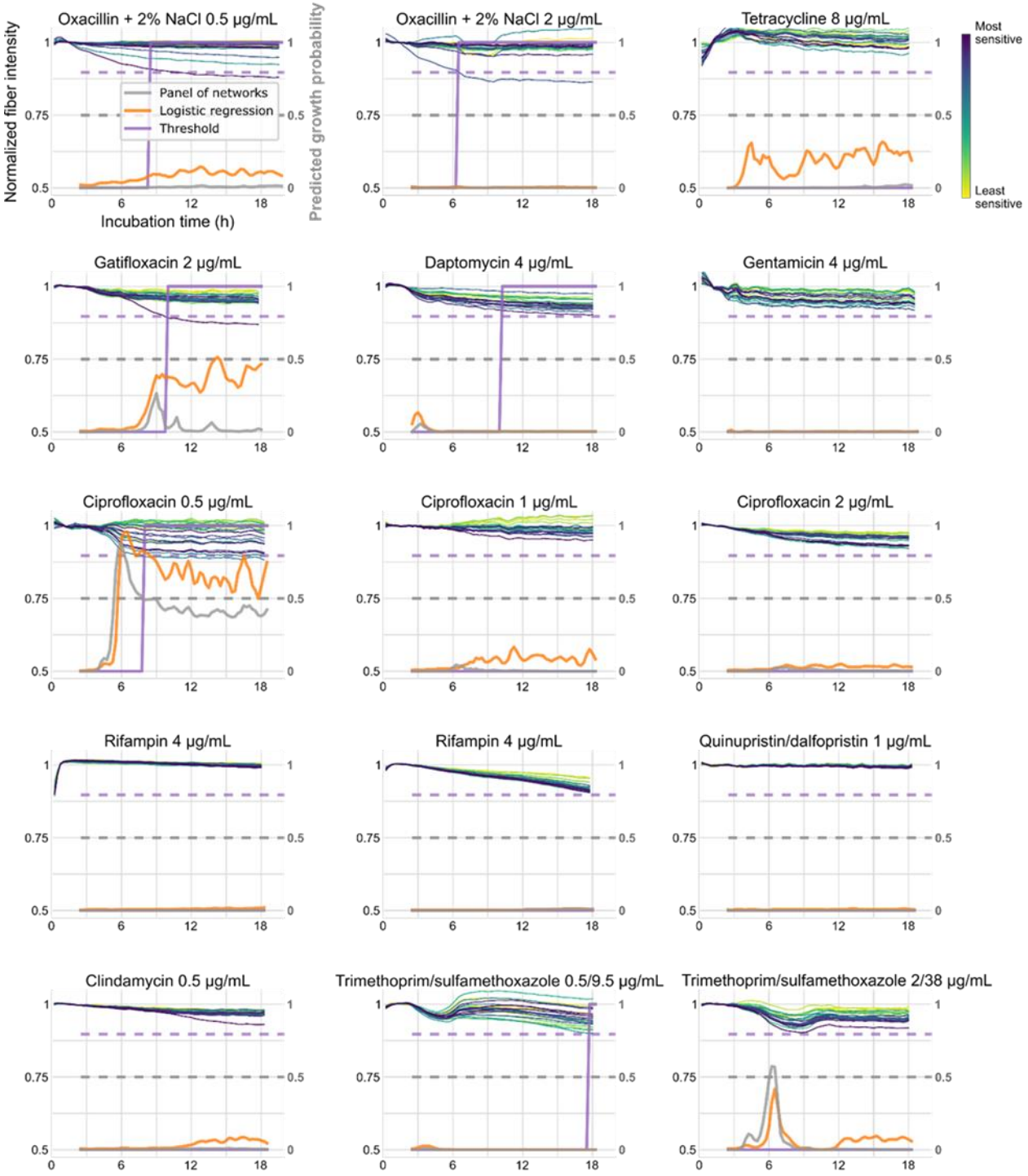

Oxacillin + 2% NaCl 0.5 µg/mL
Oxacillin + 2% NaCl 2 µg/mL
Tetracycline 8 µg/mL
Normalized fiber intensity
Predicted growth probability
Incubation time (h)
Panel of networks
Logistic regression
Threshold
Most sensitive
Least sensitive
Gatifloxacin 2 µg/mL
Daptomycin 4 µg/mL
Gentamicin 4 µg/mL
Ciprofloxacin 0.5 µg/mL
Ciprofloxacin 1 µg/mL
Ciprofloxacin 2 µg/mL
Rifampin 4 µg/mL
Rifampin 4 µg/mL
Quinupristin/dalfopristin 1 µg/mL
Clindamycin 0.5 µg/mL
Trimethoprim/sulfamethoxazole 0.5/9.5 µg/mL
Trimethoprim/sulfamethoxazole 2/38 µg/mL

Figure 2.8: Further examples of model predictions for non-turbid wells. Fiber intensities for *non-turbid* wells and our automated turbidity predictions from the panel of neural networks (gray), logistic regression (orange), and simple threshold-based approach (purple). The signature of non-turbid well varies considerably due to drift, condensation, bubble formation, and settling, often fooling the logistic regression and threshold approaches. The panel of neural networks generally provides the best accuracy, remaining *below* 50% probability of turbidity (gray dashed line) throughout incubation. The threshold-based model predicts a well to be turbid when at least one fiber drops below 0.8876 intensity (purple dashed line) in two of the previous three images.

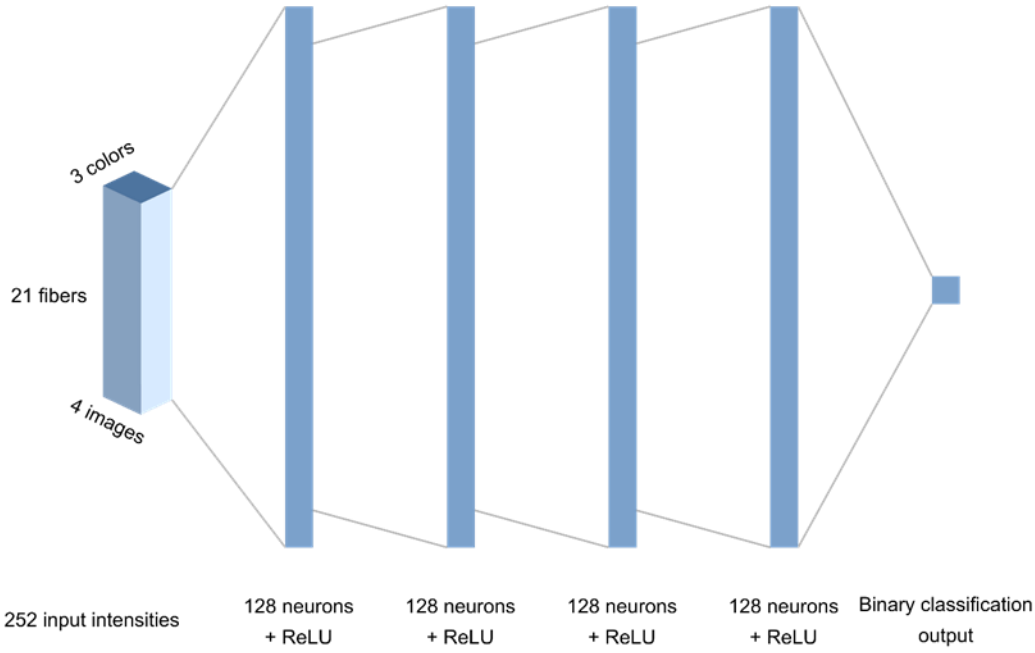


Figure 2.9: Neural network architecture. All layers are fully connected (no convolutional layers). The network was trained with additional batch normalization and dropout layers after each 128-neuron hidden layer.

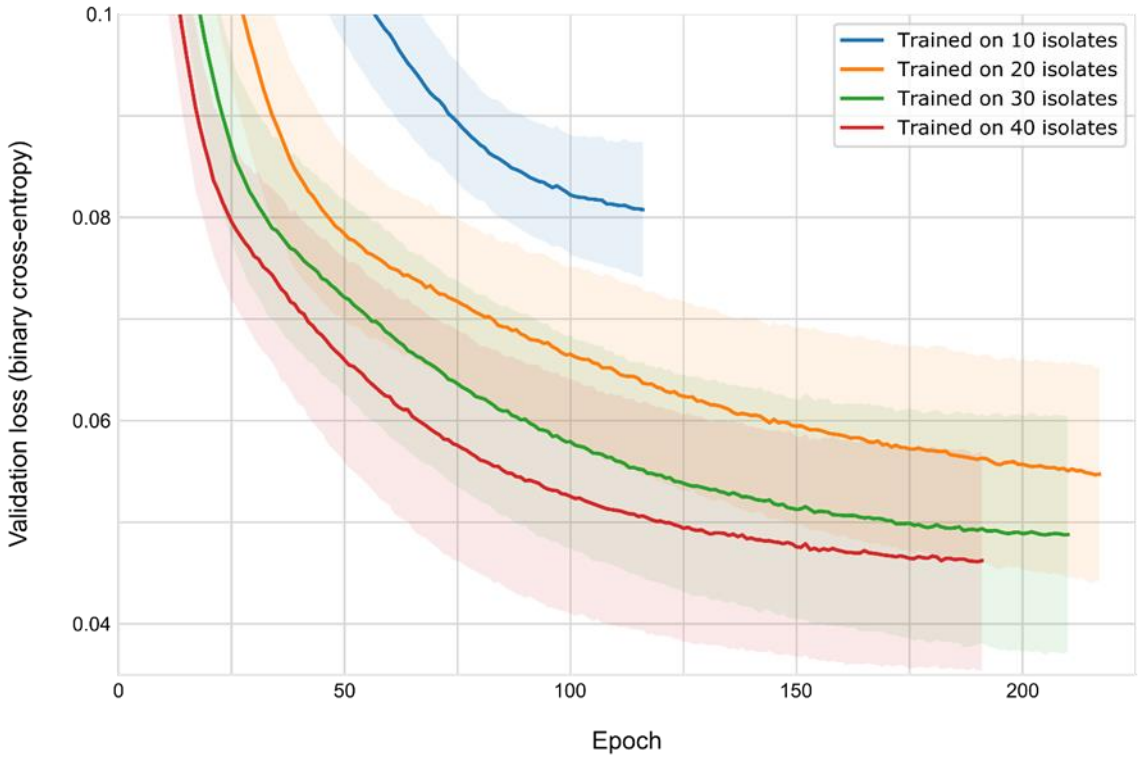


Figure 2.10: Neural network validation loss as a function of the epoch number during training. Standard deviations were calculated over 10 repetitions for each curve. The validation data was the same for all four networks plotted, but the size of the training data ranged from 10 to 40 plates containing patient isolates of *S. aureus*. With additional training data, the minimum validation loss that is achieved decreases.

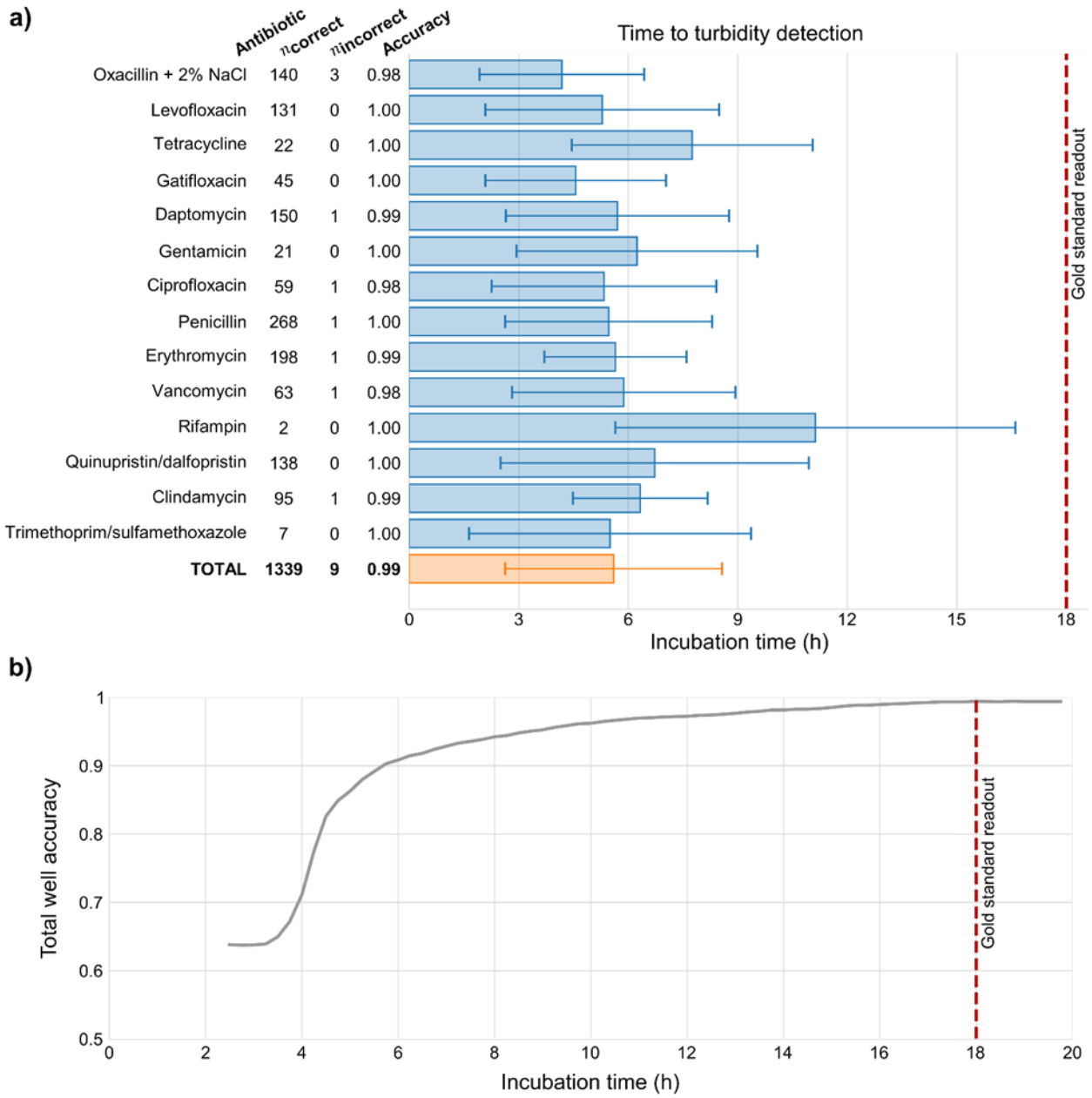


Figure 2.11: Time savings for turbidity detection on validation isolates. (a) Average time required for the panel of neural networks to make a correct turbidity prediction for each drug on validation isolates of *Staphylococcus aureus*. 99.33% of all turbid wells were correctly identified by the network, with the average turbid well requiring 5.60 h of incubation to detect. (b) Average well accuracy over the course of incubation. 90% of all wells were correctly classified after 5.75 h, and 95% after 8.75 h.

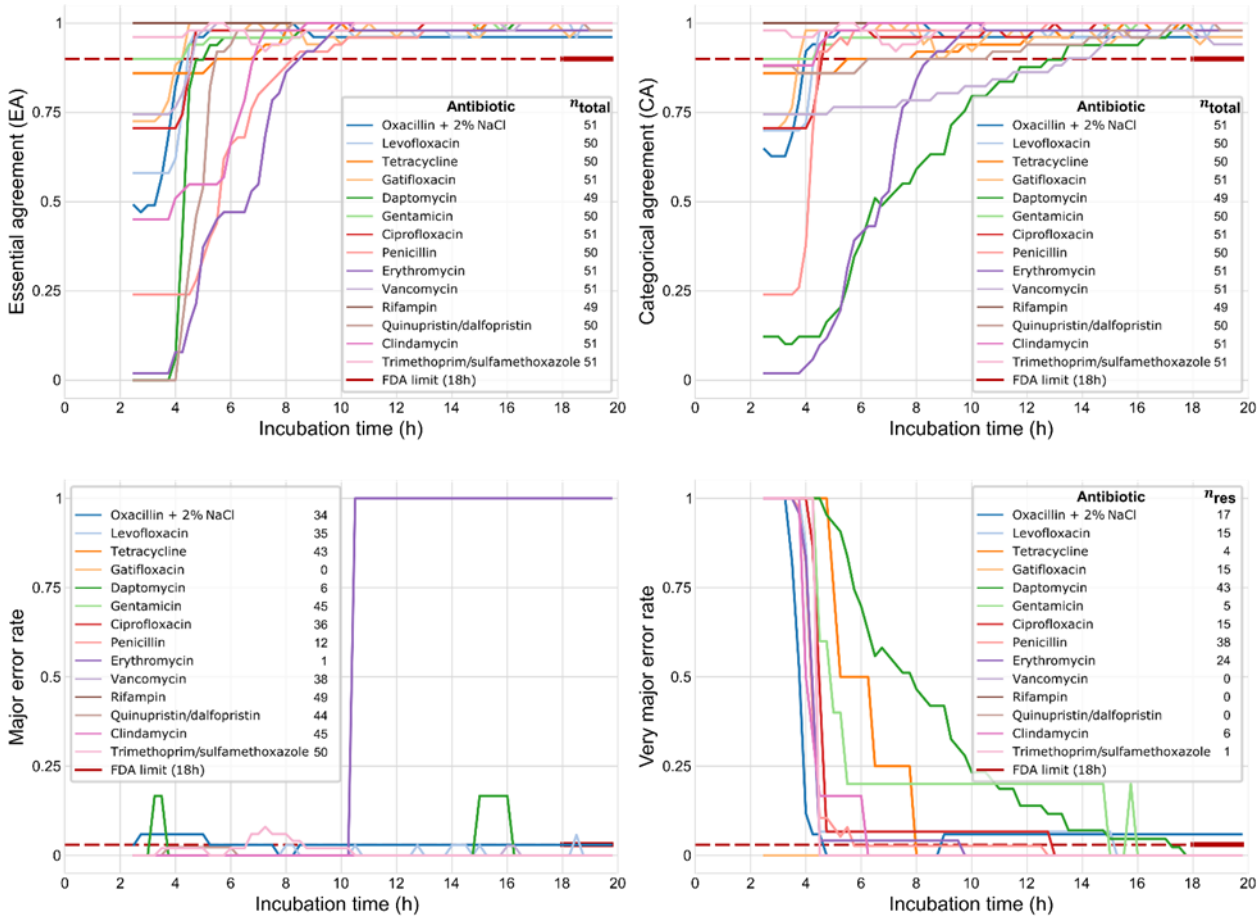


Figure 2.12: Aggregate performance metrics over time on validation isolates for each drug. Essential agreement (EA), categorical agreement (CA), major error (maj) rate, and very major error (vmj) rate as a function of different antibiotics over the course of incubation for validation isolates of *Staphylococcus aureus*. The second column in each plot legend indicates the number of samples for the corresponding metric (total number of valid samples for EA/CA, number of susceptible samples for maj rate, and number of resistant samples for vmj rate). EA and CA surpass the FDA limit of 90% for all 14 drugs well before the end of the incubation period. Maj rate remained below the FDA maximum of 3% for 12 of 13 possible drugs and vmj rate dropped below the FDA maximum of 3% before the end of incubation for 10 of 11 possible drugs.

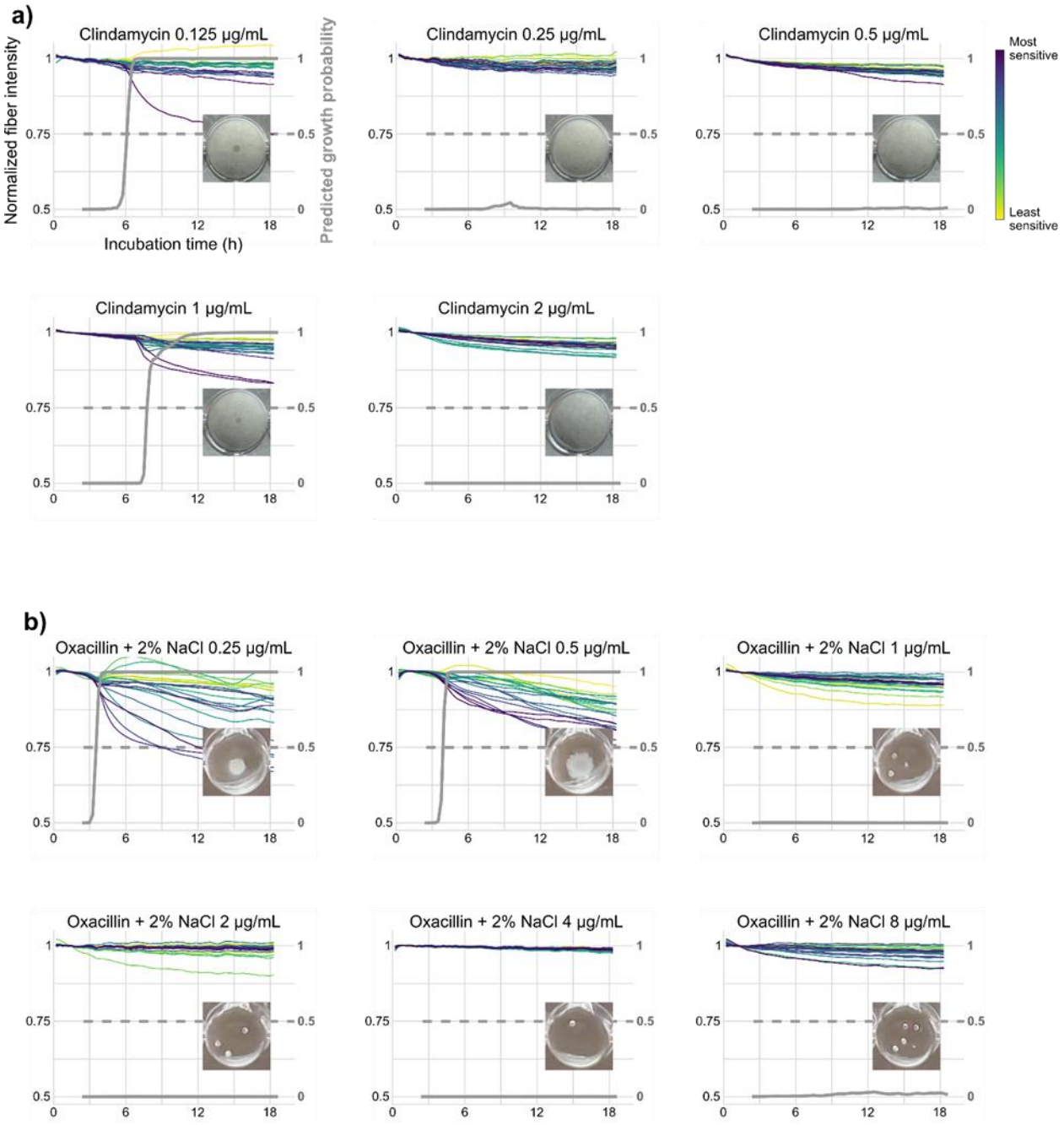


Figure 2.13: Examples of drugs with unusual behavior. (a) Drug for which growth occurs at a *higher* concentration than wells with no growth. "Skipped wells" of this nature may be due to problems with the drug or pipetting. The panel of networks correctly predicted each well and did not return an MIC as it is undefined. (b) Drug for which many bubbles formed in the wells. This unpredictable bubble formation can cause drops in fiber intensities, but *the panel of networks correctly ignores the bubble artifacts when making its predictions*.

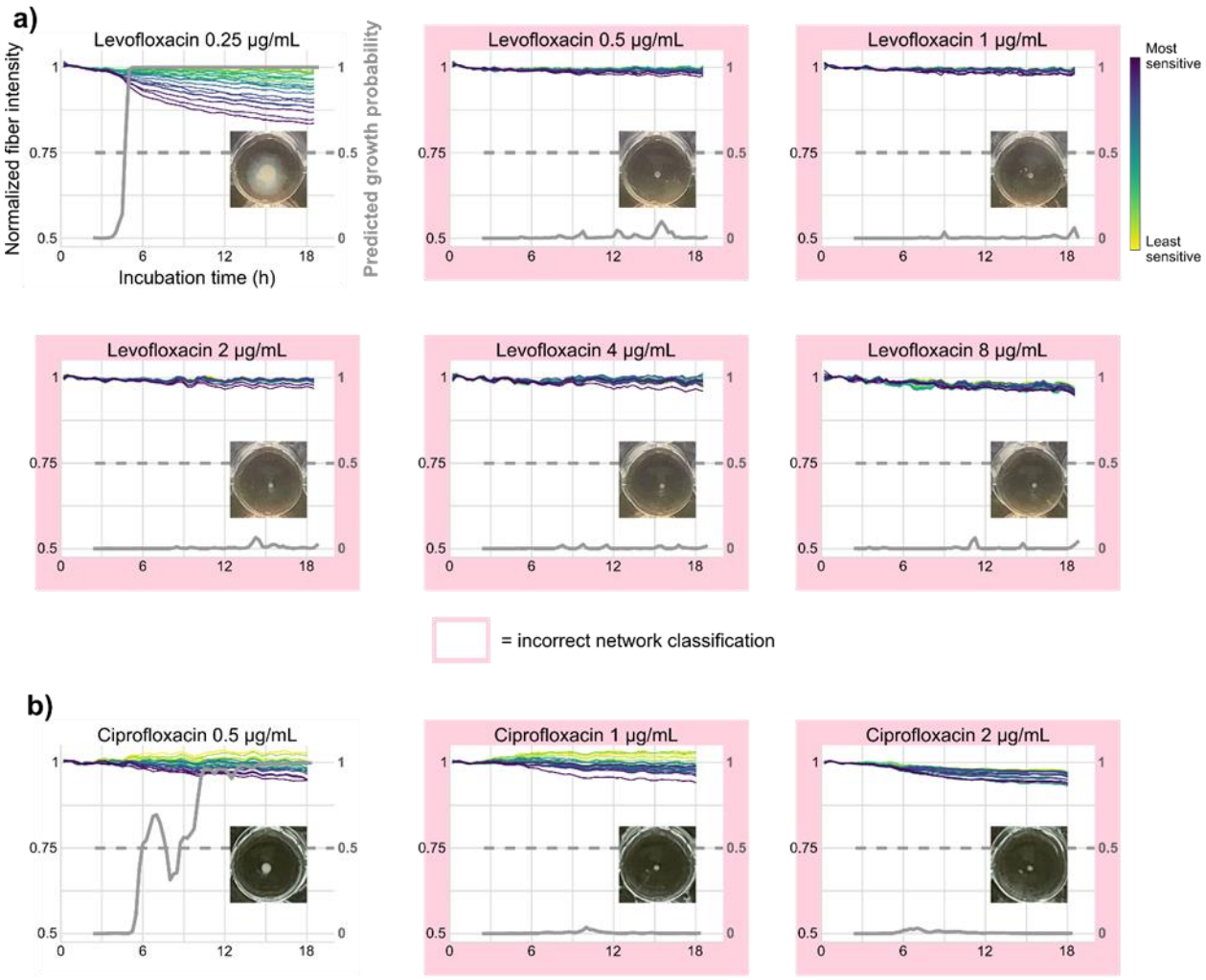


Figure 2.14: Examples of two drugs for which the panel of neural networks made very major errors. Wells for which the panel made an incorrect turbidity prediction at the end of incubation have pink shading. Images of the wells captured by a smartphone are included in the inset of each plot. (a) Drug for which every well was labeled turbid by the ground truth, but the panel of networks only classified the first well as turbid. From the images, it is clear that the remaining wells contained very weak growth. (b) Drug for which every well was again labeled turbid by the ground truth, but the panel of networks only classified the first well as turbid. Once again, the photos show that growth was weak in the wells that were predicted to be non-turbid.

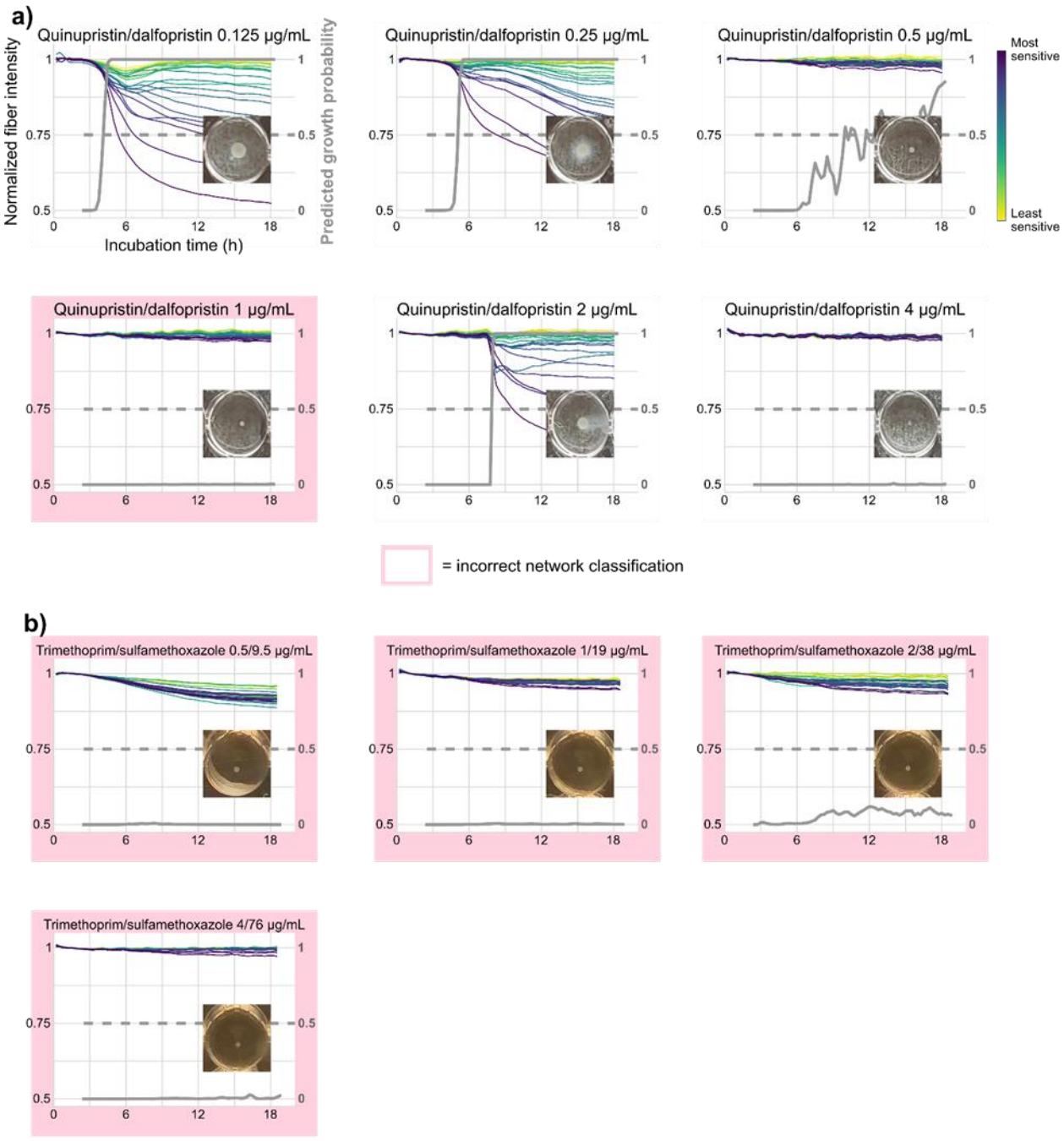


Figure 2.15: Examples of two drugs for which the panel of neural networks made very major errors. In combination with the previous figure (Figure 2.14), these wells constitute the only four very major errors on the blind testing isolates. Wells for which the panel made an incorrect turbidity prediction at the end of incubation have pink shading. Images of the wells captured by a smartphone are included in the inset of each plot. (a) Drug for which the panel of networks predicts a "skipped well," i.e. a well with no growth followed by a well with strong growth at a higher concentration of drug. In this case, the panel of networks raised an error because the MIC was undefined. (b) Drug for which every well was again labeled turbid by the ground truth, but the panel of networks classified all wells as non-turbid. The photos show that growth was weak in all wells.

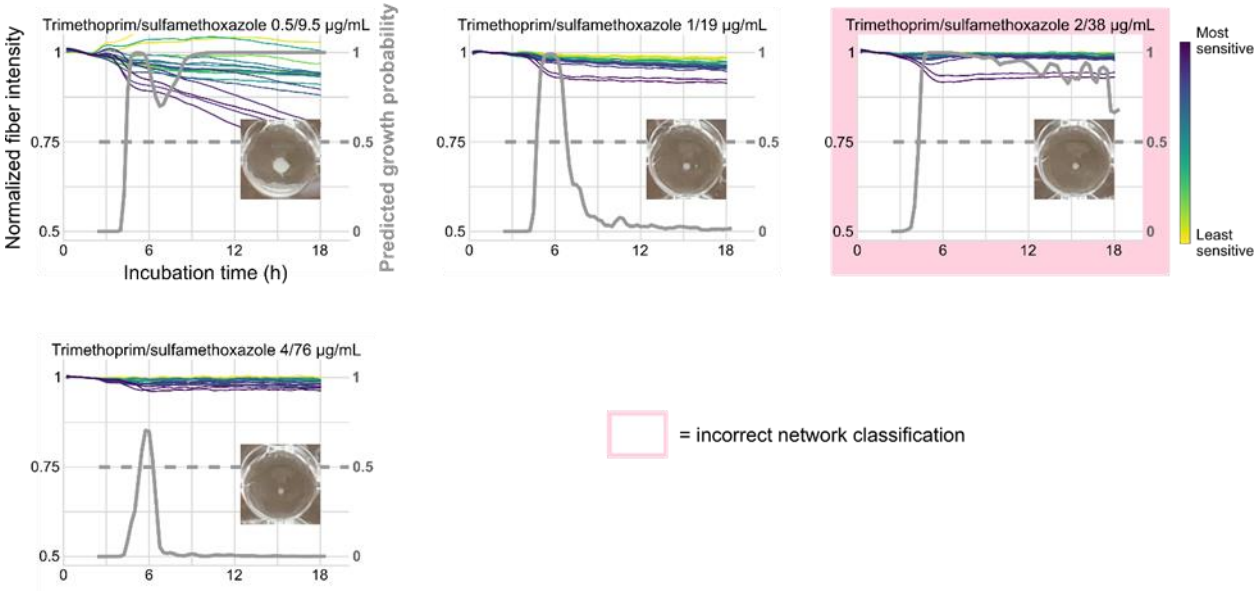


Figure 2.16: Fiber intensities and panel of networks' predictions for the drug that caused the one major error in the blind testing isolates. Wells for which the panel made an incorrect turbidity prediction at the end of incubation have pink shading. Images of the wells captured by a smartphone are included in the inset of each plot. There appears to be weak growth in the second, third, and fourth wells.

Supplementary tables

Isolate	Erythromycin	Clindamycin	Quinupristin/ dalfopristin	Daptomycin	Vancomycin	Tetracycline	Gentamicin	Rifampin	Levofloxacin	**Linezolid**	Penicillin	Ciprofloxacin	Trimethoprim/ sulfamethoxazole	Gatifloxacin	Oxacillin + 2% NaCl	Neg control	Pos control
1	8+	4+	1	1	2	2-	32+	0.5-	0.5	**8**	16+	0.5-	0.5/9.5-	1	16+	-	+
2	8+	4+	2	1	4	2-	32+	0.5-	0.5	**8**	16+	1	0.5/9.5-	1	16+	-	+
3	8+	4+	2	1	2	2-	32+	0.5-	0.5	**8**	16+	1	0.5/9.5-	1	16+	-	+
4	8+	4+	2	1	2	2-	32+	0.5-	0.5	**8**	16+	0.5-	0.5/9.5-	1	16+	-	+
5	8+	4+	2	2	4	2-	32+	0.5-	0.5	**8**	16+	0.5-	0.5/9.5-	1	16+	-	+
6	8+	4+	2	2	2	2-	32+	0.5-	0.5	**8**	16+	1	0.5/9.5-	1	16+	-	+
7	8+	4+	2	1	4	2-	32+	0.5-	0.5	**8**	16+	0.5-	0.5/9.5-	1	16+	-	+
8	8+	4+	2	2	2	2-	32+	0.5-	0.5	**8**	16+	1	0.5/9.5-	1	16+	-	+
9	8+	4+	2	1	2	2-	32+	0.5-	0.5	**8**	16+	0.5-	0.5/9.5-	1	16+	-	+
10	8+	4+	2	2	4	2-	32+	0.5-	0.5	**8**	16+	0.5-	0.5/9.5-	1	16+	-	+
11	8+	4+	2	2	4	2-	32+	0.5-	0.25-	**8**	16+	0.5-	0.5/9.5-	1	16+	-	+
12	8+	4+	2	1	4	2-	32+	0.5-	0.5	**8**	16+	1	0.5/9.5-	1	16+	-	+
13	8+	4+	2	1	2	2-	32+	0.5-	0.5	**8**	16+	0.5-	0.5/9.5-	1	16+	-	+
14	8+	4+	1	1	4	2-	32+	0.5-	0.5	**8**	16+	0.5-	0.5/9.5-	1	16+	-	+
15	8+	4+	2	1	4	2-	32+	0.5-	0.5	**8**	16+	0.5-	0.5/9.5-	1	16+	-	+
16	8+	4+	2	2	2	2-	32+	0.5-	0.5	**8**	16+	0.5-	0.5/9.5-	1	16+	-	+
17	8+	4+	2	1	4	2-	32+	0.5-	0.5	**8**	16+	0.5-	0.5/9.5-	1	16+	-	+
18	8+	4+	2	2	4	2-	32+	0.5-	0.5	**8**	16+	1	0.5/9.5-	1	16+	-	+
19	8+	4+	2	1	4	2-	32+	0.5-	0.5	**8**	16+	0.5-	0.5/9.5-	1	16+	-	+
20	8+	4+	2	1	2	2-	32+	0.5-	0.5	**8**	16+	0.5-	0.5/9.5-	1	16+	-	+
21	8+	4+	2	1	4	2-	32+	0.5-	0.5	**8**	16+	1	0.5/9.5-	1	16+	-	+
22	8+	4+	2	2	4	2-	32+	0.5-	0.5	**8**	16+	1	0.5/9.5-	1	16+	-	+
23	8+	4+	2	1	2	2-	32+	0.5-	0.5	**8**	16+	0.5-	0.5/9.5-	1	16+	-	+
24	8+	4+	2	2	2	2-	32+	0.5-	0.5	**8**	16+	0.5-	0.5/9.5-	1	16+	-	+
25	8+	4+	2	1	4	2-	32+	0.5-	0.5	**8**	16+	1	0.5/9.5-	1	16+	-	+
26	8+	4+	2	1	2	2-	32+	0.5-	0.5	**8**	16+	0.5-	0.5/9.5-	1	16+	-	+
27	8+	4+	2	1	2	2-	32+	0.5-	0.5	**4**	16+	0.5-	0.5/9.5-	1	16+	-	+
28	8+	4+	2	2	4	2-	32+	0.5-	0.5	**8**	16+	1	0.5/9.5-	1	16+	-	+
29	8+	4+	2	2	2	2-	32+	0.5-	0.5	**8**	16+	1	0.5/9.5-	1	16+	-	+
30	8+	4+	2	1	2	2-	32+	0.5-	0.5	**8**	16+	0.5-	0.5/9.5-	1	16+	-	+
31	8+	4+	2	1	2	2-	32+	0.5-	0.5	**8**	16+	0.5-	0.5/9.5-	1	16+	-	+
32	8+	4+	2	1	2	2-	32+	0.5-	0.5	**8**	16+	1	0.5/9.5-	1	16+	-	+
33	8+	4+	2	1	4	2-	32+	N/A	0.5	**8**	16+	0.5-	0.5/9.5-	1	16+	-	+
34	8+	4+	2	2	4	2-	32+	0.5-	0.5	**8**	16+	1	N/A	1	16+	-	+
35	8+	4+	2	4	2	N/A	32+	0.5-	0.5	**8**	16+	1	0.5/9.5-	1	16+	-	+
36	8+	4+	2	2	8	2-	32+	0.5-	1	**8**	16+	2	N/A	1	16+	-	+
37	8+	4+	2	2	4	2-	32+	0.5-	0.5	**8**	16+	0.5-	0.5/9.5-	1	16+	-	+
38	8+	4+	2	2	2	2-	32+	0.5-	0.5	**8**	16+	1	0.5/9.5-	1	16+	-	+
39	8+	4+	2	1	2	2-	32+	0.5-	0.5	**8**	16+	0.5-	0.5/9.5-	1	16+	-	+
40	8+	4+	2	1	4	2-	32+	0.5-	0.5	**8**	16+	0.5-	0.5/9.5-	1	16+	-	+
41	8+	4+	2	1	2	2-	32+	0.5-	0.5	**8**	16+	1	0.5/9.5-	1	16+	-	+
42	8+	4+	2	2	4	2-	32+	0.5-	0.5	**8**	16+	0.5-	0.5/9.5-	2	16+	-	+
43	8+	4+	1	1	4	2-	32+	0.5-	0.5	**8**	16+	1	0.5/9.5-	1	16+	-	+
44	8+	4+	2	2	2	2-	32+	N/A	0.5	**8**	16+	0.5-	0.5/9.5-	1	16+	-	+
45	8+	4+	2	2	2	2-	32+	0.5-	0.5	**8**	16+	0.5-	0.5/9.5-	1	16+	-	+
46	8+	4+	1	1	4	2-	32+	0.5-	0.5	**8**	16+	0.5-	0.5/9.5-	1	16+	-	+
47	8+	4+	2	2	2	2-	32+	0.5-	0.5	**8**	16+	0.5-	0.5/9.5-	1	16+	-	+

Table 2.2: Minimum inhibitory concentrations (MICs) for each 96-well plate run with the quality control strain of *Staphylococcus aureus.* The MICs were repeatable across the 47 plates that were run with the quality control strain of *S. aureus* in our AST system. Neighboring wells contain twofold dilutions of the drugs, so MICs are always a power of 2. With the exception of the drug linezolid, the MICs fell within the expected ranges for the quality control strain, confirming that the periodic illumination of the plate does not cause the bacteria to undergo phototoxicity. Because linezolid consistently showed MICs outside the expected range (possibly due to a manufacturing defect), those wells were not used in this study.

isolate	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12
1	1	1	1	1	1	1	0	0	1	0	0	0	1	1	0	0	0	0	1	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0
2	1	1	1	-	-	1	0	0	-	0	0	0	1	1	-	0	0	0	1	1	1	0	0	0	1	-	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0
3	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	-	0	0	1	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0
4	1	1	1	0	0	1	1	0	0	0	0	0	1	1	1	1	0	0	1	1	1	0	0	0	1	0	0	0	0	0	0	0	1	-	-	0	1	0	0	0	0	0	0	0	0	0	0	0
5	1	1	1	1	1	1	-	-	-	0	0	0	1	1	1	-	0	0	1	1	1	0	0	0	1	0	0	0	0	0	0	0	1	1	1	1	1	1	1	-	0	0	0	0	0	0	0	0
6	1	1	1	1	1	1	-	0	0	0	0	0	1	1	1	0	0	0	1	1	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0
7	1	1	1	-	-	1	1	-	0	0	0	0	1	1	-	-	-	-	1	1	1	0	0	0	1	1	0	0	0	0	0	0	-	-	-	-	1	1	1	1	1	1	1	0	0	0	0	0
8	1	1	1	1	0	1	1	0	0	0	0	0	1	1	-	-	0	0	-	-	-	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
9	1	1	1	0	0	1	1	0	0	0	0	0	1	1	1	0	0	0	1	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0
10	1	1	1	0	0	1	-	0	0	0	0	0	1	1	1	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0
11	1	1	1	0	0	1	1	0	0	0	0	0	1	1	1	0	0	0	1	1	1	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
12	1	1	1	0	0	1	1	0	0	0	0	0	1	1	-	0	0	0	1	1	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
13	1	1	1	0	0	1	-	0	0	0	0	0	1	1	1	0	0	0	1	1	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
14	1	1	1	0	0	1	1	0	0	0	0	0	1	1	1	0	0	0	1	1	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
15	1	1	1	0	0	1	1	0	0	0	0	0	1	1	1	0	0	0	1	1	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0
16	1	1	1	0	0	1	1	0	0	0	0	0	1	1	1	0	0	0	1	1	1	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0
17	1	1	1	0	0	1	1	0	0	0	0	0	1	1	0	0	0	0	1	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	1	1	1	0	0	1	0	0	0	0	0	0	1	1	1	0	0	0	1	1	1	0	0	0	1	1	0	0	0	0	0	0	-	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
19	1	1	1	0	0	1	1	0	0	0	0	0	1	1	0	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0
20	1	1	1	0	0	1	0	0	0	0	0	0	1	1	1	0	0	0	1	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0
21	1	1	1	0	0	1	1	0	0	0	0	0	1	1	0	0	0	0	1	1	1	0	0	0	1	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
22	1	1	1	1	1	1	1	0	0	0	0	0	1	1	1	0	0	0	1	1	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0
23	1	1	1	0	0	1	0	0	0	0	0	0	1	1	0	0	0	0	1	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0
24	1	1	1	1	0	1	1	0	0	0	0	0	1	1	0	0	0	0	1	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0
25	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	0	0	1	1	1	0	0	0	1	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0
26	1	1	0	0	0	1	1	0	0	0	0	0	1	1	0	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0
27	1	1	1	1	1	1	1	0	0	0	0	0	1	1	0	0	0	0	1	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0
28	1	1	1	0	0	1	1	0	0	0	0	0	1	1	1	0	0	0	1	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0
29	1	1	1	0	0	1	1	0	0	0	0	0	1	1	0	0	0	0	1	1	1	0	0	0	1	1	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
30	1	1	1	1	1	1	1	1	-	-	0	0	1	1	1	-	0	0	1	1	1	0	0	0	1	0	0	0	0	0	0	0	-	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0
31	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0	0	0	0	0	-	-	-	1	1	1	1	1	1	1	1	0	0	0	0	0
32	1	1	1	1	1	1	1	0	0	0	0	0	1	1	0	0	0	0	1	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0
33	1	1	1	-	0	1	1	1	0	0	0	0	1	1	0	0	0	0	1	1	1	0	0	0	1	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0

isolate	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12
1	0	1	0	0	0	0	0	1	1	1	1	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	1
2	0	1	0	0	0	0	0	1	1	1	1	0	0	1	1	1	1	1	0	0	0	-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	1
3	0	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	1	1	0	0	1	1	1	0	0	0	1	1	1	1	1	0	1	1	1	1
4	0	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	-	1	1	0	1	1	1	0	0	0	0	1	1	1	0	0	1	-	0	0	0	0	0	1	1	1	1
5	0	1	0	0	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	-	1	-	0	0	0	1	0	0	0	0	0	1	1	1	1
6	0	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	1	1	-	0	1	1	1	0	0	1	1	1	1	1	1	-	0	1	1	1
7	0	1	-	-	-	-	-	1	1	1	1	-	0	1	1	1	1	1	1	1	1	-	0	-	0	-	-	-	0	0	0	0	-	-	-	0	0	-	1	0	0	0	0	0	0	1	1	1	1
8	0	1	1	1	1	1	-	1	1	1	1	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	1	0	0	0	0	0	1	1	0	0	0	-	1	0	0	0	0	0	0	1	1	1
9	0	1	1	1	1	1	1	1	1	1	1	-	0	1	1	1	0	0	0	0	0	1	1	1	0	-	-	-	1	1	0	0	1	1	1	1	0	-	1	1	1	1	1	1	0	1	1	1
10	0	1	0	0	0	0	0	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	1
11	0	1	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-	-	0	0	0	0	1	1	1	1
12	-	1	-	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	1	1	-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	1
13	0	1	0	0	0	0	0	1	1	1	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	1
14	0	1	0	0	0	0	0	1	1	1	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	1
15	0	1	0	0	0	0	0	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	1
16	0	1	0	0	0	0	0	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	1	1	1
17	0	1	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	1	1	1
18	0	1	1	0	0	0	0	1	1	1	1	-	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	1	1	1	1
19	0	0	0	0	0	0	0	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	1
20	0	1	0	0	0	0	0	1	1	1	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	1	1	1
21	0	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	1	1	0	0	1	1	1	1	0	0	1	1	1	1	0	0	0	1	1	1
22	0	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	0	1	1	1	1
23	1	1	1	0	0	0	0	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	1	1	1	1
24	0	1	0	0	0	0	0	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	1
25	0	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	1	1	0	0	1	1	1	0	0	0	1	1	1	1	1	0	1	1	1	1
26	0	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	0	0	1	1	0	0	0	0	0	1	1	1	1
27	0	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	1	1	0	0	1	1	1	0	0	0	1	1	1	1	1	0	1	1	1	1
28	0	1	0	0	0	0	0	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	1
29	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	0	0	1	1	1	0	0	0	1	1	1	1	1	0	1	1	1	1
30	0	1	1	0	0	0	0	1	1	1	1	0	0	1	1	1	1	1	0	0	0	1	0	0	0	-	-	-	0	0	0	0	-	0	0	0	0	0	1	-	1	0	0	0	0	1	1	1	1
31	0	1	1	1	1	1	1	1	1	1	1	-	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	0	1	1	1	1
32	0	1	1	1	1	1	1	1	1	1	-	0	0	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	1	0	0	0	0	0	1	1	1	1
33	0	1	0	0	0	0	0	1	1	1	1	-	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	1

Table 2.3: Ground truth labels for all 96 wells for 33 blind testing isolates of *S. aureus*. 1 corresponds to a turbid label from both human readers, 0 corresponds to a non-turbid reading from both human readers, and – corresponds to a discrepancy between the two labels. The two medical technologists' turbidity readings differed for 86 of 3168 wells (2.7%). Interestingly, many of the discrepancies span several neighboring wells, which would cause a large change in the resulting minimum inhibitory concentration (MIC). An automated readout such as that of our system could enable more consistent readings across human readers/staff and laboratories.

Drug	Essential Agreement	Categorical Agreement	n	Major error rate	n	Very major error rate	n
Oxacillin + 2% NaCl	4.25	4	51	5.25	34	N/A	17
Levofloxacin	4.5	4.5	50	18.75	35	15.25	15
Tetracycline	5.5	5.5	50	2.5	43	8	4
Gatifloxacin	4.25	3.75	51	-	0	2.5	15
Daptomycin	5.25	13.5	49	16.25	6	17.25	43
Gentamicin	2.5	2.5	50	2.5	45	16	5
Ciprofloxacin	4.75	4.75	51	2.5	36	13	15
Penicillin	8.25	4.5	50	2.5	12	5.75	38
Erythromycin	8.5	8.5	51	N/A	1	9.75	24
Vancomycin	4.75	13.5	51	2.5	38	-	0
Rifampin	2.5	2.5	49	2.5	49	-	0
Quinupristin/ dalfopristin	5.5	6.25	50	2.5	44	-	0
Clindamycin	7	4.5	51	2.5	45	6.25	6
Trimethoprim/ sulfamethoxazole	2.5	2.5	51	8.75	50	4.5	1
AVERAGE	**5**	**5.77**		**5.75**		**9.83**	

Table 2.4: Incubation time (h) required to meet FDA criteria by drug for validation isolates of S. aureus.

SENSITITRE™ GRAM POSITIVE PLATE FORMAT

Plate Code: **GPN3F**

	1	2	3	4	5	6	7	8	9	10	11	12
A	ERY	ERY	ERY	ERY	ERY	CLI	CLI	CLI	CLI	CLI	GEN	STR
	0.25	0.5	1	2	4	0.12	0.25	0.5	1	2	500	1000
B	SYN	SYN	SYN	SYN	SYN	SYN	DAP	DAP	DAP	DAP	DAP	DAP
	0.12	0.25	0.5	1	2	4	0.25	0.5	1	2	4	8
C	VAN	VAN	VAN	VAN	VAN	VAN	VAN	VAN	TET	TET	TET	TET
	1	2	4	8	16	32	64	128	2	4	8	16
D	AMP	AMP	AMP	AMP	AMP	AMP	AMP	AMP	GEN	GEN	GEN	GEN
	0.12	0.25	0.5	1	2	4	8	16	2	4	8	16
E	RIF	LEVO	LEVO	LEVO	LEVO	LEVO	LEVO	LZD	LZD	LZD	LZD	LZD
	0.5	0.25	0.5	1	2	4	8	0.5	1	2	4	8
F	RIF	PEN	PEN	PEN	PEN	PEN	PEN	PEN	PEN	CIP	CIP	CIP
	1	0.06	0.12	0.25	0.5	1	2	4	8	0.5	1	2
G	RIF	SXT	SXT	SXT	AXO	AXO	AXO	AXO	GAT	GAT	GAT	GAT
	2	1/19	2/38	4/76	8	16	32	64	1	2	4	8
H	RIF	SXT	OXA+	OXA+	OXA+	OXA+	OXA+	OXA+	NEG	POS	POS	POS
	4	0.5/9.5	0.25	0.5	1	2	4	8				

ANTIMICROBICS

ERY	Erythromycin
CLI	Clindamycin
SYN	Quinupristin / dalfopristin
DAP	Daptomycin
VAN	Vancomycin
TET	Tetracycline
AMP	Ampicillin
GEN	Gentamicin
LEVO	Levofloxacin
LZD	Linezolid
AXO	Ceftriaxone
STR	Streptomycin
PEN	Penicillin
RIF	Rifampin
GAT	Gatifloxacin
CIP	Ciprofloxacin
SXT	Trimethoprim / sulfamethoxazole
OXA+	Oxacillin+2%NaCL
POS	Positive Control
NEG	Negative Control

Table 2.5: 96-well plate layout of antibiotics and concentrations used in this work.

References

1. O'Neill, J. Tackling Drug-Resistant Infections Globally: Final Report and Recommendations. https://amr-review.org/Publications.html (2016).

2. Antibiotic Resistance Threats in the United States. https://www.cdc.gov/drugresistance/biggest-threats.html (2019).

3. Llor, C. & Bjerrum, L. Antimicrobial resistance: risk associated with antibiotic overuse and initiatives to reduce the problem. Ther Adv Drug Saf **5**, 229–241 (2014).

4. Ramachandran, P. et al. Implications of Overprescription of Antibiotics: A Cross-Sectional Study. J Pharm Bioallied Sci **11**, S434–S437 (2019).

5. Martin, M. J., Thottathil, S. E. & Newman, T. B. Antibiotics Overuse in Animal Agriculture: A Call to Action for Health Care Providers. Am J Public Health **105**, 2409–2410 (2015).

6. McKenna, M. Antibiotics set to flood Florida's troubled orange orchards. Nature **567**, 302–303 (2019).

7. Humphries, R. M. & Hindler, J. A. Emerging Resistance, New Antimicrobial Agents … but No Tests! The Challenge of Antimicrobial Susceptibility Testing in the Current US Regulatory Landscape. Clin Infect Dis **63**, 83–88 (2016).

8. M100-S25 | Performance Standards for Antimicrobial Susceptibility Testing; Twenty-Fifth Informational Supplement. (2015).

9. VITEK® 2. bioMérieux Clinical Diagnostics https://www.biomerieux-usa.com/vitek-2.

10. Byarugaba, D. K. Antimicrobial resistance in developing countries and responsible risk factors. International Journal of Antimicrobial Agents **24**, 105–110 (2004).

11. Bonetta, L. Whole-Genome Sequencing Breaks the Cost Barrier. Cell **141**, 917–919 (2010).

12. Zankari, E. et al. Genotyping using whole-genome sequencing is a realistic alternative to surveillance based on phenotypic antimicrobial susceptibility testing. J Antimicrob Chemother **68**, 771–777 (2013).

13. Gordon, N. C. et al. Prediction of Staphylococcus aureus Antimicrobial Resistance by Whole-Genome Sequencing. Journal of Clinical Microbiology **52**, 1182–1191 (2014).

14. Meehan, C. J. et al. Whole genome sequencing of Mycobacterium tuberculosis : current standards and open issues. Nature Reviews Microbiology **17**, 533–545 (2019).

15. Wielders, C. L. C., Fluit, A. C., Brisse, S., Verhoef, J. & Schmitz, F. J. mecA Gene Is Widely Disseminated in Staphylococcus aureus Population. J Clin Microbiol **40**, 3970–3975 (2002).

16. Khan, Z. A., Siddiqui, M. F. & Park, S. Current and Emerging Methods of Antibiotic Susceptibility Testing. Diagnostics **9**, 49 (2019).

17. Ellington, M. J. et al. The role of whole genome sequencing in antimicrobial susceptibility testing of bacteria: report from the EUCAST Subcommittee. Clinical Microbiology and Infection **23**, 2–22 (2017).

18. Tagini, F. & Greub, G. Bacterial genome sequencing in clinical microbiology: a pathogen-oriented review. Eur J Clin Microbiol Infect Dis **36**, 2007–2020 (2017).

19. Boedicker, J. Q., Li, L., Kline, T. R. & Ismagilov, R. F. Detecting bacteria and determining their susceptibility to antibiotics by stochastic confinement in nanoliter droplets using plug-based microfluidics. Lab Chip **8**, 1265–1272 (2008).

20. Kinnunen, P. et al. Self-Assembled Magnetic Bead Biosensor for Measuring Bacterial Growth and Antimicrobial Susceptibility Testing. Small **8**, 2477–2482 (2012).

21. Sinn, I. et al. Asynchronous Magnetic Bead Rotation Microviscometer for Rapid, Sensitive, and Label-Free Studies of Bacterial Growth and Drug Sensitivity. Anal. Chem. **84**, 5250–5256 (2012).

22. Kalashnikov, M., C. Lee, J., Campbell, J., Sharon, A. & F. Sauer-Budge, A. A microfluidic platform for rapid, stress-induced antibiotic susceptibility testing of Staphylococcus aureus. Lab on a Chip **12**, 4523–4532 (2012).

23. Hou, Z. et al. Time lapse investigation of antibiotic susceptibility using a microfluidic linear gradient 3D culture device. Lab Chip **14**, 3409–3418 (2014).

24. Choi, J. et al. A rapid antimicrobial susceptibility test based on single-cell morphological analysis. Science Translational Medicine **6**, 267ra174 (2014).

25. Kim, S. et al. Miniaturized Antimicrobial Susceptibility Test by Combining Concentration Gradient Generation and Rapid Cell Culturing. Antibiotics **4**, 455–466 (2015).

26. Avesar, J. et al. Rapid phenotypic antimicrobial susceptibility testing using nanoliter arrays. PNAS **114**, E5787–E5795 (2017).

27. Baltekin, Ö., Boucharin, A., Tano, E., Andersson, D. I. & Elf, J. Antibiotic susceptibility testing in less than 30 min using direct single-cell imaging. PNAS **114**, 9170–9175 (2017).

28. Fredborg, M. et al. Real-Time Optical Antimicrobial Susceptibility Testing. Journal of Clinical Microbiology **51**, 2047–2053 (2013).

29. Price, C. S., Kon, S. E. & Metzger, S. Rapid antibiotic susceptibility phenotypic characterization of Staphylococcus aureus using automated microscopy of small numbers of cells. Journal of Microbiological Methods **98**, 50–58 (2014).

30. Fredborg, M. et al. Rapid antimicrobial susceptibility testing of clinical isolates by digital time-lapse microscopy. Eur J Clin Microbiol Infect Dis **34**, 2385–2394 (2015).

31. Quach, D. T., Sakoulas, G., Nizet, V., Pogliano, J. & Pogliano, K. Bacterial Cytological Profiling (BCP) as a Rapid and Accurate Antimicrobial Susceptibility Testing Method for Staphylococcus aureus. EBioMedicine **4**, 95–103 (2016).

32. Leonard, H., Halachmi, S., Ben-Dov, N., Nativ, O. & Segal, E. Unraveling Antimicrobial Susceptibility of Bacterial Networks on Micropillar Architectures Using Intrinsic Phase-Shift Spectroscopy. ACS Nano **11**, 6167–6177 (2017).

33. Pancholi, P. et al. Multicenter Evaluation of the Accelerate PhenoTest BC Kit for Rapid Identification and Phenotypic Antimicrobial Susceptibility Testing Using Morphokinetic Cellular Analysis. Journal of Clinical Microbiology **56**, e01329-17 (2018).

34. Yu, H. et al. Phenotypic Antimicrobial Susceptibility Testing with Deep Learning Video Microscopy. Anal. Chem. **90**, 6314–6322 (2018).

35. Accelerate Pheno™ system. Accelerate Diagnostics https://acceleratediagnostics.com/products/accelerate-pheno-system/.

36. oCelloScope. BioSense Solutions https://biosensesolutions.dk/ocelloscope/.

37. Mudanyali, O. et al. Compact, light-weight and cost-effective microscope based on lensless incoherent holography for telemedicine applications. Lab on a Chip **10**, 1417–1428 (2010).

38. Tseng, D. et al. Lensfree microscopy on a cellphone. Lab on a Chip **10**, 1787–1792 (2010).

39. Greenbaum, A. et al. Imaging without lenses: achievements and remaining challenges of wide-field on-chip microscopy. Nature Methods **9**, 889–895 (2012).

40. Mudanyali, O., Bishara, W. & Ozcan, A. Lensfree super-resolution holographic microscopy using wetting films on a chip. Opt. Express, OE **19**, 17378–17389 (2011).

41. Wang, H. et al. Early-detection and classification of live bacteria using time-lapse coherent imaging and deep learning. arXiv:2001.10695 [physics] (2020).

42. Berg, B. et al. Cellphone-Based Hand-Held Microplate Reader for Point-of-Care Testing of Enzyme-Linked Immunosorbent Assays. ACS Nano **9**, 7857–7866 (2015).

43. Feng, S., Tseng, D., Di Carlo, D., Garner, O. B. & Ozcan, A. High-throughput and automated diagnosis of antimicrobial resistance using a cost-effective cellphone-based micro-plate reader. Scientific Reports **6**, 39203 (2016).

44. Antimicrobial Susceptibility Test (AST) Systems - Class II Special Controls Guidance for Industry and FDA. (2018).

45. McBirney, S. E., Trinh, K., Wong-Beringer, A. & Armani, A. M. Wavelength-normalized spectroscopic analysis of Staphylococcus aureus and Pseudomonas aeruginosa growth rates. Biomed. Opt. Express, BOE **7**, 4034–4042 (2016).

46. Stefani, S., Bongiorno, D., Mongelli, G. & Campanile, F. Linezolid Resistance in Staphylococci. Pharmaceuticals **3**, 1988–2006 (2010).

47. Cui, Z., Chen, W. & Chen, Y. Multi-Scale Convolutional Neural Networks for Time Series Classification. arXiv:1603.06995 [cs] (2016).

48. Breiman, L. Bagging predictors. Mach Learn **24**, 123–140 (1996).

49. Moretti, F., Pizzuti, S., Panzieri, S. & Annunziato, M. Urban traffic flow forecasting through statistical and neural network bagging ensemble hybrid modeling. Neurocomputing **167**, 3–7 (2015).

Chapter 3 Deep learning-based on-chip spectroscopy framework

This chapter contains material that is currently under peer review and has been submitted in pre-print as follows: Calvin Brown, Artem Goncharov, Zachary Ballard, Mason Fordham, Ashley Clemens, Yunzhe Qiu, Yair Rivenson, Aydogan Ozcan. arXiv:2012.00878

Conventional spectrometers are limited by trade-offs set by size, cost, signal-to-noise ratio (SNR), and spectral resolution. Here, we demonstrate a deep learning-based spectral reconstruction framework, using a compact and low-cost on-chip sensing scheme that is not constrained by the design trade-offs inherent to grating-based spectroscopy. The system employs a plasmonic spectral encoder chip containing 252 different tiles of nanohole arrays fabricated using a scalable and low-cost imprint lithography method where each tile has a unique geometry and, thus, a unique optical transmission spectrum. The illumination spectrum of interest directly impinges upon the plasmonic encoder, and a CMOS image sensor captures the transmitted light, without any lenses, gratings, or other optical components in between, making the entire hardware highly compact, light-weight and field-portable. A trained neural network then reconstructs the unknown spectrum using the transmitted intensity information from the spectral encoder in a feed-forward and non-iterative manner. Benefiting from the parallelization of neural networks, the average inference time per spectrum is ~28 μs, which is orders of magnitude faster compared to other computational spectroscopy approaches. When blindly tested on unseen new spectra (N = 14,648) with varying complexity, our deep learning-based system identified 96.86% of the spectral peaks with an average peak localization error, bandwidth error, and height error of 0.19 nm, 0.18 nm, and 7.60%, respectively. This system is also highly tolerant to fabrication defects that may

arise during the imprint lithography process, which further makes it ideal for applications that demand cost-effective, field-portable and sensitive high-resolution spectroscopy tools.

Introduction

Spectral analysis is used in a wide array of applications in the fields of chemistry, physics, and biomedical sensing, among others. Optical spectra are conventionally recorded with spectrometers that separate light into its spectral components via a diffraction grating. The intensity of each component is recorded by a photodetector array, e.g., a complementary metal–oxide–semiconductor (CMOS) imager, to translate these intensities into the optical spectrum of the illumination beam (covering e.g., 400–750 nm). The Czerny-Turner configuration, for example, is one of the most commonly used methods for optical spectroscopy, employing two collimating mirrors to fold the optical path while partially compensating for optical aberrations[1]. Though elegant and robust, grating-based designs present two key performance trade-offs. Firstly, increasing the spectral resolution generally comes at the cost of decreasing the signal-to-noise ratio (SNR). For example, narrowing the entrance slit width, decreasing the period of the grating, or decreasing the pixel size of the sensor all improve spectral resolution at the expense of signal strength[2]. These methods also necessitate more expensive components and more precise instrument alignment. Such trade-offs can be prohibitive for low-light, low-cost, or field-based applications that still demand high spectral resolution. Secondly, increasing the spectral resolution may require a longer optical path between the grating and the photosensor array[3]. This is typically achieved with physically larger instruments (benchtop-sized), which are less suitable for mobile spectroscopy applications. In addition, a longer path length can degrade performance due to even minor ambient temperature fluctuations[2]. Therefore, traditional spectrometer designs present a compromise among resolution, cost, size, and SNR.

Computational sensing schemes have been proposed as a promising alternative to conventional grating-based spectrometers, presenting a variety of hardware and software solutions[4–11]. Instead of relying on diffraction gratings, some of these earlier systems work by encoding the incident spectra over a set of unique filter functions. The encoded information is then interpreted by a spectral reconstruction algorithm that employs precise a priori knowledge of the filter functions or leverages some calibration data to map the encoding operation to the target spectral measurement. Benefiting from the computational sensing paradigm, these emerging designs do not share the same size, throughput, and resolution trade-offs inherent to grating-based spectrometers. The quality of the spectral reconstruction is not explicitly linked to the optical path length or the spectral resolution of the detection scheme since the encoding operation does not divide the incident light into its narrowband spectral components, instead samples the input spectra with filters that can exhibit broadband transmission[5,12,13]. Performance of these computational schemes for spectroscopy therefore depends on the robustness and spectral diversity of the encoding operation as well as on the accuracy and speed of the employed algorithm to solve the underdetermined reconstruction problem[4,12,14,15].

A number of different hardware approaches have been proposed for the spectral encoding operation including variable filters in the form of liquid crystals and Fabry-Perot cavities as well as fixed filter configurations like ring resonators, Mach-Zehnder Interferometers (MZIs), photonic crystals and plasmonic filters[7–9,16–25]. Each encoding element, which may range from a narrowband spectral filter to a broadband filter function with multiple local extrema, samples the input spectrum $I(\lambda)$ using the filter functions of the spectral encoder. Reconstruction algorithms are therefore tasked to recover the incident spectrum from the raw data sampled by each encoder. The most common approach to algorithmic reconstruction is to use a priori information of the encoding

operation and spectral sensitivity of the photodetectors to define a transformation, $T_i(\lambda)$, between the target spectrum, $I(\lambda)$, and raw measurements, S_i, i.e. $S_i = T_i(\lambda)I(\lambda)$ for each i^{th} encoding operation. By expressing this transformation operation over all the encoding elements, a least-squares problem can be defined, and a solution for $I(\lambda)$ can be obtained by minimizing *e.g.*, $\|S - TI\|_2^2$. Regularization terms based on the L1 norm (least absolute shrinkage operator, LASSO[26]) and the L2 norm (Tikhonov regularization[27]), among others, are also commonly used to solve this minimization problem, necessitating iterative reconstruction algorithms that overcome the limitations of the standard least-square solution to this underdetermined problem[5–7,9,10,15,16,18,19,24,28,29]. However, given this body of work, a data-driven non-iterative spectral reconstruction approach, without the need for a priori knowledge of the specific filter functions, has yet to be demonstrated.

Here we report a deep learning-based on-chip spectrometer (Figure 3.1) that utilizes a flat spectral encoding chip (fabricated through a scalable and low-cost imprint lithography process) to filter the incident light using an array of 252 nanostructured plasmonic tiles, where each tile has a unique transmission spectrum. The transmitted light through all these tiles is acquired in parallel using a conventional CMOS image sensor that is axially positioned at ~3mm away from the plasmonic encoder, recording the free-space diffraction patterns of the plasmonic encoder without any lenses or optical components, using a compact and field-portable design (Figure 3.1b). A trained neural network is used to reconstruct the unknown input spectra from the lensfree diffraction images in a feed-forward (i.e. *non-iterative*) manner without the need for a priori information on the encoding operation or the input illumination (Figure 3.1d,e). The network generates spectral reconstructions in ~28 μs per spectrum—orders-of-magnitude faster than other computational spectroscopy methods—by leveraging batch computation. When blindly tested on

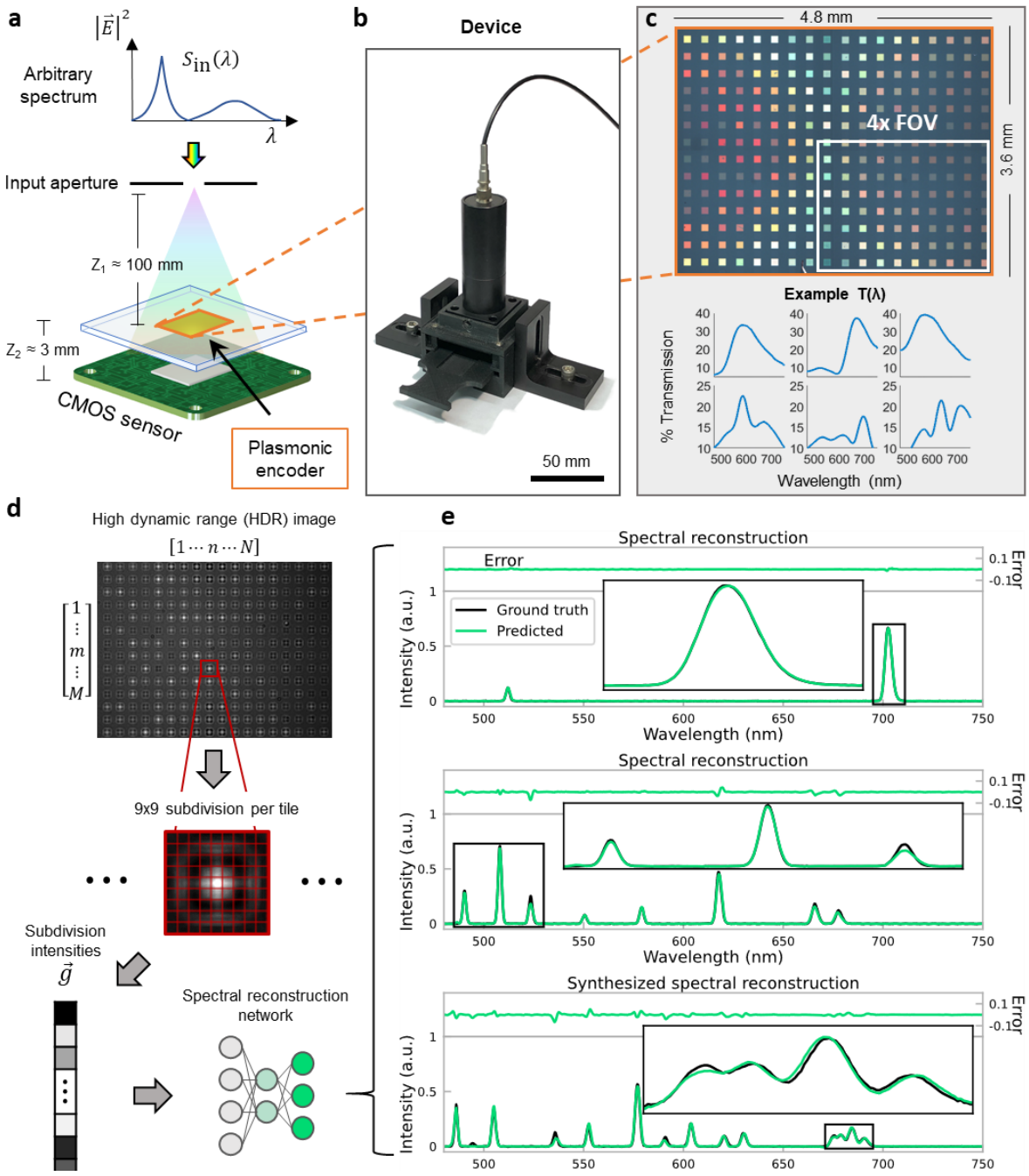


Figure 3.1: Neural network-based on-chip spectroscopy. a Schematic of the optical setup. Plasmonic encoder is located between a CMOS image sensor and the input aperture. b Photo of optical setup. c Brightfield microscope image of the plasmonic encoder chip showing example transmission spectra T(λ) below. d Workflow of spectral reconstructions. Regions of the HDR image corresponding to each tile are used as inputs to the spectral reconstruction neural network. e Spectra reconstructed during blind testing. Error is shown above each plot on the same y-scale. The network was trained only on spectra with up to 8 peaks, yet it successfully reconstructs a spectrum with 14 peaks.

new input spectra of varying complexity (N = 14,648) captured *after* the training phase, the deep

learning-based on-chip spectrometer correctly identified 96.86% of the spectral peaks with a peak localization error of 0.19 nm, a peak height error of 7.60%, and a peak bandwidth error of 0.18 nm. These performance metrics demonstrate significant improvements compared to earlier generations of computational spectrometers and were achieved despite visible fabrication defects in the plasmonic encoder chip, illustrating the robustness of our neural network-based spectral reconstruction method.

Taken together, the presented on-chip plasmonic spectroscopy design is highly cost-effective, compact, field-portable and requires no mechanical scanning components (Figure 3.1). The methods and the device design that are at the heart of this computational on-chip spectrometer can find unique applications in various fields that demand compact and sensitive high-resolution spectroscopy tools.

Results

On-chip spectroscopy framework and experimental setup

Our deep learning-based spectral reconstruction framework (Figure 3.1a,b) uses a spectral encoding chip comprising 14 × 18 = 252 unique plasmonic 'tiles', where each tile covers a region of 100 × 100 μm, defined by a unique nanohole array structure (Figure 3.2). Importantly, the encoder chip is fabricated through a scalable imprint lithography process (Figure 3.2a) that can replicate nanostructures indefinitely from a silicon 'master' chip (see Methods). As a result, our encoding chip is low-cost and with the exception of a metal deposition step, can be fabricated without the need for clean room instrumentation or processes. Each plasmonic tile, i, serves as a unique spectral filter described by a transmission function, $T_i(\lambda)$, where the local maxima and their

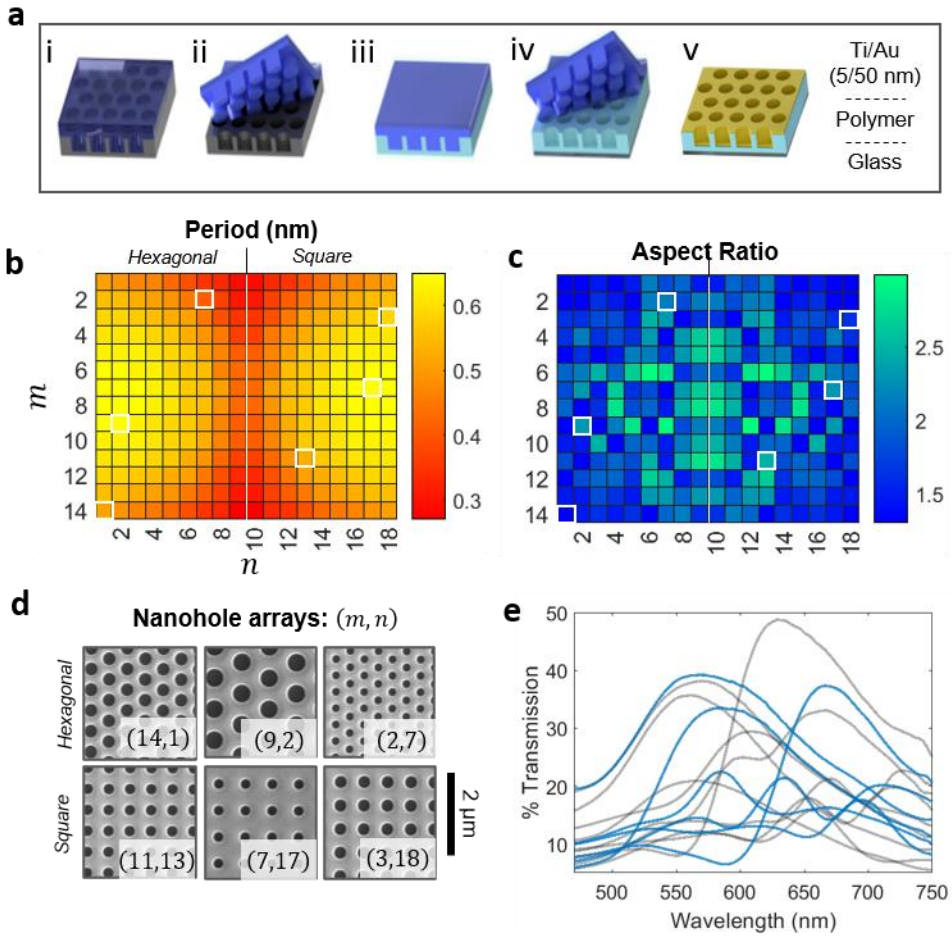


Figure 3.2: Spectral encoder chip. a Soft lithography process for molding low-cost replica of nano structures from a silicon master. Steps *i-ii* show the initial molding process with the silicon master and the initial UV curable polymer. Steps *iii-iv* show the secondary molding process followed by the metal deposition in *v*. The b Period and c Aspect Ratio of the nanohole array for each tile in the encoder are shown using a heatmap. d SEM images of example plasmonic nanohole arrays, corresponding to the outlined white boxes in b and c and example spectra in Figure 3.1c. e Example transmission spectra, where the blue lines correspond to the plasmonic tiles shown in the SEM images and in Fig 1c. Other example transmission spectra are shown in grey.

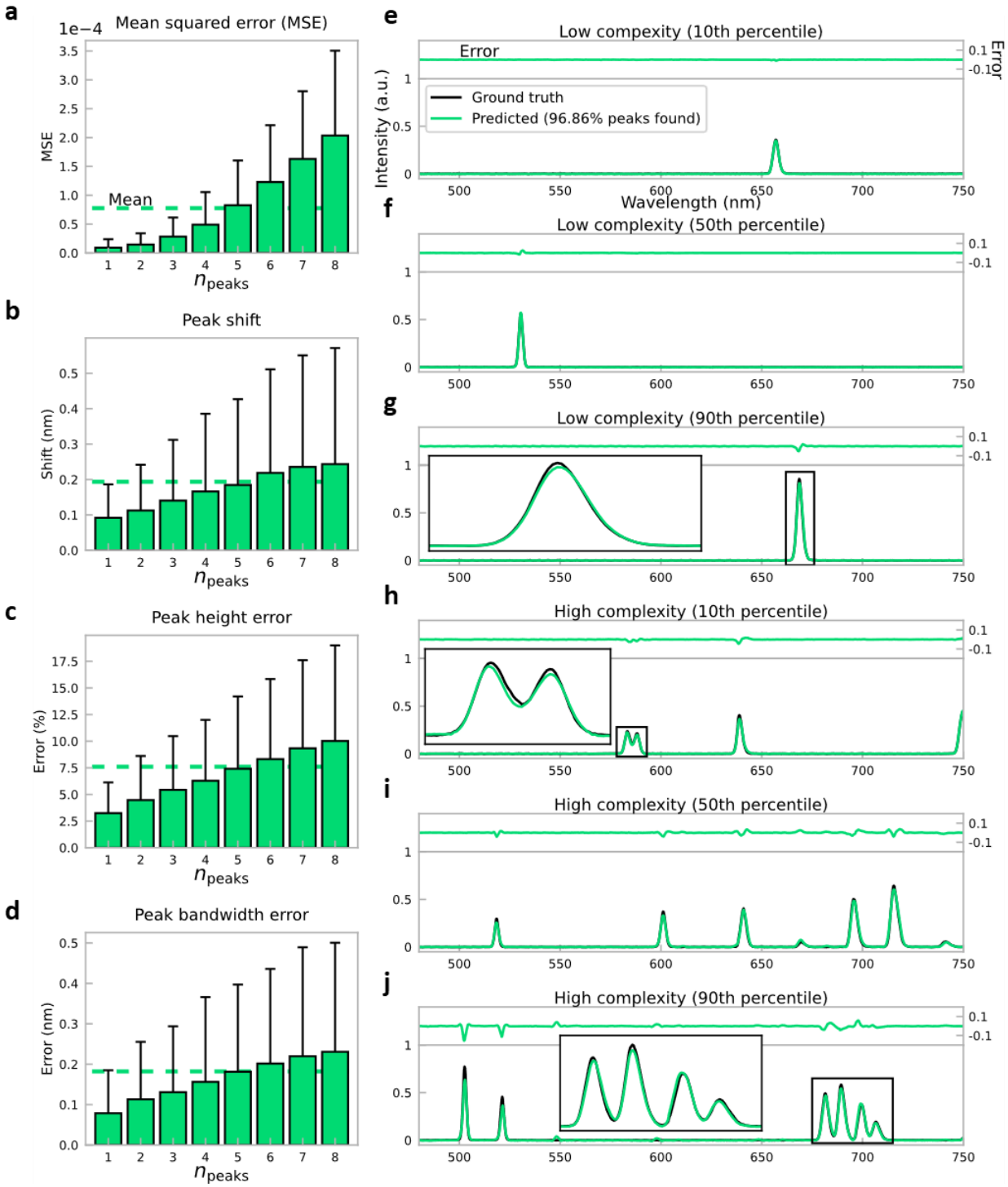
corresponding bandwidths result from the plasmonic modes supported by the dielectric and metal nanostructures (see Figure 3.2d,e).

For each illumination spectrum under test, the lensfree diffraction images of the input radiation were captured at multiple exposures to create a high dynamic range (HDR) image to limit

pixel saturation effects. Lensfree images of the encoder chip corresponding to spectral peaks from 480–750 nm are shown in Supplementary Movie 1. Each of the 252 tiles is automatically segmented and further subdivided into a 9 × 9 grid of 81 sub-tiles. The average pixel intensities of all 252 × 81 = 20,412 sub-tiles serve as the input to a trained neural network, which rapidly reconstructs the unknown illumination spectrum without any iterations, in a single feed-forward manner (Figure 3.1c,d). The spectral reconstruction neural network comprises three fully-connected layers of 2,048 neurons each, and an output layer with 1,185 nodes, representing the spectral intensities over a target spectral range of 480–750 nm, with a spectral spacing of 0.229 nm. The network was trained using a mean squared error (MSE) loss function between the reconstructed spectra (network output) and the ground truth spectra, measured by a commercially available spectrometer (see Methods). To train the network, 50,352 random spectra were generated by a programmable supercontinuum laser, with an additional 8,824 spectra used for validation data (see Methods). The resulting neural network was blindly tested on 14,648 unseen new spectra generated by the same experimental set-up. Because the network requires no further training or iterations, it is able to reconstruct a new, unseen spectrum in ~28 μs in a single feedforward manner using a desktop computer (see the Methods section).

Blindly tested spectral reconstructions

Figure 3.3 illustrates the success of the trained reconstruction network to accurately recover unknown spectra using lensfree diffracted images that are acquired by our compact set-up (Figure 3.1). The average MSE, peak localization error, peak intensity error, and bandwidth estimation error on the blindly tested spectra were 7.77e-5, 0.19 nm, 7.60%, and 0.18 nm, respectively. Overall, our experimental results reveal that 96.86% of the peaks in the ground truth spectra were correctly reconstructed by the network. Figure 3.3e-j further show reconstructed spectra and



ground truth spectra for both lower complexity (one peak) and higher complexity (4–8 peaks) spectra for various performance percentiles. These percentiles refer to the MSE loss of the network output reconstruction, where 10th percentile implies a relatively good fit (best 10% loss), 50th percentile implies the median fit, and 90th percentile implies a poor fit (worst 10%). Even for higher

Figure 3.3 (previous page): Blind testing performance. a Average MSE, b peak shift/localization error, c peak height error, and d peak bandwidth error for spectra containing 1-8 peaks. Average over all spectra is shown as a horizontal line in each plot. Reconstructions for lower complexity (1 peak) spectra in the e 10^{th}, f 50^{th}, and g 90^{th} percentile of MSE. Reconstructions for higher complexity (4-8 peaks) spectra in the h 10^{th}, i 50^{th}, and j 90^{th} percentile of MSE. 10^{th}, 50^{th}, and 90^{th} percentiles correspond to best 10%, median, and worst 10% fits, respectively. Error is shown above each plot on the same y-scale.

complexity spectra, the 90^{th} percentile network output results are rather accurate, closely matching the ground truth spectra acquired with a benchtop spectrometer. Additional examples of blind spectral reconstructions obtained at the network output are shown in Figure 3.7–Figure 3.14 to demonstrate the repeatability and success of this blind spectral inference process.

We also evaluated the peak localization and bandwidth estimation error, on blindly tested new spectra, each with a 3 nm-bandwidth peak, ranging from 480–750 nm with a step size of 1 nm. For these 271 new spectra, all the peak localization errors were within ± 0.32 nm, and all the bandwidth estimation errors were within ± 0.178 nm, significantly surpassing the performance of earlier on-chip spectroscopy results.

Spectral inference stability as a function of time

Because the training spectra were captured *before* the blind testing, one would expect some level of performance degradation in spectral inference due to e.g., temperature fluctuations, illumination source stability/coupling changes, or mechanical vibrations, especially as the time period between the capture of the training and testing data increases. The performance stability of the inference of the trained network over the course of the blind testing data capture (~15 h of *continuous* operation) is evaluated in Figure 3.4. All performance metrics remained fairly stable, with no significant difference between their values at the start and end of the 15 h continuous testing period (Figure 3.4).

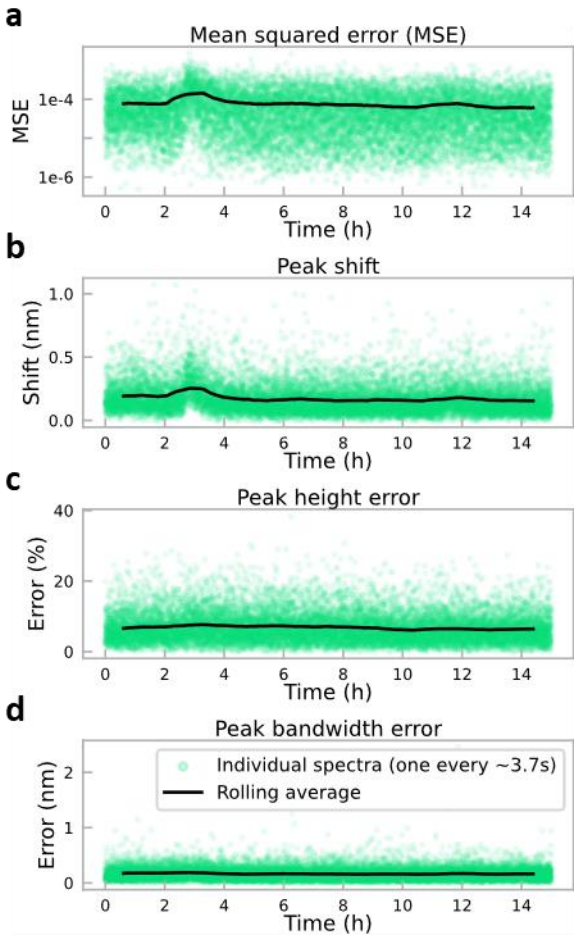
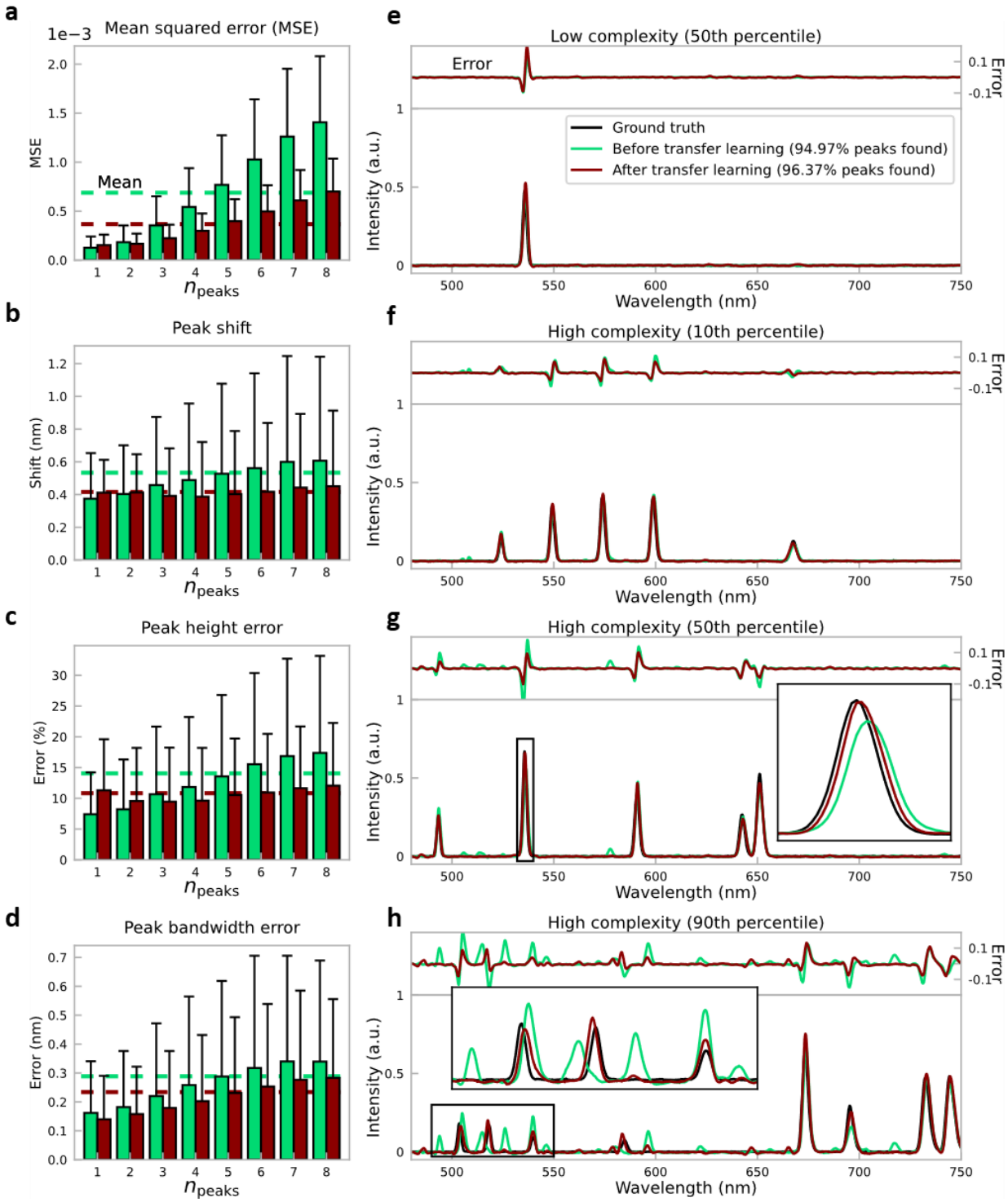


Figure 3.4: Stability of inference performance over time. a Average MSE, b peak shift/localization error, c peak height error, and d peak bandwidth error for spectra over the course of blind testing data capture. 14,648 blind testing spectra are represented in the plots, captured over ~15 h of continuous operation of the system.

To further investigate the performance stability over time, an additional 21,296 new spectra were captured ~5.8 days *after* the last training/validation spectrum. Compared with the earlier blind inference results, the performance relatively degraded on these new spectra, as shown in Figure 3.5 (green curves). The average MSE, peak localization error, peak intensity error, and bandwidth estimation error on these later-acquired unknown spectra were 6.89e-4, 0.53 nm,



14.06%, and 0.29 nm, respectively, with 94.97% of spectral peaks correctly identified. As a means to re-calibrate the reconstruction network and overcome this relative performance degradation over time, we implemented a transfer learning approach, where the weights of the previously-trained

Figure 3.5 (previous page): Performance on blind testing spectra captured ~5.8 days after training. a Average MSE, b peak shift/localization error, c peak height error, and d peak bandwidth error for spectra containing 1-8 peaks. Average over all spectra is shown as a horizontal line in each plot. Transfer learning considerably improves the inference performance, for all metrics. Reconstruction for lower complexity (1 peak) spectrum in the e 50^{th} percentile of MSE. Reconstructions for higher complexity (4-8 peaks) spectra in the f 10^{th}, g 50^{th}, and h 90^{th} percentile of MSE. Error is shown above each plot on the same y-scale.

neural network were adjusted through further training on a small fraction of the spectra captured at the start of the new measurement period (i.e., ~5.8 days *after* the last training phase). The performance metrics and spectral reconstructions after this transfer learning step are shown alongside those of the original network in Figure 3.5. All performance metrics are significantly improved after the transfer learning step: average MSE, peak localization error, peak intensity error, and bandwidth estimation error improve to 3.68e-4, 0.42 nm, 10.83%, and 0.23 nm, respectively, with 96.37% of the peaks correctly identified. Figure 3.5 further illustrates that, in addition to these considerable improvements in spectral inference metrics, background spectral noise and erroneous peaks are also suppressed well, after the transfer learning step.

It is important to emphasize that the amount of data and the computation time required for this transfer learning step are rather small; even using just 100 new spectra (requiring ~6 min to capture) and training the existing neural network for 100 epochs (requiring < 1 min on a desktop computer) shows marked improvements in the blind spectral reconstructions after 5.8 days (Figure 3.18). Therefore, transfer learning can be an effective software-based calibration tool for our data-driven computational on-chip spectrometer, as demonstrated here.

Speed of spectral reconstructions

Unlike optimization-based approaches to spectral reconstruction[5,7,9,10,15,16,18,19,24,28,29], our neural network-based inference does not require iterative computation to predict each unknown

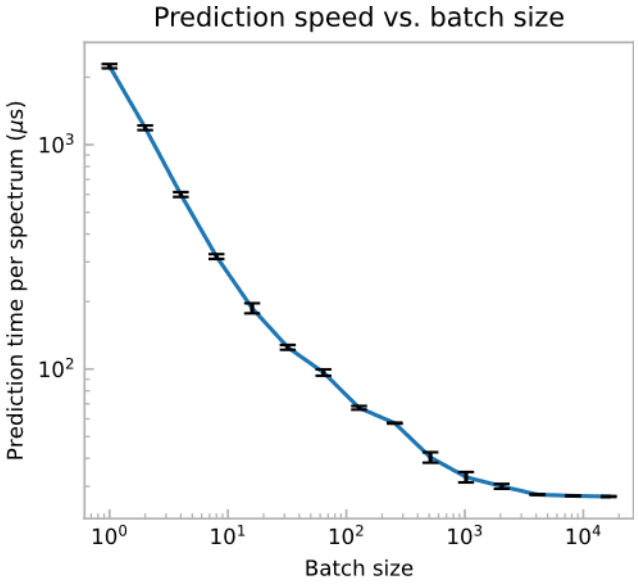


Figure 3.6: Prediction speed vs. batch size for blind testing spectra. Inference time per spectrum considerably decreases with increasing batch size due to the highly parallelizable nature of the neural network computation. An average inference time of ~28 μs is obtained for a batch sizes of 4096 or more spectra. Error bars were generated from 7 repeated trials. All predictions were performed on a desktop computer (see Methods).

spectrum. Once the network has been trained, it can perform reconstruction of unseen spectra rapidly and in parallel. The average prediction time per spectrum for different batch sizes are shown in Figure 3.6. All calculations were performed on a desktop computer (see Methods). For a batch size of 4096 spectra, the network is able to reconstruct an unknown spectrum in ~28 μs on average, providing orders of magnitude inference speed advantage compared to optimization-based iterative reconstruction methods. This parallel computation capability could be particularly beneficial for e.g., significantly increasing the speed of hyperspectral imaging systems, where a unique spectrum needs to be reconstructed for each hyperspectral pixel.

Compared to other state of the art neural networks used for image processing and enhancement tasks[30–35], the spectral reconstruction network employed in this work is compact and

shallow, comprising three hidden layers and no convolutional layers. This enabled rapid prediction on an unseen spectrum in ~ 43 μs, without requiring an iterative minimization in the blind inference phase. To further increase the speed of prediction, we investigated subsampling each tile into a 7 × 7 sub-grid instead of 9 × 9 (Figure 3.19). While the coarser subsampling (7 × 7) causes a modest degradation in prediction performance, the network inference time further decreased to ~18 μs per spectrum. It is also important to note that the spectral reconstruction network yielded the best performance when trained with a dropout probability of 3%, much lower than the typical values (10–50%) used in many common neural networks employed in other applications[30–32,36]. Because the neural network does not use convolutional layers and is relatively shallow and wide, even a small dropout probability gave us a strong regularization effect.

Network generalization to new spectral distributions not included in the training

In addition to investigating the spectral reconstruction network's ability to generalize over time, we also tested its ability to generalize to new spectral distributions that were not included in the training phase. While the network was trained on random spectral distributions containing 1–8 peaks, we synthesized more complex spectra that had more number of peaks, not represented within the training phase. The network was tested on these synthesized spectra without any retraining or transfer learning, and some examples of the synthesized spectra and the corresponding spectral reconstructions are shown in Figure 3.1e and Figure 3.15. Despite using spectral distributions that were never represented during the training phase, the reconstructions still identify 90.54% of all peaks and do not suffer from any obvious artifacts or noise. These results demonstrate the presented framework's ability to generalize to more complex spectral distributions that were not included in the training data. Additionally, we tested the same neural network with new spectra that had larger bandwidths per peak compared to the training spectra.

These new test spectra with broader bandwidths therefore serve as another example of new spectral distributions that are successfully reconstructed using the network (Figure 3.16) even though they were not represented during the training.

Discussion

The performance of the presented spectral reconstruction framework demonstrates its potential to enable new modalities of spectroscopy and hyperspectral imaging. When our deep learning-based spectrometer was blindly tested on unseen spectra, it correctly identified 96.86% of the spectral peaks, with a peak localization error of 0.19 nm, a peak height error of 7.60%, and a peak bandwidth error of 0.18 nm. The presented spectral reconstructions (Figure 3.3, Figure 3.7–Figure 3.14) match the ground truth spectra quite well, both quantitatively and qualitatively, with the peak localization error and the peak bandwidth error on the order of the resolution (0.229 nm) of the spectrometer used to capture the ground truth training data. Additionally, there is no observed bias in these performance metrics over the operational range (480 nm – 750 nm) which can be partly attributed to (1) the lack of wavelength dependency in the MSE loss function, and (2) the spectral diversity of the filter functions, $T_i(\lambda)$, of the plasmonic encoder chip (see Figure 3.2e). It is also important to emphasize that as a unique aspect of this work, successful reconstruction performance was demonstrated over a large blind testing dataset (14,648 unique spectra) containing varying degrees of complexity, including non-sparse examples with overlapping spectral features (e.g., Figure 3.3h,j). When compared to a linear regression model with L2-norm regularization trained on the same data set, the trained neural network achieved nearly a 5-fold reduction in average MSE (from 3.85e-4 to 7.77e-5). The predicted spectra resulting from the regularized linear model contain significant noise artefacts (see Figure 3.17) not seen in our deep-learning based reconstructions, suggesting that the hidden layers and the inherent

nonlinearities within the neural network play an important role in de-noising the spectral features identified through linear operations, yielding robust performance over complex and non-sparse inputs.

Avoiding overfitting of the spectral reconstruction algorithm is critical for any computational spectrometer that uses a data-driven approach. This is especially important for complex models such as neural networks that contain a large number of trainable parameters and non-linear activation functions. Reconstruction algorithms that exhibit overfitting to a particular training set, can fail to appropriately interpret minute changes in experimental system alignment, temperature, vibrations, or other unforeseen noise sources, potentially leading to significant changes in the output and overall reconstruction performance. The presented system showed very good stability over the course of ~15h *continuous* experimentation during which the 14,648 blind testing spectra were captured (Figure 3.4); however its blind inference performance relatively degraded after ~5.8 days, likely due to uncontrolled factors, e.g., temperature, vibrations. These spectral reconstructions after ~5.8 days were improved considerably by transfer learning on newly-captured data, which amounts to a simple calibration step, similar to what is typically used for some other measurement instruments.

Another important aspect of this spectral reconstruction framework is the use of a spectral encoder chip, fabricated through scalable imprint lithography. The fabrication of the encoding chip does not require cleanroom-based lithography or other processes that require an advanced fabrication infrastructure, except the metal deposition step, which is relatively much simpler and cheaper. While this low-cost and rapid imprint lithography process can introduce point imperfections in the encoder chip, as evident in Supplementary Movie 1, the data-driven spectral reconstruction network demonstrates robustness to these defects. Due to the minimal cost and

scalability of the imprint lithography, large area encoders for hyperspectral imaging could be fabricated, in which an ensemble of optimal filter functions could be grouped into a single hyperspectral pixel that is tiled across the encoder chip. While the present encoder contains 252 spectrally overlapping broadband filters, further optimization in its design can be achieved: application-specific feature reduction approaches can be used to select, in a data-driven manner, a statistically optimal sub-set of tiles[13,37]. In Figure 3.20, we explored network performance when using random subsets of the 252 tiles, demonstrating that blind spectral reconstructions using just 49 plasmonic tiles are still quite competitive. This trade-off between reconstruction performance and the number of encoder elements would be critical for designing future computational spectrometers.

Compared to traditional grating-based spectrometer designs, the presented spectral reconstruction framework offers several unique features. First, the compact nature of the on-chip spectroscopy system (Figure 3.1a,b) could enable inexpensive, lightweight designs with large fields-of-view for, e.g. remote, airborne, or even disposable sensing needs in field settings. Because the encoder chip can be manufactured at low cost over large areas with the imprinting process, an array of spectrometers or a hyperspectral imaging grid could be fabricated without the need for cleanroom-based lithography tools. Since the presented device bins the neighbouring pixels, spectrometers using large-pixel size sensors or, conversely, spectrometers with even a smaller footprint (via less pixel binning) could be designed as well. Second, the traditional trade-off between the spectral resolution and SNR that is common in grating-based spectrometers is now pushed to a different optimum point: the resolution of our spectral reconstruction network is primarily limited by the spectral resolution of the instrument used for ground truth measurements

of the training data, and the individual filters of the encoder chip do not need to be narrowband to match the ground truth resolution as demonstrated in this work.

The data-driven approach utilized in this work also offers key advantages when compared to common spectral reconstruction algorithms based on e.g., least-squares minimization employed in other computational spectrometer systems. Although the training process requires a large amount of measurements to be obtained, this is a one-time effort, and it yields a forward model that can blindly recover unknown spectra from raw sensing signals in ~28 μs, orders of magnitude shorter than the time required to solve iterative minimization problems, employed in earlier spectral reconstruction methods. Spectral reconstruction timing can be important for various applications such as hyperspectral imaging, that may demand a spectral recovery across a large sequence of images each with a large number of individual pixels. Additionally, some of the iterative reconstruction algorithms used earlier employ a 'smoothness' constraint in their optimization process, based on the second derivative of the target spectra[5]. Although this may improve some spectral reconstructions, the selection of a singular weighting parameter on this constraint introduces a trade-off in performance between narrow-band and broad-band spectral reconstructions. Lastly, instead of using training data, these iterative reconstruction methods rely on precise measurements of the encoding operation and the spectral response of the underlying photosensors, which are both used as a priori information. This presents a separate array of challenges, because the reconstruction performance relies on how precisely one can characterize the underlying hardware. All of these challenges are considerably mitigated or eliminated using the presented deep learning-based spectroscopy approach, which also lends itself to a highly compact, field-portable, sensitive and high-resolution spectrometer design that can be used in various targeted applications, in both sensing and imaging.

Methods

Plasmonic encoder

The plasmonic encoder measures 4.8 × 3.6 mm and consists of 252 (14 × 18) tiles, with each tile covering 100 × 100 µm. Each one of these tiles consists of a nanohole array with a unique combination of periodicity (square or hexagonal), period (280–650 nm), and aspect ratio (period divided by hole diameter, spanning 1.5–3.0) (Figure 3.2b,c). As a result, the 252 plasmonic tiles support distinctive plasmon resonances in the visible range of optical spectrum, manifesting as unique filter functions for the incident light.

The embedded nanostructures in the encoder are molded from a silicon 'master' chip that contains the desired nanohole array designs. The silicon master was fabricated using Electron-beam lithography (Raith EBPG5000 ES) with a ZEP520A resist (Figure 3.21). After the resist was exposed and developed, a Chlorine etcher (ULVAC NE 500 with 5 sccm Ar, 20 sccm Cl_2) was used to create the nanohole arrays in the silicon. After the production of the master, plasmonic encoder chips were then fabricated using an imprint molding process described earlier[38]. The final encoder chip is comprised of a UV-curable polymer NoA-81 (Norland Products, Inc.) backed by a standard microscope slide, with 50 nm of gold and 5 nm Titanium adhesion layer deposited via Electron Beam Evaporation (CHA Solution).

Experimental procedures

Optical spectra were generated by a programmable supercontinuum laser (~3 nm bandwidth) with up to eight independent emission channels (Fianium, United Kingdom). Random spectra were created by turning on a random number of channels between 1 and 8. For each

channel, the center wavelength was set randomly between 480 and 750 nm, and the power was set randomly between 0.1 and 0.7 (a.u.). All experiments were performed with random spectra, with the exception of the spectra with broader bandwidths (Figure 3.16), which were manually created by overlapping the 8 channels to form a single, broader peak. The output from the laser was coupled to a 50/50 2 × 1 fiber splitter (OZ Optics), with one arm coupled to the input aperture of our device and the other arm coupled to a conventional spectrometer (Ocean Optics HR+) to capture a ground truth spectrum for each measurement. For each spectrum, images were captured by the CMOS image sensor (daA1280-54um by Basler AG, Germany) at ten different exposure times (increasing in length by a factor of two each time) and the resulting images were combined into a single HDR image. Each spectrum was captured by the ground truth spectrometer five times and the resulting spectra were averaged to minimize the effects of noise. Spectra that were over- or under-saturated (in either the ground truth spectrum or the captured HDR image) due to randomness of peak location and power were removed from the dataset.

Training and validation spectra were captured over the course of ~3.75 days. The training dataset consisted of 50,352 spectra, while the validation dataset consisted of 8,824 spectra. Data for blind testing were captured immediately afterward, consisting of 14,648 spectra captured over 15 h of continuous operation of the system. Additionally, another blind testing dataset was captured starting ~5.8 days after the last training/validation spectra were captured.

Spectral reconstruction network

All images were registered to an orthogonal grid to account for gradual drift of the encoder chip relative to the CMOS active area. The 252 × 9 × 9 = 20,412 sub-tile intensities were combined into a vector to serve as the input to the neural network. The spectral reconstruction network

comprises three fully-connected layers of 2,048 neurons each, and an output layer with 1,185 nodes (shown in Figure 3.22). Batch normalization and dropout layers (with a dropout probability of 3%) were used after each fully-connected layer to prevent overfitting to the training spectra. The network was trained using the Adam optimizer[39] with a learning rate of 1e-5.

Appendix

Supplementary figures

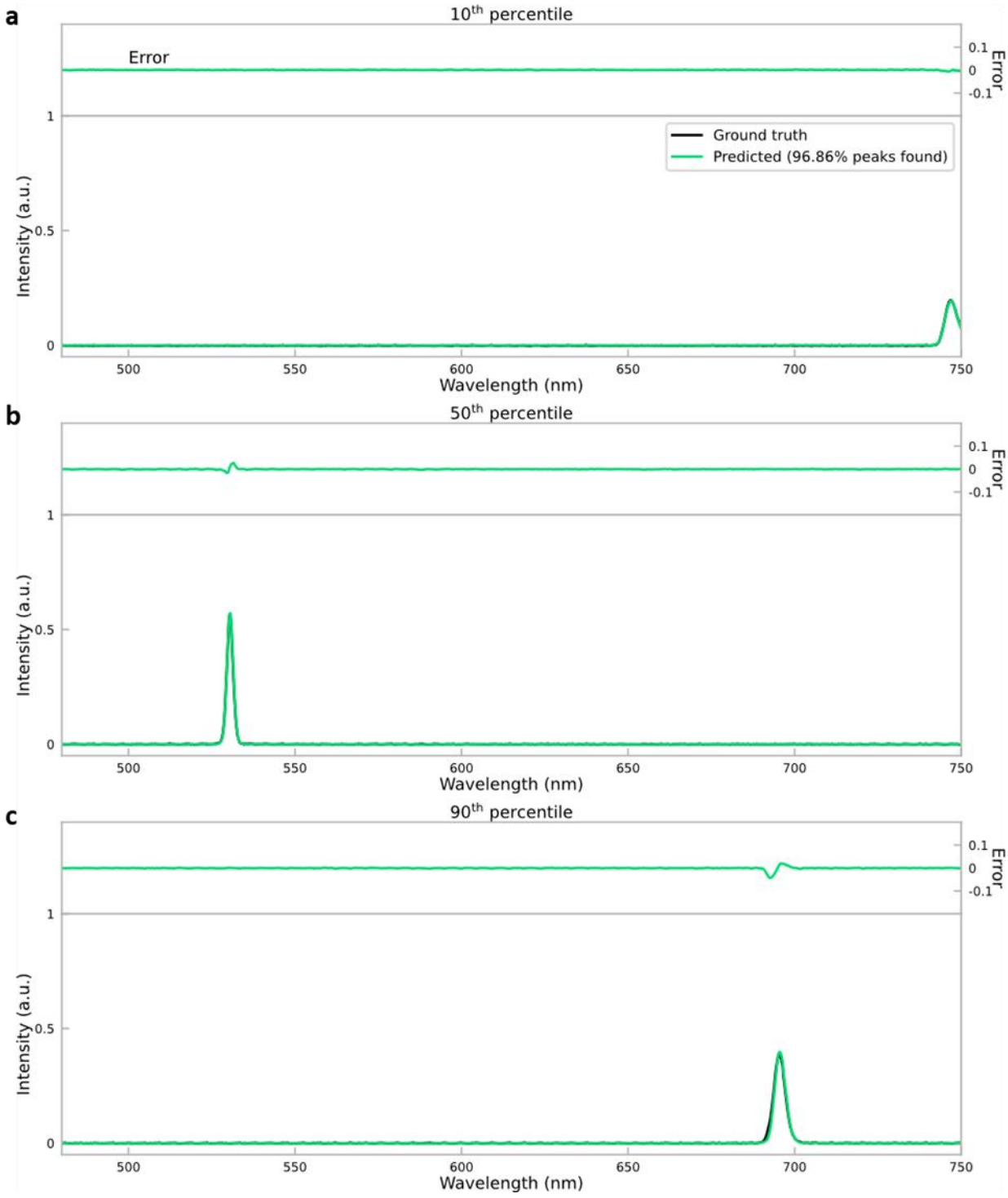


Figure 3.7: Blind reconstructions for spectra with one peak. Spectral reconstructions for a 10th, b 50th, and c 90th percentile of MSE. 10th, 50th, and 90th percentiles correspond to best 10%, median, and worst 10% fits, respectively. Error is shown above each plot on the same y-scale.

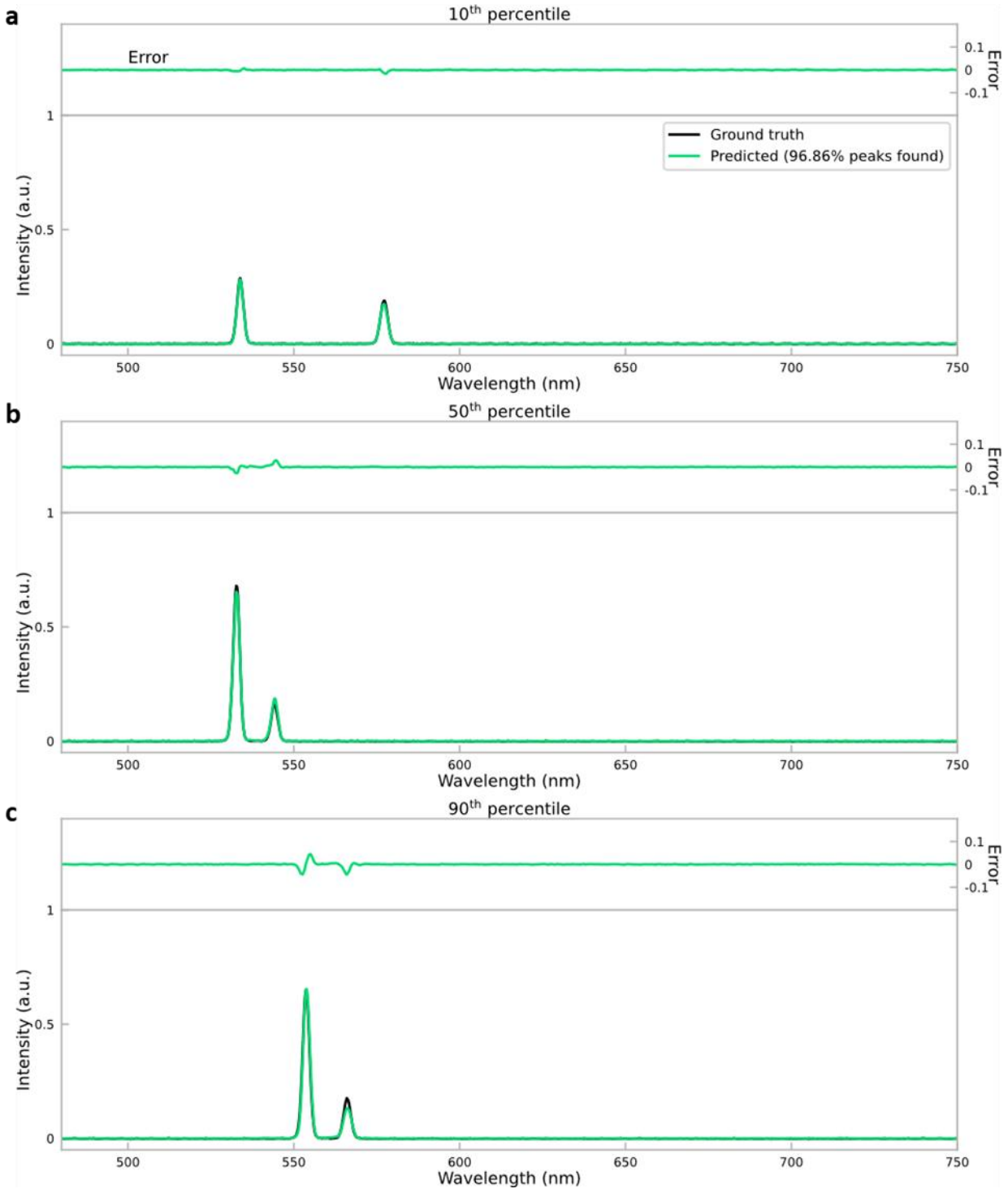


Figure 3.8: Blind reconstructions for spectra with two peaks. Spectral reconstructions for a 10th, b 50th, and c 90th percentile of MSE. 10th, 50th, and 90th percentiles correspond to best 10%, median, and worst 10% fits, respectively. Error is shown above each plot on the same y-scale.

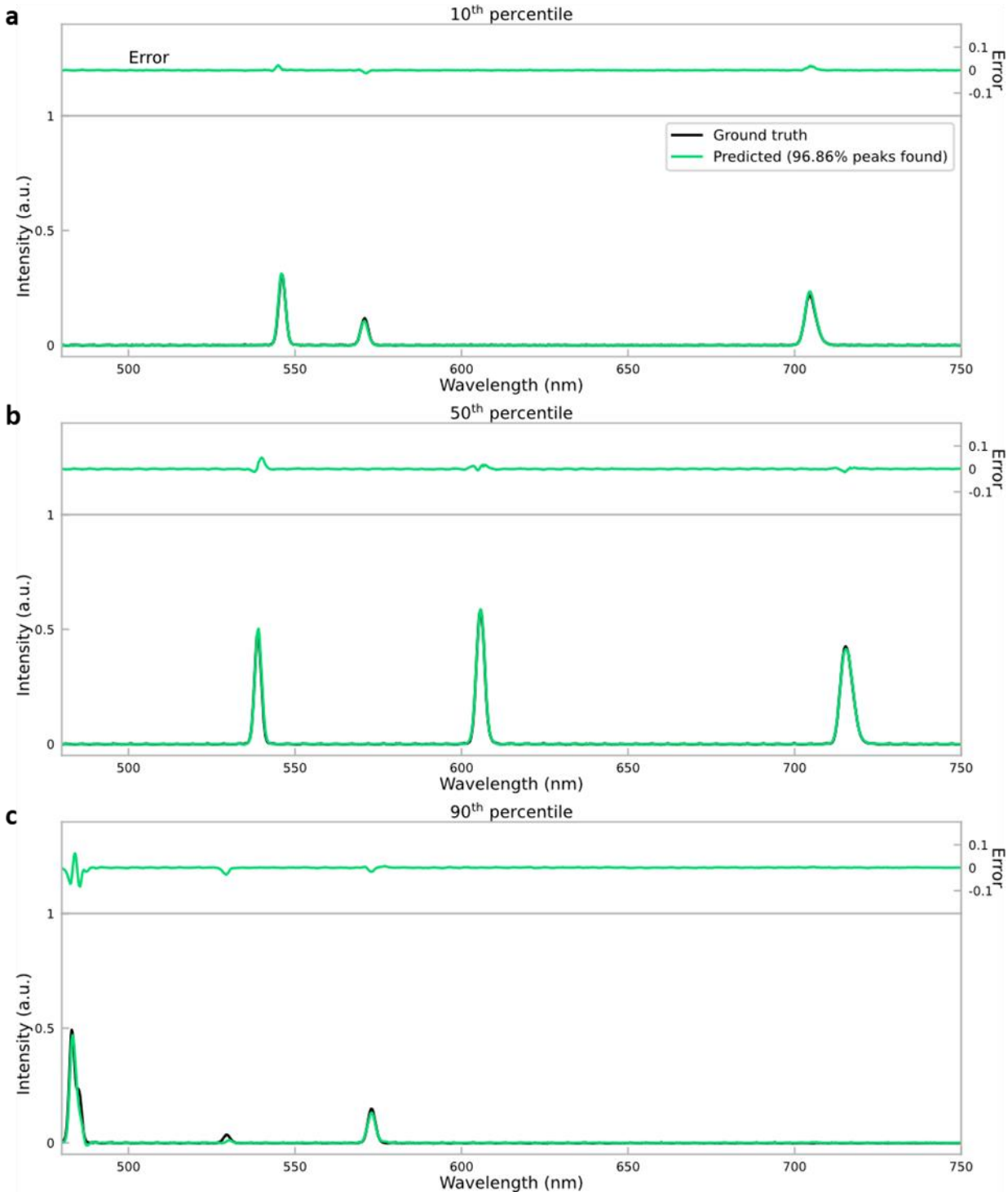


Figure 3.9: Blind reconstructions for spectra with three peaks. Spectral reconstructions for a 10th, b 50th, and c 90th percentile of MSE. 10th, 50th, and 90th percentiles correspond to best 10%, median, and worst 10% fits, respectively. Error is shown above each plot on the same y-scale.

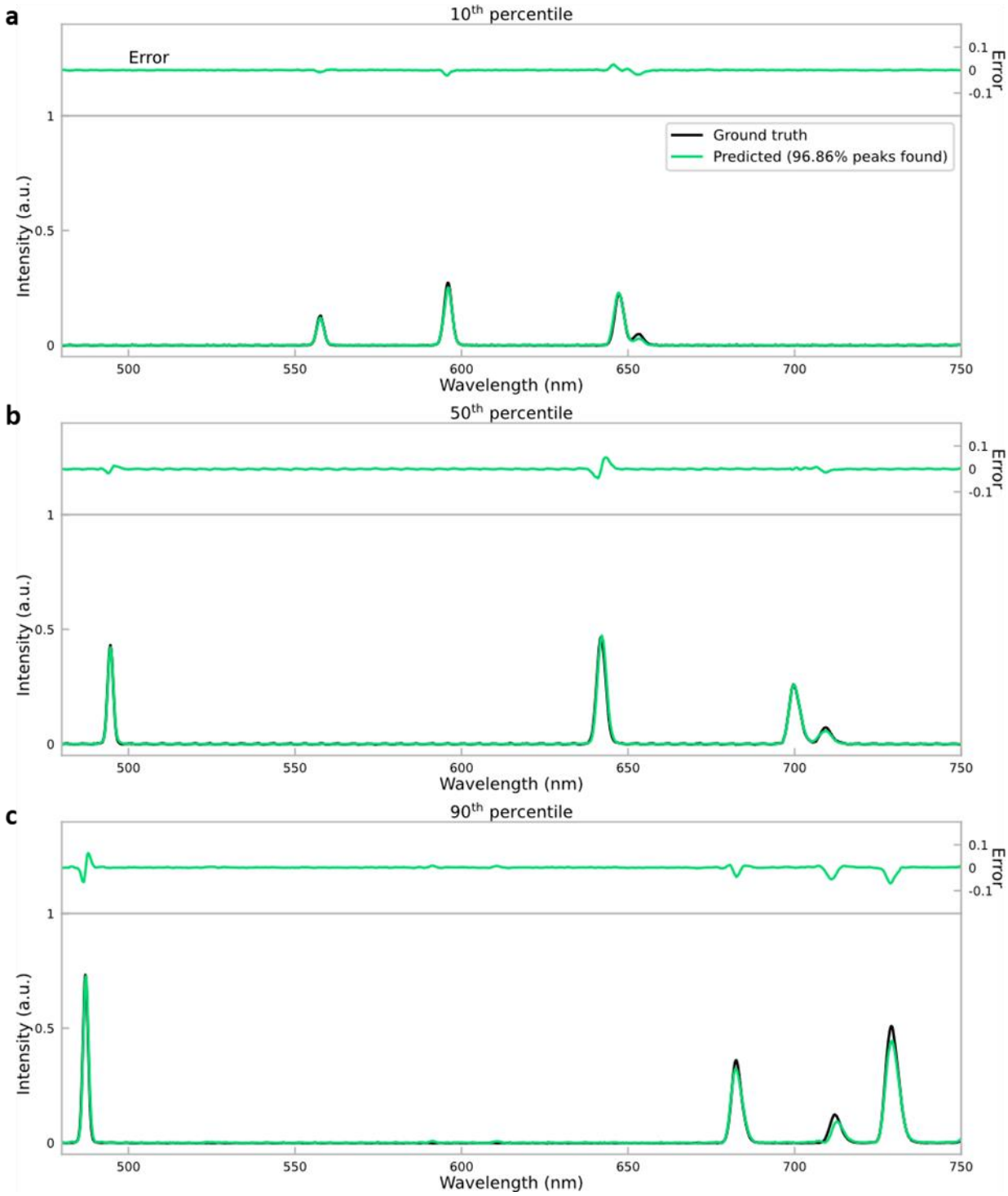


Figure 3.10: Blind reconstructions for spectra with four peaks. Spectral reconstructions for a 10th, b 50th, and c 90th percentile of MSE. 10th, 50th, and 90th percentiles correspond to best 10%, median, and worst 10% fits, respectively. Error is shown above each plot on the same y-scale.

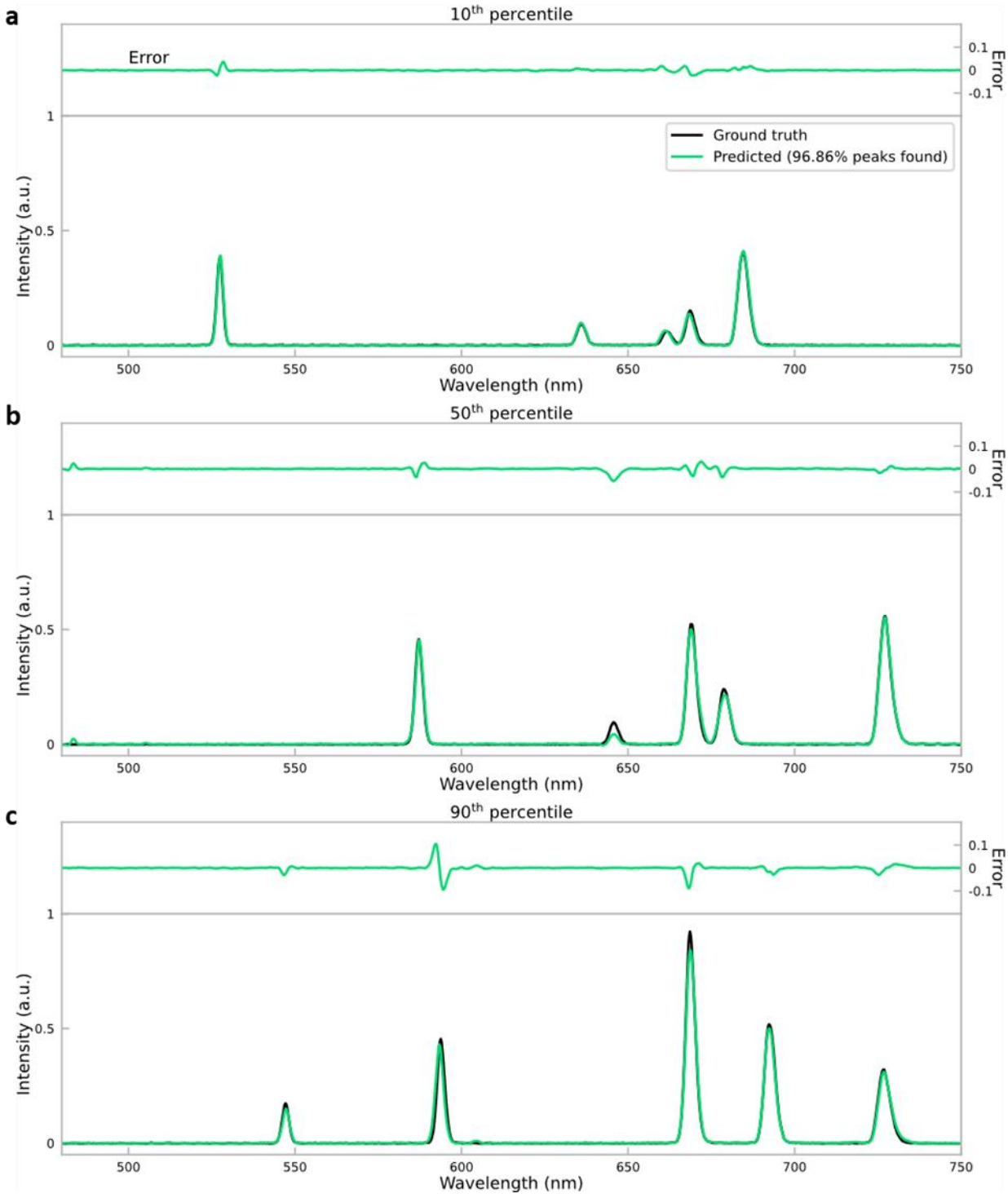


Figure 3.11: Blind reconstructions for spectra with five peaks. Spectral reconstructions for a 10th, b 50th, and c 90th percentile of MSE. 10th, 50th, and 90th percentiles correspond to best 10%, median, and worst 10% fits, respectively. Error is shown above each plot on the same y-scale.

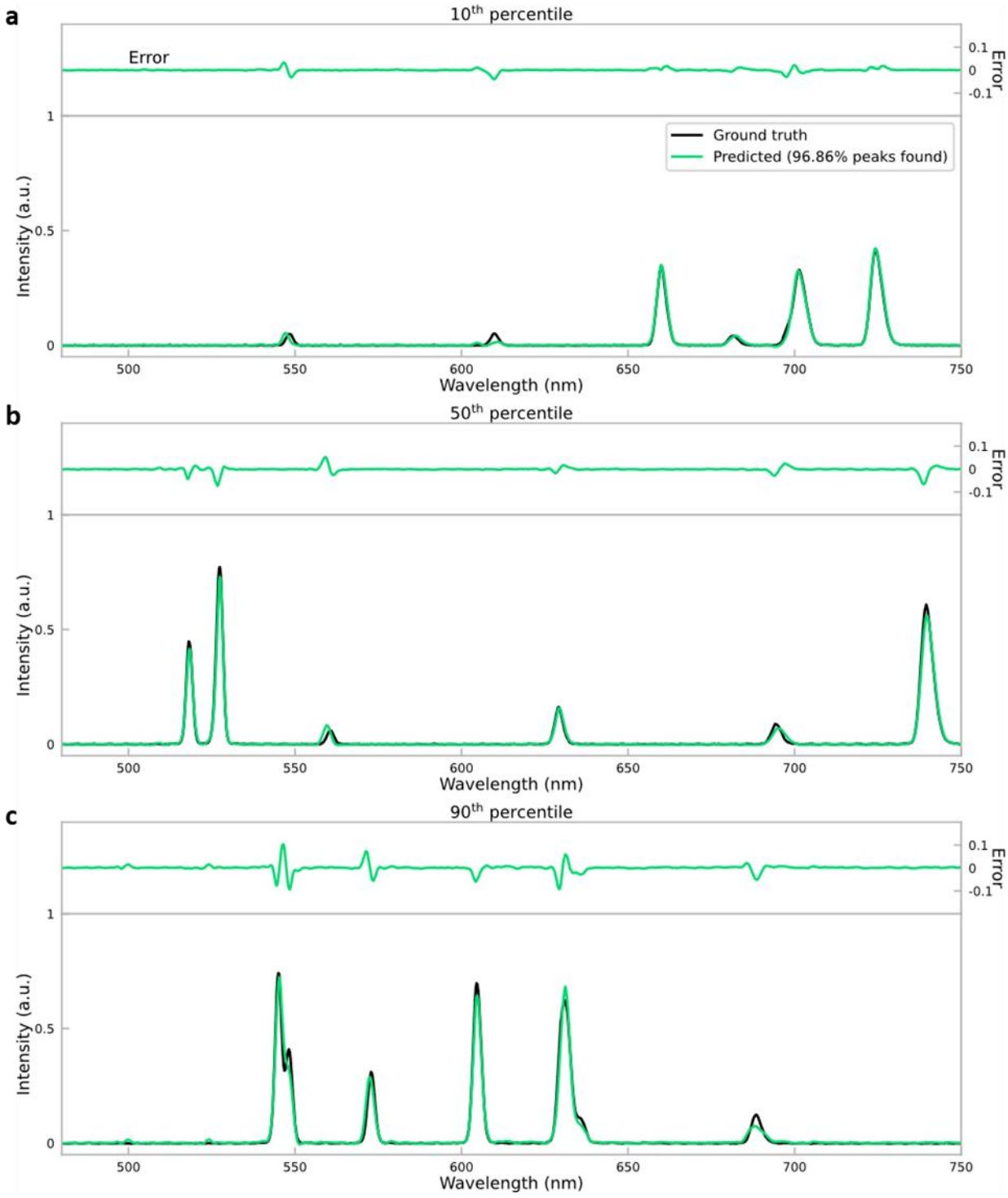


Figure 3.12: Blind reconstructions for spectra with six peaks. Spectral reconstructions for a 10th, b 50th, and c 90th percentile of MSE. 10th, 50th, and 90th percentiles correspond to best 10%, median, and worst 10% fits, respectively. Error is shown above each plot on the same y-scale.

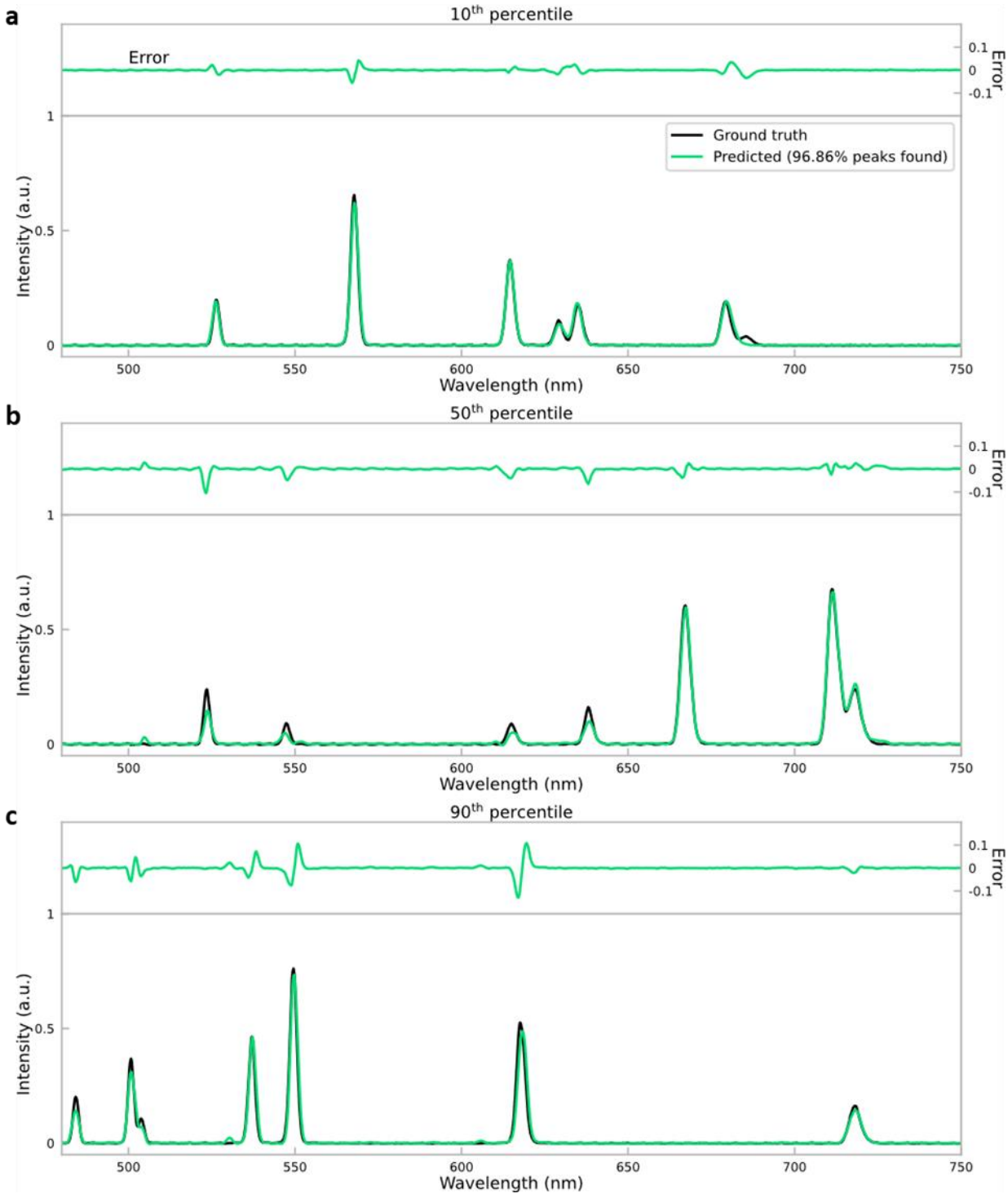


Figure 3.13: Blind reconstructions for spectra with seven peaks. Spectral reconstructions for a 10th, b 50th, and c 90th percentile of MSE. 10th, 50th, and 90th percentiles correspond to best 10%, median, and worst 10% fits, respectively. Error is shown above each plot on the same y-scale.

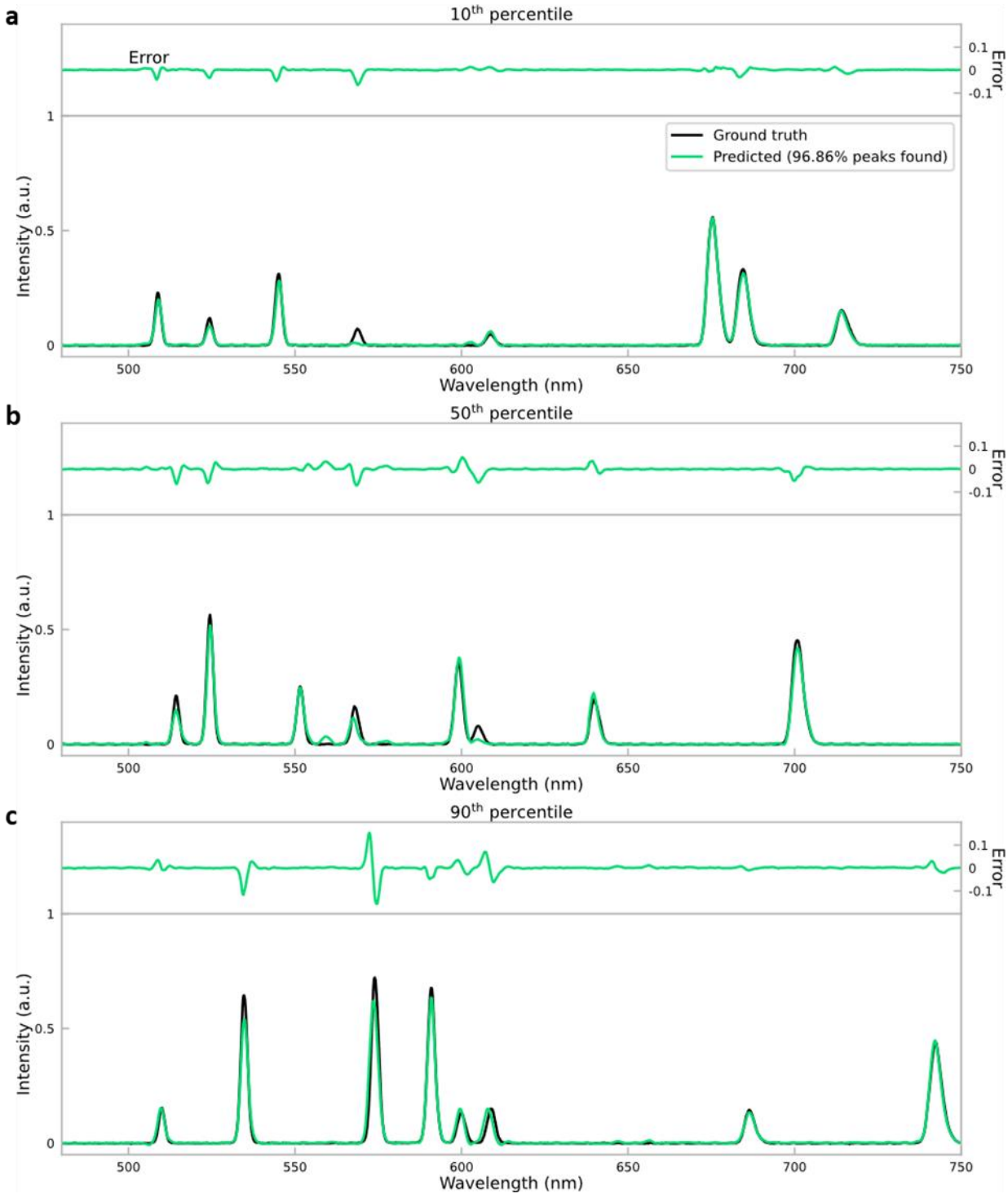


Figure 3.14: Blind reconstructions for spectra with eight peaks. Spectral reconstructions for a 10th, b 50th, and c 90th percentile of MSE. 10th, 50th, and 90th percentiles correspond to best 10%, median, and worst 10% fits, respectively. Error is shown above each plot on the same y-scale.

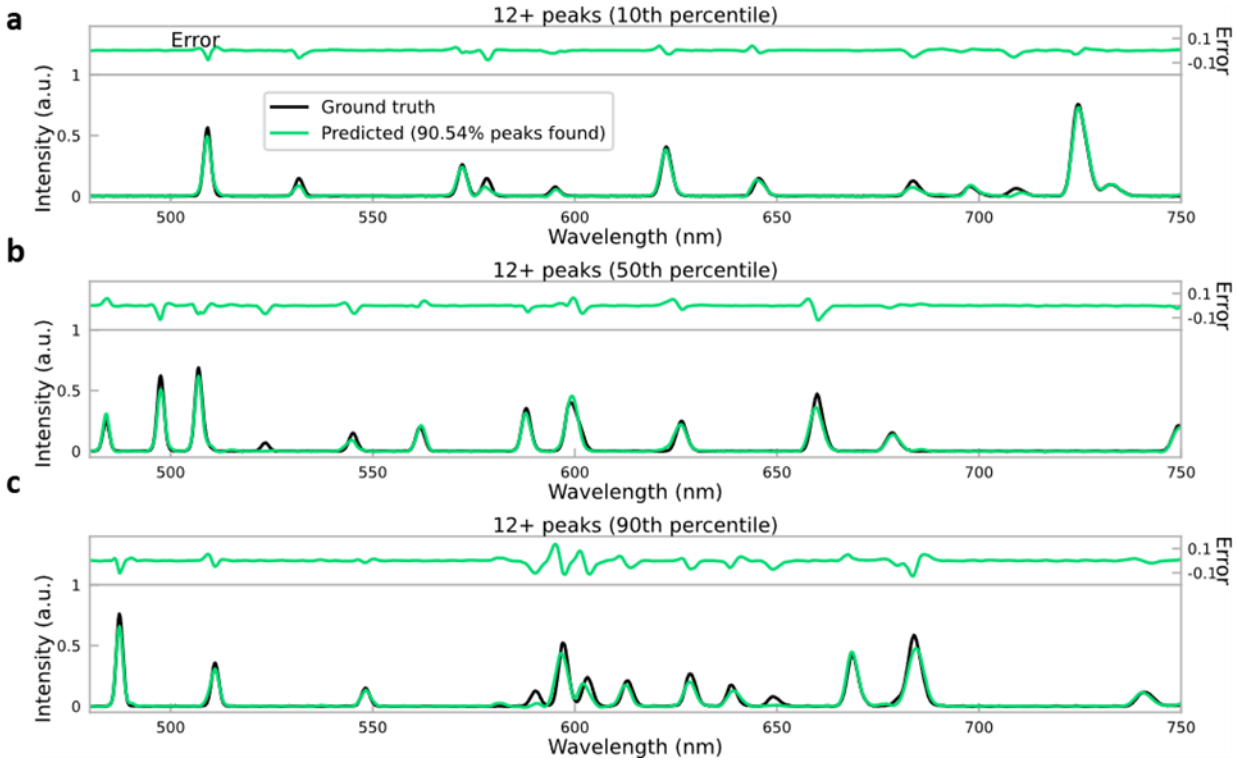


Figure 3.15: Performance on synthesized spectra. Reconstructions for synthesized spectra with 12 or more peaks in the **a** 10th percentile, **b** 50th, and **c** 90th percentile of MSE. While the network was only trained on spectra with up to 8 peaks, it faithfully reconstructs spectra with more peaks. Error is shown above each plot on the same y-scale.

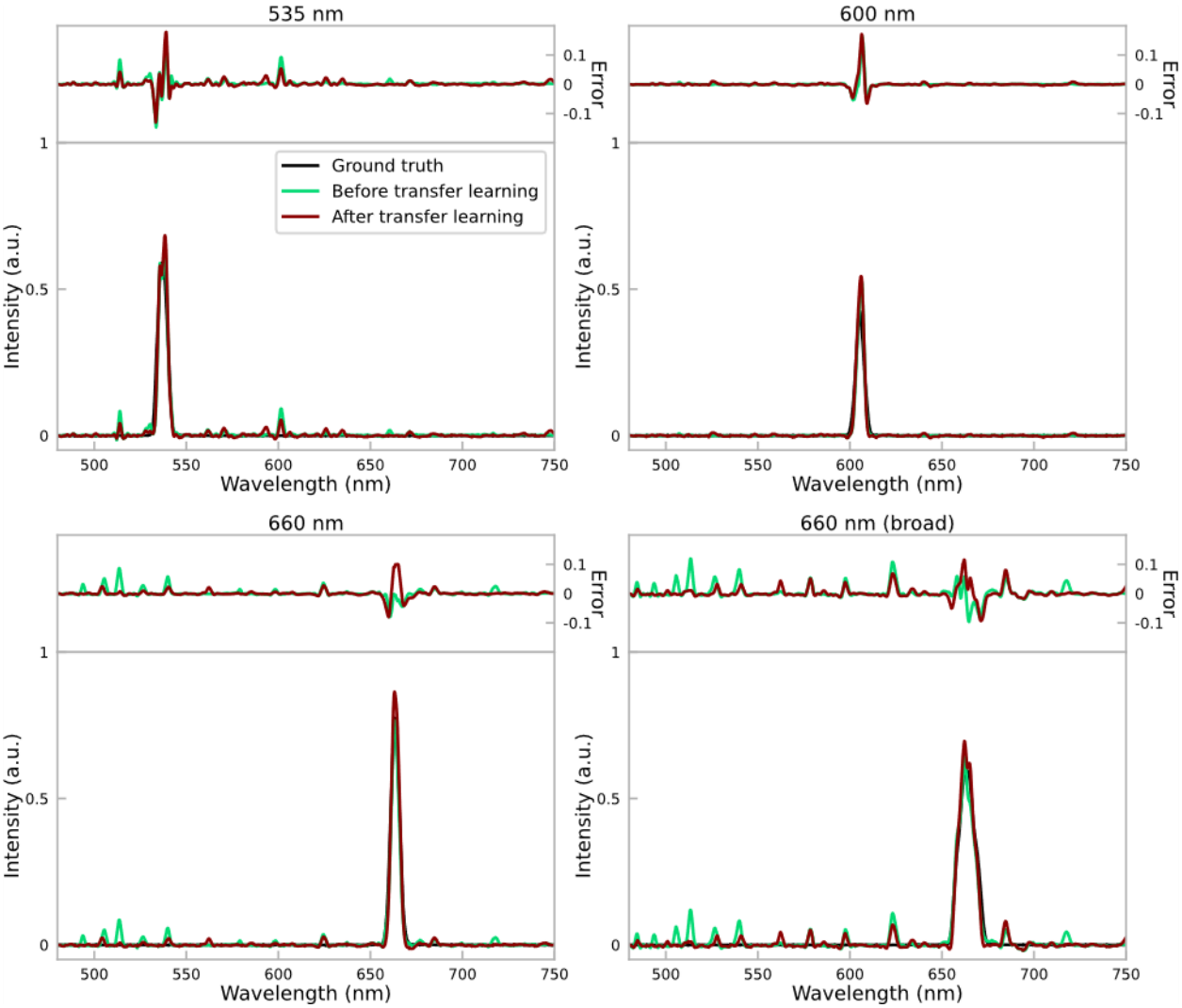


Figure 3.16: Blind reconstruction of spectra with larger bandwidths (never seen during the training phase). a-d Reconstructions of 4 different spectra. These broader-bandwidth spectra were generated by aligning 8 channels of the supercontinuum laser to adjacent wavelengths.

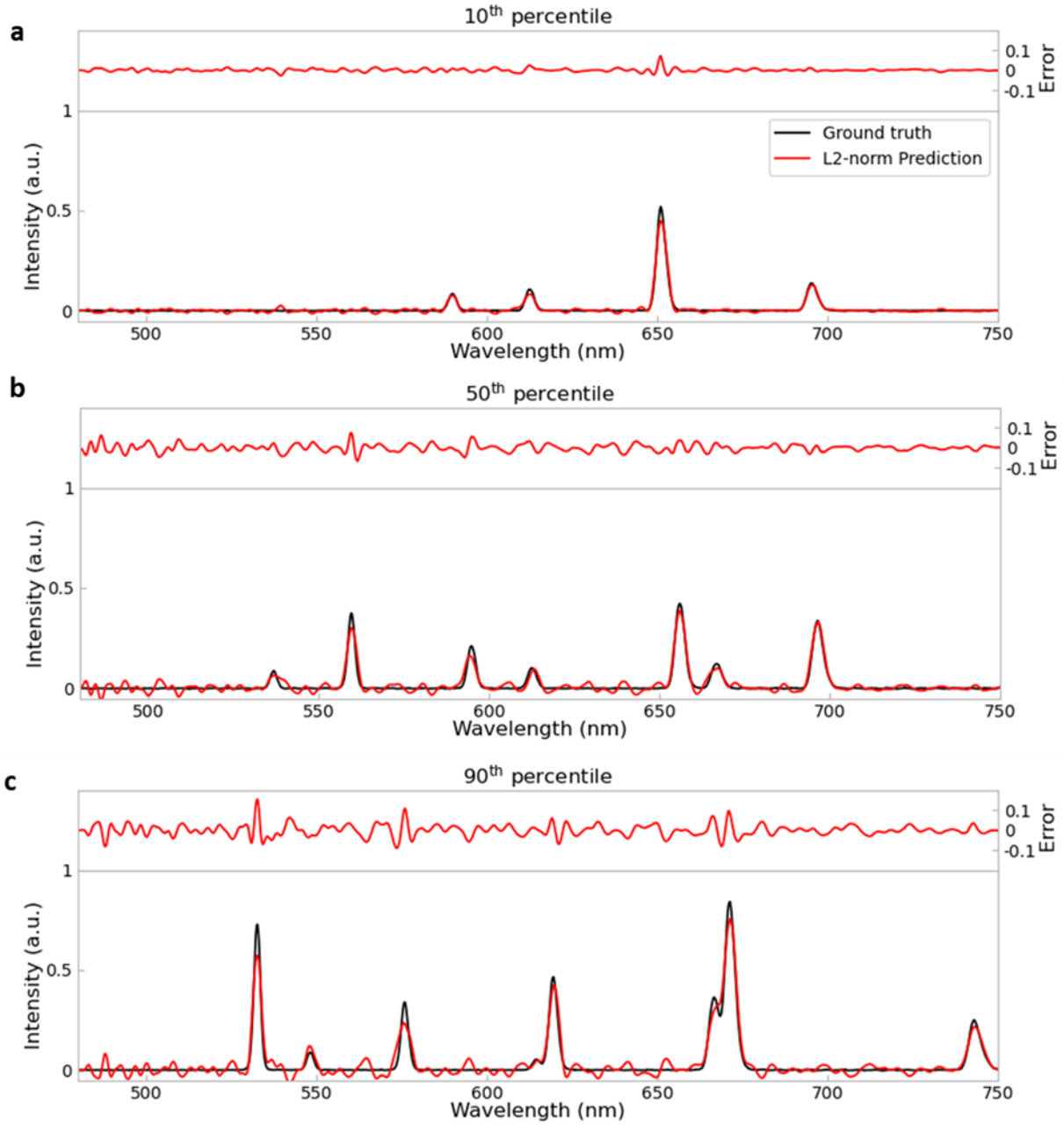


Figure 3.17: Blind reconstructions for spectra using Tikhonov (L_2-norm) regularization (used for comparison to neural network-based reconstructions). Spectral reconstructions for a 10^{th}, b 50^{th}, and c 90^{th} percentile of MSE. 10^{th}, 50^{th}, and 90^{th} percentiles correspond to best 10%, median, and worst 10% fits, respectively. Error is shown above each plot on the same y-scale.

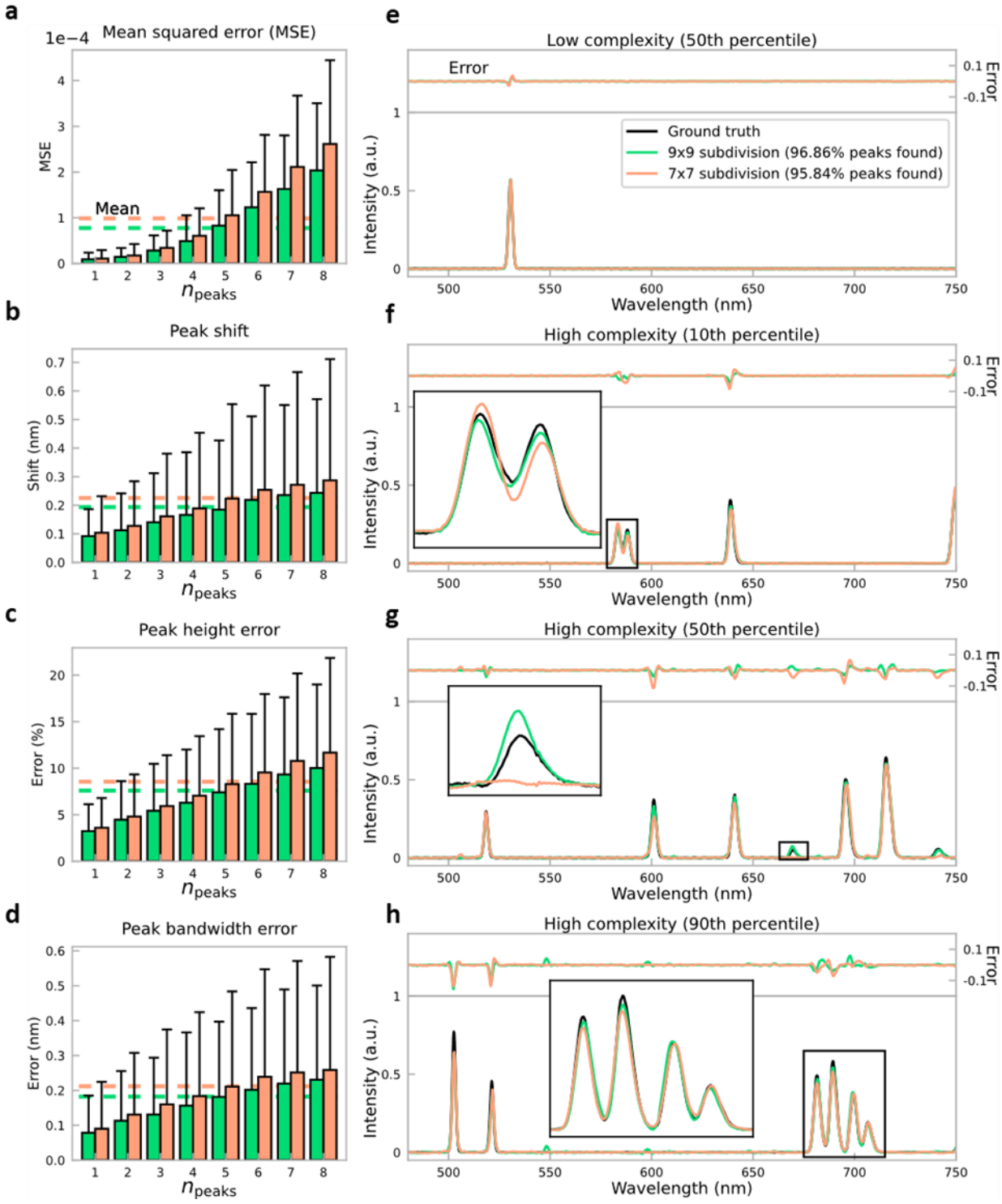
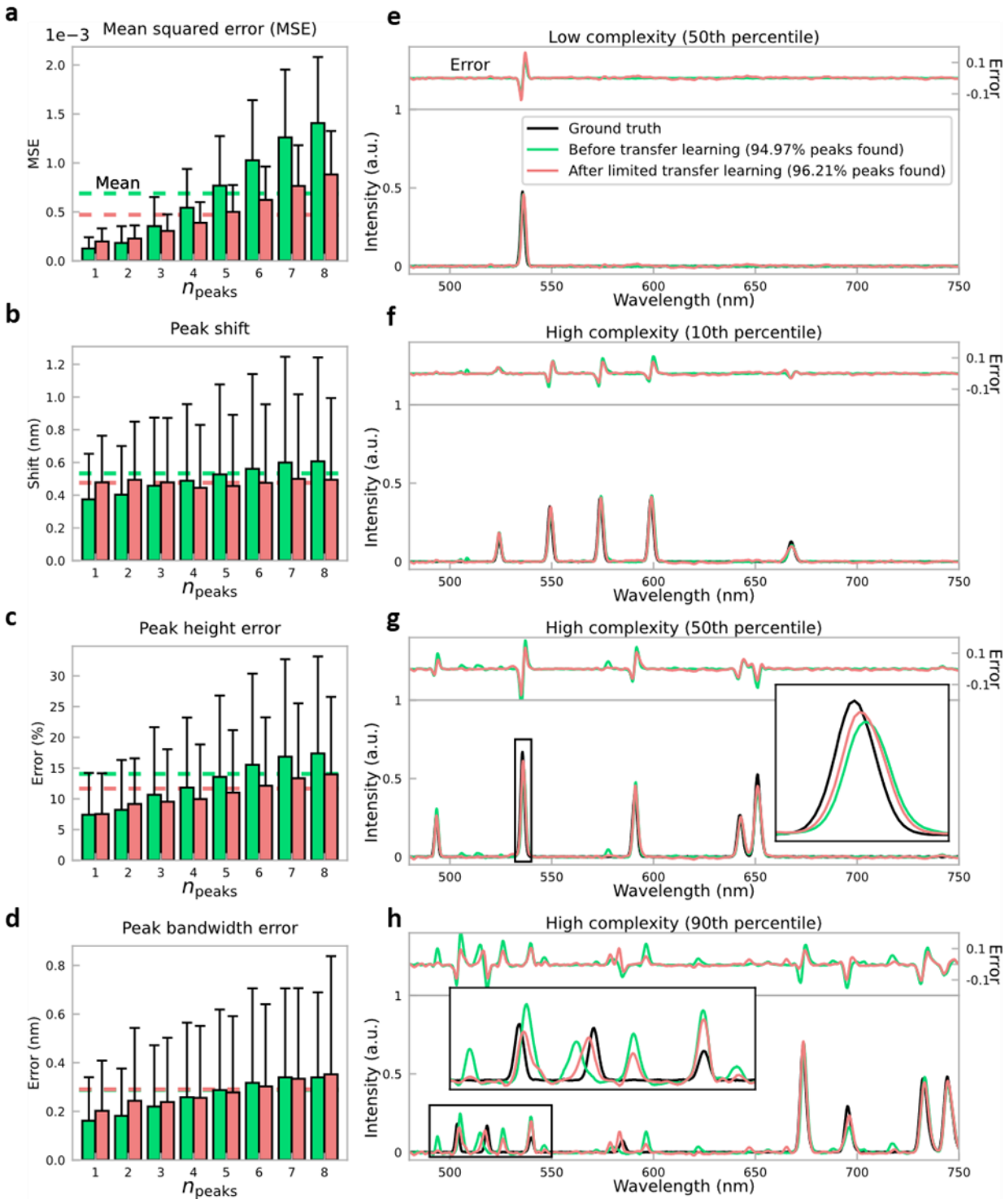

a
Mean squared error (MSE)
1e−4
MSE
Mean
n_{peaks}
b
Peak shift
Shift (nm)
n_{peaks}
c
Peak height error
Error (%)
n_{peaks}
d
Peak bandwidth error
Error (nm)
n_{peaks}
e
Low complexity (50th percentile)
Error
Ground truth
9x9 subdivision (96.86% peaks found)
7x7 subdivision (95.84% peaks found)
Intensity (a.u.)
Wavelength (nm)
f
High complexity (10th percentile)
g
High complexity (50th percentile)
h
High complexity (90th percentile)

Figure 3.19 (previous page): Comparison of 9x9 and 7x7 tile subdivisions. a Average MSE, b peak shift/localization error, c peak height error, and d peak bandwidth error for spectra containing 1-8 peaks. Average over all spectra shown as horizontal line in each plot. The 9x9 tile subdivision improves over the 7x7 for all metrics. Reconstruction for low complexity (1 peak) spectrum in the e 50th percentile of MSE. Reconstructions for high complexity (4-8 peaks) spectra in the f 10th, g 50th, and h 90th percentile of MSE. Error is shown above each plot on the same y-scale.

Figure 3.18 (next page): Performance on blind testing spectra captured 5.8 days after training with limited transfer learning. a Average MSE, b peak shift/localization error, c peak height error, and d peak bandwidth error for spectra containing 1-8 peaks. Transfer learning on just 100 spectra (~0.2% of initial training spectra, requiring just ~6 min to capture) improves MSE, peak localization error, and peak height error, and suppresses noise/spurious peaks. Reconstruction for low complexity (1 peak) spectrum in the e 50th percentile of MSE. Reconstructions for high complexity (4-8 peaks) spectra in the f 10th, g 50th, and h 90th percentile of MSE. Error is shown above each plot on the same y-scale.


a
Mean squared error (MSE)
1e−3
MSE
Mean
n_{peaks}
b
Peak shift
Shift (nm)
n_{peaks}
c
Peak height error
Error (%)
n_{peaks}
d
Peak bandwidth error
Error (nm)
n_{peaks}
e
Low complexity (50th percentile)
Error
Ground truth
Before transfer learning (94.97% peaks found)
After limited transfer learning (96.21% peaks found)
Intensity (a.u.)
Wavelength (nm)
f
High complexity (10th percentile)
g
High complexity (50th percentile)
h
High complexity (90th percentile)

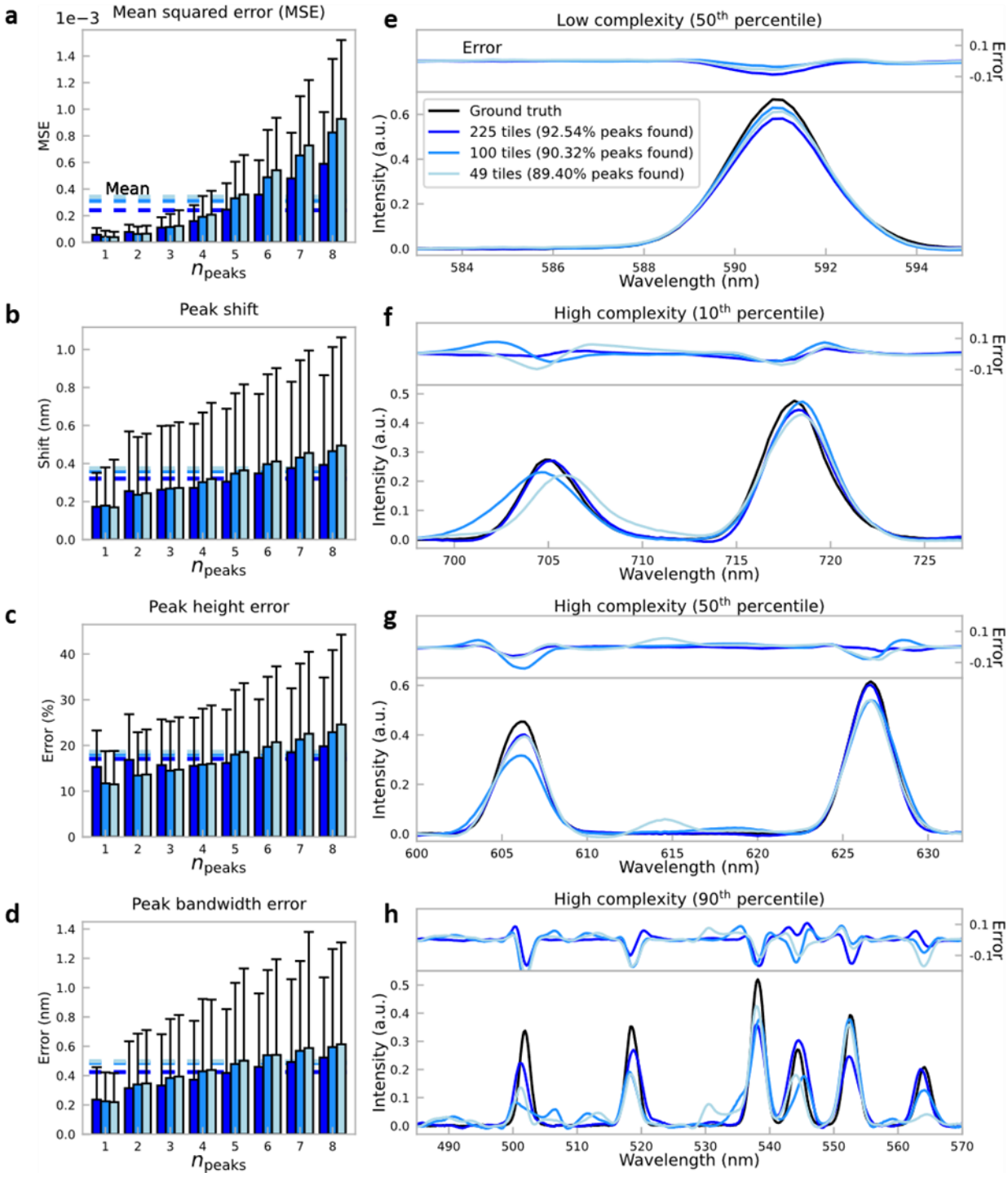

a
Mean squared error (MSE)
1e−3
MSE
Mean
n_{peaks}
b
Peak shift
Shift (nm)
n_{peaks}
c
Peak height error
Error (%)
n_{peaks}
d
Peak bandwidth error
Error (nm)
n_{peaks}
e
Low complexity (50th percentile)
Error
Ground truth
225 tiles (92.54% peaks found)
100 tiles (90.32% peaks found)
49 tiles (89.40% peaks found)
Intensity (a.u.)
Wavelength (nm)
f
High complexity (10th percentile)
Intensity (a.u.)
Wavelength (nm)
g
High complexity (50th percentile)
Intensity (a.u.)
Wavelength (nm)
h
High complexity (90th percentile)
Intensity (a.u.)
Wavelength (nm)

Figure 3.20 (previous page): Comparison between models using different numbers of plasmonic tiles. a Average MSE, b peak shift/localization error, c peak height error, and d peak bandwidth error for spectra containing 1-8 peaks. Average over all spectra shown as horizontal line in each plot. Using more tiles improves performance, but even using a fraction of the tiles allows reconstruction of most peaks. It must be noted that the networks shown here were of simplified architecture to speed training of many models. Reconstruction for low complexity (1 peak) spectrum in the e 50^{th} percentile of MSE. Reconstructions for high complexity (4-8 peaks) spectra in the f 10^{th}, g 50^{th}, and h 90^{th} percentile of MSE. Error is shown above each plot on the same y-scale.

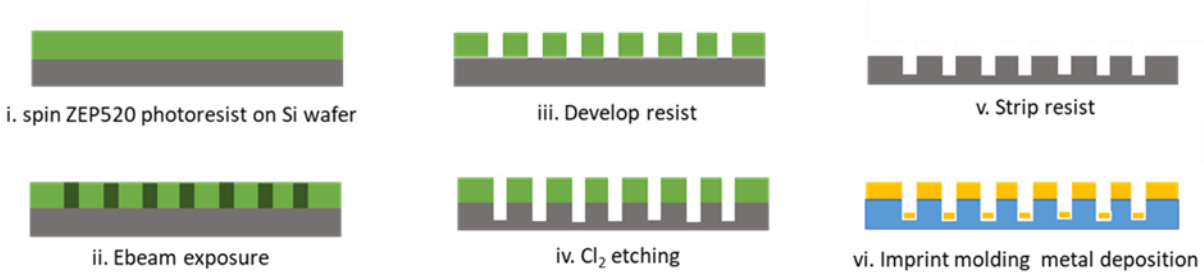


Figure 3.21: Spectral encoder fabrication process.

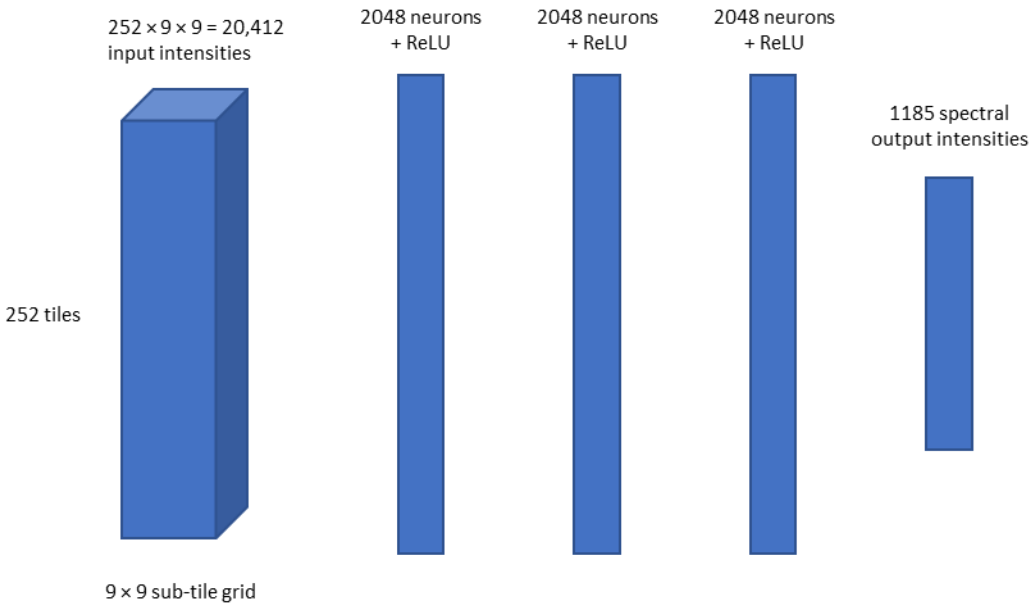


Figure 3.22: Spectral reconstruction network architecture. All layers are dense/fully-connected layers. The network was trained with additional batch normalization and dropout layers after each 2048-neuron hidden layer. Additional networks were trained with fewer than 252 tiles or fewer than 9 × 9 sub-tiles, in which case the input layer was scaled accordingly while the rest of the network remained unchanged.

References

1. M. Czerny & Turner, A. F. Uber den Astigmatismus bei Spiegelspektrometern. **61**, 792–797 (1930).

2. McClelland, A. & Mankin, M. *Optical measurements for scientists and engineers: a practical guide*. (Cambridge University Press, 2018).

3. Palmer, C. Diffraction Grating Handbook. 271.

4. Bao, J. & Bawendi, M. G. A colloidal quantum dot spectrometer. *Nature* **523**, 67–70 (2015).

5. Wang, Z. *et al.* Single-shot on-chip spectral sensors based on photonic crystal slabs. *Nature Communications* **10**, 1020 (2019).

6. Song, H., Zhang, W., Li, H., Liu, X. & Hao, X. Review of compact computational spectral information acquisition systems. *Front Inform Technol Electron Eng* (2020) doi:10.1631/FITEE.1900266.

7. Wang, Z. & Yu, Z. Spectral analysis based on compressive sensing in nanophotonic structures. *Opt. Express, OE* **22**, 25608–25614 (2014).

8. Jang, W.-Y. *et al.* Experimental Demonstration of Adaptive Infrared Multispectral Imaging using Plasmonic Filter Array. *Scientific Reports* **6**, 34876 (2016).

9. Redding, B., Liew, S. F., Sarma, R. & Cao, H. Compact spectrometer based on a disordered photonic chip. *Nature Photonics* **7**, 746–751 (2013).

10. Degraux, K., Cambareri, V., Geelen, B., Jacques, L. & Lafruit, G. Multispectral Compressive Imaging Strategies Using Fabry–Pérot Filtered Sensors. *IEEE Transactions on Computational Imaging* **4**, 661–673 (2018).

11. Chang, C.-C. & Lee, H.-N. On the estimation of target spectrum for filter-array based spectrometers. *Opt. Express, OE* **16**, 1056–1061 (2008).

12. Oliver, J., Lee, W.-B. & Lee, H.-N. Filters with random transmittance for improving resolution in filter-array-based spectrometers. *Opt. Express, OE* **21**, 3969–3989 (2013).

13. Paskaleva, B., Hayat, M. M., Wang, Z., Tyo, J. S. & Krishna, S. Canonical Correlation Feature Selection for Sensors With Overlapping Bands: Theory and Application. *IEEE Transactions on Geoscience and Remote Sensing* **46**, 3346–3358 (2008).

14. Xia, L., Zhang, A., Li, T. & Zou, Y. Deep-learning-assisted on-chip Fourier transform spectrometer. in *Integrated Optics: Devices, Materials, and Technologies XXIV* vol. 11283 1128305 (International Society for Optics and Photonics, 2020).

15. Paudel, U. Ultra-high resolution and broadband chip-scale speckle enhanced Fourier-transform spectrometer. in *OSA Advanced Photonics Congress (AP) 2019 (IPR, Networks, NOMA, SPPCom, PVLED) (2019), paper IW2A.4* IW2A.4 (Optical Society of America, 2019).

16. August, Y. & Stern, A. Compressive sensing spectrometry based on liquid crystal devices. *Opt. Lett., OL* **38**, 4996–4999 (2013).

17. Yang, Z. *et al.* Single-nanowire spectrometers. *Science* **365**, 1017–1020 (2019).

18. Sarwar, T., Cheekati, S., Chung, K. & Ku, P.-C. On-chip optical spectrometer based on GaN wavelength-selective nanostructural absorbers. *Appl. Phys. Lett.* **116**, 081103 (2020).

19. Zhu, Y., Lei, X., Wang, K. X. & Yu, Z. Compact CMOS spectral sensor for the visible spectrum. *Photon. Res., PRJ* **7**, 961–966 (2019).

20. Fathy, A., Sabry, Y. M., Nazeer, S., Bourouina, T. & Khalil, D. A. On-chip parallel Fourier transform spectrometer for broadband selective infrared spectral sensing. *Microsystems & Nanoengineering* **6**, 1–9 (2020).

21. Xia, Z. *et al.* High resolution on-chip spectroscopy based on miniaturized microdonut resonators. *Opt. Express, OE* **19**, 12356–12364 (2011).

22. Liapis, A. C., Gao, B., Siddiqui, M. R., Shi, Z. & Boyd, R. W. On-chip spectroscopy with thermally tuned high-Q photonic crystal cavities. *Appl. Phys. Lett.* **108**, 021105 (2016).

23. Li, Z., Palacios, E., Butun, S. & Aydin, K. Visible-Frequency Metasurfaces for Broadband Anomalous Reflection and High-Efficiency Spectrum Splitting. *Nano Lett.* **15**, 1615–1621 (2015).

24. Huang, E., Ma, Q. & Liu, Z. Etalon Array Reconstructive Spectrometry. *Scientific Reports* **7**, 40693 (2017).

25. Zheng, S. N. *et al.* Microring resonator-assisted Fourier transform spectrometer with enhanced resolution and large bandwidth in single chip solution. *Nature Communications* **10**, 2349 (2019).

26. Tibshirani, R. Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)* **58**, 267–288 (1996).

27. Tikhonov, A. N. On the solution of ill-posed problems and the method of regularization. *Dokl. Akad. Nauk SSSR* **151**, 501–504 (1963).

28. Kurokawa, U., Choi, B. I. & Chang, C.-C. Filter-Based Miniature Spectrometers: Spectrum Reconstruction Using Adaptive Regularization. *IEEE Sensors Journal* **11**, 1556–1563 (2011).

29. Kita, D. M. *et al.* High-performance and scalable on-chip digital Fourier transform spectroscopy. *Nature Communications* **9**, 4405 (2018).

30. Hinton, G. E., Srivastava, N., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R. R. Improving neural networks by preventing co-adaptation of feature detectors. *arXiv:1207.0580 [cs]* (2012).

31. Krizhevsky, A., Sutskever, I. & Hinton, G. E. ImageNet Classification with Deep Convolutional Neural Networks. in *Advances in Neural Information Processing Systems 25* (eds. Pereira, F., Burges, C. J. C., Bottou, L. & Weinberger, K. Q.) 1097–1105 (Curran Associates, Inc., 2012).

32. Simonyan, K. & Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv:1409.1556 [cs]* (2015).

33. Wu, Y. *et al.* Three-dimensional virtual refocusing of fluorescence microscopy images using deep learning. *Nature Methods* **16**, 1323–1331 (2019).

34. Wang, H. *et al.* Deep learning enables cross-modality super-resolution in fluorescence microscopy. *Nat Methods* **16**, 103–110 (2019).

35. Haan, K. de, Rivenson, Y., Wu, Y. & Ozcan, A. Deep-Learning-Based Image Reconstruction and Enhancement in Optical Microscopy. *Proceedings of the IEEE* **108**, 30–50 (2020).

36. Barbastathis, G., Ozcan, A. & Situ, G. On the use of deep learning for computational imaging. *Optica, OPTICA* **6**, 921–943 (2019).

37. Anzanello, M. J. *et al.* A genetic algorithm-based framework for wavelength selection on sample categorization. *Drug Testing and Analysis* **9**, 1172–1181 (2017).

38. Ballard, Z. S. *et al.* Computational Sensing Using Low-Cost and Mobile Plasmonic Readers Designed by Machine Learning. *ACS Nano* **11**, 2266–2274 (2017).

39. Kingma, D. P. & Ba, J. Adam: A Method for Stochastic Optimization. *arXiv:1412.6980 [cs]* (2017).

Chapter 4 Conclusion

In this dissertation, I have presented two novel frameworks for computational sensing in two markedly different applications: a deep learning-enabled, automated AST system and an on-chip framework for designing custom, compact spectrometers. Both systems were designed from the ground up with both hardware sensing and software reconstruction algorithms in mind, leading to novel architectures with unique capabilities that are not available in the systems conventionally used for these tasks.

The AST framework uses cost-effective components and machine learning to address the primary drawbacks of the gold standard method of testing for resistance: the need to wait 18–24 h for incubation and the need for a trained medical technologist (who inevitably is susceptible to human error). Conventional imaging approaches, such as scanning microscopy over the entire 96-well plate, are not well suited to the task and would require bulkier, more expensive hardware, as well as calibration and most likely a trained technician. Similarly, the on-chip plasmonic spectroscopy framework also avoids many of the pitfalls that plague conventional system design (in this case, grating-based spectrometers). By partially avoiding the trade-off between spectral resolution and signal strength as well as employing a neural network to drastically improve reconstruction speed compared to other computational spectroscopy approaches, the plasmonic approach demonstrates the potential for the design of sensitive, high resolution spectrometers for various applications.

Rethinking conventional sensing and imaging systems—armed with the tools of machine learning and compressive sensing—has already led to new designs in applications beyond those I have mentioned, in which more compact, lower cost, higher sensitivity, lower power consumption,

remote operation, or other characteristics are needed. I believe new designs will continue to emerge, and crucially, those that have already emerged will continue to mature and improve due to the ease of iterative design and testing cycles. We have only just begun down the road of computational system design, and there remains much work to do in combining these newfound computational tools with established optical systems.