DEEP-LEARNING-ASSISTED FOURIER TRANSFORM IMAGING SPECTROSCOPY FOR HYPERSPECTRAL FLUORESCENCE IMAGING

A PREPRINT

Cory Juntunen
College of Engineering and Applied Science
University of Wisconsin-Milwaukee
Milwaukee, WI 53211
juntune3@uwm.edu

Isabel M. Woller
College of Health Sciences
University of Wisconsin-Milwaukee
Milwaukee, WI 53211
imwoller@uwm.edu

Andrew R. Abramczyk
College of Engineering and Applied Science
University of Wisconsin-Milwaukee
Milwaukee, WI 53211
ara@uwm.edu

Yongjin Sung
College of Engineering and Applied Science
University of Wisconsin-Milwaukee
Milwaukee, WI 53211
ysung4@uwm.edu

August 18, 2021

ABSTRACT

Hyperspectral fluorescence imaging is widely used when multiple fluorescent probes with close emission peaks are required. In particular, Fourier transform imaging spectroscopy (FTIS) provides unrivaled spectral resolution; however, the imaging throughput is very low due to the amount of interferogram sampling required. In this work, we apply deep learning to FTIS and show that the interferogram sampling can be drastically reduced by an order of magnitude without noticeable degradation in the image quality. For the demonstration, we use bovine pulmonary artery endothelial cells stained with three fluorescent dyes. Further, we show that the deep learning approach is more robust to the translation stage error and environmental vibrations. Thereby, the He-Ne correction, which is typically required for FTIS, can be bypassed, thus reducing the cost, size, and complexity of the FTIS system. Finally, we construct neural network models using Hyperband, an automatic hyperparameter selection algorithm, and compare the performance with our manually-optimized model.

Keywords Deep learning · Fourier transform imaging spectroscopy · Fluorescence imaging

1 Introduction

Fluorescence imaging allows for direct observation of various organelles in a biological specimen with high resolution and contrast. It typically uses fluorescent dyes which bind to different key targets/organelles of the sample or fluorescent proteins (FPs) that are fused to protein targets in live-cells[13]. Each fluorophore requires the use of a dichroic filter tuned for the characteristic excitation and emission bands of the fluorophore. For the observation of more than one fluorophore, a dichroic filter with multiple passbands or a set of filters mounted on a filter wheel is typically adopted. Many FPs commonly used in live-cell imaging have overlapping emission spectra, which limit the number of FPs that can be used simultaneously[27]. Hyperspectral imaging, which combines imaging and spectroscopy, allows for using a multitude of fluorophores with close emission peaks[24, 23]. It also allows for accurate detection and quantification of target fluorescence signals in a tissue with highly autofluorescent background[16].

For hyperspectral imaging with the spectral resolution of 10 nm or below, an acousto-optic tunable filter (AOTF), a liquid crystal tunable filter (LCTF), or Fourier transform spectroscopy (FTS) is typically combined with an imaging device. AOTF enables us to scan the entire visible wavelength range in several seconds or randomly access to any

wavelength in the range[25]. LCTF is slower but can provide superior image quality[1]. AOTF provides a narrower spectral bandwidth (1.5 – 4.1 nm) than LCTF (4.5 – 19 nm)[1]. FTS is a preferred method when a high spectral resolution is required. For FTS, a Michelson interferometer or a Sagnac interferometer is typically used, which splits the interrogated light into two, and a series of intensity data is acquired for varying optical path differences (OPDs)[9]. The Fourier transform of the interferogram can be related to the spectral profile of the interrogated light. Fourier transform imaging spectroscopy (FTIS) combines FTS with an imaging device which is used to measure the spectrum of each pixel of an image. FTIS has been used for spectral karyotyping, which uses a combination of 4-5 fluorophores to label 24 human chromosome pairs[11]. FTIS has also been used to classify seven fluorophores with overlapping emission spectra in immunofluorescence-stained tissue samples[24]. One key disadvantage of FTIS is low throughput due to the large number of interferogram images that need to be collected to reconstruct the spectrum. Typically, more than 1000 images are acquired, which can take tens of seconds even with a high-speed camera. Several efforts have been made to reduce sampling number with methods such as compressed sensing[4], which has also been demonstrated in fluorescence microscopy for biological and hyperspectral imaging[22, 5, 15]. Since the interferogram measurements of FTIS are taken in the Fourier space, the signal measurement procedure for FTIS satisfies the incoherence property which is a requirement of compressed sensing[6]. Here, we show that deep learning can reduce the required FTIS sampling number to about 50. We also show that deep-learning-assisted FTIS is more robust to the translation stage error and environmental vibrations than conventional FTIS. Thereby, the optical components typically needed for the correction (called He-Ne correction) can be removed. Deep learning-based approaches have shown to outperform traditional methods in a wide array of fields including imaging and natural language processing[10, 19]. For example, deep learning has been applied to improve the resolution of optical microscopy[20], scanning electron microscopy[3], and multispectral imaging[26]. The enhanced imaging performance has been used to accurately identify components of images that are indicative of specific diagnosis[10, 17, 7].

2 Materials and Methods

2.1 FTIS Optical System

For the data collection, we have built a wide-field epi-fluorescence microscope equipped with a laboratory-built FTS module. Figure 1(a) shows a schematic diagram of the system. For the light source (LS), three individually-controlled light-emitting diodes (Thorlabs, M385L2, M505L4, and M565L3) were combined using two dichroic beam splitters (Thorlabs, DMLP425R and DMLP550R). The excitation light was introduced to the sample through a triple-band filter set (Semrock, DA/FI/TX-3X). The fluorescence light emitted from the sample was collected by the objective lens (OL) (Olympus, UPLFLN 100X) with the numerical aperture adjusted to 0.6. The tube lens (TL) of 100 mm focal length created an image at the intermediate image plane, which coincides with the input plane for the FTS module. The FTS module was built upon a Michelson interferometer with one of the mirrors mounted on a motorized translation stage (Physik Instrumente, M-227.25). Two lenses (L1 and L2) of the same focal length (150 mm) in a 4-f telecentric configuration relayed the image from the intermediate image plane to the image plane where the camera was located. To record the images, we used an electron-multiplying charge-coupled device (EMCCD) camera (C) (Andor, iXon Ultra 888). The FTS module records a multitude of images for varying OPDs, i.e., for different locations of the translating mirror. For accurate reconstruction of the spectrum, it is crucial to know the OPD values for which the images were acquired. The actual mirror position, however, can be different from the target position due to a translation stage error or vibrations from the surrounding environment. To correct for this, a separate interferogram is recorded using a reference laser, typically a He-Ne laser. Because the wavelength of a He-Ne laser fell in one of the fluorescence emission bands, we used a diode laser of the wavelength 488 nm (Coherent, OBIS 488 nm LS 60 mW). The intensity of the reference interferogram was monitored using a silicon photodiode (Thorlabs, SM1PD1A). To minimize contamination of the fluorescence images by scattered light photons of 488 nm, we operated the diode laser at a minimum power level (1 mW) and installed a notch filter (Thorlabs, NF488-15) in front of the EMCCD camera. The rejected band did not overlap with the emission bands of the fluorescence filter set. The control of the translation stage as well as the acquisition of the fluorescence images and the reference interferogram were controlled using a Labview program (National Instruments, version 15).

2.2 Data Acquisition

To demonstrate the proposed technique, we imaged bovine pulmonary artery endothelial cells labeled with three fluorescent dyes (MitoTracker Red CMXRos, Alexa Fluor 488 Phalloidin, and DAPI). The sample slide was purchased from Thermo Fisher (F36924). Before the experiment, the zero OPD position, where the interferogram had the maximum value, was found. The power of each LED was adjusted to produce about the same fluorescence intensity levels for all the fluorophores. For each field of view (FOV), 1000 interferogram images were recorded while moving the translation stage in steps of 50 nm. This step size corresponds with 100 nm in terms of the OPD. The images were

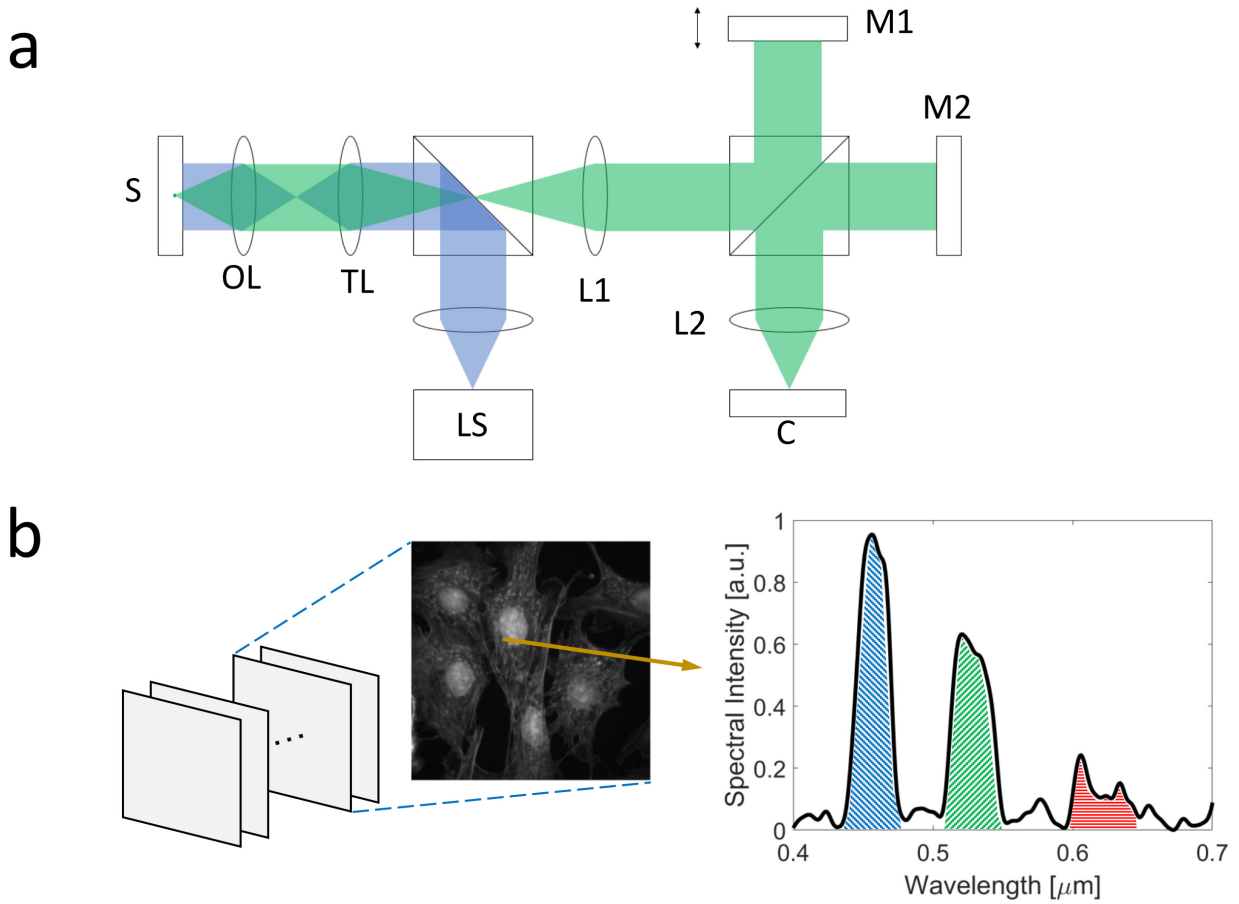

Figure 1: Schematic diagram of the optical system showing the excitation beam path (blue) and emission beam path (green). LS: Light Source, S: Sample, OB: Objective Lens, TL: Tube Lens, L1 and L2: Lenses, M1: Moving Mirror, M2: Stationary Mirror, and C: Camera.

recorded with the camera EM gain of 300 and the exposure time of 0.01s. Simultaneously capturing the images with the camera, the reference interferogram was collected with a photodiode. Each measurement took 23 seconds per set of 1000 images. A total of 30 sets of images (i.e., FOVs) were collected.

2.3 Data Preprocessing

A flow chart of the data processing procedure is shown in Figure 2. The training data flow is shown on the left side of the figure. For each FOV, we collect raw interferogram images, then extract the interferogram arrays from the sample region using a binary mask. The binary mask was obtained by applying a threshold to the maximum projection of the raw interferogram images. The extracted interferogram arrays were normalized to be used as training data for the neural network (NN). In this work, we train the NN to predict the normalized channel intensity, the area under each fluorescent band divided with the total area for all the three emission bands, which is shown in Figure 1(b). The ground truth labels for the training dataset were computed from the interferograms using the conventional FTS method (including He-Ne correction). For each channel (i.e., emission band), 20,000 interferograms with the highest label values were selected, which resulted in 60,000 interferograms from each FOV. When all the training data samples have been processed, the training data was randomly mixed so that each mini batch contained data from multiple FOVs, multiple locations on each sample, and different fluorescent signal types. The validation and test data were processed using the same procedure, where one entire FOV for the validation set and one entire FOV for the test set were set aside.

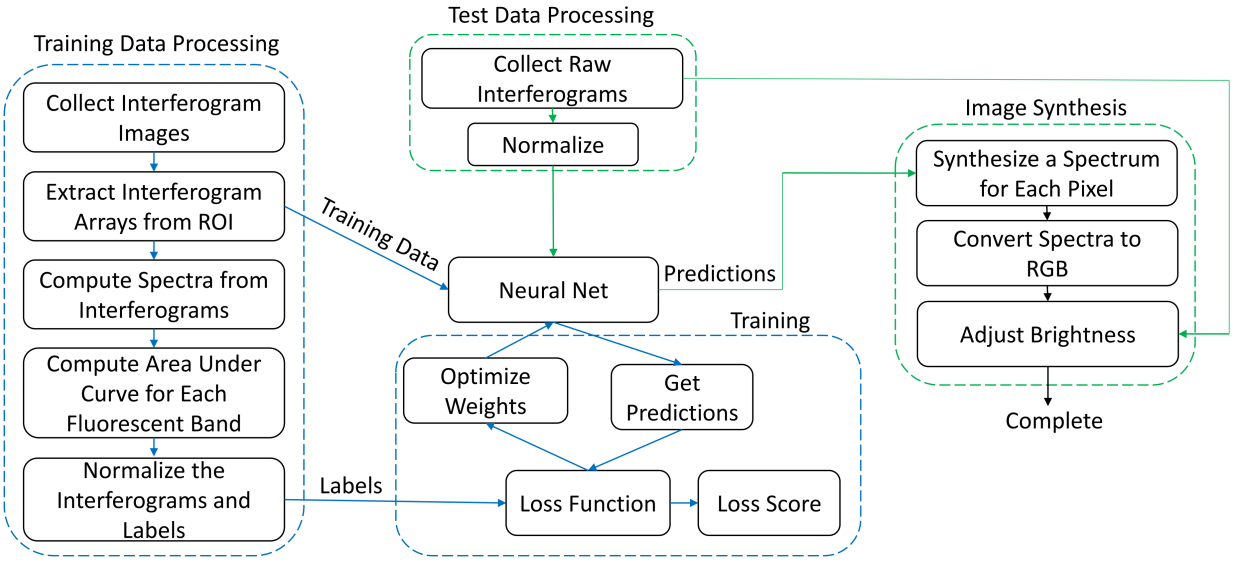

Figure 2: Data processing flowchart starting with collected interferogram images. Training data is processed by computing the spectra, and finding the area under the curve of each fluorescent band. The neural network is trained, and then used to predict and synthesize an image from an unknown sample. Training process is shown in blue, while the testing process is shown in green. Validation process followed training process except with a sample which was not trained on.

2.4 Deep Learning Model

Our NN model is built upon a 1D convolutional neural network (1D CNN) as shown in Figure 3. After each convolutional block, there is a max pooling layer which outputs the maximum of each neighborhood. After the max pooling layer of the final convolutional block, the model is flattened and connected to one or more fully connected "dense" layers providing enough capacity for the model to correctly represent the function before reaching the output layer. The output layer consists of 3 fully connected nodes with sigmoid activation to predict the normalized channel intensity for each of 3 fluorescence emission bands. Although we tested several configurations, using several layers of small kernels allowed us to detect more complex features at a lower cost than using larger kernels, which is consistent with the previous works, for example, Simonyan & Zisserman[21]. The commonly used ReLU activation function was applied at each of the hidden layers. ReLU activation is widely used for deep CNNs as it introduces non-linearity while being more computationally efficient than other non-linear activation functions such as tanh[14, 2]. Regularization reduces validation and test loss while sacrificing training loss leading to better generalization. Here, we selected to use dropout and L2 regularization, which is often referred to as "weight decay", with a constant value in each layer. L2 regularization penalizes large weights without leading to additional model sparsity, which results from L1 regularization[8]. The Adam adaptive learning rate optimization algorithm was selected for weight optimization, which is known to perform well with stochastic gradient descent[12]. The loss for the cost function was selected to be based on mean absolute error (MAE) instead of mean squared error (MSE), because MAE-based loss punishes smaller errors in prediction harder than MSE-based loss, which disproportionately punishes larger prediction errors. With our dataset, the models trained with MAE-based loss tended to synthesize images closer to the ground truth.

2.5 Training of NN with 1D interferogram

The 1D CNN was trained using the standard training procedure shown in the lower middle area of Figure 2. Training data is fed to the NN, which outputs predictions. These predictions are compared with the ground-truth labels to compute the loss, and the weights of the 1D CNN are optimized to minimize the loss. Our training set contained 1,680,000 interferograms and labels from 28 FOVs. Each epoch, or time the 1DCNN trains on the mini batches covering the entire data set, the MAE and loss for training data were calculated and written to a separate file; the same was done for the validation set which consisted of 30,000 (10,000 of the highest label for each channel) interferograms and labels from a separate FOV. These values are evaluated by the early stopping algorithm. Early stopping is important as it functions as a source of regularization in a synergistic way alongside L2 regularization[20]. The early stopping algorithm observes the loss of the validation data and the epoch number. The training stops if either of the following

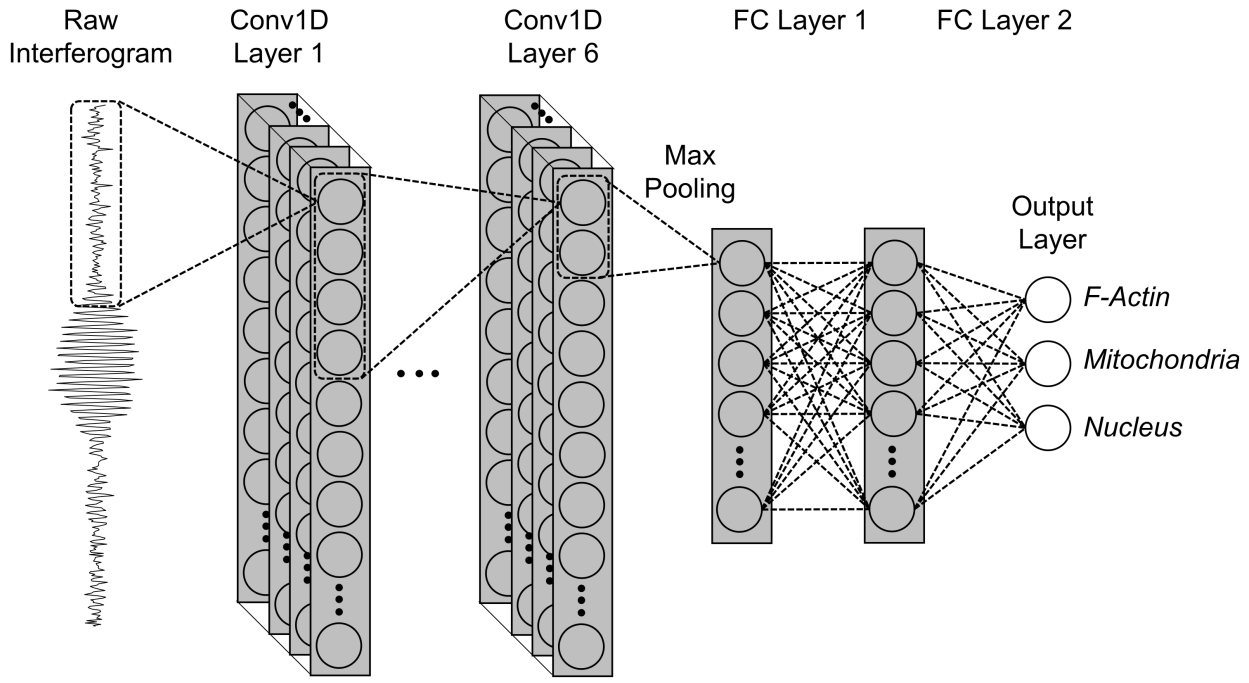

Figure 3: Schematic of 1DCNN used for this study, which is comprised of six 1D convolutional layers followed by a max pooling layer, and two fully connected layers. The raw input data consists of an interferogram with 1000-50 sampled points each. The values at the output predict the amount of each type of fluorescent signal in the pixel.

conditions have been met: 1) the current epoch has reached the maximum number of desired epochs 2) the loss at the current epoch has not decreased below the minimum recorded loss in a set number of epochs, which is referred to as the patience. We used the maximum iteration number of 100 and patience of 5. When an early stopping condition has been met, the training stopped, and the model resulting in the lowest validation MAE was saved. For the training and testing of NN, we used a workstation with two GPUs (NVIDIA Quadro P6000). The MirroredStrategy function built in to TensorFlow was used for synchronous training across the multiple GPUs on a single workstation.

2.6 Hyperparameter Selection

The hyperparameters for our NN models were manually selected while monitoring the MAE values for the training data and validation data as shown in Figure 4. For each interferogram sampling case, the model capacity (the number of trainable weights and biases) was increased until overfitting was observed. Once the model with suitable capacity was selected, L2 regularization was increased until the difference between training and validation loss was minimized. For comparison, we have also generated NNs using Hyperband, an automatic hyperparameter selection algorithm which uses random search and successive halving[18]. The process starts with a multitude of NN architectures trained in search space for a small number of epochs. The best performing networks are trained further while the poor performing networks are abandoned. This process continues until the best network and corresponding set of hyperparameters are chosen.

2.7 Image Synthesis

For the FTS-computed spectrum, the cell image was synthesized by converting each wavelength to an RGB value and then taking the sum of the RGB values weighted by the spectral intensity. For the NN output, the average spectrum for each fluorophore was obtained from the training data. The outputs from the NN, which is labeled as “Predictions” on the flow chart in Figure 2, represent signal weight between 0 and 1 of each fluorescent band region. These predictions were each multiplied with their respective average spectrum and combined, resulting in a synthesized spectrum for each pixel. These spectra were converted to RGB values the same way as described earlier, resulting in an RGB image.

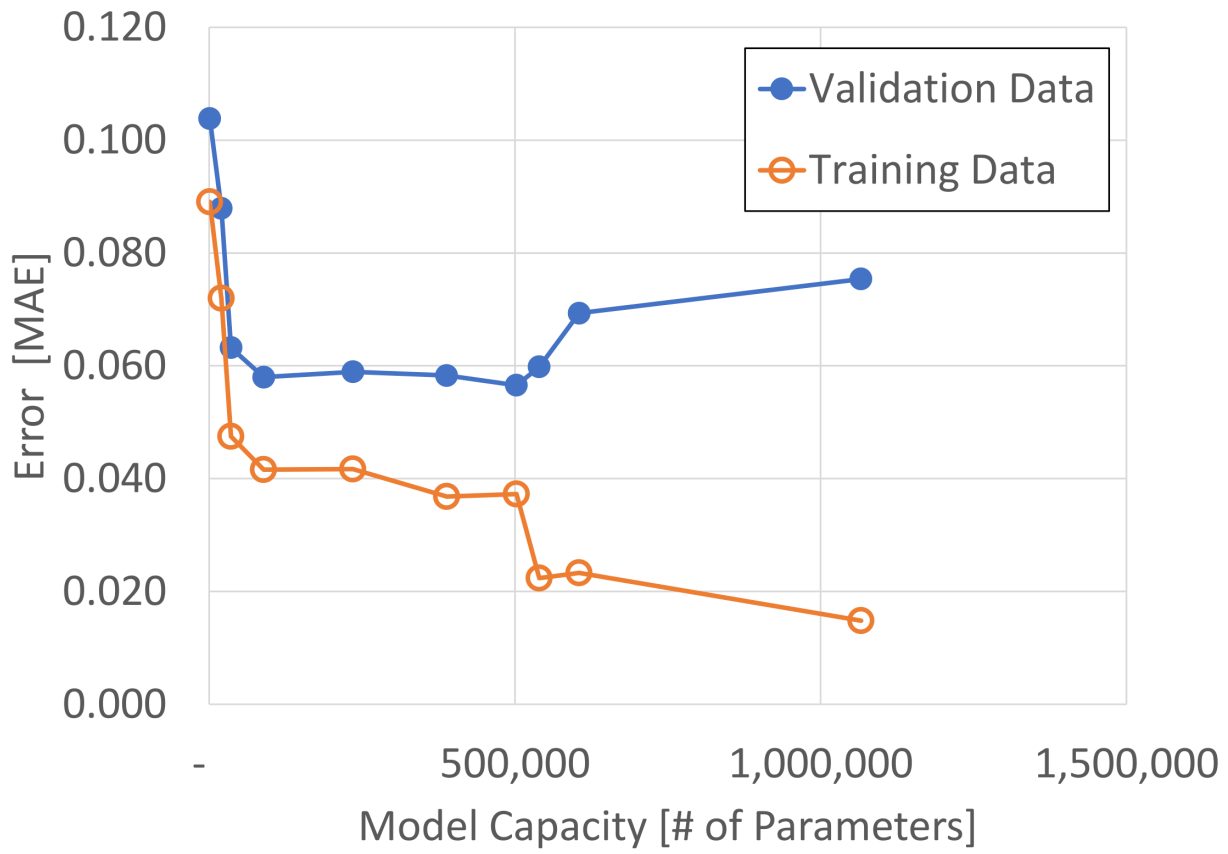

Figure 4: Error vs model capacity of 1DCNNs with manually selected hyperparameters, without regularization.

3 Results and Discussion

Providing the fluorescence spectrum for each pixel, hyperspectral imaging allows us to use various fluorophores with close emission spectra and distinguish target fluorescence signals from autofluorescence background. With the superior spectral resolution of FTIS, we can increase the number of fluorophores that can be simultaneously used, thereby increasing the imaging throughput and obviating the need for sample washing. Here we show our deep learning-based approach can increase the throughput by an order of magnitude while minimally sacrificing the accuracy of measurement.

Figure 5 shows the images synthesized with FTS for different sampling numbers: (a) 1000, and (b) 50. The raw interferogram for each pixel was corrected with the He-Ne data before being processed with the FTS algorithm. Figure 5(a) is the ground truth that we use for comparison. Figure 5(b), reconstructed from 50 sampled points, shows the image completely lost the ability to distinguish between the blue DAPI (nucleus) and green Alexa Fluor™ 488 Phalloidin (F-actin) fluorescent dyes. Figure 6 shows the FTS-synthesized images without the He-Ne correction. For N = 1000 (Figure 6(a)), the nucleus is shown in an incorrect color, and the F-actin and mitochondria are indistinguishable. For N=50 (Figure 6(b)), all three fluorescent dyes are indistinguishable. Comparing the images with those in Figure 5, the importance of He-Ne correction in the conventional FTS can be clearly seen. Figure 7 shows the images synthesized with the NN described earlier.

The interferograms used for the training, validation, and testing have not been corrected with the He-Ne data. Even though the He-Ne correction was not applied, the NN-synthesized image shown in Figure 7(a), which corresponds to N=1000, looks very similar to the ground truth. The image shown in Figure 7(b) is also almost the same as the ground truth except for some speckles. This is remarkable considering that it was synthesized from 20 times less data than the ground truth and without the He-Ne correction.

Figure 8 shows the MSE with varying sampling amounts for the various synthesis methods: FTS without He-Ne correction, FTS with He-Ne correction, and NN without He-Ne correction. The FTS without He-Ne correction method

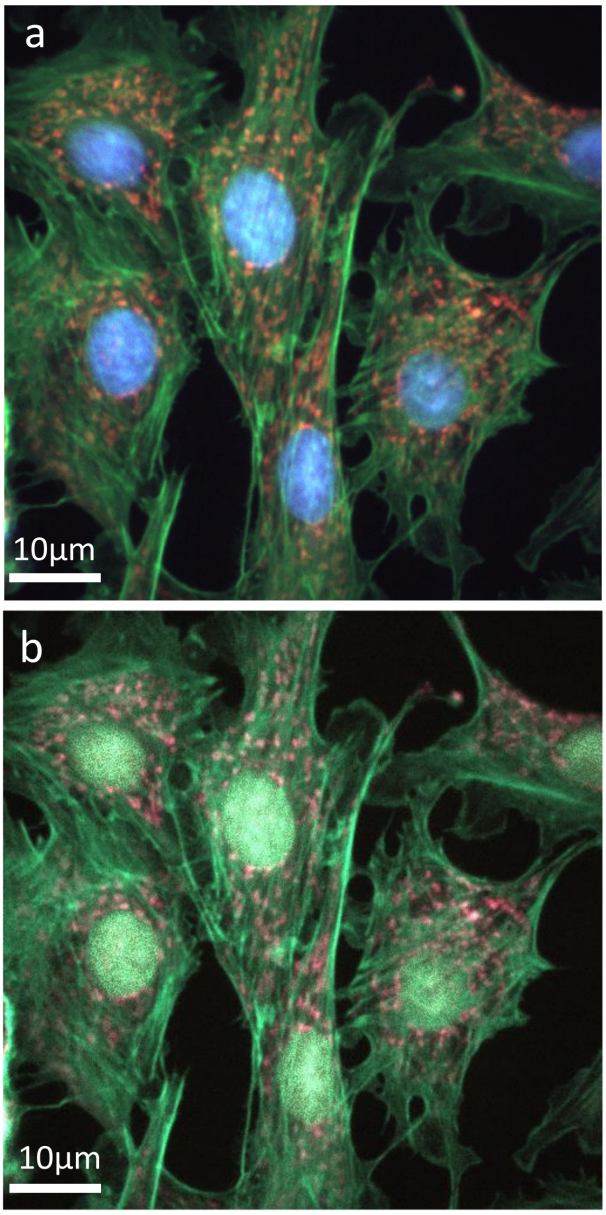

Figure 5: Synthesized fluorescent images using conventional FTS with He-Ne correction for various sampling numbers (N): (a) N = 1000, and (b) N=50. Scale bars: 10 μm.

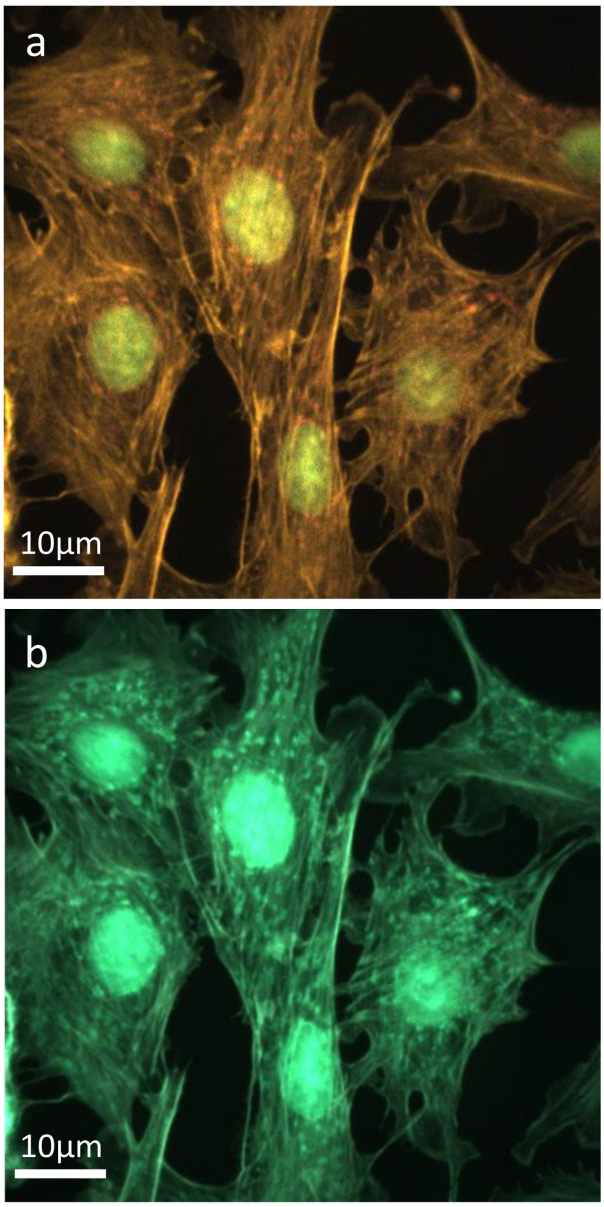

Figure 6: Synthesized fluorescent images using conventional FTS without He-Ne correction for various sampling numbers (N): (a) N = 1000, and (b) N=50. Scale bars: 10 μm.

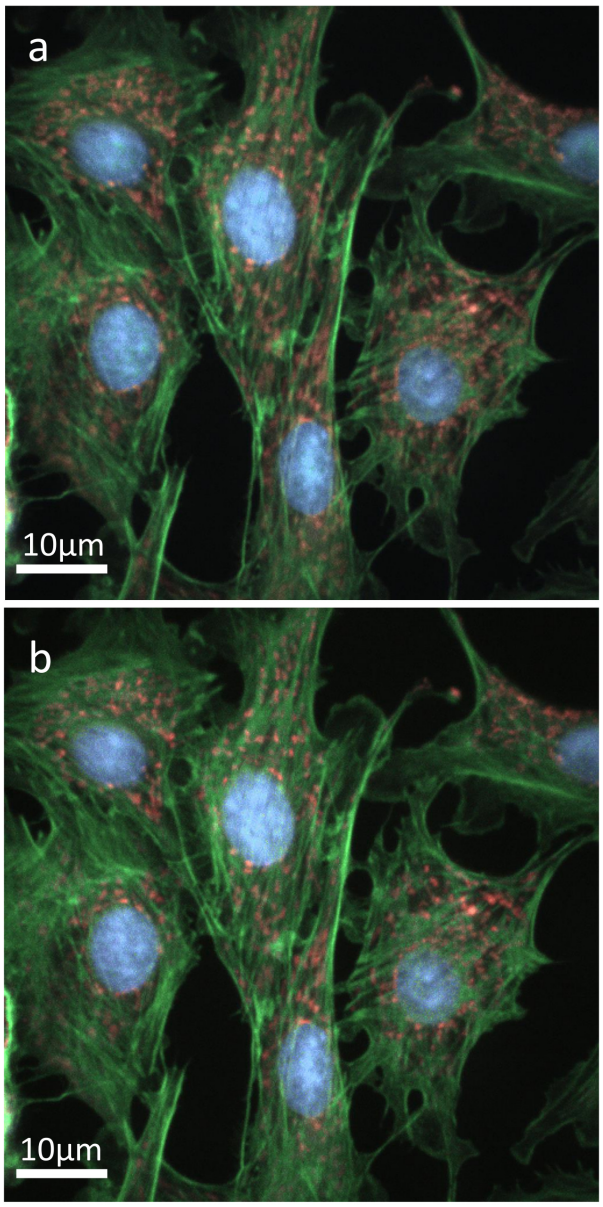

Figure 7: Synthesized fluorescent images using deep learning without He-Ne correction for various sampling numbers (N): (a) N = 1000, and (b) N=50. Scale bars: 10 μm.

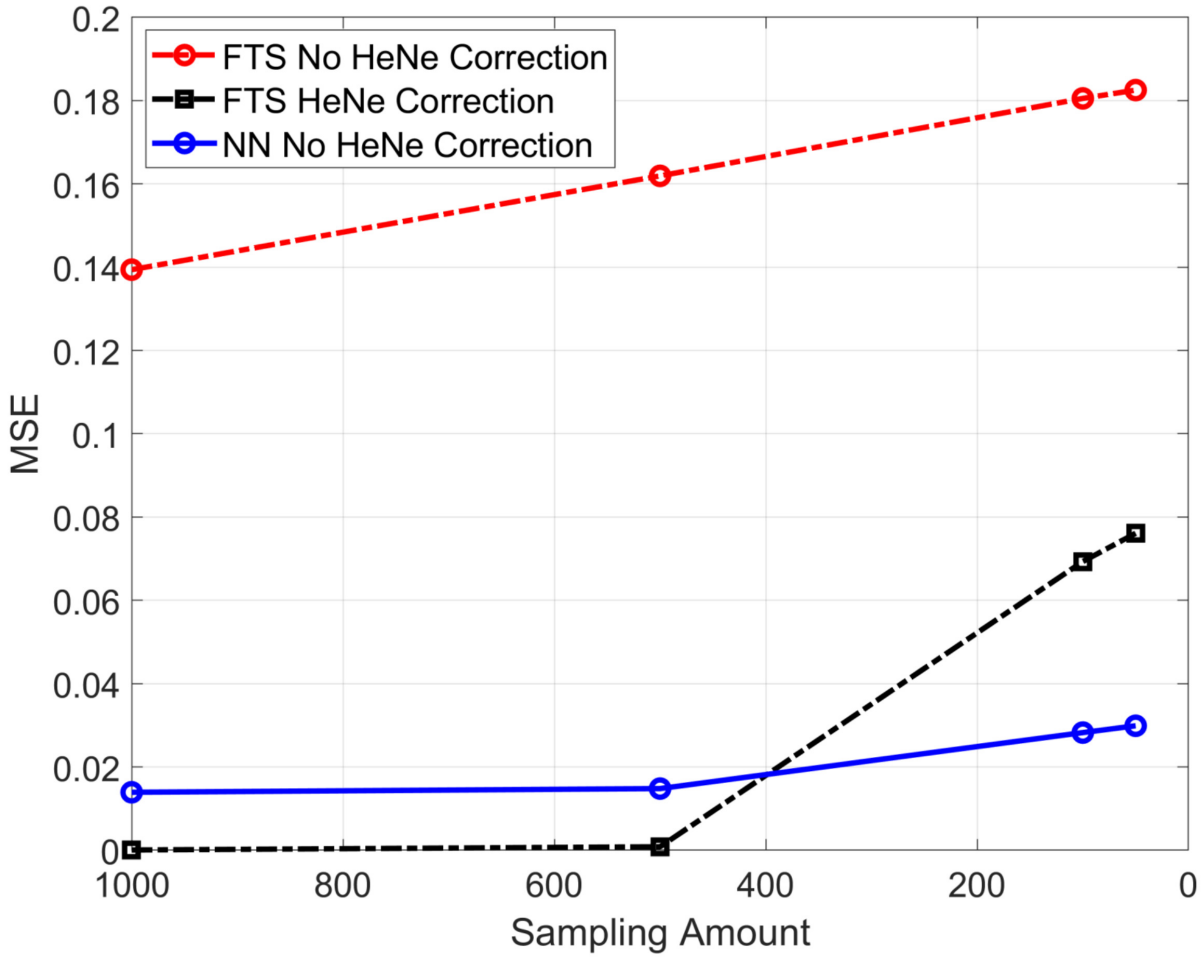

Figure 8: MSE of reconstructed images vs sampling number of traditional FTS with and without He-Ne correction and DLFITS without He-Ne correction.

starts heavily degraded with a high MSE and continues to increase. This is observed by comparing the FTS-synthesized image without He-Ne correction images (Figure 6) with the ground truth image in Figure 5(a). This image degradation with reduced sampling number is also clearly seen even with the He-Ne corrected. We see that the NN without He-Ne correction has some small MSE at full interferogram sampling of N=1000; however, the MSE stays below 0.04 as the sampling number decreases. In contrast, the MSE of traditional FTS reconstruction without He-Ne correction remains very high, above 0.14, while the MSE of traditional FTS reconstruction with He-Ne correction quickly degrades to near 0.08 as the sampling number decreases. For N=50, the MSE for the NN without He-Ne correction is less than half of the corresponding FTS with He-Ne correction. While the images computed by FTS with and without He-Ne correction are heavily degraded with less sampling, the images synthesized by the NN remain intact at the sampling amount of N=50. The limit of our approach appears to be at N=50, which is reducing the sampling number by 95%.

The results presented earlier were obtained using a manually optimized NN. For comparison, we built NNs using Hyperband, an automatic hyperparameter selection algorithm. Figure 9 shows the hyperparameters of the top 3 Hyperband models, which have a few key differences. Model 1 does not use pooling, which may reduce the ability to detect translations in patterns[8], but uses the highest dropout rate of 20%. Model 2 is the only model to use L2 regularization without dropout. To compensate for omitting dropout, the L2 regularization in Model 2 is two orders of magnitude higher than the other models which also feature dropout. Model 3 features 2 standard convolutional blocks, each with max pooling which may lead to the ability to detect pattern translation[8]. The manually selected model uses only one convolutional block consisting of 4 convolutional layers and a max pooling layer with the largest pool size of the 4 models. It also uses significantly fewer nodes in the fully connected layers. Using the models shown in Figure 9, we applied k-fold cross validation with 10 folds using only the training and validation sets. Table 1 shows a table of

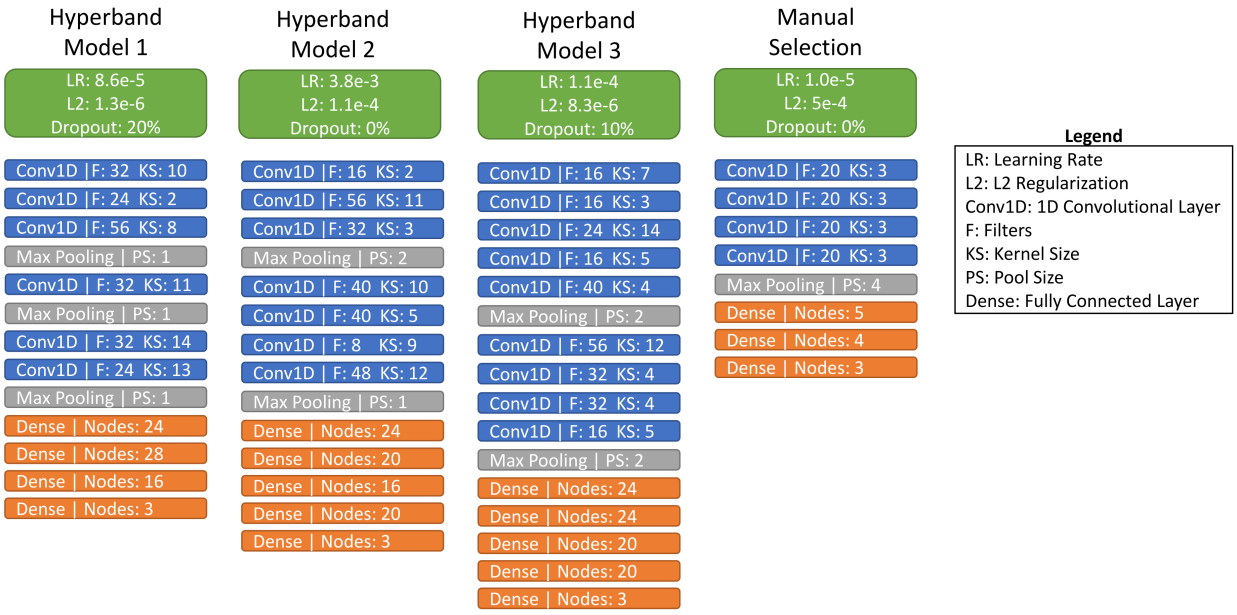

Figure 9: Architectures of the top 3 performing models selected by Hyperband and the architecture resulting from manual hyperparameter selection.

Table 1: K-fold cross validation results (MAE) of the top 3 performing model architectures selected by the Hyperband hyperparameter optimization algorithm compared with the model with manually selected hyperparameters.

	Hyperband (1,083 – 1,015,075 parameters)			Manual selection
	Model 1	Model 2	Model 3	
Number of parameters	86,879	73,855	57,903	4,460
Validation error (MAE)	0.0405 ± 0.0017	0.0432 ± 0.0027	0.0393 ± 0.0014	0.0589 ± 0.0006
Test error (MAE)	0.0976	0.1157	0.0852	0.0920

k-fold cross validation results for the N=50 case with the models found using Hyperband compared to the manually selected model. From Table 1, we see that Model 3 has the lowest validation error (MAE) of all the models. For this N=50 case, the Hyperband model search space was between 1,083 and 1,015,075 trainable parameters. The model resulting from manual hyperparameter selection consisted of only 4,460 trainable parameters, an order of magnitude below all three models selected by Hyperband. This manually selected model resulted the highest mean k-fold cross validation error out of the 4 models; however, its test error is greater than the best performing model, Model 3, only by 8%. In contrast, Models 1 and 2 produced about the same validation error as Model 3 but significantly higher test error: 15% and 36%, respectively. The good generalization of the manually selected model is also reflected in the smallest standard deviation for the validation error and may be attributed to the small number of parameters (i.e., low capacity).

The reduction of interferogram sampling of 95% by our approach leaves us the need to sample only 1/20 of the data. This is slightly better than the compressed sensing approaches which require 1/16 of data traditionally needed[5] or 1/9 of random samples from the original dataset[15]. By attempting to collect our data uniformly while allowing for some random experimental error, we eliminate the need for He-Ne correction, eliminating several optical elements which reduces the cost, size, and complexity of the FTIS system. Reducing the sampling size allows FTIS-based approaches to be readily used for hyperspectral fluorescence imaging, allowing more fluorescent dyes to be used including dyes with close emission spectra.

4 Conclusion

In this paper, we have demonstrated hyperspectral fluorescence imaging by combining deep learning and FTIS. The image synthesized by the NN with a 20 times reduction in sampling accurately matched the ground truth image. Using triple-labeled bovine pulmonary artery endothelial cells, we demonstrated the capabilities of our approach. While greatly reducing the required sampling, we also bypass the need for He-Ne correction, eliminating several optical

elements which reduces the cost, size, and complexity of the FTIS system. The developed system can be used in a wide range of applications where several fluorescent dyes with close emission spectra must be used, with much higher throughput.

5 Acknowledgements

This research was funded by the Research Growth Initiative of the University of Wisconsin-Milwaukee, the National Science Foundation (1808331), and the National Institute of General Medical Sciences of the National Institutes of Health (R21GM135848). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Institutes of Health.

References

[1] Ramy Abdlaty, Samir Sahli, Joseph Hayward, and Qiyin Fang. Hyperspectral imaging: comparison of acousto-optic and liquid crystal tunable filters. In *Medical Imaging 2018: Physics of Medical Imaging*, volume 10573, page 105732P. SPIE, March 2018.

[2] Charu C. Aggarwal. Neural networks and deep learning. *Springer*, 10:978–3, 2018. Publisher: Springer.

[3] Kevin de Haan, Zachary S. Ballard, Yair Rivenson, Yichen Wu, and Aydogan Ozcan. Resolution enhancement in scanning electron microscopy using deep learning. *Sci. Rep.*, 9(1):12050, August 2019. Number: 1 Publisher: Nature Publishing Group.

[4] D.L. Donoho. Compressed sensing. *IEEE Trans. Inf. Theory*, 52(4):1289–1306, April 2006. Conference Name: IEEE Transactions on Information Theory.

[5] Josef A. Dunbar, Derek G. Osborne, Jessica M. Anna, and Kevin J. Kubarych. Accelerated 2D-IR Using Compressed Sensing. *J. Phys. Chem. Lett.*, 4(15):2489–2492, August 2013. Number: 15.

[6] Yonina C. Eldar and Gitta Kutyniok. *Compressed Sensing: Theory and Applications*. Cambridge University Press, May 2012. Google-Books-ID: 9ccLAQAAQBAJ.

[7] Andre Esteva, Brett Kuprel, Roberto A. Novoa, Justin Ko, Susan M. Swetter, Helen M. Blau, and Sebastian Thrun. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639):115–118, February 2017. Number: 7639 Publisher: Nature Publishing Group.

[8] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, November 2016. Google-Books-ID: omivDQAAQBAJ.

[9] Peter R. Griffiths and James A. De Haseth. *Fourier Transform Infrared Spectrometry*. John Wiley & Sons, March 2007. Google-Books-ID: C_c0GVe8MX0C.

[10] Philip Hunter. The advent of AI and deep learning in diagnostics and imaging. *EMBO Rep.*, 20(7):e48559, July 2019. Number: 7 Publisher: John Wiley & Sons, Ltd.

[11] George Imataka and Osamu Arisaka. Chromosome Analysis Using Spectral Karyotyping (SKY). *Cell Biochem. Biophys.*, 62(1):13–17, January 2012. Number: 1.

[12] Diederik P. Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization. *ICLR*, January 2017. arXiv: 1412.6980.

[13] Gert-Jan Kremers, Sarah G. Gilbert, Paula J. Cranfill, Michael W. Davidson, and David W. Piston. Fluorescent proteins at a glance. *J. Cell Sci.*, 124(15):2676–2676, August 2011. Number: 15.

[14] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. ImageNet classification with deep convolutional neural networks. *Commun. ACM*, 60(6):84–90, May 2017. Number: 6.

[15] Bernd Kästner, Franko Schmähling, Andrea Hornemann, Georg Ulrich, Arne Hoehl, Mattias Kruskopf, Klaus Pierz, Markus B. Raschke, Gerd Wübbeler, and Clemens Elster. Compressed sensing FTIR nano-spectroscopy and nano-imaging. *Opt. Express*, 26(14):18115, July 2018. Number: 14.

[16] Silas J. Leavesley, Naga Annamdevula, John Boni, Samantha Stocker, Kristin Grant, Boris Troyanovsky, Thomas C. Rich, and Diego F. Alvarez. Hyperspectral imaging microscopy for identification and quantitative analysis of fluorescently-labeled cells in highly autofluorescent tissue. *J. Biophotonics*, 5(1):67–84, 2012. Number: 1 Publisher: Wiley Online Library.

[17] June-Goo Lee, Sanghoon Jun, Young-Won Cho, Hyunna Lee, Guk Bae Kim, Joon Beom Seo, and Namkug Kim. Deep Learning in Medical Imaging: General Overview. *Korean J Radiol*, 18(4):570–584, 2017. Number: 4.

[18] Lisha Li, Kevin Jamieson, Giulia DeSalvo, Afshin Rostamizadeh, and Ameet Talwalkar. Hyperband: A Novel Bandit-Based Approach to Hyperparameter Optimization. *J Mach Learn Res*, page 52, 2018.

[19] Michael C. Mozer, Michael I. Jordan, and Thomas Petsche. *Advances in Neural Information Processing Systems 9: Proceedings of the 1996 Conference*. MIT Press, 1997. Google-Books-ID: QpD7n95ozWUC.

[20] Yair Rivenson, Zoltán Göröcs, Harun Günaydin, Yibo Zhang, Hongda Wang, and Aydogan Ozcan. Deep learning microscopy. *Optica*, 4(11):1437–1443, November 2017. Number: 11 Publisher: Optical Society of America.

[21] Karen Simonyan and Andrew Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. *cs.CV*, April 2015. arXiv: 1409.1556.

[22] Vincent Studer, Jérome Bobin, Makhlad Chahid, Hamed Shams Mousavi, Emmanuel Candes, and Maxime Dahan. Compressive fluorescence microscopy for biological and hyperspectral imaging. *Proc. Natl. Acad. Sci.*, 109(26):E1679–E1687, June 2012. Number: 26 Publisher: National Academy of Sciences Section: PNAS Plus.

[23] Hiromichi Tsurui, Jeremy M. Lerner, Kuniaki Takahashi, Sachiko Hirose, K. Mitsui, Ko Okumura, and Toshikazu Shirai. Hyperspectral imaging of pathology samples. In *Three-Dimensional and Multidimensional Microscopy: Image Acquisition and Processing VI*, volume 3605, pages 273–281. SPIE, May 1999.

[24] Hiromichi Tsurui, Hiroyuki Nishimura, Susumu Hattori, Sachiko Hirose, Ko Okumura, and Toshikazu Shirai. Seven-color Fluorescence Imaging of Tissue Samples Based on Fourier Spectroscopy and Singular Value Decomposition. *J Histochem Cytochem.*, 48(5):653–662, May 2000. Number: 5 Publisher: Journal of Histochemistry & Cytochemistry.

[25] E. S. Wachman, W. Niu, and D. L. Farkas. AOTF microscope for imaging with increased speed and spectral versatility. *Biophys J*, 73(3):1215–1222, September 1997.

[26] Mingde Yao, Zhiwei Xiong, Lizhi Wang, Dong Liu, and Xuejin Chen. Spectral-depth imaging with deep learning based reconstruction. *Opt. Express*, 27(26):38312, December 2019. Number: 26.

[27] Timo Zimmermann, Jens Rietdorf, and Rainer Pepperkok. Spectral imaging and its applications in live cell microscopy. *FEBS Lett.*, 546(1):87–92, 2003. Number: 1 Publisher: Wiley Online Library.