Digital Staining of High-Definition FTIR Images Using Deep Learning

Mahsa Lotfollahi[1], **Sebastian Berisha**[1], **Davar Daeinejad**[1], and **David Mayerich**[1]

[1]Department of Electrical and Computer Engineering, University of Houston

Abstract

Histological stains, such as hematoxylin and eosin (H&E), are routinely used in clinical diagnosis and research. While these labels offer a high degree of specificity, throughput is limited by the need for multiple samples. Traditional histology stains, such as immunohistochemical labels, also rely only on protein expression and cannot quantify small molecules and metabolites that may aid in diagnosis. Finally, chemical stains and dyes permanently alter the tissue, making downstream analysis impossible. Fourier transform infrared (FTIR) spectroscopic imaging has shown promise for label-free characterization of important tissue phenotypes, and can bypass the need for many chemical labels. FTIR classification commonly leverages supervised learning, requiring human annotation that is tedious and prone to errors. One alternative is digital staining, which leverages machine learning to map infrared spectra to a corresponding chemical stain. This replaces human annotation with computer-aided alignment. Previous work relies on alignment of adjacent serial tissue sections. Since the tissue samples are not identical at the cellular level, this technique cannot be applied to high-definition FTIR images. In this paper, we demonstrate that cellular-level mapping can be accomplished using identical samples for both FTIR and chemical labels. In addition, higher-resolution results can be achieved using a deep convolutional neural network that integrates spatial and spectral features.

Keywords

Introduction

Histology relies heavily on expert interpretation of staining patterns produced by chemical contrast agents. Tissue staining is a critical component of modern clinical practice, since it is extremely difficult to differentiate structure and morphology in unstained tissue. Accurate disease diagnosis, prognosis, and prediction therefore requires reliable labeling.

The most common clinical labels include hematoxylin and eosin (H&E),[1–3] periodic acid-Schiff (PAS), and Masson's trichrome. Selecting the appropriate label largely depends on tissue type, and panels requiring 3+ labels of serial tissue sections is common clinical

Corresponding author: David Mayerich, mayerich@uh.edu.

practice.[4,5] Although staining protocols have been widely adopted, variability makes automated analysis difficult.[6–8] Clinical diagnosis therefore still relies on manual assessment by experts, increasing cost and reducing throughput.

Vibrational spectroscopy provides a label-free alternative for pathological assessment that is comparable to traditional histology for many applications.[9–13] Raman and infrared (IR) spectroscopy are highly specific techniques for molecular identification. However, Raman spectroscopy is limited by the weak signal.[14–17] Infrared (IR) absorbance spectroscopy has routinely demonstrated the ability to differentiate tissues relevant to disease[18–22] without significantly perturbing the sample,[23] providing the potential for augmenting traditional histology. IR spectroscopy can also probe changes associated with abnormal tissue that are difficult to determine using histology alone.[24–28]

Motivation

While chemical staining is the current standard, it is non-quantitative and prone to human error and inter-observer disagreement.[16] Our goal is to leverage the quantitative infrared spectroscopy to provide consistent molecular maps that aid pathological inspection. The most common approach leverages annotated IR images to train supervised machine learning (ML) algorithms for automated classification.[29–33] This approach would require additional training for experts to integrate quantitative IR maps with existing clinical practice. For example, immunohistochemical labels frequently rely on counterstains to label tissue landmarks that are currently difficult to identify in IR images. This landmark challenge also effects the annotated ground truth used for ML training, since experts are likely to limit annotations to pixels that have a high degree of certainty. Critically, this introduces a bias in validation, making classifiers appear more accurate than they actually are.

Related Work

Many research works have been conducted to classify tissues in IR images, including unsupervised techniques such as K-means clustering[6,34] or hierarchical cluster analysis[34] and supervised techniques such as random forests,[33,35] Bayesian,[30,32] artificial neural network (ANN)[32,36] and support vector machines (SVMs)[32,37].

We have previously shown that many traditional stains and counterstains can be digitally applied using artificial neural networks (ANNs) to map molecular spectra to RGB values.[36] This technique limits spatial resolution based on two critical factors: First, training requires aligned input/target pairs. Previous work uses adjacent sections, making cellular-level mapping impossible because adjacent images are fundamentally different at this scale. Second, an ANN uses pixel-independent spectra to determine the output color. Since a large amount of molecular information is embedded within the fingerprint region (500 to 1500 cm^{-1}), the resulting digital stain is diffraction limited to this region (6.5 to 20μm).

In this paper, we demonstrate digital staining of high-definition FTIR images to provide cellular-level labeling that mimics traditional histology.[38] Our proposed algorithm relies on improved standard protocols for imaging, labeling, and aligning the sample, allowing accurate generation of input/target training pairs that reflect micrometer-scale features. We use a convolutional neural network (CNN)[39] to leverage both spatial and spectral image

features, allowing us to take advantage of scattering at higher wavenumbers to overcome limitations in chemical resolution.

Materials and methods

Sample preparation

Tissue samples, consisting of pig kidney, pig skin, mouse kidney and mouse brain, were fixed for 24 hours in 10% neutral buffered formalin. The tissue was then embedded in paraffin using standard clinical techniques, which include dehydration in graded ethanol baths, clearing with xylene, and finally infiltration with paraffin at 6°C. The tissue sections were cut into thin sections (2 to 5μm) and mounted on calcium fluoride (CaF_2) substrates. The tissue was then deparaffinized with xylene and graded ethanol, followed by re-hydration using phosphate buffered saline.[40]

After imaging, samples were labeled with either hematoxylin and eosin (H&E) or 4,6-diamidino-2-phenylindole (DAPI) to provide chemical contrast.

Data collection

Deparaffinized sections were imaged using a Cary 600 FTIR spectroscopic microscope (Agilent Technologies) in the range of 800 to 4000 cm^{-1} at 8 cm^{-1} spectral resolution and focal plane array with a size of 128×128 . Imaging was performed using a 15X 0.62NA Cassegrain objective with a projected pixel size of 1.1 μm for high magnification and 5.5 μm for standard magnification in transmission mode. A smaller projected pixel size allows diffraction-limited in high-wavenumber regions. After staining, samples were imaged using a Nikon optical microscope with a 10X 0.3NA objective. H&E stains were imaged in brightfield mode, while DAPI was imaged using a 390 ± 9nm excitation filter and 460 ± 30nm emission filter.

Data preprocessing

FTIR images were preprocessed using standard protocols,[40] including piece-wise linear baseline correction and band normalization to Amide I (1650 cm^{-1}). These methods help to mitigate spatial effects introduced by scattering within heterogeneous samples, however peak shifts introduced at interfaces with differing refractive index can still impact the positions of resonance peaks. Principle component analysis (PCA) was employed to reduce the number of bands to 60 components representing 97.7% of the overall variance. Dimension reduction was necessary in order to mitigate memory constraints when using a CNN.

Image alignment

In order to perform semantic (pixel-wise) segmentation, IR spectra must be assigned input and target values for training. The input was selected from the IR image at position $\mathbf{p} = [x, y]^T$ and the target color is selected from the associated location in the stained image. While the tissue samples in both images are identical, the staining process introduces deformation and tearing. Both images must be aligned using an elastic deformation transform $\mathbf{T}$ such that the position $\mathbf{x}$ within the absorbance image $\mathbf{A}(\mathbf{p})$ maps to the stained image $S(\mathbf{Tp})$ at a

transformed position **Tp**. This requires pixel-level alignment of both IR and brightfield images to extract the corresponding positions. Using the same tissue sample for both images facilitated alignment, however distortions were introduced through chemical staining, due to required rehydration and chemical manipulation of the sample. Alignment was performed manually using the GIMP open-source editing software (Figure 1) to apply affine transformations to sub-regions of the sample used for training and validation. Regions that had undergone significant distortion (e.g. tearing) were excluded.

Convolutional neural network architecture

Our CNN architecture is composed of alternating convolutional and pooling layers followed by a fully connected ANN.

Convolutional layer: Convolutional layers apply a kernel with a small receptive field[41,42] to extract spatial features. Required hyper-parameters for optimization include filter size, padding, and stride. The filter size is equivalent to the spatial region considered to contain a viable image feature. The output of a convolutional layer is always smaller than the input by the filter size. The input is therefore padded with zeros at the borders to preserve spatial extent (width and height). The stride hyper-parameter controls the shifting step size for the convolution filters.

A point-wise nonlinear activation function is applied to the output of each convolution layer. The activity of feature map j in layer l is given by:

$$\mathbf{X}_j^l = g\left(\sum_i^{N_{l-1}} \mathbf{k}_{i,j}^l * \mathbf{X}_i^{l-1} + \mathbf{b}_j^l \right),$$

where N_{l-1} indicates the number of feature maps in layer $(l-1)$, $\mathbf{X}_i^{l-1}$ is the feature map i in layer $(l-1)$, $\mathbf{k}_{i,j}^l \in \mathbb{R}^{k \times k}$ is a convolutional kernel applied to $\mathbf{X}_i^{l-1}$, and $\mathbf{b}_j^l$ is a bias term that facilitates learning. The transfer function $g(\ )$ is a nonlinear activation function, such as a sigmoid, tanh, or a rectified linear unit (relu). The * denotes the discrete convolution.

Local response normalization: In case of using the relu activation function, local response normalization (LNR) is employed to mitigate problems introduced with an unbounded activation function by enforcing local competition among features in different feature maps at the same spatial location.[43] The normalization response, $b_{x,y}^i$, of a neuron activity at location (x, y) after applying kernel i and activation function $g(.)$, denoted as $a_{x,y}^i$ is computed as[39]

$$b_{x,y}^i = a_{x,y}^i / \left(k + \alpha \sum_{j=\max(0, i-n/2)}^{\min(N-1, i+n/2)} \left(a_{x,y}^j \right)^2 \right)^\beta,$$

where n is the number of adjacent feature maps around that kernel i. k, n, α and β are constant hyper-parameters. They are chosen to be $k = 1$, $n = 5$, $\alpha = 1$ and $\beta = 0.5$.

Pooling layer: The normalized output of each convolution layer is passed through a pooling layer to summarize local features.[44] By using a stride (stepping size over input pixels) greater than 1, the pooling layer sub-samples the feature maps to merge semantically similar features to provide invariant representation.[45,46] This further improves the computational time by reducing the number of parameters. The Pooling Layer operates independently on every depth slice of the input and resizes it spatially. The depth dimension remains unchanged. If the feature map volume size is $m \times n \times d$, The pooling layer with filter dimension of q and stride s produces output with volume size of $m' \times n' \times d$ where:

$$m' = \frac{1}{s}(m - q) + 1$$

$$n' = \frac{1}{s}(n - q) + 1$$

Fully connected layer: Our CNN is followed by a fully connected layer composed of a traditional artificial neural network (ANN) to facilitate deep learning.[47] The input of fully connected layer is the output of previous layer flatten into a feature vector ($\mathbf{X}^{l-1}$).[48] The output vector of the fully connected layer l is:

$$X^l = f(\mathbf{W}^l \mathbf{X}^{l-1} + \mathbf{b}^l)$$

Connections between neurons in the layer $l - 1$ and l are stored in the weight matrix $\mathbf{W}^l$ with bias vector $\mathbf{b}^l$.

Loss function: The resulting supervised optimization minimizes the difference between network predictions and target values provided through a training set. For regression with linear output, the mean square error is used as a loss function:

$$L = \sum \frac{1}{2}(t - p)^2$$

where t is the expected output and p is the value predicted by the most recent model.

Training: Training the network model starts with weight initialization in convolutional and fully-connected layers. This step is important because bad initialization leads to a slow training or no convergence training.[49] Then the network is fed with training set to determine the activation of each neuron(forward propagation). The loss function compares the network output with the actual value and is optimized to adjust the weights so that the loss decreases(back propagation). Different gradient descent based optimization algorithms, such as stochastic gradient descent, Adam,[50] and Adadelta,[51] have been proposed and employed for ANN optimization.[52] Our method uses the Adam optimizer for back-propagation.

Regularization: A regularization term with/without dropout prevents the model from over-fitting on training data and provides better generalization on test data.[53] We use ℓ_1 regularization without dropout.

Batch training: Most neural network training is based on mini-batch stochastic gradient optimization. In this optimization method, training data set is split into small batches to calculate the loss function and update the parameters (weights and biases). Small batch training provides more robust convergence by avoiding local minima and also facilitate training on system with small memory.

Implementation

The network's input $\mathbf{X} \in \mathbb{R}^{m \times n \times b}$ is a hyperspectral image with a spatial size of m × n and b spectral bands, which in our case are composed of the 60 principle components chosen from PCA. The trained CNN takes an input image to create a chemical map. In order to satisfy CNN requirements, we crop an $s \times s$ area around each pixel of $\mathbf{X}$ to create a $s \times s \times b$ patch volume centered at p_i. We use a spatial crop size of 9 × 9 around each pixel. The intensity value I_i of pixel p_i in the color image is assigned to this volume path.

The experimental design of our CNN consists of four convolution layers with 32, 32, 64, 64 feature maps, respectively, and 3 × 3 kernel windows with a stride of 1 for all convolution layers. Every other convolution layer is followed by a pooling layer with a kernel size of 2 × 2 to reduce the spatial dimension by the factor of 2. The network is followed by two fully connected layers: one with 128 neurons and the other with three neurons that provide the color channels for the output color image. An overview of the proposed CNN architecture used is shown in Figure 2.

This network requires the adjustment of many hyperparameters, including the choice of optimizer, learning rate, and regularization factor, which are tuned on a validation set. Training is carried out using batch optimization with a batch size of 128. We use the Adam[50] optimizer with a learning rate of 0.01 to minimize our mean squared error (MSE) performance metric.[52] To reduce the generalization error, we add L_2 regularization to the weights, with the parameter set to 0.001. We use the softplus activation function for each layer except the output, which is a smoother version of relu. If x is the output of a neuron, then softplus is defined as:

$$softplus(x) = \ln(1 + e^x).$$

The output layer passes through the linear function that returns the incoming tensor without changes.

We initialized the weights in our model using a normal distribution with a 0 mean and standard deviation of 0.02.

Results

Our network is used for pixel-level mapping of large high- resolution images of normal pig kidney. In order to mitigate over-fitting due to biological variation, random sampling was used to select training data from larger images without the need to use every pixel in the training data. We used 16000 samples, each with the tensor size of $9 \times 9 \times 60$ data points. Training and testing were performed using a Tesla k40M GPU. The average training time for 7 epochs was 33 seconds, and the CNN required 20 seconds to digitally stain a $512 \times 512 \times 60$ pixels hyperspectral image.

MSE was computed after each iteration (mini-batches) to evaluate model convergence on the training data (Figure 3a), and after each epoch for random samples on validation data (Figure 3b).

Model performance was evaluated using the coefficient of determination $R^2 \in [0, 1]$ to measure the *goodness of fit* on an independent validation set. This metric indicates better performance as R^2 approaches 1:

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (\hat{y}_i - y_i)^2}{\sum_{i=1}^{N} (y_i - \bar{y})^2}$$

where y and $\hat{y}$ are target and predicted images respectively, $\bar{y} = \frac{1}{N} \sum_{i=1}^{N} y_i$ is mean value of target image over all pixels(N). The parameter R^2 is computed after each epoch for validation data (Figure 3c).

The network was tested on independent images and compared to previously published algorithms[36] (Figure 4) using the same imaging system. Applying a pixel level ANN (1 hidden layer with 7 nodes) to the high-magnification data does not provide a significant improvement in resolution, likely due to reliance on chemical information in the fingerprint region of the spectrum (900 to 1500cm^{-1}) where the diffraction limited resolution is comparable to a standard-magnification optical path (Figures 4a and 4b). The introduction of spatial features available at higher wavenumbers using a CNN provides significantly better spatial resolution in the output image (Figure 4c) compared to the chemically stained tissue (Figure 4d).

We presented the result of the digital staining networks qualitatively. The generated images from proposed CNN simply resemble the ground truth images visually. Due to color variation and local deformation happening during physical staining (ground truth), the synthesized images cannot be fully overlapped with the ground truths and quantitative evaluation is challenging. We quantified the visibility of differences between the digitally generated staining patterns and the chemically stained images using structural similarity (SSIM) index (Table 1). The SSIM index measures the similarity in case of luminance, contrast and structure.[54,55] SSIM between two images x and y is defined as:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (1)$$

where μ_x, μ_y, σ_x, σ_y and σ_{xy} are the local means, standard deviation, and cross-covariance for images x and y. C_1 and C_2 are regularization constants to avoid division by zero where the local means or standard deviations are close to zero.

In addition, the proposed CNN was trained and tested on different mouse kidney, mouse brain, and pig skin tissue samples to digitally mimic DAPI staining (Figure 5) to demonstrate that high-magnification infrared can capture the morphological features of cell nuclei and often individual cell positions.

Conclusion

We propose a method to reproduce high-resolution staining patterns based on direct molecular imaging using mid- infrared spectroscopy. These digital stains can be applied without the use of chemical labels or dyes and provide results similar to those used in traditional histology. We have developed a tissue preparation protocol that can be used to acquire input/target pairs for training. Unlike previous methods, our approach relies on the same tissue sample as the source of both input and target data. While these protocols allow a more accurate mapping between IR spectra and the desired output, a large amount of the relevant lportion of the spectrum. When relying only on spectral features, the low spatial resolution of the fingerprint region in the absorbance spectrum limits the corresponding spatial resolution of the output image. This is likely due to the abundance of biological information encoded in wavenumbers below 1800 cm^{-1}. CNNs are able to combine the chemistry within the fingerprint region with spatial features from higher wavenumbers that provide greater spatial resolution. This approach clearly produces a more accurate duplication of the desired staining patterns, based on a subjective assessment (Figure 4).

While we are able to capture broad features produced by clusters of nuclei, and often individual cells, the diffraction limit imposed by IR is not yet competitive with bright field or fluorescence imaging. Other spectroscopic contrast mechanisms that do not perturb the sample, such as second harmonic generation (SHG) may also be used to extract more accurate target values. In addition, our current analysis relies on subjective criteria to evaluate output quality. Clinical application will require extensive validation, potentially as a double-blind experiment coordinated with trained histologists. Given the increase in image acquisition time over brightfield microscopy, it will also be important to identify spectral characteristics provided by FTIR that are unavailable or inaccessible to standard clinical practice. Finally, the higher resolution output achieved in our experiments relies on high spatial frequencies encoded in bands collected at higher wavenumber. Therefore, the results described here are likely incompatible with imaging systems using *quantum cascade laser* (QCL) sources.

Acknowledgements

References

1. Bayramoglu N, Kaakinen M, Eklund L et al. Towards virtual h&e staining of hyperspectral lung histology images using conditional generative adversarial networks. In ICCV Workshops. pp. 64–71.

2. Nguyen L, Tosun AB, Fine JL et al. Spatial statistics for segmenting histological structures in h&e stained tissue images. IEEE transactions on medical imaging 2017; 36(7): 1522–1532. [PubMed: 28328502]

3. Pilling MJ, Henderson A, Shanks JH et al. Infrared spectral histopathology using haematoxylin and eosin (h&e) stained glass slides: a major step forward towards clinical translation. Analyst 2017; 142(8): 1258–1268. [PubMed: 27921102]

4. Puebla-Osorio N, Sarchio SN, Ullrich SE et al. Detection of infiltrating mast cells using a modified toluidine blue staining. In Fibrosis Springer, 2017 pp. 213–222.

5. Mittal S, Yeh K, Leslie LS et al. Simultaneous cancer and tumor microenvironment subtyping using confocal infrared microscopy for all-digital molecular histopathology. Proceedings of the National Academy of Sciences 2018; : 201719551.

6. Demir C and Yener B. Automated cancer diagnosis based on histopathological images: a systematic survey. Rensselaer Polytechnic Institute, Tech Rep 2005;.

7. AlZubaidi AK, Sideseq FB, Faeq A et al. Computer aided diagnosis in digital pathology application: Review and perspective approach in lung cancer classification. In New Trends in Information & Communications Technology Applications (NTICT), 2017 Annual Conference on. IEEE, pp. 219–224.

8. Xing F and Yang L. Robust nucleus/cell detection and segmentation in digital pathology and microscopy images: a comprehensive review. IEEE reviews in biomedical engineering 2016; 9: 234–263. [PubMed: 26742143]

9. Benard A, Desmedt C, Smolina M et al. Infrared imaging in breast cancer: automated tissue component recognition and spectral characterization of breast cancer cells as well as the tumor microenvironment. Analyst 2014; 139(5): 1044–1056. [PubMed: 24418921]

10. Farah I, Nguyen TNQ, Groh A et al. Development of a memetic clustering algorithm for optimal spectral histology: application to ftir images of normal human colon. Analyst 2016; 141(11): 3296–3304. [PubMed: 27110605]

11. Nallala J, Piot O, Diebold MD et al. Infrared and raman imaging for characterizing complex biological materials: a comparative morpho-spectroscopic study of colon tissue. Applied spectroscopy 2014; 68(1): 57–68. [PubMed: 24405955]

12. Varma VK, Kajdacsy-Balla A, Akkina S et al. Predicting fibrosis progression in renal transplant recipients using laser- based infrared spectroscopic imaging. Scientific reports 2018; 8(1): 686. [PubMed: 29330374]

13. Fernandez DC, Bhargava R, Hewitt SM et al. Infrared spectroscopic imaging for histopathologic recognition. Nature biotechnology 2005; 23(4): 469.

14. Atkins CG, Buckley K, Blades MW et al. Raman spectroscopy of blood and blood components. Applied spectroscopy 2017; 71(5): 767–793. [PubMed: 28398071]

15. Yang Y, Chen L and Ji M. Stimulated raman scattering microscopy for rapid brain tumor histology. Journal of Innovative Optical Health Sciences 2017; 10(05): 1730010.

16. Bunaciu AA, Aboul-Enein HY and Fleschin Ș. Vibrational spectroscopy in clinical analysis. Applied Spectroscopy Reviews 2015; 50(2): 176–191.

17. Orringer DA, Pandian B, Niknafs YS et al. Rapid intraoperative histology of unprocessed surgical specimens via fibre-laser-based stimulated raman scattering microscopy. Nature biomedical engineering 2017; 1(2): 0027.

18. Depciuch J, Kaznowska E, Zawlik I et al. Application of raman spectroscopy and infrared spectroscopy in the identification of breast cancer. Applied spectroscopy 2016; 70(2): 251–263. [PubMed: 26903561]

19. Talari ACS, Martinez MAG, Movasaghi Z et al. Advances in fourier transform infrared (ftir) spectroscopy of biological tissues. Applied Spectroscopy Reviews 2017; 52(5): 456–506.

20. Seo J, Hoffmann W, Warnke S et al. An infrared spectroscopy approach to follow β-sheet formation in peptide amyloid assemblies. Nature chemistry 2017; 9(1): 39.

21. Ellis DI and Goodacre R. Metabolic fingerprinting in disease diagnosis: biomedical applications of infrared and raman spectroscopy. Analyst 2006; 131(8): 875–885. [PubMed: 17028718]

22. Wray S, Cope M, Delpy DT et al. Characterization of the near infrared absorption spectra of cytochrome aa3 and haemoglobin for the non-invasive monitoring of cerebral oxygenation. Biochimica et Biophysica Acta (BBA)- Bioenergetics 1988; 933(1): 184–192. [PubMed: 2831976]

23. Pounder FN and Bhargava R. Toward automated breast histopathology using mid-ir spectroscopic imaging. Vibra- tional Spectroscopic Imaging for Biomedical Appications 2010; : 1–26.

24. Leslie LS, Wrobel TP, Mayerich D et al. High definition infrared spectroscopic imaging for lymph node histopathology. PLoS One 2015; 10(6): e0127238. [PubMed: 26039216]

25. Krafft C, Schmitt M, Schie IW et al. Label-free molecular imaging of biological cells and tissues by linear and nonlinear raman spectroscopic approaches. Angewandte Chemie International Edition 2017; 56(16): 4392–4430. [PubMed: 27862751]

26. Nazeer SS, Sreedhar H, Varma VK et al. Infrared spectroscopic imaging: Label-free biochemical analysis of stroma and tissue fibrosis. The international journal of biochemistry & cell biology 2017; 92: 14–17. [PubMed: 28888785]

27. Mankar R, Walsh MJ, Bhargava R et al. Selecting optimal features from fourier transform infrared spectroscopy for discrete-frequency imaging. Analyst 2018; 143(5): 1147–1156. [PubMed: 29404544]

28. Fabian H, Lasch P, Boese M et al. Mid-ir microspectroscopic imaging of breast tumor tissue sections. Biopolymers 2002; 67(4–5): 354–357. [PubMed: 12012465]

29. Gemignani J, Middell E, Barbour RL et al. Improving the analysis of near-spectroscopy data with multivariate classifi- cation of hemodynamic patterns: a theoretical formulation and validation. Journal of neural engineering 2018; 15(4): 045001. [PubMed: 29616976]

30. Bhargava R, Fernandez DC, Hewitt SM et al. High throughput assessment of cells and tissues: Bayesian classification of spectral metrics from infrared vibrational spectroscopic imag- ing data. Biochimica et Biophysica Acta (BBA) - Biomembranes 2006; 1758(7): 830–845. DOI:10.1016/j.bbamem.2006.05.007. URL http://linkinghub.elsevier.com/retrieve/pii/S0005273606001854. [PubMed: 16822477]

31. Tiwari S and Bhargava R. Extracting knowledge from chemical imaging data using computational algorithms for digital cancer diagnosis. The Yale journal of biology and medicine 2015; 88(2): 131–143. [PubMed: 26029012]

32. Mu X, Kon M, Ergin A et al. Statistical analysis of a lung cancer spectral histopathology (shp) data set. Analyst 2015; 140(7): 2449–2464. [PubMed: 25664623]

33. Mayerich DM, Walsh M, Kadjacsy-Balla A et al. Breast histopathology using random decision forests-based classification of infrared spectroscopic imaging data. In Proc. SPIE–Int. Soc. Opt. Eng, volume 9041 p. 904107.

34. Pezzei C, Pallua JD, Schaefer G et al. Characterization of normal and malignant prostate tissue by fourier transform infrared microspectroscopy. Molecular BioSystems 2010; 6(11): 2287–2295. [PubMed: 20871936]

35. Großerueschkamp F, Kallenbach-Thieltges A, Behrens T et al. Marker-free automated histopathological annotation of lung tumour subtypes by ftir imaging. Analyst 2015; 140(7): 2114–2120. [PubMed: 25529256]

36. Mayerich D, Walsh MJ, Kadjacsy-Balla A et al. Stain-less staining for computed histopathology. Technology 2015; 3(01): 27–31. [PubMed: 26029735]

37. Tian P, Zhang W, Zhao H et al. Intraoperative diagnosis of benign and malignant breast tissues by fourier transform infrared spectroscopy and support vector machine classification. International journal of clinical and experimental medicine 2015; 8(1): 972. [PubMed: 25785083]

38. Reddy RK, Walsh MJ, Schulmerich MV et al. High-definition infrared spectroscopic imaging. Applied spectroscopy 2013; 67(1): 93–105. [PubMed: 23317676]

39. Krizhevsky A, Sutskever I and Hinton GE. Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems pp. 1097–1105.

40. Baker MJ, Trevisan J, Bassan P et al. Using fourier transform ir spectroscopy to analyze biological materials. Nature protocols 2014; 9(8): 1771. [PubMed: 24992094]

41. Dong C, Loy CC, He K et al. Image super-resolution using deep convolutional networks. IEEE transactions on pattern analysis and machine intelligence 2016; 38(2): 295–307. [PubMed: 26761735]

42. Mobiny A, Moulik S, Gurcan I et al. Lung cancer screening using adaptive memory-augmented recurrent networks. arXiv preprint arXiv:171005719 2017;.

43. Jarrett K, Kavukcuoglu K, LeCun Y et al. What is the best multi-stage architecture for object recognition? In Computer Vision, 2009 IEEE 12th International Conference on. IEEE, pp. 2146–2153.

44. Hinton GE, Srivastava N, Krizhevsky A et al. Improving neural networks by preventing co-adaptation of feature detectors. arXiv preprint arXiv:12070580 2012;.

45. LeCun Y, Bengio Y and Hinton G. Deep learning. Nature 2015; 521(7553): 436–444. DOI: 10.1038/nature14539. URL http://www.nature.com/articles/nature14539. [PubMed: 26017442]

46. Dominik Scherer A and Behnke S. Evaluation of pooling operations in convolutional architectures for object recognition. Artificial Neural Networks (ICANN)-Lecture Notes in Computer Science 2010; 6354: 92–101.

47. Deng L, Yu D et al. Deep learning: methods and applications. Foundations and Trends® in Signal Processing 2014; 7(3–4): 197–387.

48. Cruz-Roa A, Basavanhally A, González F et al. Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. In Medical Imaging 2014: Digital Pathology, volume 9041 International Society for Optics and Photonics, p. 904103.

49. Glorot X and Bengio Y. Understanding the difficulty of training deep feedforward neural networks. In Proceedings of the thirteenth international conference on artificial intelligence and statistics pp. 249–256.

50. Kingma DP and Ba J. Adam: A method for stochastic optimization. arXiv preprint arXiv:14126980 2014;.

51. Zeiler MD. ADADELTA: an adaptive learning rate method. CoRR 2012; abs/1212.5701. URL http://arxiv.org/abs/1212.5701.1212.5701.

52. Ruder S An overview of gradient descent optimization algorithms. CoRR 2016; abs/1609.04747. URL http://arxiv.org/abs/1609.04747.1609.04747.

53. Srivastava N, Hinton G, Krizhevsky A et al. Dropout: a simple way to prevent neural networks from overfitting. The Journal of Machine Learning Research 2014; 15(1): 1929–1958.

54. Wang Z, Bovik AC, Sheikh HR et al. Image quality assessment: from error visibility to structural similarity. IEEE transactions on image processing 2004; 13(4): 600–612. [PubMed: 15376593]

55. Bhargava R Towards a practical Fourier transform infrared chemical imaging protocol for cancer histopathology. Analytical and Bioanalytical Chemistry 2007; 389(4): 1155–1169. DOI:10.1007/s00216-007-1511-9. URL http://link.springer.com/10.1007/s00216-007-1511-9. [PubMed: 17786414]

56. Berisha S, Chang S, Saki S et al. Siproc: an open-source biomedical data processing platform for large hyperspectral images. Analyst 2017; 142(8): 1350–1357. [PubMed: 27924319]
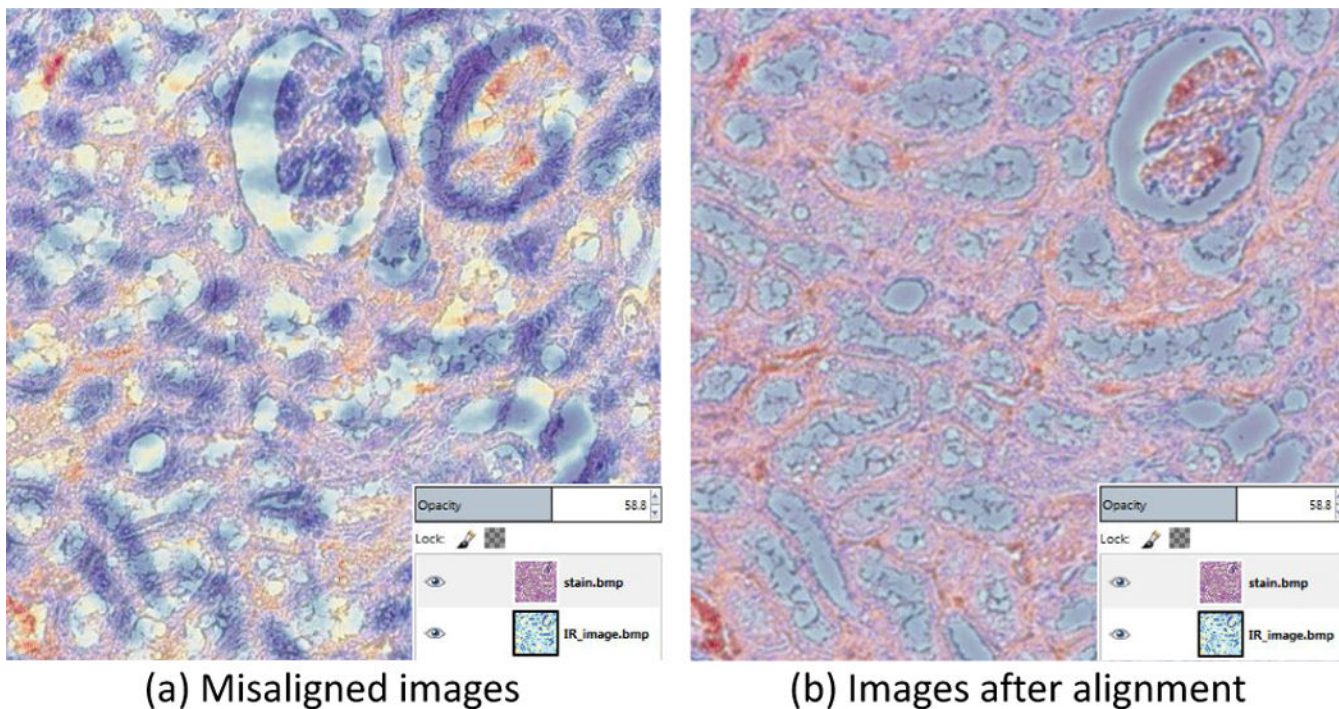
(a) Misaligned images

(b) Images after alignment

Figure 1.
Image alignment using GIMP. A single FTIR band was used for manual alignment. A transparency value was assigned to the H&E image in order to visualize the degree of misalignment (a), which was then corrected using affine transformations (b).
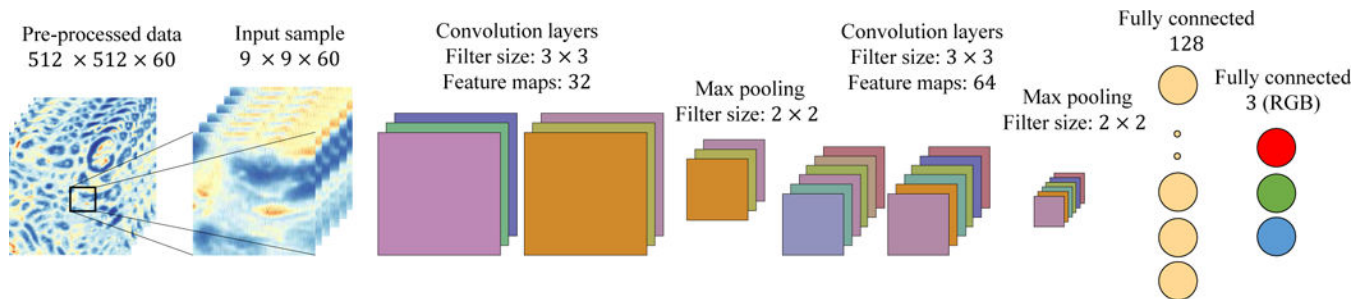
Figure 2.
Schematic presentation of our CNN model composed of convolution and pooling layers followed by two fully connected layers (a). The network is fed by a data cube of size $9 \times 9 \times 60$, which is a spatial crop around each pixel.
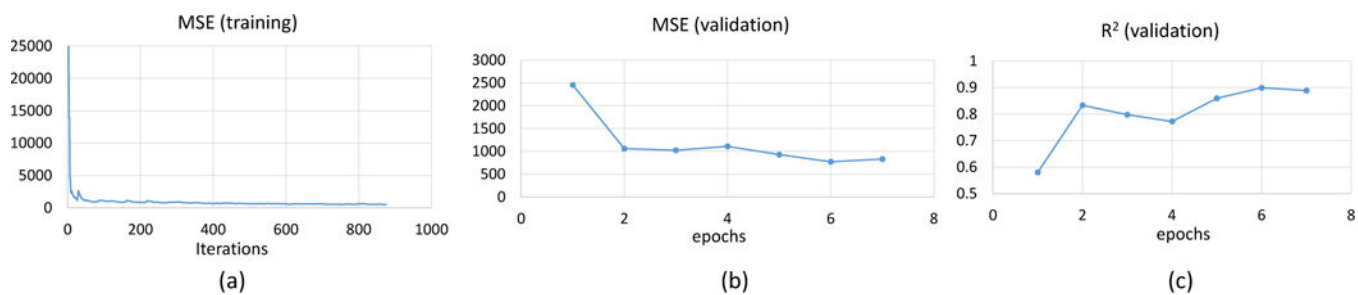
Figure 3.
MSE plot convergence (a) with respect to iteration numbers during training on training data (b) with respect to epoch numbers for validation data. (c) Coefficient of determination (R^2) of validation data during training process computed after each epoch
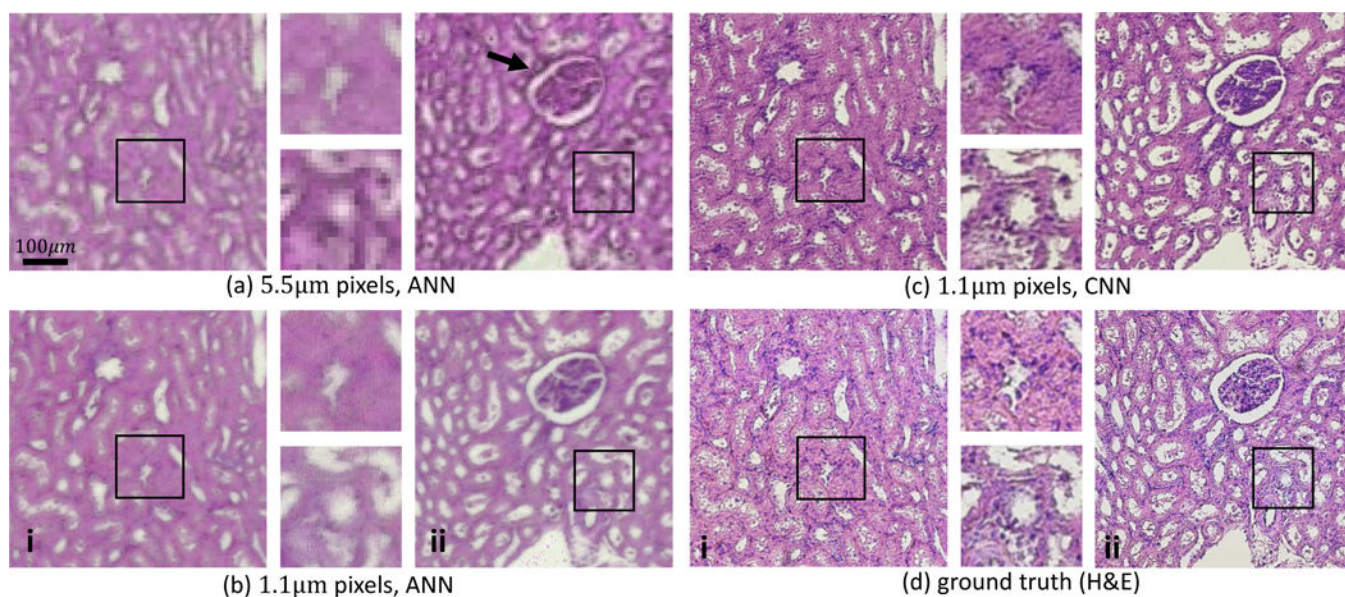
Figure 4.
Molecular imaging reproduced by chemical imaging using different digital staining methods and varying spatial resolutions for two pig kidney tissue sections i and ii. Standard digital H&E staining is shown using an ANN[36] on (a) standard-magnification (5.5 μm pixel size) and (b) high-magnification (1.1 μm pixel size) images. (c) Digital H&E using the proposed CNN framework on high-magnification data shows greater cellular-level detail when compared to the ground truth (d) chemically-stained H&E images collected with a bright-field microscope. Improved spatial resolution provided by the CNN is clear in the magnified insets, as well as in the kidney glomerulus (black arrow), where boundries are more clearly defined. Note that some deformations were introduced between the FTIR (a-c) and bright-field (d) images because chemical staining was performed between these two steps.
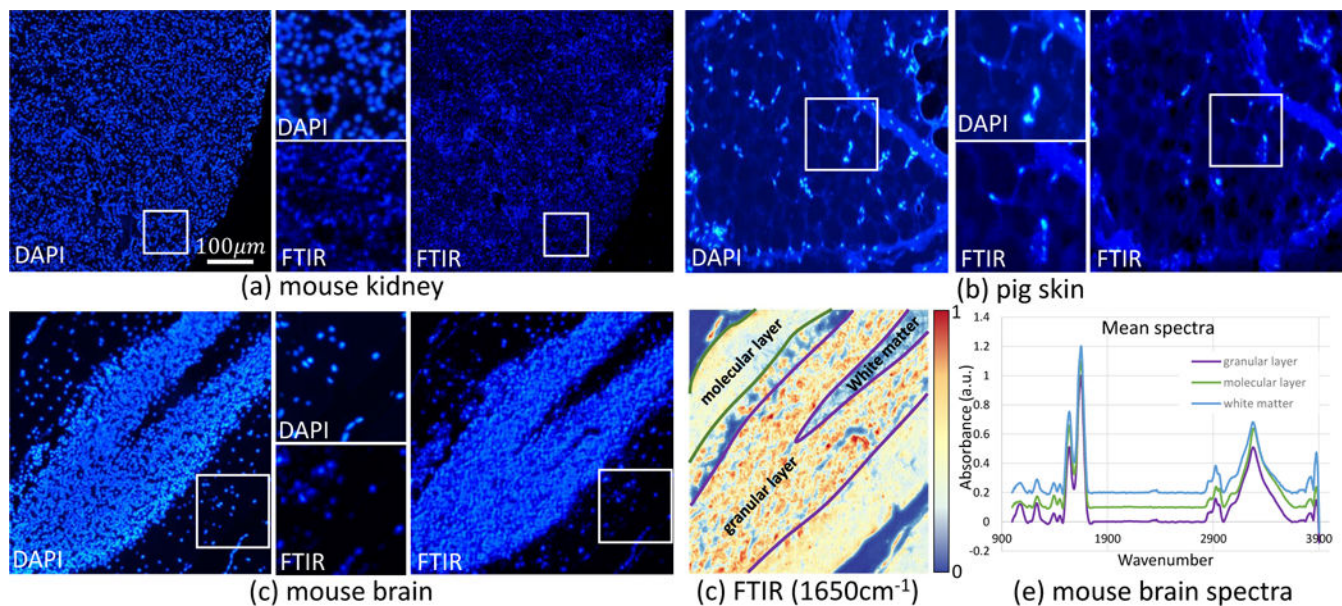
Figure 5.
Digital DAPI staining of (a) mouse kidney, (b) pig skin, and (c) mouse brain respectively
using our CNN framework. Digitally stained FTIR images are shown on the right, with
close-ups on the bottom. (d) FTIR absorbance image at 3300 cm^{-1} indicating the regions
used to collect average spectra. (e) Mean spectra were extracted using SIproc[56] from masks
of the molecular layer (green), granular layer (purple), and white matter (blue) of the
cerebellum. Spectra are baseline corrected, normalized to Amide I, and offset for clarity.

Table 1.

SSIM index for the different digital staining method outputs shown in Figure 4.

SSMI	Low mag. ANN	High mag. ANN	High mag. CNN
Test Image (i)	0.61	0.62	0.70
Test Image (ii)	0.58	0.63	0.71