Fluorescence Image Histology Pattern Transformation Using Image Style Transfer

Mohammadhassan Izadyyazdanabadi[1,2], Evgenii Belykh[2], Xiaochun Zhao[2],
Leandro Borba Moreira[2], Sirin Gandhi[2], Claudio Cavallo[2], Jennifer Eschbacher[3],
Peter Nakaji[3], Mark C. Preul[2] and Yezhou Yang[1]*

[1] School of Computing, Informatics, and Decision System Engineering, Arizona State University, Tempe, AZ, United States,
[2] The Loyal and Edith Davis Neurosurgery Research Laboratory, Department of Neurosurgery, St. Joseph's Hospital and
Medical Center, Barrow Neurological Institute, Phoenix, AZ, United States, [3] Department of Neuropathology, St. Joseph's
Hospital and Medical Center, Barrow Neurological Institute, Phoenix, AZ, United States

Confocal laser endomicroscopy (CLE) allow on-the-fly *in vivo* intraoperative imaging in
a discreet field of view, especially for brain tumors, rather than extracting tissue for
examination *ex vivo* with conventional light microscopy. Fluorescein sodium-driven CLE
imaging is more interactive, rapid, and portable than conventional hematoxylin and eosin
(H&E)-staining. However, it has several limitations: CLE images may be contaminated
with artifacts (motion, red blood cells, noise), and neuropathologists are mainly trained
on colorful stained histology slides like H&E while the CLE images are gray. To improve the
diagnostic quality of CLE, we used a micrograph of an H&E slide from a glioma tumor
biopsy and image style transfer, a neural network method for integrating the content
and style of two images. This was done through minimizing the deviation of the target
image from both the content (CLE) and style (H&E) images. The style transferred images
were assessed and compared to conventional H&E histology by neurosurgeons and
a neuropathologist who then validated the quality enhancement in 100 pairs of original
and transformed images. Average reviewers' score on test images showed 84 out of 100
transformed images had fewer artifacts and more noticeable critical structures compared
to their original CLE form. By providing images that are more interpretable than the original
CLE images and more rapidly acquired than H&E slides, the style transfer method allows
a real-time, cellular-level tissue examination using CLE technology that closely resembles
the conventional appearance of H&E staining and may yield better diagnostic recognition
than original CLE grayscale images.

Keywords: confocal laser endomicroscopy, fluorescence, digital pathology, brain tumor imaging, deep learning,
image style transfer

INTRODUCTION

Confocal laser endomicroscopy (CLE) is undergoing rigorous assessment for its potential to assist
neurosurgeons to examine tissue *in situ* during brain surgery (1–5). The ability to scan tissue or
surgical resection bed on-the-fly essentially producing optical biopsies, compatibility with different
fluorophores, miniature size of the probe and the portability of the system are essential features
of this promising technology. Currently, the most frequent technique used for neurosurgical

intraoperative diagnosis is examination of frozen section hematoxylin and eosin (H&E)-stained histology slides.

Figure 1A shows an example image from a glioma acquired by CLE (left) and a micrograph of an H&E slide (right), acquired by conventional light microscopy. Although generating CLE images is much faster than H&E slides (1 s per image compared to about 20 min per slide), many CLE images may be non-optimal and can be obscured with artifacts including background noise, blur, and red blood cells (6). The histopathological features of gliomas are often more identifiable in the H&E slide images compared to the CLE images generated using non-specific fluorescent dyes such as fluorescein sodium (FNa). Neuropathologists as well are used to evaluating detailed histoarchitecture colorfully stained with H&E for diagnoses, especially for frozen section biopsies. Fluorescent images from intraoperative neurosurgical application present a new digital gray scale (monochrome) imaging environment to the neuropathologist for diagnosis that may include hundreds of images from one case. Recently, the U.S. FDA has approved a blue laser range CLE system primarily utilizing FNa for use in neurosurgery.

Countervailing these diagnostic and visual deficiencies in CLE images requires a rapid, automated transformation that can: (1) remove the occluding artifacts, and (2) add (amplify) the histological patterns that are difficult to recognize in the CLE images. Finally, this transformation should avoid removing the critical details (e.g., cell structures, shape) or adding unreal patterns to the image, to maintain the integrity of the image content. Such a method may present "transformed" CLE images to the neuropathologist and neurosurgeon that may resemble images based on familiar and standard, even colored, appearances from histology stains, such as H&E.

One method for implementing this transformation could be supervised learning, however, supervised learning requires paired images (from the same object and location) to learn the mapping between the two domains (CLE and H&E). Creating a dataset of colocalized H&E and CLE images is infeasible because of problems in exact co-localization and intrinsic tissue movements, although small, and artifacts introduced during H&E slide preparation. "Image style transfer," first introduced by Gatys et al. (7), is an image transformation technique that blends

generated from non-specific FNa application during glioma surgery appear like standard H&E-stained histology and evaluate the accuracy and usefulness. We used the image style transfer method since it extracts abstract features from the CLE and H&E image that are independent of their location in the image and thus can operate on the images that are not from the same location. More details about the image style transfer algorithm and the quality assessment protocol follow in section Methods. Our results from a test dataset showed that on average, the diagnostic quality of stylized images was higher than the original CLE images, although there were some cases where the transformed image showed new artifacts.

METHODS

Image Style Transfer

Image style transfer takes a content and style image as input and produces a target image that shows the structures of the content image and the general appearance of the style image. This is achieved through four main components: (1) a pretrained CNN that extracts feature maps from source and target images, (2) quantitative calculation of the content and style representations for source and target images, (3) a loss function to capture the difference between the content and style representation of source and target images, and (4) an optimization algorithm to minimize the loss function. In contrast to CNN supervised learning, where the model parameters are updated to minimize the prediction error, image style transfer modifies the pixel values of the target image to minimize the loss function while the model parameters are fixed.

A 19-layer visual geometry group network (VGG-19), that is pretrained on ImageNet dataset, extracts feature maps from CLE, H&E, and target images. Feature maps in layer "Conv4_2" of VGG-19 are used to calculate the image content representation, and a list of gram matrices from feature maps of five layers ("ReLU1_1", "ReLU2_1", …, "ReLU5_1") are used to calculate the image style representation. To examine the difference between the target and source images, the following loss function was used:

$$Loss_{Total} = \underbrace{\frac{1}{2} \sum \left(C_{CLE} - C_{Target} \right)^2}_{\substack{\text{Content Loss: difference} \\ \text{between content} \\ \text{representation of the} \\ \text{CLE and target image}}} + \underbrace{\alpha \times \sum_{i=1}^{5} w^i \times \sum \left(S_{H\&E}^i - S_{Target}^i \right)^2}_{\substack{\text{Style Loss: difference} \\ \text{between style} \\ \text{representation of the H\&E} \\ \text{and target image}}}$$

the content and style of two separate images to produce the target image. This process minimizes the distance between feature maps of the source and target images using a pretrained convolutional neural network (CNN).

In this study, we aimed to remove the inherent occlusions and enhance the structures that were problematical to recognize in CLE images. Essentially, we attempted to make CLE images

C_{CLE} and C_{Target} are the content representations of the CLE and target image, $S_{H\&E}^i$ and S_{Target}^i are the style representations of the H&E and target image based on the feature maps of the ith layer, and w^i (weight of ith layer in the style representation) equals 0.2. The parameter α determines relative weight of style loss in the total loss and is set to 100. A limited memory optimization algorithm [L-BFGS (8)] minimizes this loss.
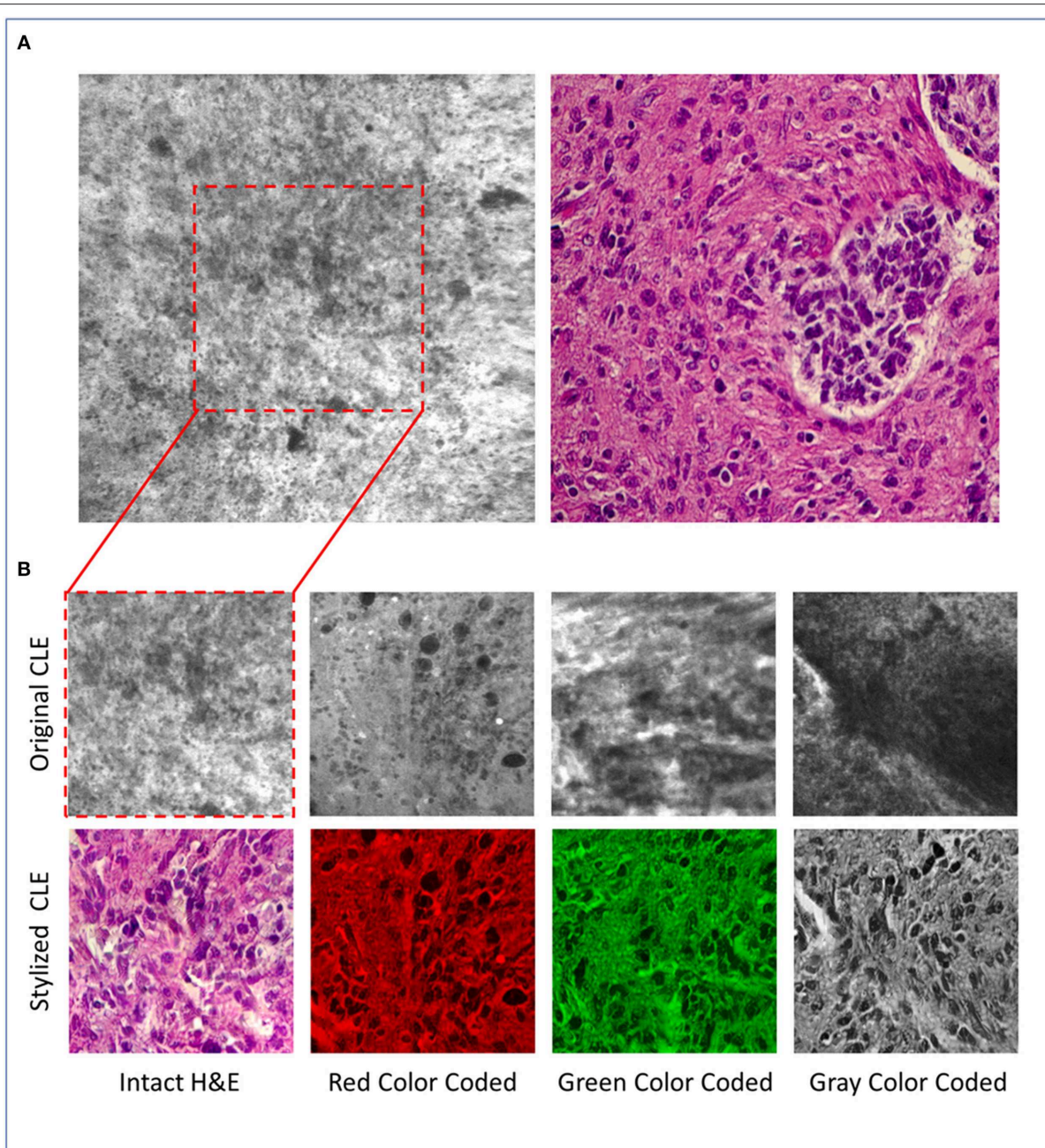
FIGURE 1 | (A) Representative CLE (Optiscan 5.1, Optiscan Pty., Ltd.) and H&E images from glioma tumors. **(B)** Original and stylized CLE images from glioma tumors, in 4 color coding: gray, green, red, intact H&E.

For the experiment, 100 CLE images [from a recent study by Martirosyan et al. (5)] were randomly selected from 15 subjects with glioma tumors as content images. A single micrograph of an H&E slide from a glioma tumor biopsy of a different patient (not one of the 15 subjects) was used as the style image (**Figure 1A**, right). For each CLE image, the optimization process was run for 1,600 iterations and the target image was saved for evaluation and referred to as the "stylized image" in the following sections.

Evaluation

Although the stylized images presented the same histological patterns as H&E images and seemed to contain similar structures to those present in the corresponding original CLE images, a quantitative image quality assessment was performed to rigorously evaluate the stylized images. Five neurosurgeons independently assessed the diagnostic quality of the 100 pairs of original and stylized CLE images. For each pair, the reviewers sought to examine two properties in each stylized image and

provided a score for each property based on their evaluation: (1) whether the stylization process removed any critical structures (negative impact) or artifacts (positive impact) that were present in the original CLE image, and (2) whether the stylization process added new structures that were not present (negative impact) or were difficult to notice (positive impact) in the original CLE image. The scores are between 0 and 6 with the following annotations: 0, extreme negative impact; 1, moderate negative impact; 2, slight negative impact; 3, no significant impact; 4, slight positive impact; 5, moderate positive impact; and 6, extreme positive impact (Further information and instructions about the quality assessment survey is available in the **Supplementary Materials**).

Since the physicians were more familiar with the H&E than the CLE images, it was possible that the reviewers would overestimate the quality of stylized images merely due to their color resemblance to H&E images (during style transfer the color of CLE image is also changed to pink and purple). To explore how the reviewers' scores would change if the stylized images were presented in a different color other than the pink and purple (the common color for H&E images), the stylized images were processed in four different ways: (I) 25 stylized images were converted into gray-scale images (averaging the three red, green, and blue channels), (II) 25 stylized images were color-coded in green (first converted the image to gray-scale and then set red and blue channels to zero), (III) 25 stylized images were color-coded as red (similar approach), and (IV) 25 stylized images were used without any further changes (intact H&E). Since there are too many structures in each CLE image, and to examine the images more precisely, we used the center-crop of each original CLE and its stylized version for evaluation. **Figure 1B** shows some example stylized images used for evaluation.

RESULTS AND ANALYSIS

Figure 2A shows a histogram of all reviewers' scores for the removed artifacts (blue bars) and added structures (orange bars) in the stylized images with different colors. Overall, the number of stylized CLE images that have higher diagnostic quality than the original images (score >3) was significantly larger than those with equal or lower diagnostic quality for both removed artifacts and added structures scores (one-way chi square test $p < 0.001$). Results from stylized images that were color-coded (gray, green, red) showed the same trend for the added structures scores, indicating that the improvement was not just because of color resemblance.

There was significant difference between how much the model added structures and removed artifacts. For all the color-coded and intact stylized images, the average of added structures scores was larger than the removed artifacts scores (t-test $p < 0.001$). This suggests that the model was better at enhancing the structures that were challenging to recognize than removing the artifacts.

Figure 2B shows the frequency of different combinations of scores for removed artifacts and added structures in an intensity map. Each block represents how many times a rater scored an
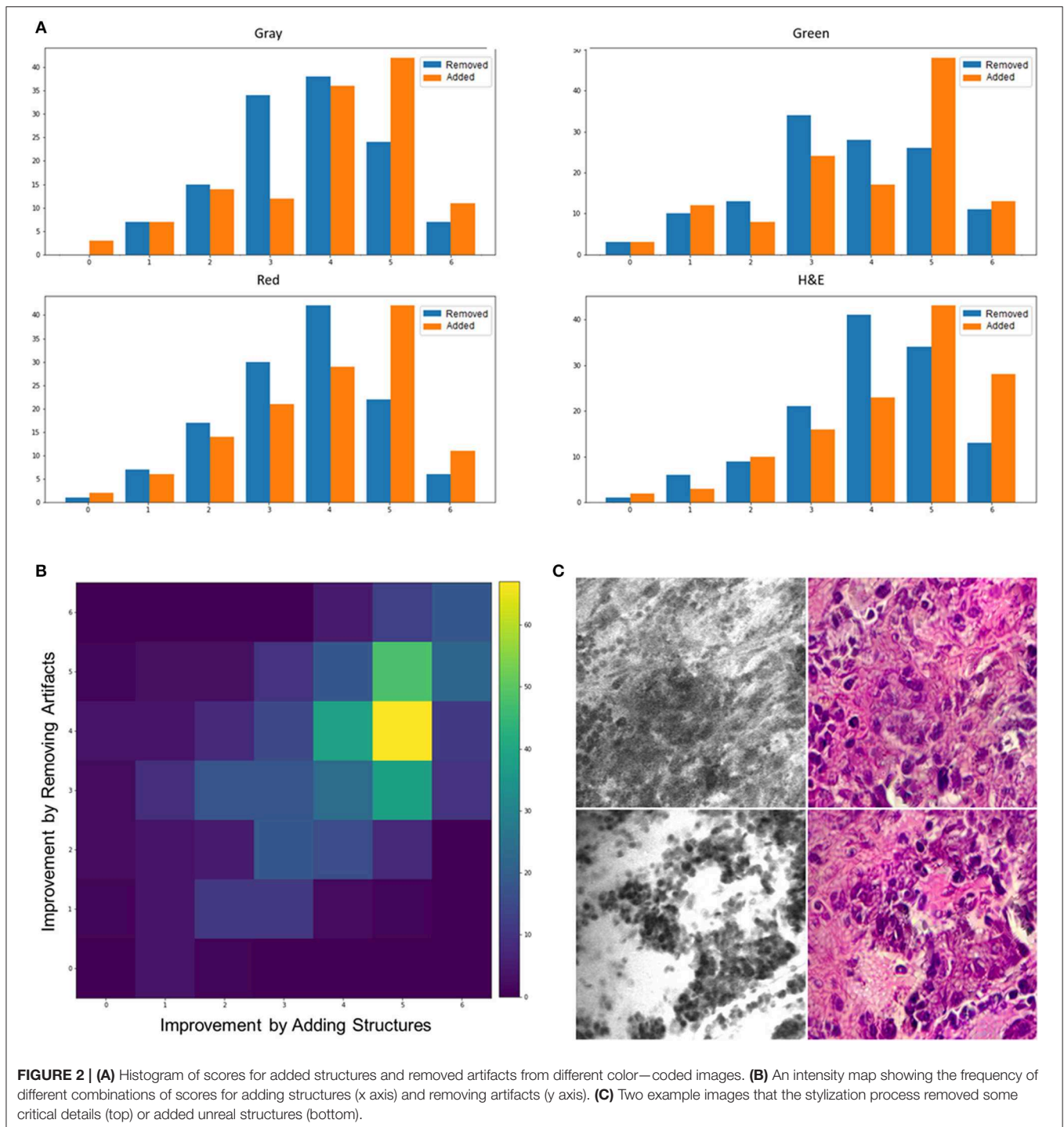
image with the corresponding values on the x (improvement by added structures) and y (improvement by removed artifacts) axes for that block. The most frequent incident across all the stylized images is the coordinates (5,4), which means moderately adding structures and slightly removing artifacts, followed by (5,5) meaning moderately adding structures and removing artifacts. Although the intensity maps derived from different color-coded images were not exactly similar, the most frequent combination in each group still indicated positive impact in both properties. The most frequent combination of scores, for each of the color-coded images, was as follows: gray = (5,4), green = (5,5), red = (5,4), and intact = (5,4).

As a further analysis, we counted the number of images that had an average score of below three to see how often the algorithm removed critical structures or added artifacts that were misleading to the neurosurgeons. From the 100 tested images, 3 images had only critical structures removed, 4 images had only artifacts or unreal added structures, and 2 images had both artifacts added and critical structures removed. On the contrary, 84 images showed improved diagnostic quality through both removed artifacts and added structures that were hard to recognize, 6 images had only artifacts removed, and 5 images had only critical structures added. **Figure 1B** shows some example stylized images with improved quality compared to the original CLE, and **Figure 2C** shows two stylized images with decreased diagnostic quality through removed critical structures (top) and added artifacts (bottom).

DISCUSSION

In this study, image style transfer method was used to improve the diagnostic quality of CLE images from glioma tumors by transforming the histology patterns in CLE images of fluorescein sodium-stained tissue into the ones in conventional H&E-stained tissue slides. From the 100 test images, 90 images showed reduced artifacts and 89 images showed improvement in difficult structures (more identifiable) after the transformation, confirmed by five neurosurgeon reviewers. Similar results in color coded images (gray, green, red) suggested that the improvement was not solely because of the color resemblance of stylized images to the H&E slide images. The same method could be used to transform CLE videos as shown in the **Supplementary Materials**.

In a related study by Gu et al. (9), cycle-consistent generative adversarial network (CycleGAN) was used as a data synthesis method to overcome the limited size of dataset in probe-based CLE (pCLE). A real dataset of pCLE images and histology slide images were used to train a generative model that transformed the histology slide images into pCLE. The pCLE dataset was then augmented with the synthesized images to train a computer aided diagnostic system for classification of breast lesions that are stained with acriflavine hydrochloride. However, for use in the *in vivo* brain, fluorophores are currently extremely limited—only three are approved: 5-ALA, indocyanine green, and fluorescein sodium. Currently CLE in the brain only uses fluorescein sodium. Each of these fluorophores yields completely unique unrelated appearances. However, our aim was to evaluate the diagnostic

FIGURE 2 | (A) Histogram of scores for added structures and removed artifacts from different color—coded images. **(B)** An intensity map showing the frequency of different combinations of scores for adding structures (x axis) and removing artifacts (y axis). **(C)** Two example images that the stylization process removed some critical details (top) or added unreal structures (bottom).

quality of images that are obtained using fluorescein sodium stained biopsy tissue of brain lesions acquired during intracranial surgery. Fluorescein sodium is a non-specific stain which collects in the tissue background or extracellular space, thus does not produce images such as are stained with acriflavine, which is more specific and stains intracellular structures. Such a situation using fluorescein is thus even more of a diagnostic challenge and

for transforming patterns from the CLE image domain to one of the familiar H&E histopathology images.

Although other studies have used CycleGAN for medical domain adaptation (9–11), none has applied "image style transfer" method for transforming CLE to H&E. Despite CycleGAN requirement of many images from both modalities (CLE and H&E) to train the model, image style transfer

can produce transformed images with a single image of each modality. This aspect of efficiency may have an important impact in surgery or pathology interpretation where there is a limited number of images available. Such an aspect of efficiency may lend itself better for inclusion into the CLE operating system for automated decision-making.

The purpose of our study was to evaluate the impact of image style transfer on the diagnostic quality of CLE brain and brain tumor (glioma) images from the neurosurgeons' perspectives. While studies like (9) explore the impact of synthesized data on computer-aided diagnostic methods, our study focused on improving the quality of CLE images from the physician point of view. Due to the intrinsic differences and mechanical limitations between the intraoperative CLE with a field of view of only 400 microns, neurosurgical biopsy instruments that acquire about 1 mL of tissue at minimum, normal physiologic brain movements, and light microscopy imaging of histology sections and slides, it is nearly impossible to image and correlate the same exact location with the two modalities. This limited us to compare the transformed CLE images with the potential real H&E stained sections of the same area, using measures such as structural similarity as performed by Shaban et al. (11) in their study. Indeed, the only available ground truth is the original CLE image (before transformation) which could be used to examine if any critical structures are removed or if unreal cellular architecture is added.

CONCLUSIONS

In this study, image style transfer was applied to CLE images from gliomas to enhance their diagnostic quality. Style transfer with an H&E-stained slide image had an overall positive impact on the diagnostic quality of CLE images. The improvement was not solely because of the colorization of CLE images; even the stylized images that were converted to gray, red, and green, reported improved diagnostic quality compared to the original CLE images. Employment of more specific clinical tasks to explore the advantage of stylization in diagnosing gliomas and other tumor types is underway based on this preliminary success. The fact that the style transfer is based on permanent histology H&E, provides an intraoperative advantage. Initial pathology diagnosis for brain tumor surgery is usually based on frozen section histology, with formal diagnosis awaiting the inspection of permanent histology slides requiring one to several days. The style transfer is based on rapidly acquired, on-the-fly (i.e., real time) *in vivo* intraoperative CLE images that most resemble the permanent histology; therefore, it is a significant advantage for interpretation. Frozen section histology often involves freezing artifact, cutting problems, and may have inconsistent staining for histological characteristics that are important for diagnosis. Style transferred CLE images are then most comparable to the permanent histology, and may be even better because CLE is imaging live tissue without such architectural disturbance.

In the future, application of more advanced methods for transferring patterns in the histology slides to the CLE images will be used to improve their interpretability. Because of the high number of CLE images acquired during a single case, style transfer could add value to such fluorescence images, and allows for computer-aided techniques to play a meaningful, convenient, and efficient role to aid the neurosurgeon and neuropathologist in analysis of CLE images and to more rapidly determine diagnosis.

DATA AVAILABILITY

The datasets generated for this study are available on request to the corresponding author.

AUTHOR'S NOTE

This research article was part of a Ph.D. dissertation (12).

AUTHOR CONTRIBUTIONS

MI, EB, YY, and MP conceived and wrote the article. MI and YY formulated computer learning systems. LM, XZ, SG, CC, and EB assessed the quality of images. PN acquired images from neurosurgical cases. JE examined the CLE and histology images.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2019.00519/full#supplementary-material

REFERENCES

1. Martirosyan NL, Georges J, Eschbacher JM, Cavalcanti DD, Elhadi AM, Abdelwahab MG, et al. Potential application of a handheld confocal endomicroscope imaging system using a variety of fluorophores in experimental gliomas and normal brain. *Neurosurg Focus*. (2014) 36:E16. doi: 10.3171/2013.11.FOCUS13486

2. Foersch S, Heimann A, Ayyad A, Spoden GA, Florin L, Mpoukouvalas K, et al. Confocal laser endomicroscopy for diagnosis and histomorphologic imaging of brain tumors *in vivo*. *PLoS ONE*. (2012) 7:e41760. doi: 10.1371/journal.pone.0041760

3. Belykh E, Miller EJ, Patel AA, IzadyYazdanabadi M, Martirosyan NL, Yagmurlu K, et al. Diagnostic accuracy of the confocal laser endomicroscope for *in vivo* differentiation between normal and tumor tissue during fluorescein-guided glioma resection: laboratory investigation.

World Neurosurg. (2018) 115:e337–48. doi: 10.1016/j.wneu.2018.04.048

4. Izadyyazdanabadi M, Belykh E, Mooney M, Martirosyan N, Eschbacher J, Nakaji P, et al. Convolutional neural networks: ensemble modeling, fine-tuning and unsupervised semantic localization for neurosurgical CLE images. *J Vis Commun Image Represent*. (2018) 54:10–20. doi: 10.1016/j.jvcir.2018.04.004

5. Martirosyan NL, Eschbacher JM, Kalani MYS, Turner JD, Belykh E, Spetzler RF, et al. Prospective evaluation of the utility of intraoperative confocal laser endomicroscopy in patients with brain neoplasms using fluorescein sodium: experience with 74 cases. *Neurosurg Focus*. (2016) 40:E11. doi: 10.3171/2016.1.FOCUS 15559

6. Izadyyazdanabadi M, Belykh E, Martirosyan N, Eschbacher J, Nakaji P, Yang Y, et al. Improving utility of brain tumor confocal laser endomicroscopy: objective value assessment and diagnostic frame detection with convolutional neural networks. In: *Progress in Biomedical Optics and Imaging - Proceedings of SPIE* (Orlando, FL), (2017).

7. Gatys LA, Ecker AS, Bethge M. Image Style Transfer using convolutional neural networks. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (Las Vegas, NV), (2016).

8. Zhu C, Byrd RH, Lu P, Nocedal J. Algorithm 778: L-BFGS-B: fortran subroutines for large-scale bound-constrained optimization. *ACM Trans Math Softw*. (1997) 23:550–60. doi: 10.1145/279232.279236

9. Gu Y, Vyas K, Yang J, Yang G. Transfer recurrent feature learning for endomicroscopy image recognition. *IEEE Trans Med Imaging*. (2019) 38:791–801. doi: 10.1109/TMI.2018.2872473

10. Rivenson Y, Wang H, Wei Z, de Haan K, Zhang Y, Wu Y, et al. Virtual histological staining of unlabelled tissue-autofluorescence images via deep learning. *Nat Biomed Eng*. (2019) 3:466–77. doi: 10.1038/s41551-019-0362-y

11. Shaban MT, Baur C, Navab N, Albarqouni S. Staingan: stain style transfer for digital histological images. *arXiv [Preprint]. arXiv:180401601*. (2018). Available online at: https://arxiv.org/abs/1804.01601

12. Yazdanabadi MI. *Confocal laser endomicroscopy image analysis with deep convolutional neural networks* (Doctoral dissertation). Arizona State University, Tempe, AZ, United States (2019).