Computers in Biology and Medicine 180 (2024) 108958

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Deep learning for rapid virtual H&E staining of label-free glioma tissue from hyperspectral images

Ruohua Zhu [a,1], Haiyang He [a,1], Yuzhe Chen [a], Ming Yi [a], Shengdong Ran [a], Chengde Wang [c,**], Yi Wang [a,b,*]

[a] National Engineering Research Center of Ophthalmology and Optometry, School of Biomedical Engineering, Eye Hospital, Wenzhou Medical University, Xueyuan Road 270, Wenzhou, 325027, China
[b] Wenzhou Institute, University of Chinese Academy of Sciences, Jinlian Road 1, Wenzhou, 325001, China
[c] Department of Neurosurgery, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou 325000, China

ARTICLE INFO

Keywords:
Hematoxylin-eosin staining
Virtual staining
Deep learning
Generative adversarial network
Hyperspectral imaging technology

ABSTRACT

Hematoxylin and eosin (H&E) staining is a crucial technique for diagnosing glioma, allowing direct observation of tissue structures. However, the H&E staining workflow necessitates intricate processing, specialized laboratory infrastructures, and specialist pathologists, rendering it expensive, labor-intensive, and time-consuming. In view of these considerations, we combine the deep learning method and hyperspectral imaging technique, aiming at accurately and rapidly converting the hyperspectral images into virtual H&E staining images. The method overcomes the limitations of H&E staining by capturing tissue information at different wavelengths, providing comprehensive and detailed tissue composition information as the realistic H&E staining. In comparison with various generator structures, the Unet exhibits substantial overall advantages, as evidenced by a mean structure similarity index measure (SSIM) of 0.7731 and a peak signal-to-noise ratio (PSNR) of 23.3120, as well as the shortest training and inference time. A comprehensive software system for virtual H&E staining, which integrates CCD control, microscope control, and virtual H&E staining technology, is developed to facilitate fast intra-operative imaging, promote disease diagnosis, and accelerate the development of medical automation. The platform reconstructs large-scale virtual H&E staining images of gliomas at a high speed of 3.81 mm^2/s. This innovative approach will pave the way for a novel, expedited route in histological staining.

1. Introduction

Glioma is the most common primary intracranial tumor in adults, accounting for 81 % of malignant brain tumors. It occurs in any part of the central nervous system and has a high mortality and incidence rate. Currently, the principal imaging examinations for gliomas include computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET) [1], which can reveal brain anatomical features, tumor lesion location, size, surrounding edema, and blood-brain barrier disruption. Other methods such as fluorescence [2] and Raman [3,4] spectroscopy have also been reported for the detection and imaging of cancer tissues. However, these devices are bulky and not easily suitable for real-time intraoperative detection. Moreover, imaging methods often result in blurred tumor borders, making it challenging to accurately determine glioma lesion boundaries and categorize gliomas.

The diagnosis and grading of brain gliomas require specimens obtained through tumor resection surgery or biopsy surgery, followed by an integrated histopathological and molecular pathological diagnosis to determine tumor grade and stage. Histopathological analysis involves microscopic observation of tissue-stained sections using stains such as H&E stains. H&E staining can differentiate cell nuclei in blue-purple and cytoplasm in pink, aiding in the distinction of different cell and tissue structures. WHO CNS5 classifies gliomas into six types based on histological morphology [5], immunophenotype, molecular diagnosis, etc. H&E staining is used to observe histological morphology in the report,

* Corresponding author. National Engineering Research Center of Ophthalmology and Optometry, School of Biomedical Engineering, Eye Hospital, Wenzhou Medical University, Xueyuan Road 270, Wenzhou, 325027, China.
** Corresponding author.
E-mail addresses: yihe723@126.com (C. Wang), yiwang@wmu.edu.cn (Y. Wang).
[1] Equal contribution authors.

https://doi.org/10.1016/j.compbiomed.2024.108958
Received 22 March 2024; Received in revised form 2 July 2024; Accepted 26 July 2024
Available online 1 August 2024

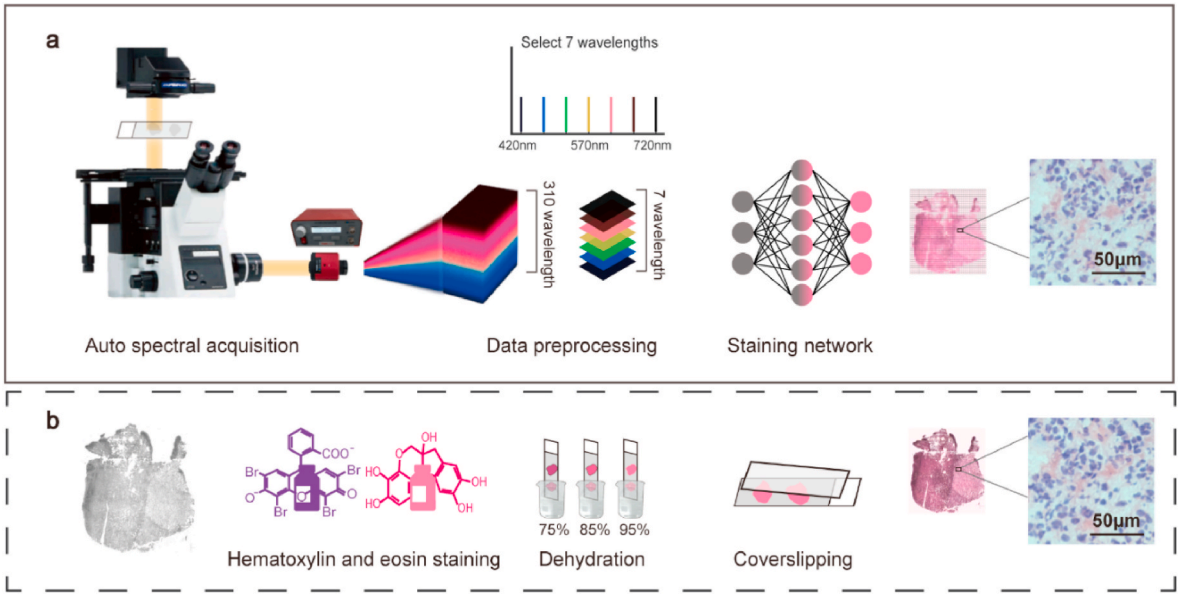

Fig. 1. Comparisons of staining processes: (a) The deep learning method employs a developed automatic slice scanning system to obtain high-spectral images of glioma tissues. Data from seven wavelength channels are extracted and then fed into a pre-trained neural network for virtual staining. (b) The traditional H&E staining process.

allowing precise determination of tumor boundaries at the cellular level. However, chemical staining is a time-consuming and complex process that is susceptible to variability in manual operation, resulting in issues such as weak image contrast and unclear tissue structures. Additionally, it is irreversible, non-repeatable, and may lead to staining contamination.

In recent years, deep learning has shown a wide range of applications in the field of image processing, for example, stereo image segmentation based on convolutional neural networks (CNN) [6], multi-level fusion, and attention-guided CNN image dehazing methods [7], deep learning-based face recognition algorithms [8], and cartoon image pixelation [9]. Moreover, deep learning has demonstrated remarkable performance in medical image processing [10–21]. Helong Yu et al. proposed the TinyBERT-CNN fusion model for intent classification of "Treatise on Febrile diseases" [22]. Tao Zhou et al. summarized the latest deep learning-based image fusion methods [23]. Generative Adversarial Networks (GANs) are a typical medical image generation network [21, 24–27], and Pix2Pix GAN is a variant form of conditional adversarial networks. It utilizes the Unet structure as the generator and Patch GAN as the discriminator, thereby enabling the generation of realistic images. This network exhibits excellent realism in virtual staining of pathological tissues, making it challenging for even experienced pathologists to distinguish [12].

Deep learning-based virtual staining approaches can be categorized into two types: virtual staining of unstained tissues and transformations between various histological stains. Mainstream virtual histological staining utilizes spontaneous fluorescence images of label-free tissues as network inputs. For instance, Rivenson et al. employed a generative adversarial network to convert label-free images of human tissue samples into multiple stains, including H&E (salivary gland and thyroid), Jones stain (kidney), and Masson's trichrome (liver and lung) [12]. Yuzhu Li et al. employed neural networks to conduct virtual staining on autopsy tissues. This virtual autopsy staining technique was capable of producing artifact-free H&E staining even under severe autolysis and cell death conditions [28]. Ming Y. Lu et al. introduced a visual-linguistic foundation model called CONCH, which achieved state-of-the-art performance in tissue pathology image classification, segmentation, captioning, text-to-image, and image-to-text retrieval tasks [29]. More sophisticated optical techniques have demonstrated effectiveness in virtual staining of label-free tissues, such as two-photon excitation microscopy (TPEF) [13], stimulated Raman scattering (SRS) microscopy [14], reflective confocal microscopy (RCM) [15], quantitative phase imaging (QPI) [30], ultraviolet photoacoustic microscopy (UV-PAM) [16], and ultraviolet photoacoustic remote sensing microscopy (UV-PARS) [31]. Furthermore, Yijie et al. combined virtual staining of label-free tissues with a digital staining matrix, achieving a diverse set of staining on the same tissue section [32]. The alternative approach involved utilizing unsupervised training networks, like the cycle generative adversarial network (CycleGAN [27,33]), to reconstruct images of stained tissue in diverse staining styles. For instance, de Haan et al. trained a deep neural network to convert H&E images of human kidney samples into specific stains, including Jones silver, MT, and PAS stains [34]. Levy et al. used an H&E stain of human liver tissue to generate a virtual trichrome stain for investigating the staging of liver fibrosis [17]. Additionally, there are endeavors to generate immunohistochemical (IHC) staining from H&E images for cancer diagnosis [35].

Deep learning-based virtual staining methods allow for rapid and accurate tissue virtual staining without the need for dyes or specialized pathologists, which does not cause damage to the original tissues. These methods can provide clinicians with quick diagnostic assistance in clinical settings, offering more possibilities for disease diagnosis and treatment. However, the accuracy of virtual staining relies on the richness of information in the input image. Consequently, numerous studies leverage advanced optical setups or original tissue chemical processing techniques to enhance the information of the input images. Hyperspectral imaging technology, which captures a continuous range of spectral images, can differentiate healthy and diseased tissues based on their unique light reflection and absorption patterns across various wavelengths, thereby offering a viable virtual H&E staining approach.

This paper utilizes tunable spectral filters to capture hyperspectral images of glioma tissues and based on deep learning methods, trains a rapid virtual H&E staining model. The model is cost-effective with a small volume of tunable spectral filters, and the high-spectral imaging method is contamination-free and contactless, enabling convenient and quick virtual H&E staining of gliomas. The collection of high-spectral images through a microscope is depicted in Fig. 1a, followed by the selection of specific wavelength images that are subsequently fed into

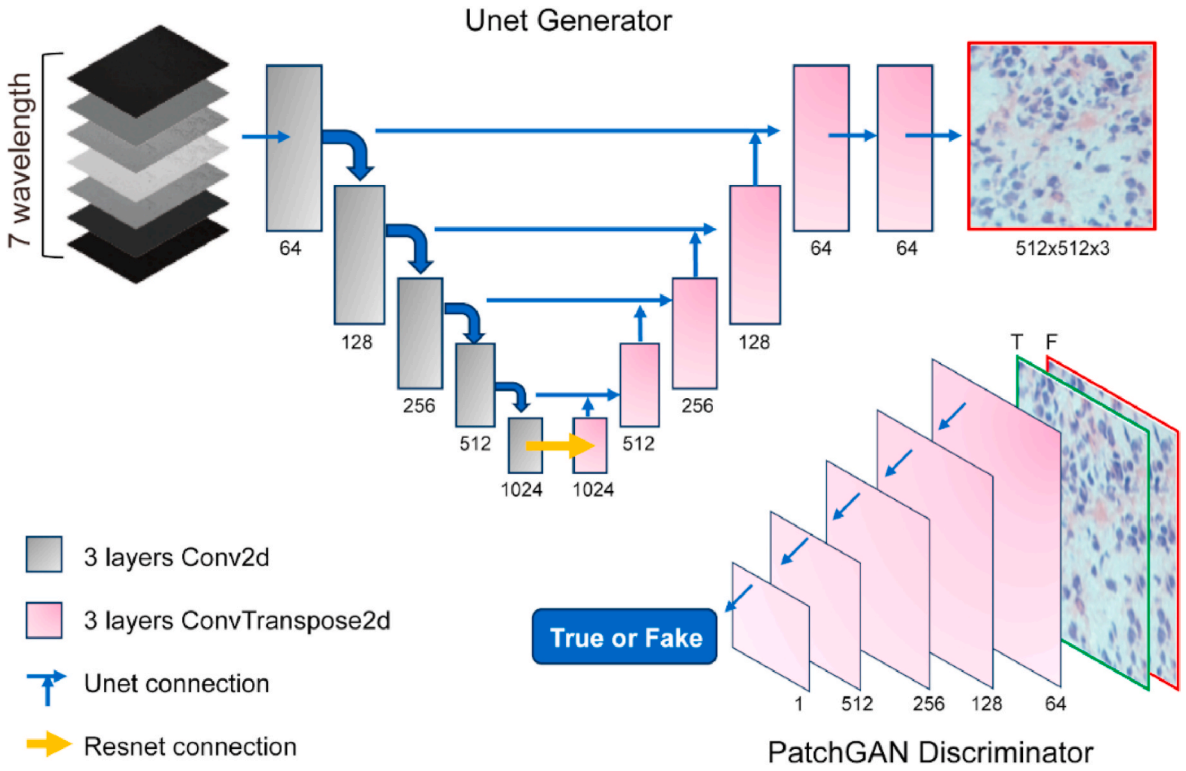

Fig. 2. The structure of the generator and the discriminator. The U-Net network includes 5 layers of encoders and decoders, which takes a 7-channel wavelength as the input data and produces a corresponding 3-channel H&E staining result. This synthesized result is then input into the discriminator alongside real staining images for classification. The discriminator is responsible for determining whether the generated image is real or fake.

the virtual staining network to generate virtual staining images. These reconstructed virtual staining images exhibit comparable image quality to the traditional H&E stained images shown in Fig. 1b, demonstrating highly similar tissue structures and color characteristics.

In this paper, our main contributions are summarized as follows.

1. Deep learning and hyperspectral imaging are utilized to create virtual H&E-stained images that closely resemble true H&E images, reducing costs, time, and potential errors in the staining process while minimizing damage to original tissue samples.
2. After evaluating different network structures, Unet has been selected as the virtual staining neural network. Additionally, the combinations of imaging wavelength are further investigated to achieve a high accuracy and efficiency of the virtual H&E staining results.
3. A comprehensive software platform has been developed to seamlessly integrate CCD control, microscope control, and virtual H&E staining technology. This system reduces manual labor and time required for traditional staining while also ensuring accurate and consistent results for pathologists and clinicians.

2. Materials and methods

2.1. Data acquisition

The glioma slices used in this study are sourced exclusively from the First Affiliated Hospital of Wenzhou Medical University. The hyperspectral images of 10 glioma tissue sections are used as the data sets of the neural network. Images are captured using an inverted microscope (IX83, Olympus) equipped with a tunable bandpass filter (KURIOS-XL1, Thorlabs). The microscope parameter "brightness" is consistently configured to 4, with a fixed exposure time of 300 ms set for the CCD (BioHD-C20, FluoCa). No additional processing is required for frozen tissue sectioning. Hyperspectral images are captured in the spectral range of 420–730 nm, with a 5 nm step, resulting in 63 spectral channels. After obtaining hyperspectral microscopic imaging data, standard H&E staining is performed on tissue sections and the slides are sealed to ensure uniform staining of all sections. Following this procedure, H&E images are captured using an inverted microscope. The image data is captured through a fully automated slide scanning system and stored in TIF format.

2.2. Data preprocessing

This study employs a Pix2Pix supervised deep learning model to achieve virtual staining, requiring precise registration between hyperspectral and H&E staining images. Due to the fact that hyperspectral images of label-free tissue sections lack high contrast when viewed under a microscope, it poses challenges in precise image registration. Therefore, rigid registration methods are chosen to select corresponding feature points between real H&E images and hyperspectral images with the highest contrast at 570 nm. Then the rotation angle and translation distance are calculated, and the registration matrix is applied to other corresponding spectral data, completing image registration of hyperspectral images and H&E images.

After obtaining pairs of registered images, overlapping regions are preserved and 7-channel spectral image data is extracted. The data is then stacked into a single .npy file for easy use as input for model training and testing. Before inputting the dataset into the neural network, the data is randomly cropped to a size of 1024×1024 and then resized to 512×512 to increase cell quantity in the receptive field. Data augmentation is achieved through rotation, flipping, and random erase. The random erase is used to simulate large blank areas and ensures that the model accurately predicts values in the blank regions. During testing, the testing dataset is cropped into sizes of 1024×1024 with a 30 % overlap between each cropped image. After obtaining the model output, the results are concatenated in the original sequence to facilitate the restoration of the full-size images.

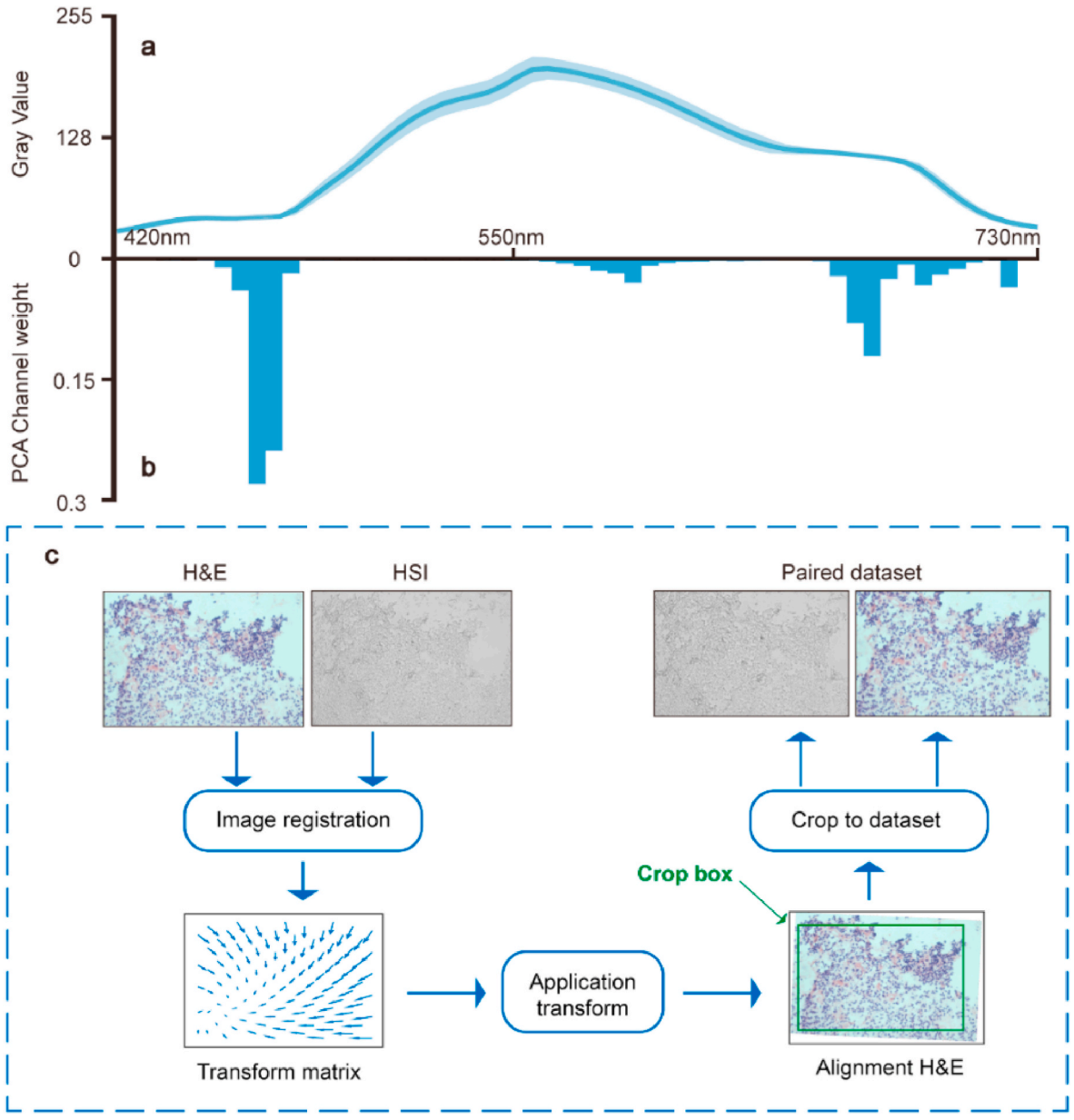

Fig. 3. Preprocessing of hyperspectral data includes the following steps: (a) Calculation of the mean and standard deviation of the spectral image at each wavelength. (b) Generation of histograms for the top three wavelength channels through dimensionality reduction (PCA). (c) Manual selection of feature points for image registration using feature-based image registration [40].

2.3. GAN architecture and training

The study uses a GAN architecture [36] for the deep learning model, which includes a generator and a discriminator. The Unet generator and discriminator structure are illustrated in Fig. 2. These two components work together, learning from each other to improve over time, ultimately creating a generator capable of producing highly convincing fake images. The loss function of the generator is defined as:

$$Ł_{generator} = \mathrm{L1}\{y, G(x)\} + (1 - D(G(x)))^2 + \mathrm{VGG}\{y, G(x)\} \quad (1)$$

Where x, y, G(.), and D(.) represent the hyperspectral data, real H&E images, generator, and discriminator, respectively. The perceptual loss (VGG Loss) [37] is used to enhance the fine details of the generated images. Additionally, it is desired that the generated images have a higher level of confusion in the discriminator, making it more likely to misclassify the generated virtual staining as real staining images. Hence, this loss is included as the latter part of the loss function.

The discriminator is equally crucial in the GAN network, and in this paper, the patchGAN structure is selected as the discriminator. In the GAN network, the discriminator is utilized to accurately distinguish between real and virtual staining. The discriminator assigns a score of 0 to virtual staining outputs and a score of 1 to real staining outputs. To minimize the loss for the discriminator, the loss function of the generator is defined as:

$$Ł_{discrimnator} = D(G(x))^2 + (1 - D(y))^2 \quad (2)$$

The generator of deep neural networks uses the Unet structure, which consists of two parts: encoder (downsampling) and decoder (upsampling). The encoder includes 4 downsampling blocks and 4 upsampling blocks, with concatenate connections connecting downsampling and upsampling. Each downsampling block consists of 3 convolutional layers, InstanceNorm normalization layers, and LeakyReLU activation layers. In the upsampling part, transposed convolution

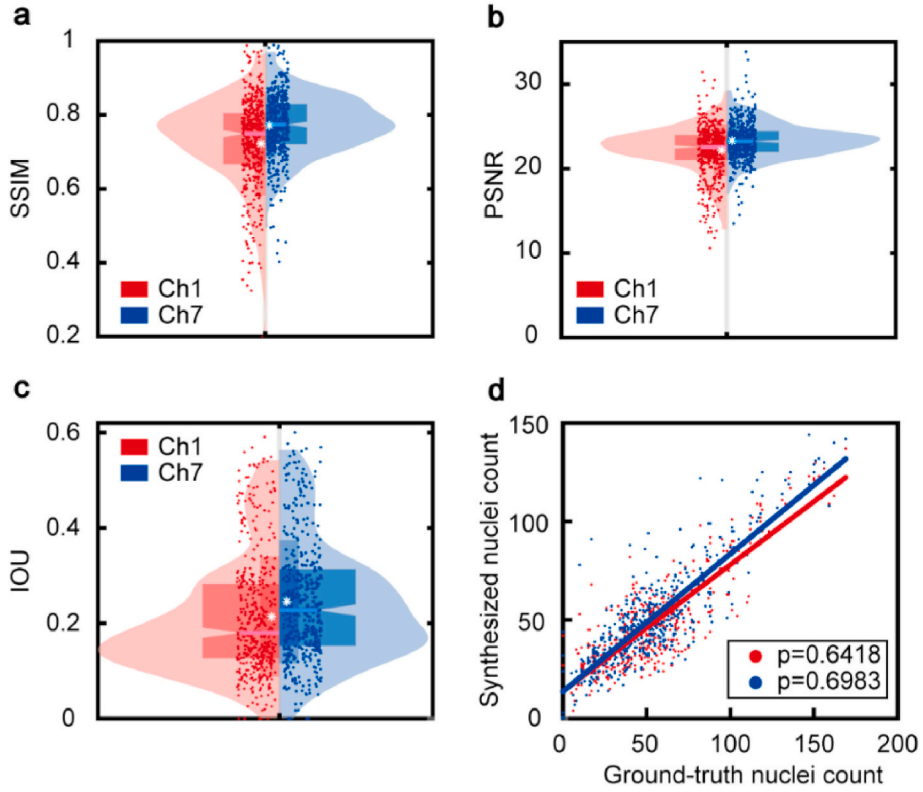

Fig. 4. Comparisons of SSIM (a), PSNR (b), and IOU (c) metrics are conducted to evaluate the impact of 1-channel (red) and 7-channel inputs (blue) on virtual staining results. The violin plots display metrics, including the mean, median, and histograms. (d) The influence of different channel inputs on the number of generated cell nuclei [43,44] is examined through cell nucleus segmentation counting using the stardist library.

(deconvolution) layers are used instead of simple upsampling for upsampling. Each upsampling block also includes 3 convolutional layers, InstanceNorm normalization layers, and LeakyReLU activation layers. The final layer is a convolutional layer designed to transform the multi-channel data into a size of batch 3 × 512 × 512. The skip connections between the downsampling and upsampling parts of Unet allow the network to preserve information from shallow layers, contributing to better image reconstruction.

The various deep neural networks mentioned in the paper are all trained for the same number of epochs (230 epochs). The convolutional layers have 3 × 3 convolutional kernels. The same dataset is used for training and testing across all networks.

2.4. Implementation details

All the deep learning neural network models in this paper are implemented using Python 3.6 and PyTorch 1.10. The computations are performed at the Sunway Supercomputing Center, Chinese Academy of Sciences, utilizing an XEON 6248R (3.0 GHz, 24 cores, Intel Corporation) and a GeForce RTX 3090 GPU (NVIDIA Corporation). The network is trained using a single GPU on the cluster for a total of 40 h, covering 230 epochs.

3. Results and discussion

3.1. Performance comparisons using various wavelength combinations

In this study, tumor sections of 10 glioma patients are collected, and spectral images with a size of 5440 × 1648 pixels and corresponding H&E staining results are collected. Finally, 627 pairs of data sets are obtained through image registration. The hyperspectral dataset contains a substantial amount of redundant information, necessitating dimensionality reduction of the raw data [38]. Principal component analysis (PCA) [39] is a commonly used dimensionality reduction technique, specifically designed to transform high-dimensional spectral data into low-dimensional data while preserving essential information as in Fig. 3a–b. In Fig. 3c, corresponding feature point coordinates between spectral images and true H&E images are identified, and a deformation matrix is calculated to accurately align the images. Subsequently, the registered image pairs undergo an identical cropping process, resulting in a paired dataset.

Fig. 3 illustrates the preprocessing of the hyperspectral datasets. Fig. 3a depicts the error band graph of the hyperspectral data, revealing significant variations in grayscale intensity. Fig. 3b utilizes the PCA method to reduce the dimensionality of data, obtaining statistical histograms for the top three influential spectral wavelengths (465, 595, and 620 nm) and they deviate from previous studies (457, 517, and 645 nm) [41,42] due to variations in tissue compositions. The wavelength distribution, which makes a significant contribution to hyperspectral data, spans the high, medium, and low ranges. Therefore, in this study, spectral images are captured from 420 nm to 730 nm with an interval of 50 nm. Among them, the 7-channel dataset is selected, which consists of seven spectral image channels with wavelengths at 420, 470, 520, 570, 620, 670, and 720 nm. In addition, the peak of spectral data is located in a region centered around a wavelength of 570 nm, and it is noteworthy that the gray mean value of the spectral image at 570 nm attains its maximum level. Hence, the label-free images captured at 570 nm serve as a 1-channel dataset.

To quantitatively analyze the performance of different methods, three evaluation metrics: PSNR (Peak Signal-to-Noise Ratio), SSIM (Structural Similarity), and intersection over union (IOU) are adopted to measure the reconstructed image quality. In Fig. 4a–c, comparisons of SSIM (a), PSNR (b), and IOU (c) metrics are conducted to evaluate the impact of 1-channel (red) and 7-channel inputs (blue) on virtual staining results, wherein the 7-channel results exhibit a more concentrated distribution of SSIM, PSNR, and IOU with higher mean values than those of

Table 1
Performance of several virtual staining neural networks with varying combinations of spectral information on the test datasets.

wavelength combinations	SSIM	Std	PSNR	Std
570	0.7366	0.1106	22.5880	2.4796
520,570,620	0.7711	0.0893	23.1124	2.2164
420,570,720	0.7454	0.1115	22.9146	2.5935
420,470	0.7114	0.1459	22.0702	2.8708
670,720	0.7276	0.1021	22.0702	2.6049
420,520,620,720	0.7561	0.1084	22.9424	2.3357
420,470,520,570,620,670,720	**0.7731**	0.0867	**23.3120**	2.2626

the 1-channel results. Additionally, in Fig. 4d, the nuclear counts of both are statistically quantified against ground truth, yielding slope P-values via linear regression. The findings indicate that the 7-channel dataset generated P-values closer to reality. The 7-channel virtual staining neural network demonstrates clear superiority over the 1-channel virtual staining neural network. This implies that a single wavelength, such as 570 nm alone, cannot encompass all the required data information for virtual staining, emphasizing the necessity of obtaining data from multiple wavelengths. Figs. S1–S3 present a statistical analysis of the virtual staining results, including nuclear area and nuclear distance. Fig. S4 and Table S4 demonstrate the influence of image brightness at different spectral wavelengths on the virtual staining neural network.

Utilizing the control variables method, the impacts of 7-channel combinations on virtual staining results are investigated, and all virtual staining networks are with the same network structures, identical initial parameters, and 230 training epochs. The SSIM and PSNR in Table 1 and Table S1 imply that 7-channel (420, 470, 520, 570, 620, 670, and 720 nm) virtual staining neural networks, which encompass a broader range of wavelengths, are of superior performance than either 1-channel or 3-channel virtual staining neural networks. Consequently, all subsequent analyses and experiments in this paper utilize data of the 7-channel wavelength (420, 470, 520, 570, 620, 670, and 720 nm) as input vectors to the virtual staining neural network.

3.2. Comparisons of methods with various generator structures

The GANs involve the iterative learning and refinement of the generator and discriminator to achieve a realistic output from the generator. However, due to the complexity of the generator task and the numerous parameters in the generator structure, training the generator is often challenging. Therefore, to investigate the influence of generator structures on the virtual staining results, six deep learning (Crossnet [45], D-LinkNet [46], HRnet [47], Resnet [48], Unet [49], and Unet3D [50]) are models trained on an identical 7-channel spectral dataset for the same number of epochs as in Tables 2 and 3.

The qualitative and quantitative comparisons of virtual staining results generated by six network models are compared in Figs. 5 and 6 and Tables S2–S3. The Crossnet, HRnet, and Unet show relatively concentrated SSIM values for the test sets, with mean values centered around 0.7731 and variances of approximately 0.08 in Fig. 5a. Fig. 5b–c depict that Unet exhibits the shortest training and testing time, with a training process taking half the time of Unet3D and a testing process requiring 1/7 the time of HRnet. The results indicate that these virtual staining methods effectively convert high-spectral images of glioma tissue into high-fidelity virtual H&E images, accurately replicating image color, nuclear distribution, and nuclear morphology as shown in Fig. 6. After

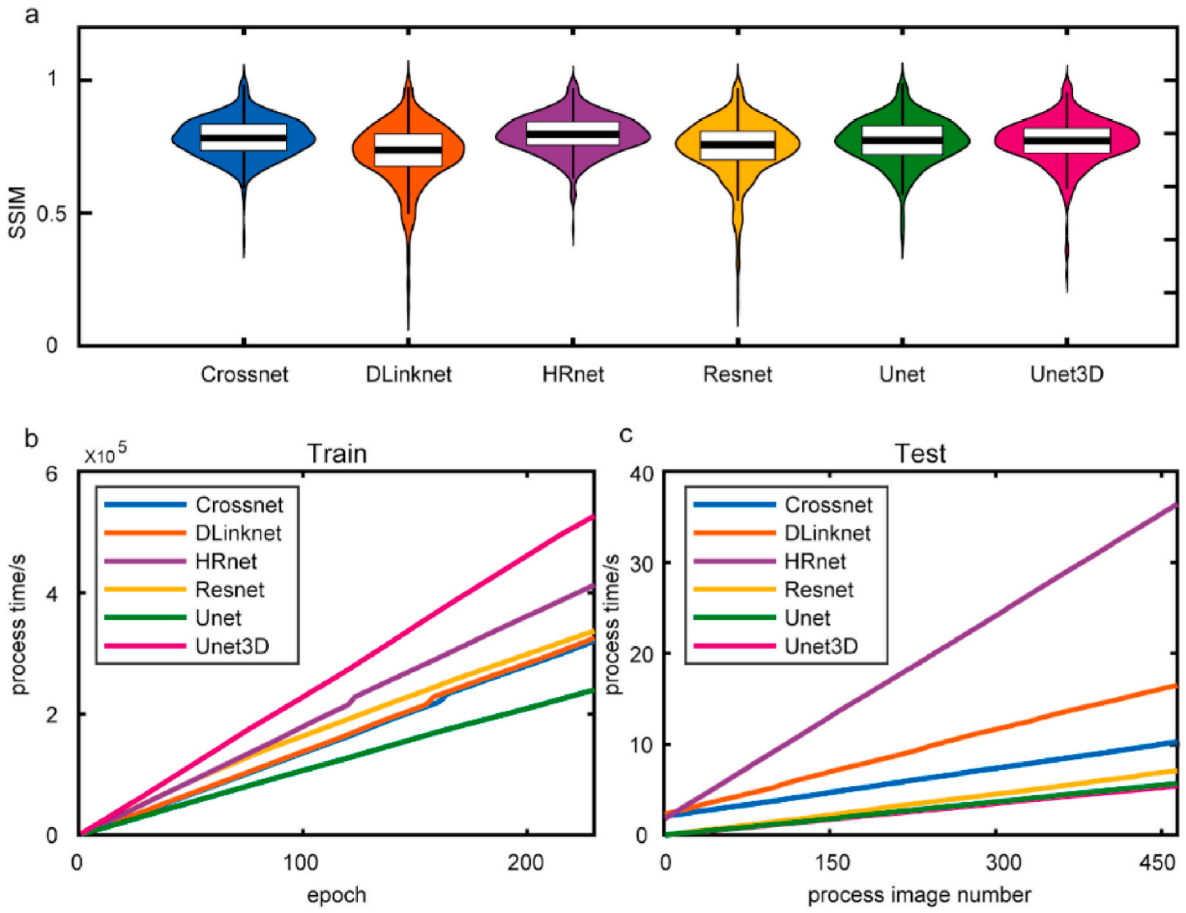

Fig. 5. Comparisons of the performance and execution time for six deep learning models: (a) The SSIM of the test results and real H&E images. (b) Cumulative training time with the same epochs and (c) cumulative testing time.

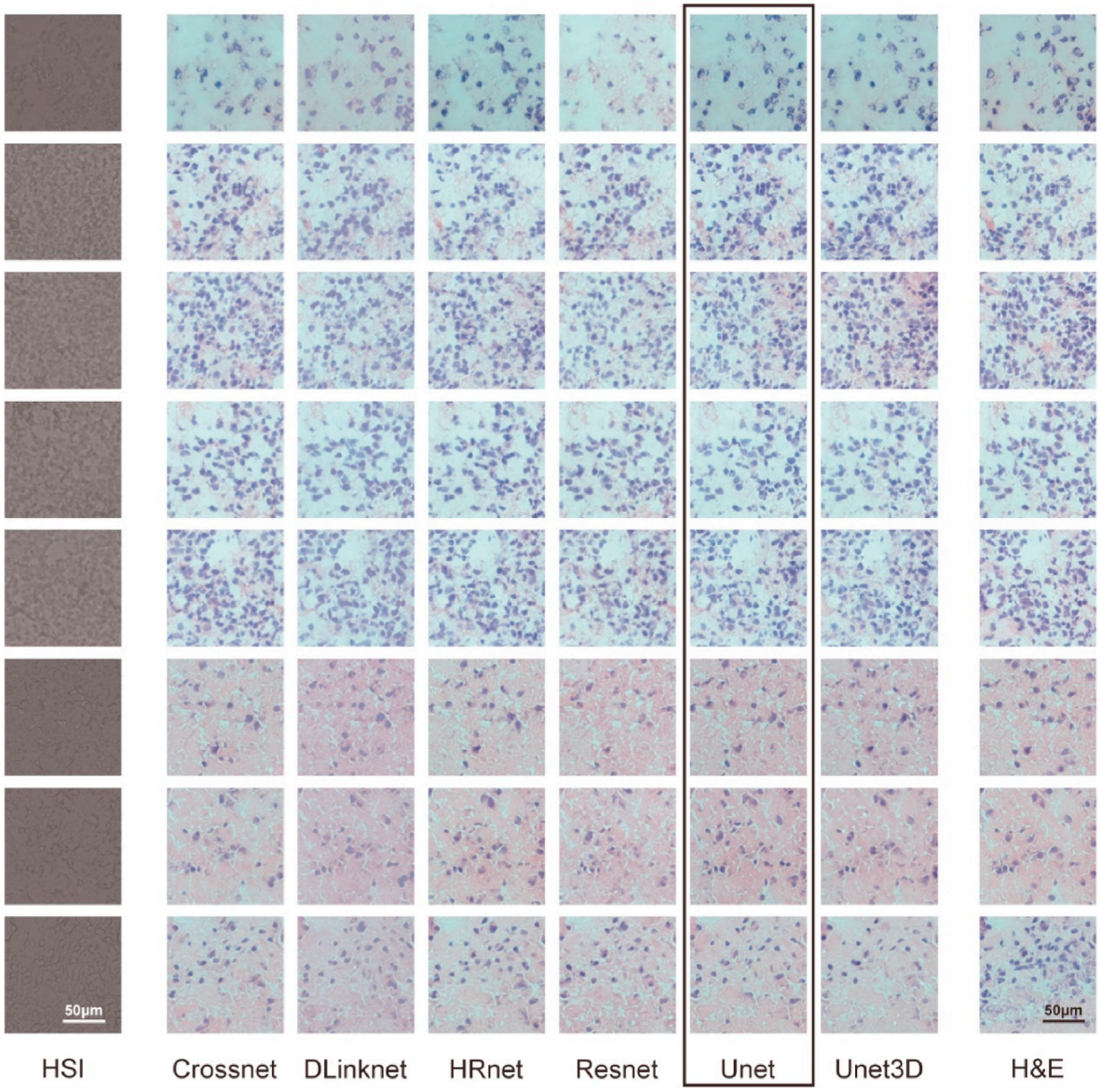

Fig. 6. Qualitative analysis of the virtual staining results generated by different models. The 1st column illustrates hyperspectral images (570 nm), which consist of 7 spectral channels (420,470,520,570,620,670,720 nm). The 2nd-7th columns demonstrate the virtual staining results generated by the CrossNet, D-LinkNet, HRnet, ResNet, Unet, and Unet3D, respectively. The 8th column exhibits corresponding true H&E image pairs.

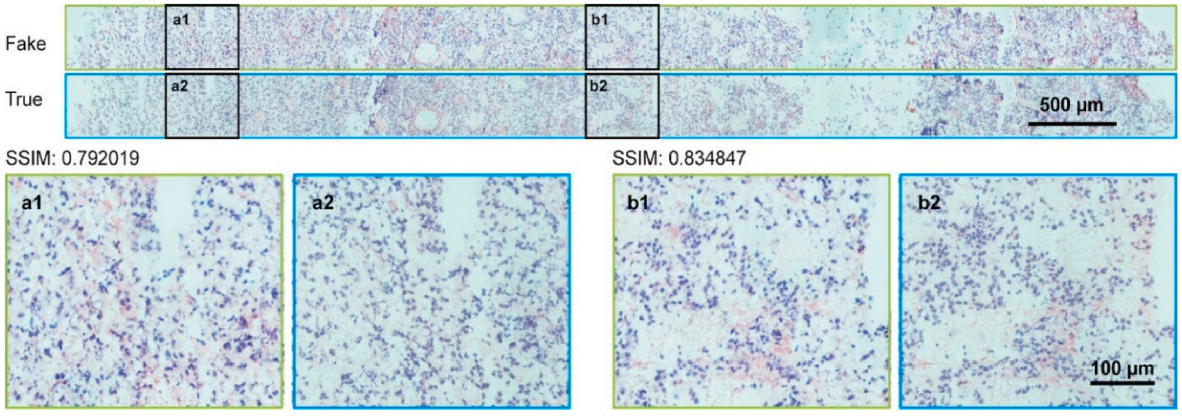

Fig. 7. The bottom section displays enlarged result comparisons in the virtual staining WSI (green box) and real H&E WSI (blue box). Fig.a1 and Fig.b1 correspond to the black insets in the virtual staining WSI, and Fig.a2 and Fig.b2 correspond to the black insets in the real H&E WSI.

comprehensive consideration of training and testing time, reconstructed image quality, network complexity, and other aspects, the Unet structure is finally selected as the generator structure in this paper.

3.3. Comprehensive software platform

The automation software, for virtual staining of whole-slide imaging (WSI) tissue sections, has been developed by integrating virtual staining functionality with hardware components, including a microscope, CCD camera, and tunable optical filters (demonstration video in SI). As shown in Fig. 7, a comparative analysis is conducted between full-size virtual staining images and real H&E staining images. Because the virtual staining neural network takes an input of a certain size, thus large-scale virtual staining images require cropping the input images into small patches (1024 × 1024) and then stitching virtual staining results together. To avoid artifacts during the image stitching process, 30 % overlaps are preserved between adjacent input patches. The virtual staining results of the overlapping regions are similar, and image fusion can reduce the edge artifacts. It can be observed that the full-size virtual staining results exhibit similar structural features (with SSIM above 0.7) and consistent staining patterns with traditional H&E staining, albeit with some inconsistent results. This indicates the success of our automated WSI staining system, showcasing staining results comparable to traditional methods, and potentially enabling faster clinical diagnostic capabilities. Additionally, it facilitates the prediction of methods.

In the virtual staining process of large-scale images, inevitable artifacts are observed, especially in blank background regions, manifesting as noticeable color differences. Additionally, foreign objects and overlapping tissue regions in the slide may lead to inconsistent results compared to the actual situation. To address the issue, enhancing the dataset with more instances of these special scenarios is suggested, and color normalization methods are hoped to achieve consistency in the training dataset [51,52].

4. Conclusion

This paper investigates the feasibility of converting HSI data into virtual H&E staining images through a combination of deep learning method and hyperspectral image technique. Multiple virtual staining neural networks are trained using various combinations of hyperspectral data, and experiments show that a virtual staining neural network trained with a wider wavelength range facilitates better performance. Wherein, the virtual neural network, which is trained with wavelength combination (420, 470, 520, 570, 620, 670, and 720 nm), achieves optimal SSIM of 0.7731 and PSNR of 23.3120. By quantitatively comparing the SSIM, PSNR, IOU, and training/inference time of six virtual staining neural networks, the Unet, characterized by its simple deep and shallow feature connection structure, achieves relatively high accuracy while ensuring considerable processing efficiency. Moreover, using a laptop (Intel i5-1135G7, NVIDIA GeForce MX450), it takes only 30 s to finish a 5440 × 3648-pixel virtual staining image (field-of-view ~ 114.31 mm^2), i.e. the imaging speed of 3.81 mm^2/s.

This study presents a virtual staining method using spectral images of tissue slices, which allows for quick virtual H&E staining results without damaging the tissues while maintaining color normalization, and provides a new auxiliary tool for physicians in clinical diagnosis and treatment. This novel deep learning-based method exhibits high professionalism, and the developed automated slide scanning virtual staining system, capable of automatically collecting tissue spectral information to infer full-size virtual staining images, significantly reduces the complexity of use, making it accessible to individuals without specialized knowledge. This virtual staining method has the potential to further evolve into a collection of various chemical stains, possibly becoming a new assistant for clinical diagnosis by doctors.

However, the generator structure used is relatively simple, and it should be customized with novel architectural designs to improve the overall efficiency and effectiveness of the generator model [24–26]. Additionally, the available clinical data is limited and may not encompass all cases, thereby restricting the generalization ability. Furthermore, the assessment of the virtual H&E staining results conducted by the pathologists would further promote the accuracy and potential clinical applications. In the future, improving the current methodology would involve collecting more unified tissue datasets, customizing a high-performance generator framework, introducing richer evaluation factors, and conducting research on transitioning from H&E staining to multiplex staining.

Code availability

The virtual H&E staining pytorch code of glioma hyperspectral images used in this study can be found on GitHub at https://github.com/Ansherp/GVS.

CRediT authorship contribution statement

Ruohua Zhu: Writing – review & editing, Validation, Software, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Haiyang He:** Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis. **Yuzhe Chen:** Software, Methodology. **Ming Yi:** Investigation, Formal analysis. **Shengdong Ran:** Data curation. **Chengde Wang:** Resources, Validation, Data curation. **Yi Wang:** Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

All authors disclosed no relevant relationships.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (Grants No. 62004062 and No. 22274116), the Basic Research Project of Wenzhou City, China (G2023056) and Wenzhou Medical University, China (KYYW202210). This research was supported by the High Performance Computing Center of Wenzhou Institute, UCAS, China. We thank the High Performance Computing Center of Wenzhou Institute, UCAS, China, for the computational resources.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2024.108958.

References

[1] T. Jiang, et al., Clinical practice guidelines for the management of adult diffuse gliomas, Cancer Lett. 499 (2021) 60–72.
[2] Z. Hu, et al., First-in-human liver-tumour surgery guided by multispectral fluorescence imaging in the visible and near-infrared-I/II windows, Nat. Biomed. Eng. 4 (2020) 259–271.
[3] C. Wang, L. Huang, S. Wang, L. Wu, Y. Wang, J. Dong, A distinction of gliomas at cellular and tissue level by surface-enhanced Raman scattering spectroscopy, Chin. Chem. Lett. 35 (2024) 109383.
[4] L. Huang, et al., Rapid, label-free histopathological diagnosis of liver cancer based on Raman spectroscopy and deep learning, Nat. Commun. 14 (2023) 48.
[5] D.N. Louis, et al., The 2021 WHO classification of tumors of the central nervous system: a summary, Neuro Oncol. 23 (2021) 1231–1251.
[6] X. Li, H. Huang, H. Zhao, Y. Wang, M. Hu, Learning a convolutional neural network for propagation-based stereo image segmentation, Vis. Comput. 36 (2020) 39–52.
[7] X. Zhang, T. Wang, W. Luo, P. Huang, Multi-level fusion and attention-guided CNN for image dehazing, IEEE Trans. Circ. Syst. Video Technol. 31 (2021) 4162–4173.
[8] Z. Mu, et al., Algorithm analysis of face recognition robot based on deep learning, Int. J. Pattern Recogn. Artif. Intell. 37 (2023) 2356004.
[9] P. Lei, S. Xu, S. Zhang, An art-oriented pixelation method for cartoon images, Vis. Comput. 40 (2024) 27–39.

[10] A. Saber, M. Sakr, O.M. Abo-Seida, A. Keshk, H. Chen, A novel deep-learning model for automatic detection and classification of breast cancer using the transfer-learning technique, IEEE Access 9 (2021) 71194–71209.
[11] J. Lian, et al., Parrot optimizer: algorithm and applications to medical problems, Comput. Biol. Med. (2024) 108064.
[12] Y. Rivenson, et al., Virtual histological staining of unlabelled tissue-autofluorescence images via deep learning, Nat. Biomed. Eng. 3 (2019) 466–477.
[13] N. Borhani, A.J. Bower, S.A. Boppart, D. Psaltis, Digital staining through the application of deep neural networks to multi-modal multi-photon microscopy, Biomed. Opt Express 10 (2019) 1339–1350.
[14] D.A. Orringer, et al., Rapid intraoperative histology of unprocessed surgical specimens via fibre-laser-based stimulated Raman scattering microscopy, Nat. Biomed. Eng. 1 (2017) 27.
[15] J. Li, et al., Biopsy-free in vivo virtual histology of skin using deep learning, Light Sci. Appl. 10 (2021) 233.
[16] T.T.W. Wong, et al., Fast label-free multilayered histology-like imaging of human breast cancer by photoacoustic microscopy, Sci. Adv. 3 (2017) e1602168.
[17] J.J. Levy, et al., A large-scale internal validation study of unsupervised virtual trichrome staining technologies on nonalcoholic steatohepatitis liver biopsies, Mod. Pathol. 34 (2021) 808–822.
[18] B. Schmauch, et al., A deep learning model to predict RNA-Seq expression of tumours from whole slide images, Nat. Commun. 11 (2020) 3877.
[19] X. Fan, J. Zhou, X. Jiang, M. Xin, L. Hou, CSAP-UNet: convolution and self-attention paralleling network for medical image segmentation with edge enhancement, Comput. Biol. Med. 172 (2024) 108265.
[20] B.K. Fu, Y.S. Peng, J.J. He, C. Tian, X.H. Sun, R.P. Wang, HmsU-Net: a hybrid multi-scale U-net based on a CNN and transformer for medical image segmentation, Comput. Biol. Med. 170 (2024) 108013.
[21] Y.H. Ma, et al., DSFF-GAN: a novel stain transfer network for generating immunohistochemical image of endometrial cancer, Comput. Biol. Med. 170 (2024) 108046.
[22] H. Yu, et al., An intent classification method for questions in "Treatise on Febrile diseases" based on TinyBERT-CNN fusion model, Comput. Biol. Med. 162 (2023) 107075.
[23] T. Zhou, Q. Cheng, H. Lu, Q. Li, X. Zhang, S. Qiu, Deep learning methods for medical image fusion: a review, Comput. Biol. Med. 160 (2023) 106959.
[24] C. Wang, Y. Wang, P. Ding, S. Li, X. Yu, B. Yu, ML-FGAT: identification of multi-label protein subcellular localization by interpretable graph attention networks and feature-generative adversarial networks, Comput. Biol. Med. 170 (2024) 107944.
[25] L. Zhao, H. Chi, T. Zhong, Y. Jia, Perception-oriented generative adversarial network for retinal fundus image super-resolution, Comput. Biol. Med. 168 (2024) 107708.
[26] K. Wei, et al., CT synthesis from MR images using frequency attention conditional generative adversarial network, Comput. Biol. Med. 170 (2024) 107983.
[27] L.W. Deng, Y.F. Ji, S.J. Huang, X. Yang, J. Wang, Synthetic CT generation from CBCT using double-chain-CycleGAN, Comput. Biol. Med. 161 (2023) 106889.
[28] Y. Li, et al., Virtual histological staining of unlabeled autopsy tissue, Nat. Commun. 15 (2024) 1684.
[29] M.Y. Lu, et al., A visual-language foundation model for computational pathology, Nat. Med. 30 (2024) 863–874.
[30] M. Schnell, S. Gupta, T.P. Wrobel, M.G. Drage, R. Bhargava, P.S. Carney, High-resolution label-free imaging of tissue morphology with confocal phase microscopy, Optica 7 (2020) 1173–1180.
[31] M.T. Martell, et al., Deep learning-enabled realistic virtual histology with ultraviolet photoacoustic remote sensing microscopy, Nat. Commun. 14 (2023) 5967.
[32] Y. Zhang, K. de Haan, Y. Rivenson, J. Li, A. Delis, A. Ozcan, Digital synthesis of histological stains using micro-structured and multiplexed virtual staining of label-free tissue, Light, Science and Applications 9 (2020), 78-78.
[33] J.Y. Zhu, T. Park, P. Isola, A.A. Efros, Unpaired image-to-image translation using cycle-consistent adversarial networks, in: 2017 IEEE International Conference on Computer Vision (ICCV), 2017, pp. 2242–2251.
[34] K. de Haan, et al., Deep learning-based transformation of H&E stained tissues into special stains, Nat. Commun. 12 (2021) 4884.
[35] S. Liu, et al., Unpaired stain transfer using pathology-consistent constrained generative adversarial networks, IEEE Trans. Med. Imag. 40 (2021) 1977–1989.
[36] A.-B. Mohamed, M. Nour, H. Hossam, Generative adversarial networks (GANs). Deep Learning Approaches for Security Threats in IoT Environments, IEEE, 2023, pp. 271–285.
[37] J. Johnson, A. Alahi, L. Fei-Fei, Perceptual losses for real-time style transfer and super-resolution, in: B. Leibe, J. Matas, N. Sebe, M. Welling (Eds.), Computer Vision – ECCV 2016, Springer International Publishing, Cham, 2016, pp. 694–711.
[38] M. Wang, A.A. Heidari, H. Chen, A multi-objective evolutionary algorithm with decomposition and the information feedback for high-dimensional medical data, Appl. Soft Comput. 136 (2023) 110102.
[39] X. Jia, B.C. Kuo, M.M. Crawford, Feature mining for hyperspectral image classification, Proc. IEEE 101 (2013) 676–697.
[40] D.L.G. Hill, P.G. Batchelor, M. Holden, D.J. Hawkes, Medical image registration, Phys. Med. Biol. 46 (2001) R1–R45.
[41] J. Salido, N. Vallez, L. Gonzalez-Lopez, O. Deniz, G. Bueno, Comparison of deep learning models for digital H&E staining from unpaired label-free multispectral microscopy images, Comput. Methods Progr. Biomed. 235 (2023) 107528.
[42] N. Bayramoglu, M. Kaakinen, L. Eklund, J. Heikkilä, Towards virtual H&E staining of hyperspectral lung histology images using conditional generative adversarial networks, in: 2017 IEEE International Conference on Computer Vision Workshops (ICCVW), 2017, pp. 64–71.
[43] U. Schmidt, M. Weigert, C. Broaddus, G. Myers, Cell detection with star-convex polygons, in: A.F. Frangi, J.A. Schnabel, C. Davatzikos, C. Alberola-López, G. Fichtinger (Eds.), Medical Image Computing and Computer Assisted Intervention – MICCAI 2018, Springer International Publishing, Cham, 2018, pp. 265–273.
[44] M. Weigert, U. Schmidt, R. Haase, K. Sugawara, G. Myers, Star-convex polyhedra for 3D object detection and segmentation in microscopy, in: 2020 IEEE Winter Conference on Applications of Computer Vision (WACV), 2020, pp. 3655–3662.
[45] Q. Yu, L. Qi, Y. Gao, W. Wang, Y. Shi, Crosslink-Net: double-branch encoder network via fusing vertical and horizontal convolutions for medical image segmentation, IEEE Trans. Image Process. 31 (2022) 5893–5908.
[46] L. Zhou, C. Zhang, M. Wu, D-LinkNet: LinkNet with pretrained encoder and dilated convolution for high resolution satellite imagery road extraction, in: IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2018, pp. 182–186.
[47] J. Wang, et al., Deep high-resolution representation learning for visual recognition, IEEE Trans. Pattern Anal. Mach. Intell. 43 (2021) 3349–3364.
[48] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778.
[49] P. Isola, J.Y. Zhu, T. Zhou, A.A. Efros, Image-to-Image translation with conditional adversarial networks, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 5967–5976.
[50] Ö. Çiçek, A. Abdulkadir, S.S. Lienkamp, T. Brox, O. Ronneberger, 3D U-net: learning dense volumetric segmentation from sparse annotation, in: S. Ourselin, L. Joskowicz, M.R. Sabuncu, G. Unal, W. Wells (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016, Springer International Publishing, Cham, 2016, pp. 424–432.
[51] G. Campanella, et al., Clinical-grade computational pathology using weakly supervised deep learning on whole slide images, Nat. Med. 25 (2019) 1301–1309.
[52] M.Z. Hoque, A. Keskinarkaus, P. Nyberg, T. Seppaenen, Stain normalization methods for histopathology image analysis: a comprehensive review and experimental comparison, Inf. Fusion 102 (2024) 101997.