2D to 3D Medical Image Colorization

Aradhya Neeraj Mathur*
IIITD
aradhyam@iiitd.ac.in

Apoorv Khattar*
IIITD
apoorv16016@iiitd.ac.in

Ojaswa Sharma*
IIITD
ojaswa@iiitd.ac.in

Abstract

Colorization involves the synthesis of colors while preserving structural content as well as the semantics of the target image. This problem has been well studied for 2D photographs with many state-of-the-art solutions. We explore a new challenge in the field of colorization where we aim at colorizing multi-modal 3D medical data using 2D style exemplars. To the best of our knowledge, this work is the first of its kind and poses challenges related to the modality (medical MRI) and dimensionality (3D volumetric images) of the data. Our approach to colorization is motivated by modality conversion that highlights its robustness in handling multi-modal data.

1. Introduction

Image colorization enables transferring radiometric characteristics of one image onto another. From an artistic perspective, it has opened new horizons in content creation and has seen rapid strides with the advent of digital media. The problem of colorization has been extensively explored for 2D photographs, however, the same approaches do not work well for 3D medical images. Our current research aims at finding a robust approach to apply colorization in the 3D domain for multi-modal medical data. Automatic and realistic colorization of 3D medical data is required to improve the volume visualization capabilities of current systems. Our approach aims at automating colorization in medical imaging visualizations that are mostly manually controlled currently. Furthermore, what makes this problem setting more unique is the ability of our approach to generate photorealistic 3D volumes that can be used for better visualizations compared with existing approaches.

We view our approach as a suitable replacement to the current colorization methods prevalent in the volume visualization community such as transfer functions (TF) based methods that are heavily used in the visualization of 3D CT and MRI modalities. A TF maps intensity values to colors, however the creation of such a mapping is mostly manual making it cumbersome and time-consuming. Automatic and semi-automatic synthesis of TFs have been explored in the volume visualization community and many such ideas are summarized by Ljung et al. [18], however, these approaches still rely on the creation of a TF. Colorization can be seen as an effective alternative to using a transfer function but it is an ill-posed problem since MRI or CT medical data is a different modality compared to a photograph.

In this work, we propose a framework to perform modality conversion of a 3D MRI to a photographic volume. While doing so, we also address the challenges of handling 3D data in medical deep learning. We show multi-modal volume colorization using only a single 2D style image. Such a colorization can be used to create better 3D visualizations without the need of creating time-consuming transfer functions. Our approach can be used to colorize other modalities like 3D CT as well.

2. Related Work

We address the problem of multi-modal colorization where the source and the target images belong to different modalities, which are 3D MRI scans and photographic volumes in our case. There are many approaches for colorization of grayscale images and they mostly differ in the additional input required along with a grayscale image for generating results.

Classical approaches to colorization

Several classical methods look into solving color transfer by treating it as a texture transfer problem. Welsh et al. [30] propose a patch matching algorithm that requires a source image from which color information has to be transferred to the target image. Their approach performs a jittered sampling on the source image to get color patches, finds the closest matching patch from the sampled data, and transfers colors onto the target image. Levin et al. [13] view colorization as an optimization problem where given some hints in the form of scribbles, colorization is performed with the constraint that similar intensities should have similar colors.

[1] * indicates equal contribution

Nie et al. [19] develop on the above idea and use quad-tree image decomposition for cost function minimization.

Automatic Colorization

With the advent of Convolutional Neural Networks (CNNs), visual tasks have seen rapid strides. Iizuka et al. [9] leverage the estimation capability of CNNs and perform colorization using both local and global information present in an image. Their approach relies on the classification of an image to generate global priors in the form of class labels that are added to the colorization network for better results. Zhang et al. [34] propose an automatic colorization algorithm where colorization is treated as a multinomial classification problem. The authors employ a CNN to perform colorization in the CIELab color space by predicting only the chroma channel. Zhang et al. [35] propose an interactive colorization algorithm that utilizes user-generated hints and demonstrate the generation of different results for the same grayscale image. Yang et al. [33] propose GANs for colorizing 3D objects. They perform voxel colorization using a GAN consisting of a 3 component architecture. The first part learns a compressed representation, the second component decompresses and adds information from the latent z vector to generate colors. The last component uses a mask for refinement and generates final colorized volume.

Exemplar Colorization

Exemplar colorization involves the transfer of colors from a source image to a target grayscale image. These methods rely on CNNs and make use of the semantic context in the source and target images to produce meaningful and coherent colored images. In one of the recent works He et al. [6] present a deep exemplar-based colorization in which they use a similarity sub-net and a colorization subnet to perform exemplar colorization. The similarity subnet computes the bidirectional similarity maps and aligned chrominance channels for the input and reference images which are then passed to the colorization network. Li et al. [14] present a cross-scale texture matching method using a multi-label graph-cut algorithm. Xiao et al. [31] propose a novel pyramidal network to perform exemplar colorization as a multinomial classification problem. They propose a hierarchical encoder-decoder filter to pass color distributions from lower to higher levels. Xu et al. [32] propose a stylization based approach for fast deep exemplar colorization. They achieve this with the help of a transfer sub-net to extract coarse chrominance information referred to as ab map for target gray image using WCT [16]. For faster feature matching they use Adaptive Instance Normalization proposed by Huang et al. [7]. They use this generated ab map for subsequent colorization using a colorization net which is optimized using Huber Loss [8].

Style transfer as a generalization

Style transfer forms an integral part of our proposed method to generate hints for colorizing a volume. Prior to the appearance of neural style transfer, a number of approaches on artistic stylization and non-photorealistic rendering were explored. Gatys et al. [3] perform style transfer by minimizing reconstruction losses for feature and style with a pre-trained VGG network [26]. Their seminal work demonstrated that CNNs are capable of extracting content information from a photograph and style information from an artwork image to produce a stylized image. Li et al. [16] perform image reconstruction along with whitening and coloring transforms (WCT). The authors use VGG convolutions as an encoder and train a decoder symmetrical to VGG to produce photorealistic stylizations. Jamriška et al. [10] propose a non deep learning based approach for stylizing videos. For an input video, they require some hand-crafted frames as stylization hints. A patch-based optimization uses a mask guide for an object of interest, a color guide for capturing color information, and a positional guide to capture structural information.

Modality conversion

For a more realistic colorization, we perform modality conversion on the input MRI volume. GANs have been explored for converting the modality of medical volumes, commonly known as medical image synthesis. Nie et al. [20] propose a method to convert MRI to CT using GANs. Their method takes an MRI image and generates the corresponding CT image using a fully convolutional network and use a gradient difference loss function. This work is further extended by Nie et al. in [21] where they combine this approach with Auto-Context [28] and create context-aware GAN to generate 7T MRI from 3T MRI and CT from MRI. In a two-step approach, Guibas et al. [5] use a DCGAN to generate synthetic segmentation maps and teach a GAN to generate photorealistic images.

3. Our approach

Our approach enables us to bypass designing a transfer function for color volume visualization. The colorization of MRI volume shows the capability of our approach in handling multi-modal data. In our three-step approach, we use a Generative Adversarial Networks (GANs) [4] to convert the modality of the input MRI volume to a corresponding gray cryosection (cryo) volume. This is followed by style transfer on selected slices in the generated gray volume, and a subsequent colorization of the entire volume. Figure 1 shows an overview of our colorization approach.

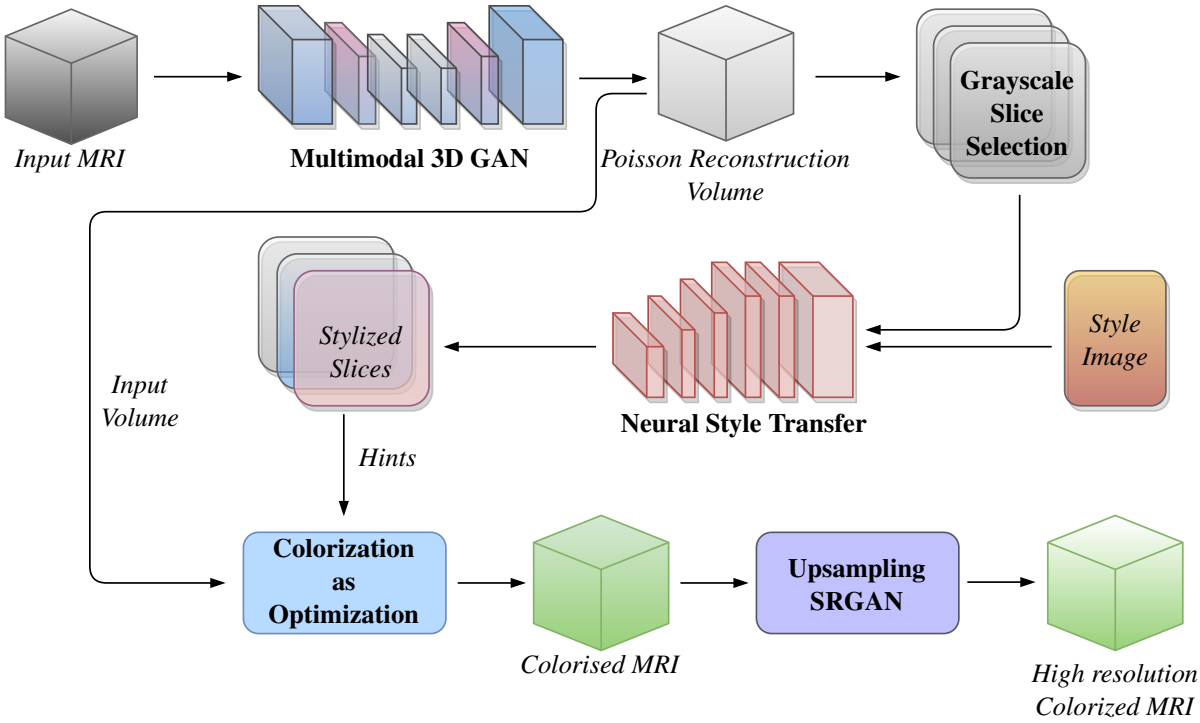


Figure 1. Our colorization framework is based on the modality conversion of input MRI. The first stage uses 3D GANs to remap the MRI intensities allowing for a more flexible colorization. Some of the selected slices from the produced output are stylized. These stylized slices serve as hints in the final stage where volumetric colorization takes place.

3.1. Dataset

We use the Visible Human 2 [1] dataset which contains unregistered 16-bit MRI and 24-bit color cryo imaging data of the human head. It is necessary to perform the registration between the two modalities in order to get correspondence between the two. We perform a rigid registration of MRI with cryo volume using fiducial markers. We also perform a marker-based segmentation of the MRI volume into foreground and background. The binary background *segmentation mask* is used in removing background for visualization purposes.

3.2. MRI to grayscale conversion

A direct colorization from MRI to color will not yield good results given substantial differences in the radiometric characteristics of MRI imaging and photographs. An MRI response of tissue is different than it is for visible light i.e. the tissues in an MRI will look different than they will in a photograph. Therefore, the first step of colorization is to convert a given MRI to a corresponding grayscale photographic volume (hereafter referred to as a *grayscale volume*). In essence, this step performs modality conversion and predicts how the grayscale intensity of the same tissue will appear under visible light.

3.2.1 MRI aligned grayscale volume for training

In a real scenario a cryo volume for an MRI will not be available, therefore to generalize our method we train a neural network that learns to convert an MRI to a grayscale volume. In order to train such a network, both an MRI and a corresponding grayscale volume are needed. Unfortunately, a grayscale cryo volume is not suitable for this purpose making the problem non-trivial due to the following reasons:

- Even after careful registration, the features in cryo and MRI do not align perfectly, and there are errors of few voxels at some places, and
- The MRI volume has a lower level of detail in comparison with the cryo volume that manifests as ambiguity in the tissue boundaries.

Poisson volume synthesis To create a grayscale volume that matches the features and scale of the input MRI volume, we perform a Poisson reconstruction [24]. Here, we reconstruct a volume I_{gray} from gradients of the MRI volume I_{mri} treating intensities from grayscale cryo volume I_{cryo} as boundary condition by solving the following Poisson equation:

$$\Delta I_{gray} = \nabla \cdot \nabla I_{mri}, \text{s.t. } I_{gray}|_{\partial\Omega} = I_{cryo}|_{\partial\Omega}.$$

This helps us obtain the radiometry of cryo volume while maintaining the geometry of the structures in MRI thus creating a new modality of photo-realistic MRI (see Figure 2(a)). However, in a real-life scenario, such correspondence might not be available thus, we require a method that can learn this conversion of modality. We achieve this with the help of 3D GAN. This modality plays a key role in contributing to the visual perception and greater stability of the GANs by providing a tighter structural correspondence to the MRI, thus reinforcing the preservation of structures in the MRI.

Segmentation mask generation For visualization, background removal is usually required. Therefore, we remove the background of the Poisson volume using a segmentation mask. The segmentation is performed by the graph cut algorithm given by Jirik et al. [11] which uses the region and boundary properties of the image to perform segmentation. The segmentation is performed using manual hints for internal and external sections of the volume.

Training data We generate about 3000 sub-volumes of size $32 \times 32 \times 32$ for training the network. These volumes are taken from the MRI and the Poisson grayscale volumes by performing random rotations and translations to augment the training dataset. The final results and illustrations are generated by passing the complete volume of size $128 \times 128 \times 88$ from the trained network. We further compute the metrics from the generated Poisson volume w.r.t to the MRI volume as shown in Table 1 which serves as the upper bound for our GAN outputs. Since the Poisson volume contains the radiometry corresponding to the cryo volume and gradients from the MRI, it serves as the ideal candidate for realistic ground truth.

3.2.2 GAN network for modality conversion

We use the sub-volumes generated from the MRI and Poisson grayscale volume to train a GAN intended to perform modality conversion of any MRI volume. The GAN learns to generate a grayscale volume corresponding to an input MRI. The generator is based on 3D-FCN (Fully Convolutional Network) without any max-pooling layers. This plays an important role in the selection of architecture for our method since the MRI and grayscale volume have a strict geometric correspondence. Avoiding the use of max-pooling layers helps the network essentially to learn the remapping of the intensity values without distorting the structures. The generator architecture consists of 8 3D convolutional layers, each subsequent convolution is followed by an activation function, in our case it being Leaky ReLU with a negative slope of 0.01 except for the last layer where sigmoid activation is used. The generator architecture can be summarized in the table below.

Layer	**Filter size**	**Stride**	**kernel size**	**activation**
Conv3D	1×64	1	3	LeakyReLU
Conv3D	64×128	1	3	LeakyReLU
Conv3D	128×256	1	3	LeakyReLU
Conv3D	256×512	1	3	LeakyReLU
Conv3D	512×256	1	3	LeakyReLU
Conv3D	256×128	1	3	LeakyReLU
Conv3D	128×64	1	3	LeakyReLU
Conv3D	64×1	1	3	Sigmoid

In order to ensure better structures we use Structural Similarity (SSIM) loss [29] and an L_1 loss term. The SSIM loss enables the network to learn corresponding cryo structures more robustly while the L_1 loss helps in further improving the intensities of the generated voxels. The SSIM loss $\mathcal{L}_s$ between two volumes I_1 and I_2 is given by,

$$\mathcal{L}_s(I_1, I_2) = 1 - \frac{(2\mu_1\mu_2 + c_1)(2\sigma_{1,2} + c_2)}{(\mu_1^2 + \mu_2^2 + c_1)(\sigma_1^2 + \sigma_2^2 + c_2)},$$

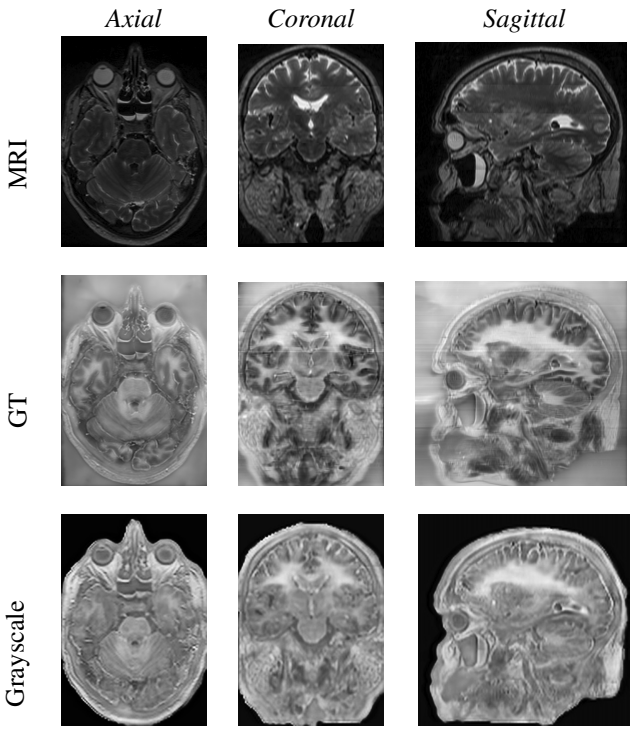
where μ_1 and σ_1^2 are the mean and variance of I_1, μ_2 and σ_2^2 are the mean and variance of I_2 and $\sigma_{1,2}$ is the covariance of I_1 and I_2. c_1 and c_2 are constants to prevent overflow in case of small denominators based on the dynamic range of voxel values. The complete generator loss can be formulated as,

$$\mathcal{L}_G = \mathcal{L}_s(G(I_{mri}), I_{gray}) + ||G(I_{mri}) - I_{gray}||_1 + \log(1 - D(G(I_{mri}))), \quad (1)$$

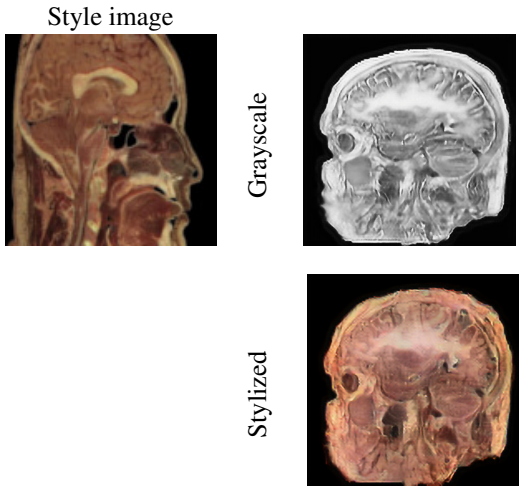
where I_{mri} is the input MRI subvolume, I_{gray} is the corresponding ground truth Poisson grayscale subvolume, $\log(1 - D(G(I_{mri})))$ [4] is the generator loss as described in equation (1) for generator G and discriminator D for input volume. The discriminator loss can be formulated as, $\mathcal{L}_D = \log(D(G(I_{mri})))$. Our discriminator architecture is as follows:

Layer	**Filter size**	**Stride**	**kernel size**
Conv3D + Batchnorm	1×256	2	3
Conv3D + Batchnorm	256×256	1	3
Conv3D + Batchnorm	256×128	2	3
Conv3D + Batchnorm	128×64	4	3

The last layer being linear of size 512×1 with Leaky ReLU as an activation function for all layers except last where sigmoid is used. The results of MRI to grayscale conversion are shown in Figure 2(a). The performance of our network can be seen by its ability to generate a grayscale volume that is as close as possible to the ground truth (GT) Poisson grayscale volume.



(a) MRI to grayscale conversion.



(b) Hint generation via stylization.

Figure 2. Modality conversion and hint slice generation for colorization.

3.3. Style transfer for generating hints

For generating planar hints, we use the style transfer method proposed by Gatys et al. [3]. The objective is to minimize the style loss and content loss, for which convolution layers of a pre-trained VGG19[26] are used. We use the features from the first five convolutions to characterize the style and content loss. Figure 2(b) shows sample colorization.

Since the VGG is not trained on medical data, it is not aware of the tissue structures present in our data. This results in subsequent irregularities in colors and color bleeding. To control this to some extent we further penalize the construction of color channels by using total variation loss [25]. This is done by converting the resultant image to YUV colorspace and applying total variation loss over the color channels U and V. This additional loss is combined with the content and style losses of the style transfer method.

When we apply Gatys style transfer on the complete volume slice by slice, it leads to incoherent results because the loss function is defined for images and it ignores the content and style information along the third dimension in the volume. Therefore, as explained next we utilize a separate colorization step that completely preserves the MRI structures while propagating colors from the hints.

3.4. Volume colorization using hints

To perform colorization we refer to the work by Levin et al. [13] that performs colorization using optimization with hints given by the user. We extend this method to 3D volume colorization and use the stylized images as 2D hints for colorization in 3D. To select slices as hints, we choose the slices along any of the axes with a high standard deviation to ensure good contrast. The colorization is solved as a constrained optimization problem where two neighboring voxels in a volume, p and q should have the same colors if their intensities are similar. The colorization is performed in the YUV color space where the Y channel is the same as the one generated using GAN and the cost function for U and V is defined as

$$J(U) = \sum_{p} \left(U(p) - \sum_{q \in N(p)} w_{pq} U(q) \right)^2$$

where $N(p)$ is a set of all voxels that lie in a $3 \times 3 \times 3$ window around p. w_{pq} is a weighting function which is large for similar values of $Y(p)$ and $Y(q)$. Here we used $w_{pq} = e^{-(Y(p)-Y(q))^2/2\sigma_p^2}$, where σ_p is the the variance along $N(p)$. Given the quadratic cost function subject to the defined constraints, we minimize $J(U)$ and $J(V)$. Since the optimization problem creates a large system of sparse equations, we use multi-grid solver [22] to obtain an optimal

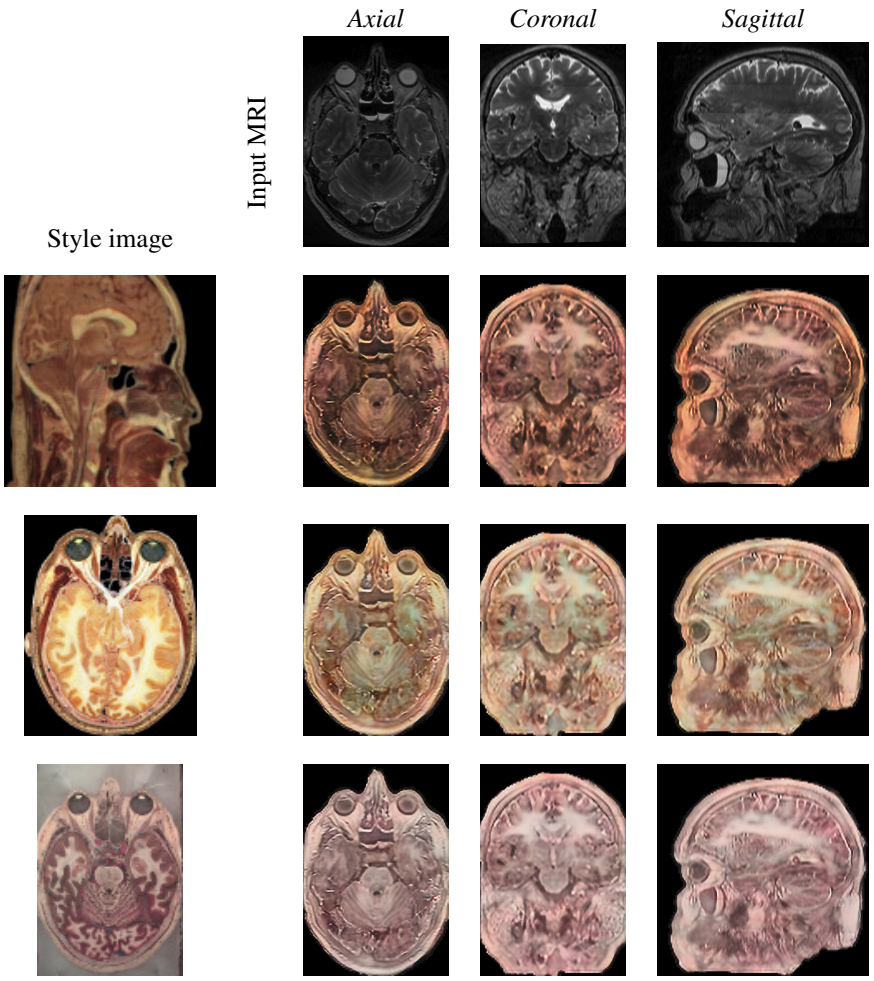


Figure 3. Colorization result with a cryo style image.

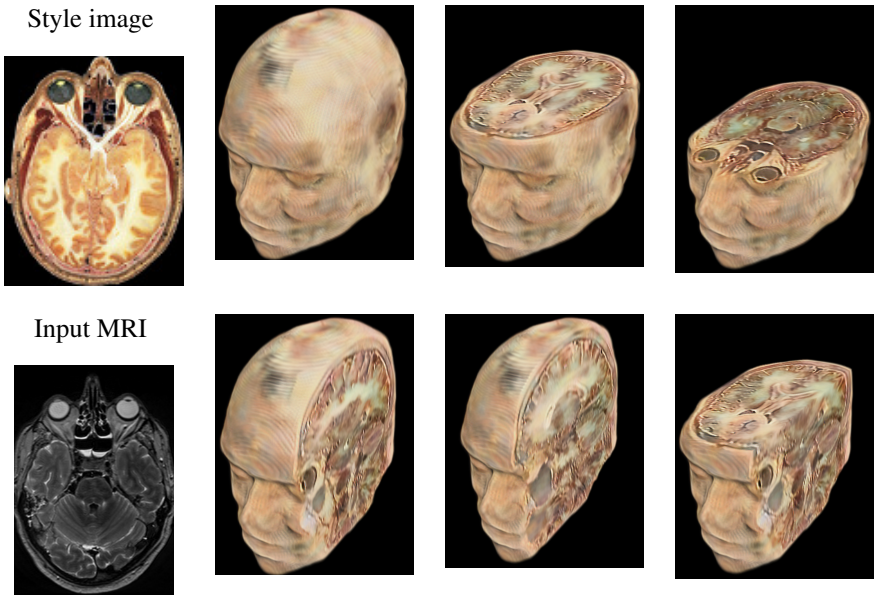


Figure 4. 3D rendering of the synthesised volume generated with the style image. RGBA volume visualized using Inviwo renderer [27].

solution.

3.5. Upsampling

Due to hardware restrictions, the maximum cube size that the GAN in our case could process is $128 \times 128 \times 88$. Therefore, we separately train an SRGAN [12] to upscale the slices from the output and obtain the high-resolution volume from the output. The SRGAN is trained to perform 8x upsampling of the image patches of cryo images of size 128×128. The final colorized volume is upsampled in a slice by slice manner along the axial view.

4. Results and comparisons

We train our method using the Visible Human 2 [1] 16-bit MRI and color cryo dataset. Our code is written in Python where we use PyTorch [23] for training the GAN and PyAMG [22] for solving the optimization problem to generate the colorized volume. Our computational setup consisted of an Intel Xeon 2.4 GHz processor with 128 GB memory and Nvidia Quadro P6000 GPU with 24 GB memory. Our code will be made available for research purposes.

Once our network is trained with $32 \times 32 \times 32$ sized sub-volumes, we pass the complete $128 \times 128 \times 88$ MRI volume for colorization at test-time, which is the maximum size of the volume that our 3D GAN network can handle due to available GPU memory resource. It took 0.165 sec for the modality conversion on the content volume of size $128 \times 128 \times 88$ by the GAN and upsampling takes approximately 1.5 seconds per image on our hardware. The high-resolution samples in the paper are of size 512×352 obtained by the trained SRGAN and the overall size of the resulting high-resolution color volume is $512 \times 512 \times 352$. In Figure 3 and Figure 6(e), it can be seen that the tissue intensities in MRI have been correctly mapped to tissue colors of the cryo style image, and structures of input MRI are maintained very well. In general MRI intensities do not map to cryo intensities due to the difference in the capturing methods, while cryo images are captured by camera containing an RGB array sensor, the MRIs are captured by magnetic fields and magnetic field gradients and our colorization takes care of such differences in the two modalities.

To the best of our knowledge, there are no existing methods that can be compared directly with our method. Most existing colorization methods work on 2D photographs and none on medical images. We first compare these methods with state-of-the-art exemplar colorization methods since that is similar to our objective. However, we observe that the modality of the data poses a major challenge for these state-of-the-art methods. In Figure 5 we show the results of exemplar colorization and observe that the results of colorization are not visually perceptive. We suspect that this is due to the different intensity distribution of MRI (w.r.t a photograph) and the dependence of these methods on a feature extraction network such as VGG. Further, these methods generate only the chroma channels without any changes to the luminosity channel which is not suitable for our objective. Hence we are motivated to compare our approach with four state-of-the-art methods for style transfer after suitable modifications needed to be applied for medical imaging data (see Figure 6). All of these methods require 2D images as input and therefore we process the MRI volume slice-by-slice along the axial plane. In practice, we suggest a full 3D processing as with our algorithm since that leads to a better coherency in the volume compared to a slice-by-slice approach.

In the comparisons, it can be seen that Huang et al. [7] is unable to maintain coherency across different slices and also adds undesirable artifacts. Li et al. [17] also do not maintain such a coherency. In addition, they perform photorealistic smoothening on the stylized image that blurs the structures of the input volume. Since Li et al. [15] operates on a video sequence, it preserves coherency across slices but does not fully transfer the style details. Results of Jamriška et al. [10] are best at reproducing the style but the method requires hints that are completely aligned with the input slices that may not be available for every style image. It also fails to preserve the structures in the input frames. We also perform a quantitative comparison of the generated images with respect to the input MRI and RGB Poisson volume. Table 1 shows average SSIM, PSNR, and MSE metrics computed for the entire volume and averaged over RGB color channels. SSIM is calculated between the generated volume and the input MRI volume to ensure that the structures present in the MRI are preserved. PSNR and MSE are calculated between the generated volume and the RGB Poisson volume to measure the color consistency and overall coherency across the volume. Please see the supplementary material for additional results.

Table 1. Quantitative comparison with recent style transfer methods. The comparisons were done on a $128 \times 128 \times 88$ cubic volume generated using each method before any upsampling and segmentation mask are applied for sake of a uniform comparison.

Method	**SSIM** $\uparrow$	**PSNR** $\uparrow$	**MSE** $\downarrow$
Li et al. [17]	0.313	7.57	0.177
Huang et al. [7]	0.313	7.02	0.201
Jamriška et al. [10]	0.314	6.03	0.255
Li et al. [15]	0.314	7.97	0.163
Ours	**0.319**	**9.72**	**0.144**

5. Discussion and Challenges

The unique setting of our problem brings forward several key challenges in medical data colorization. The 3D

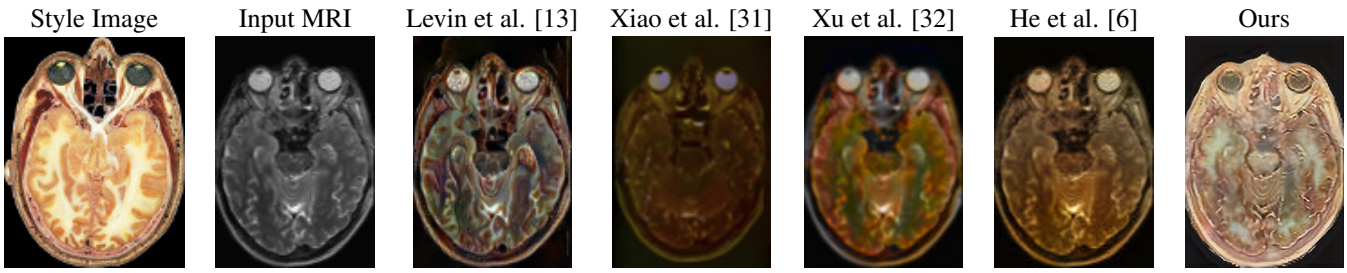


Figure 5. Comparison with exemplar colorization methods. For a particular style image, our method results in a much better visual perception. This demonstrates that exemplar colorization that cannot be performed on the MRI directly due to the nature of image modality.

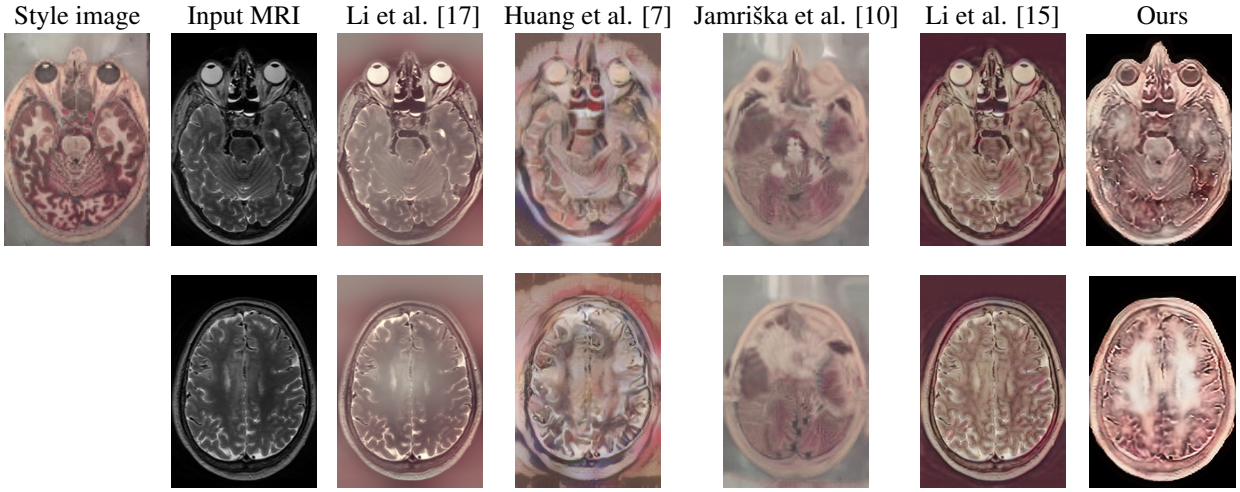


Figure 6. Comparison with style transfer methods. Our method results in the most photorealistic colors while preserving the structures of the input data. This balance of structural and radiometric reconstruction is lacking in most other methods that mostly focus on only one of these aspects.

nature of the data places severe restrictions on computational resources. The difference in modality leads to a lack of direct translation between MRI and cryo volume, thus requiring the need for grayscale synthesis. A key limitation of our method is that of the lack of semantic information for MRI volume that prevents context-aware stylization of hints. Gatys style transfer method and other stylization techniques are not suitable for medical data and do not seem to perform well leading to issues such as patchy colors and color bleeding due to lack of semantic information. Further, the lack of suitable colorization metrics for medical images make comparison and evaluation difficult.

6. Conclusion

Automatic colorization of 3D imaging datasets has not been much explored and is still dependent on highly manual methods. The lack of methods for 3D volumetric colorization can be attributed to the absence of a model such as VGG that is trained on datasets such as ImageNet [2] which provides an important backbone for a myriad of visual computing tasks. Our primary contribution is to propose a framework with the aim that it could be used on multi-modal data in 3D for performing colorization keeping in view the photorealism of the generated outputs. To do so we demonstrated how we can create a synthetic dataset to create a different modality that is much more suited to visual perception. Then we extend this method on newer unseen data using the generalization capabilities of a GAN. Overall, we build a framework that could allow us to perform colorization from contextual images and extend it to 3D. We also demonstrate how several state-of-the-art methods in 2D fail considerably on medical imaging data due to its multi-modal and 3D nature. Our approach takes into account all these factors and shows robustness both in terms of metrics and visual perception. Though our method has been demonstrated to work on 3D medical volumes, the overall framework is general enough to operate on other 3D multi-modal data. Our approach opens up many broad aspects of multi-modal colorization that has applications in high-fidelity photorealistic visualization and colorization.

7. Acknowledgement

We would like to thank reviewers for their useful comments and NVIDIA for supporting us by the NVIDIA GPU grant.

References

[1] Michael J Ackerman. The visible human project. *Proceedings of the IEEE*, 86(3):504–511, 1998.

[2] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In *CVPR09*, 2009.

[3] Leon A Gatys, Alexander S Ecker, and Matthias Bethge. Image style transfer using convolutional neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2414–2423, 2016.

[4] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680, 2014.

[5] John T Guibas, Tejpal S Virdi, and Peter S Li. Synthetic medical images from dual generative adversarial networks. *arXiv preprint arXiv:1709.01872*, 2017.

[6] Mingming He, Dongdong Chen, Jing Liao, Pedro V Sander, and Lu Yuan. Deep exemplar-based colorization. *ACM Transactions on Graphics (TOG)*, 37(4):1–16, 2018.

[7] Xun Huang and Serge Belongie. Arbitrary style transfer in real-time with adaptive instance normalization. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1501–1510, 2017.

[8] Peter J Huber. Robust estimation of a location parameter. In *Breakthroughs in statistics*, pages 492–518. Springer, 1992.

[9] Satoshi Iizuka, Edgar Simo-Serra, and Hiroshi Ishikawa. Let there be color!: joint end-to-end learning of global and local image priors for automatic image colorization with simultaneous classification. *ACM Transactions on Graphics (TOG)*, 35(4):110, 2016.

[10] Ondřej Jamriška, Šárka Sochorová, Ondřej Texler, Michal Lukáč, Jakub Fišer, Jingwan Lu, Eli Shechtman, and Daniel Sỳkora. Stylizing video by example. *ACM Transactions on Graphics (TOG)*, 38(4):1–11, 2019.

[11] M. Jirik, V. Lukes, M. Svobodova, and M. Zelezny. Image segmentation in medical imaging via graph-cuts. 2013.

[12] Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4681–4690, 2017.

[13] Anat Levin, Dani Lischinski, and Yair Weiss. Colorization using optimization. In *ACM transactions on graphics (tog)*, volume 23, pages 689–694. ACM, 2004.

[14] Bo Li, Yu-Kun Lai, Matthew John, and Paul L Rosin. Automatic example-based image colorization using location-aware cross-scale matching. *IEEE Transactions on Image Processing*, 28(9):4606–4619, 2019.

[15] Xueting Li, Sifei Liu, Jan Kautz, and Ming-Hsuan Yang. Learning linear transformations for fast image and video style transfer. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3809–3817, 2019.

[16] Yijun Li, Chen Fang, Jimei Yang, Zhaowen Wang, Xin Lu, and Ming-Hsuan Yang. Universal style transfer via feature transforms. In *Advances in neural information processing systems*, pages 386–396, 2017.

[17] Yijun Li, Ming-Yu Liu, Xueting Li, Ming-Hsuan Yang, and Jan Kautz. A closed-form solution to photorealistic image stylization. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 453–468, 2018.

[18] Patric Ljung, Jens Krüger, Eduard Groller, Markus Hadwiger, Charles D Hansen, and Anders Ynnerman. State of the art in transfer functions for direct volume rendering. In *Computer Graphics Forum*, volume 35, pages 669–691. Wiley Online Library, 2016.

[19] Dongdong Nie, Qinyong Ma, Lizhuang Ma, and Shuangjiu Xiao. Optimization based grayscale image colorization. *Pattern recognition letters*, 28(12):1445–1451, 2007.

[20] Dong Nie, Roger Trullo, Jun Lian, Caroline Petitjean, Su Ruan, Qian Wang, and Dinggang Shen. Medical image synthesis with context-aware generative adversarial networks. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 417–425. Springer, 2017.

[21] Dong Nie, Roger Trullo, Jun Lian, Li Wang, Caroline Petitjean, Su Ruan, Qian Wang, and Dinggang Shen. Medical image synthesis with deep convolutional adversarial networks. *IEEE Transactions on Biomedical Engineering*, 65(12):2720–2730, 2018.

[22] L. N. Olson and J. B. Schroder. PyAMG: Algebraic multigrid solvers in Python v4.0, 2018. Release 4.0.

[23] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. Pytorch: An imperative style, high-performance deep learning library. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems 32*, pages 8026–8037. Curran Associates, Inc., 2019.

[24] Patrick Pérez, Michel Gangnet, and Andrew Blake. Poisson image editing. *ACM Transactions on graphics (TOG)*, 22(3):313–318, 2003.

[25] Leonid I Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: nonlinear phenomena*, 60(1-4):259–268, 1992.

[26] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[27] Erik Sunden, Peter Steneteg, Sathish Kottravel, Daniel Jonsson, Rickard Englund, Martin Falk, and Timo Ropinski. Inviwo-an extensible, multi-purpose visualization framework. In *2015 IEEE Scientific Visualization Conference (SciVis)*, pages 163–164. IEEE, 2015.

[28] Zhuowen Tu and Xiang Bai. Auto-context and its application to high-level vision tasks and 3d brain image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 32(10):1744–1757, 2009.

[29] Zhou Wang, Alan C Bovik, Hamid R Sheikh, Eero P Simoncelli, et al. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004.

[30] Tomihisa Welsh, Michael Ashikhmin, and Klaus Mueller. Transferring color to greyscale images. In *ACM Transactions on Graphics (TOG)*, volume 21, pages 277–280. ACM, 2002.

[31] Chufeng Xiao, Chu Han, Zhuming Zhang, Jing Qin, Tien-Tsin Wong, Guoqiang Han, and Shengfeng He. Example-based colourization via dense encoding pyramids. In *Computer Graphics Forum*, volume 39, pages 20–33. Wiley Online Library, 2020.

[32] Zhongyou Xu, Tingting Wang, Faming Fang, Yun Sheng, and Guixu Zhang. Stylization-based architecture for fast deep exemplar colorization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.

[33] Zhenpei Yang, Lihang Liu, and Qixing Huang. Learning generative neural networks for 3d colorization. In *AAAI*, 2018.

[34] Richard Zhang, Phillip Isola, and Alexei A Efros. Colorful image colorization. In *European conference on computer vision*, pages 649–666. Springer, 2016.

[35] Richard Zhang, Jun-Yan Zhu, Phillip Isola, Xinyang Geng, Angela S Lin, Tianhe Yu, and Alexei A Efros. Real-time user-guided image colorization with learned deep priors. *arXiv preprint arXiv:1705.02999*, 2017.