Segmentation and Immersive Visualization of Brain Lesions Using Deep Learning and Virtual Reality

Brendan Kelley
Natural User Interaction Lab
Colorado State University
Fort Collins, CO, USA
brendan.kelley@colostate.edu

Lucas Plabst
Natural User Interaction Lab
Colorado State University
Fort Collins, CO, USA
Department of Human-Computer Interaction
Julius Maximilian University of Würzburg
Würzburg, Germany
lucas.plabst@colostate.edu

Lena Plabst
Natural User Interaction Lab
Colorado State University
Fort Collins, CO, USA
lena.hinzer@colostate.edu

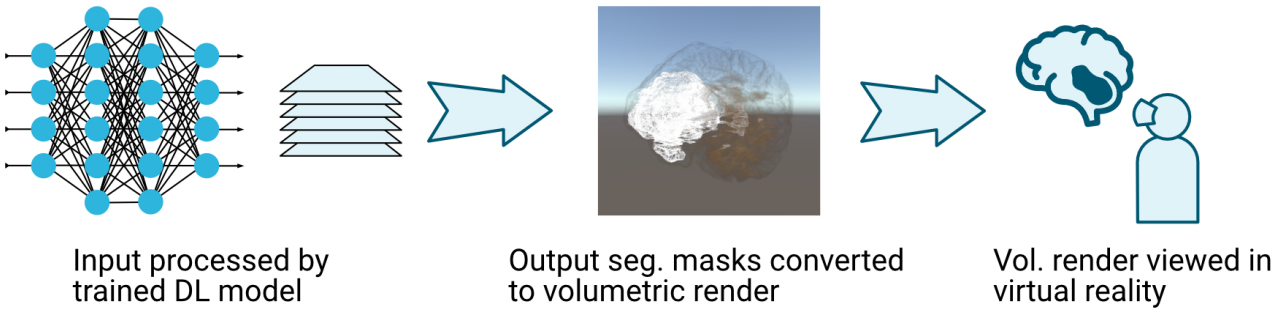


Figure 1: Overview of our visualization pipeline. MRI scan images are passed through a deep learning image segmentation model. These are then rendered volumetrically in a Unity virtual reality application along with the output segmentation. The volumetric renders can then be viewed in an immersive HMD.

Abstract

Magnetic resonance imaging (MRIs) are commonly used for diagnosing potential neurological disorders, however preparation and interpretation of MRI scans requires professional oversight. Additionally, MRIs are typically viewed as single cross sections of the affected regions which does not always capture the full picture of brain lesions and can be difficult to understand due to 2D's inherent abstraction of our 3D world. To address these challenges we propose a immersive visualization pipeline that combines deep learning image segmentation techniques using a VGG-16 model trained on MRI fluid attenuated inversion recovery (FLAIR) with virtual reality (VR) immersive analytics. Our visualization pipeline begins with our VGG-16 model predicting which regions of the brain are potentially affected by a disease. This output, along with the original scan, are then volumentrically rendered. These renders can then be viewed in VR using an head mounted display (HMD). Within the HMD users can move through the volumentric renderings to view the affected regions and utilize planes to view cross sections of the MRI scans. Our work provides a potential pipeline and tool for diagnosis and care.

CCS Concepts

• **Human-centered computing** → **Scientific visualization**; *Visualization systems and tools*; • **Computing methodologies** → **Virtual reality**; *Machine learning*.

Keywords

Deep Learning, Image Segmentation, Medical Imaging, Medical Applications, Virtual Reality, Immersive Analytics, Immersive Visualization

VRCAI '24, December 01–02, 2024, Nanjing, China

ACM ISBN 979-8-4007-1348-4/24/12
https://doi.org/10.1145/3703619.3706035

ACM Reference Format:
Brendan Kelley, Lucas Plabst, and Lena Plabst. 2024. Segmentation and Immersive Visualization of Brain Lesions Using Deep Learning and Virtual Reality. In *The 19th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry (VRCAI '24), December 01–02, 2024, Nanjing, China.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3703619.3706035

1 Introduction

Whenever there is suspected brain disease in a patient, an MRI scan is commonly conducted. This allows medical practitioners to view the patient's white and grey matter layer by layer to diagnose potential issues. It is imperative that this process be done efficiently, and accurately, to allow early intervention for potentially life-threatening diseases such as tumors, Alzheimer's, traumatic brain injuries, etc. To help with this process, deep learning techniques have been leveraged for image segmentation allowing for fast and effective detection of brain lesions [Cheng et al. 2001; Minaee et al. 2020]. However, without context, these images mean little to the individual.

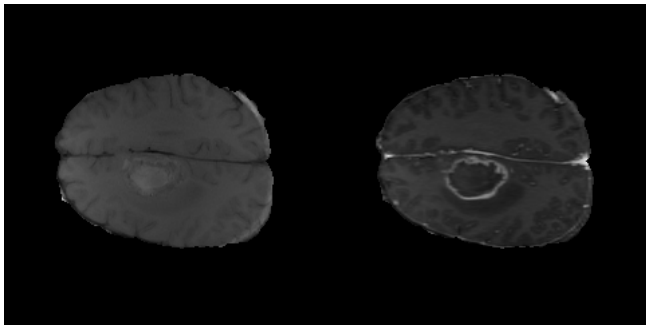
Virtual Reality (VR) allows for immersive visualization of medical imagery that would otherwise be presented using 2D image cross sections. These cross-sections are inherently abstracted from our 3D world and don't always convey the complete severity and urgency needed for understanding brain lesions. To this end, we propose a pipeline for producing MRI segmentations of brain lesions and visualizing the output in VR. A VGG-16 model was trained using distributed deep learning with FLAIR MRI data to segment lesions. The output from this model was then converted into a volumetric rendering. Finally, this render was imported into a virtual reality application for visualization.

2 Background

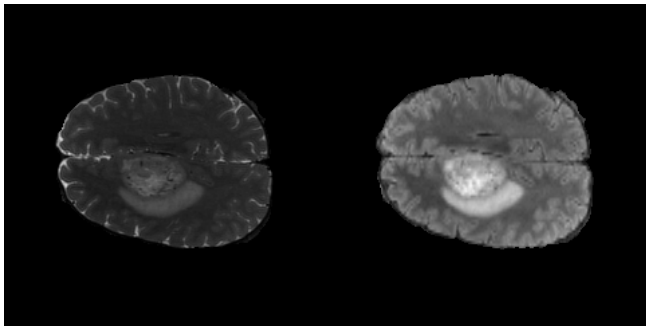
Magnetic resonance imaging or MRIs are often used to detect brain lesions and diagnose various forms of brain disease such as tumors [Ishida and Cucchiara 2021], migraines [Negm et al. 2018], glioma [Brabec et al. 2022], Alzheimer's [Tubi et al. 2020], and a variety of other cerebrovascular diseases [DeCarli and Lockhart 2014]. The body's natural magnetic properties are leveraged to create an MRI scan [Berger 2002]. During scanning, protons within the brain are exposed to a strong magnetic field. This, in turn, aligns the protons creating a magnetic vector that can be detected, read, interpreted, and transformed into imagery.

Several different forms of MRI scans exist, including T1 weighted, T1 contrast-enhanced (T1CE), T2 weighted, and fluid-attenuated inversion recovery (FLAIR) [Kates et al. 1996] (Figure 2). T1 weighted scans more closely approximate the tissue appearance (Figure 2a). Related to T1 scans, are T1CE scans (Figure 2b), which push the contrast of the tissues causing lesions to appear as dark spots. T2 scans (Figure 2c), on the other hand, show these spots as brighter. Of particular interest to this work is the use of FLAIR scans (Figure 2d), which remove cerebrospinal fluid and increase contrast between white and grey matter [Kates et al. 1996], and the aforementioned T1CE scans since both have a higher contrast of pixel values for classification.

White matter is of particular interest (as opposed to grey matter which houses neurons) for MRI scans, as the brain lesions that indicate other issues such as tumors, migraines, strokes, or other trauma appear as white matter areas of the brain. With FLAIR, these lesions will appear as bright spots, often referred to as white matter hyper-intensities or WMH. Conversely, with T1CE these lesions will frequently appear as dark spots. In either case, these serve as markers for various neurological problems. However, not all white matter is problematic as these can also include nerve fibers surrounded by myelin which are necessary for brain function [Fields 2010].



(a) T1 MRI (b) T1CE MRI



(c) T2 MRI (d) FLAIR MRI

Figure 2: Types of MRI scans with different image features. Scans shown are from the same individual and the same NII slice, but with different scan methods.

2.1 Image Segmentation with Deep Learning CNNs

To identify brain lesions, image segmentation strategies are often employed. This is the process of partitioning images into segments for identification [Minaee et al. 2020]. This can be conceptualized as pixel-by-pixel classification tasks, where a mask is produced with each resulting pixel corresponding to a classification code (i.e. 0 corresponding to the background, 1 corresponding to a person, and 2 corresponding to a dog). This mask can then be utilized for a variety of different tasks, such as autonomous vehicles, medical imaging, or object detection. A variety of different techniques from computer vision to artificial intelligence have been used for segmentation tasks [Cheng et al. 2001; Minaee et al. 2020], including machine learning and deep learning.

A variety of convolutional neural networks (CNNs), as well as their deep learning counterparts have been used for medical image segmentation, including ResNet50[He et al. 2015], VGG-16[Simonyan and Zisserman 2014], and MobleNet[Howard et al. 2017]. ResNet and VGG-19, were both used for the detection of pneumonia from chest x-rays [Victor Ikechukwu et al. 2021], both with over 90% accuracy, and have been shown to be beneficial in

a variety of different medical image processing [Xu et al. 2023]. Another study used MNet for the classification of Alzheimer's disease [S et al. 2023]. Similar techniques were also applied to the detection of Parkinson's [Ul Haq et al. 2018] and glioma [Bakas et al. 2017a; Ben Ahmed et al. 2022; Coupet et al. 2022]. The use of these techniques has helped to improve the detection and diagnosis of breast cancer, where the use of deep neural networks was able to achieve a high AUC for prediction [Wu et al. 2020]. In some cases, the predictions produced by deep learning segmentation models were even translated into 3D reconstructions, such as in [Huang et al. 2022], where the median nerve in ultrasound images was converted into a 3D model after segmentation, which may allow for better patient understanding of the size, nature, and relation of the medical data than with the traditional 2D images.

2.2 Immersive Analytics

The immersive properties of VR technology have long been shown to have a number of benefits, such as better spatial understanding, peripheral awareness, and a more realistic experience overall [Bowman and McMahan 2007]. Specifically, VR has proven useful when conveying information to users, as it uses spatial cues and provides more space to organize and place this information [Ragan et al. 2012], supporting memorization and recall [Mann et al. 2017]. A meta-analysis from 2019 evaluated the effectiveness of 2D and 3D VR visualization regarding discrete-event simulation. All 32 articles they analyzed found 3D VR representation to be more useful than 2D representation [Akpan and Shanker 2019]. Further research supports the benefits of offering 3D VR visualization in combination with the existing 2D visualization options [Akpan et al. 2020].

In the medical field, 3D visualization is an important tool to support more accurate diagnoses, as it can provide doctors with more accurate information on the size and shape of lesions, as well as spatial information on the lesions in relationship to the surrounding tissue. This is especially true when the visualization utilizes interaction to support the presentation of spatial information [Liang et al. 2011]. 3D visualization has previously been used successfully in several cases within the medical field. An evaluation of laparoscopic surgery skill training compared performance between 2D and 3D visualizations and found that 3D visualization leads to shorter task completion time and improved task precision; without an increase in side effects [Tanagho et al. 2012]. While 3D visualization using traditional displays has been utilized successfully in practice and in research, there is little research on the effectiveness of using virtual reality for 3D visualization in the medical field. Given the importance of interaction with 3D visualization and the positive results of 3D VR visualization outside of the medical field, there is a lot of potential for exploration.

3 Methods

Our MRI visualization pipeline (see Figure 1) consists of three major parts: 1) a CNN model built with the Keras ML package in python[Chollet et al. 2015; Gupta 2023] and trained using Horovod [Sergeev and Del Balso 2018] over a distributed spark cluster [Zaharia et al. 2016], 2) a volumetric rendering or model produced using the open source imaging software 3DSlicer [Fedorov et al. 2012], and 3) an immersive VR visualization system built using Unity3D [Technologies 2023]. MRI images are processed using the CNN DL model, producing a series of segmentation image predictions which can then be converted into a volumetric render and finally imported into our Unity application for viewing.

3.1 Data Set

MRI samples from the brain tumor segmentation (BraTS) 2021, '20, '19, and '18 datasets [Bakas et al. 2018, 2017a,b,c; Menze et al. 2015] were collected for training our distributed CNNs. This dataset contains NII MRI scans for 1251 patients. Each patient has four MRI scan types (T1 weighted, T1CE (contrast-enhanced), T2 weighted, and FLAIR), as well as, prepared segmentation files to serve as the ground truth for comparison. Each NII scan file contains a total of 155 images each containing a cross-section of the patient's brain (see Figure 3). Depending on the scan type, lesions will have different properties (i.e., black holes for T1CE and white hyperintensity for FLAIR). Some images consisted of entirely null space with no patient cross-section. These particular images were either at the very beginning or the very end of a scan series. Other images contained cross-sections of healthy matter as the lesion was either above or below the particular cross-section.

A total of 969525 images were extracted from the NII files with each scan type (T1, T1CE, T2, FLAIR) containing 193905 images. These images were then separated into T1, T1CE, T2, FLAIR, and segmentation subsets, however only the T1CE, FLAIR, and segmentation data was used. Each subset was then further divided into training, validation, and testing sets with a ratio of 60% training, 20% validation, and 20% testing. Each segmentation mask image consists of 240 by 240 pixels, with the pixel values representing one of 4 classes. Portions of the image classified as 0 contain no relevant medical data (i.e. background or healthy brain matter), and classes 1-3 represent a portion of the brain with detected lesions (but corresponded to different features of the lesion).

On average a full scan, with the exclusion of completely empty images, consisted of 97.44% background pixels or pixels with no relevant segmentation (lesion) data (see Table 1). The remaining 2.56% of pixels constituted the area of interest depicting the lesions. Of this 2.56%, approximately 70% of these pixels are class 2, around 21% are class 3, and the remaining 9% are class 1.

Table 1: Segmentation class composition excluding completely empty images. Per scan and within the segmentation key values are presented.

	Class 0	Class 1	Class 2	Class 3
Per-scan	97.44%	0.35%	1.59%	0.62%
Within seg.	NA	9.16%	70.19%	20.65%

3.2 Model Training

Four different models were trained using both T1CE and FLAIR datasets (custom implemented U-Net model, ResNet50[He et al. 2015], VGG-16[Simonyan and Zisserman 2014], and MobleNet[Howard et al. 2017]) for a total of eight models. Several prior studies have used ResNet, VGG, and MobleNet convolutional neural networks successfully for medical image segmentation [Huang et al. 2022;

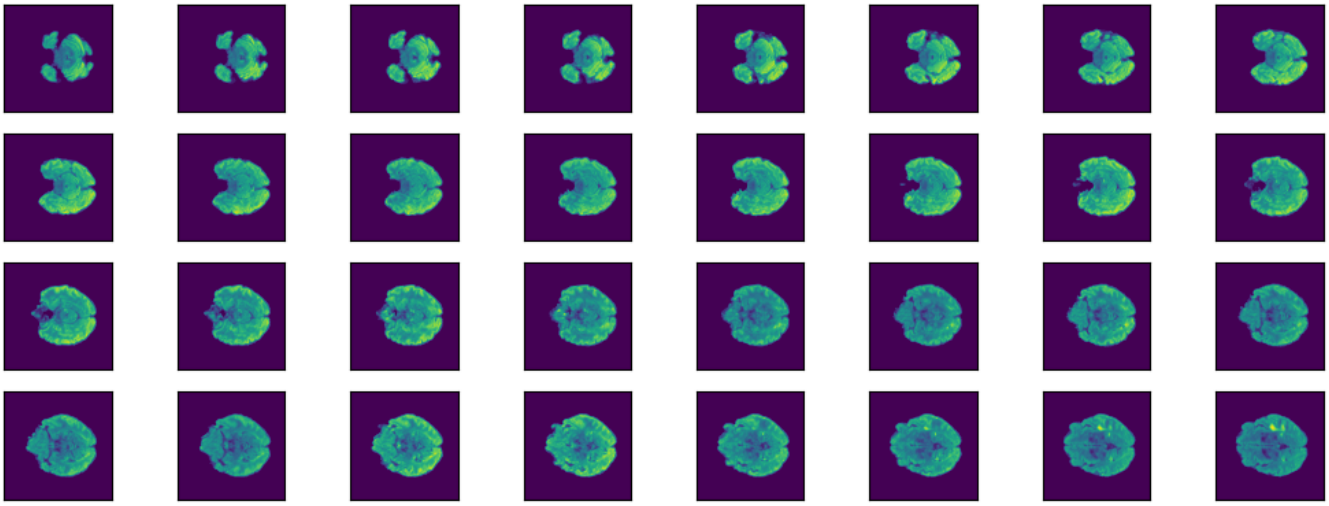


Figure 3: Series of MRI pngs contained within a NII scan file.

S et al. 2023; Victor Ikechukwu et al. 2021; Xu et al. 2023]. Model training was conducted using Keras [Chollet et al. 2015] over a distributed Apache Spark cluster [Zaharia et al. 2016] with Horovod [Sergeev and Del Balso 2018]. After conducting initial training on eight models (each using one of four architectures and trained on two distinct datasets) over five epochs, with each epoch comprising 50 steps, the visual and validation metrics consistently favored the VGG-16 architecture as the most robust choice for our task. From here on VGG-16 was used for further refinement with both the T1CE and FLAIR datasets at a variety of epochs, steps, and learning rates to fine-tune the datasets.

3.2.1 Model Refinement and Validation. To validate the two VGG-16 models, precision and recall were measured (see 4a and 4b) for each pixel in the output segmentation prediction. Precision is the number of true positives (TP) over the number of all positive classifications (both TP and false positives FP). For both TP and FP a positive classification is any classification other than 0, representing background pixels with no relevant medical data. Recall, on the other hand, is the sensitivity or true positive rate. This is the number of TP results divided by the sum of the TP and false negatives (FN), where a FN pixel is one where a background pixel is erroneously classified 1-3.

3.3 Virtual Reality Visualization Tool

To apply the segmentation output of our model to a real-world use case, such as consultation with a patient, we implemented a VR-enabled visualization tool using Unity Engine 2022.3.7f1 [Technologies 2023]. To enable the VR functionality, we used the Unity XR interaction toolkit version 2.4.3 [XRI 2024] and OculusXR [ocu 2024]. In our testing, we used a Meta Quest 3 VR headset released in 2023, featuring 2064x2208 pixels per eye resolution and 120-hertz screens. The application itself is platform-agnostic, and can be used on any recent VR headset. The built binary utilized in this work, however, is currently only configured to run on Meta headsets.

$$Precision = \frac{TP}{TP + FP} \quad \textbf{(a)}$$

$$Recall = \frac{TP}{TP + FN} \quad \textbf{(b)}$$

Figure 4: The precision and recall formulas for validation where TP is the number of true positives, FP is the number of false positives, and FN is the number of false negatives.

Our process started by using the open-source medical imaging software 3D Slicer [Fedorov et al. 2012]. Using this tool, we can create volumetric renders from the model's image series input and output. We input our scans and the VGG-16 model outputs as a series of PNG images, and use the 3D Slicer tool to create a nearby raw raster data (NRRD)-file from that series. The NRRD format is an open-source file format, designed for storing data, particularly n-dimensional raster data used in scientific visualization and image processing applications. This file can then be imported into our Unity application, which then renders the format as a 3D volume using the Unity Volume Rendering Package [Lavik 2024].

Using this, we display the initial brain scan (Figure 5), as well as the VGG-16 segmentation output. The brain scan is rendered in a gray color, while the highlights are displayed in teal. From the NRRD file, we create a 3D array, where every data point represents the density at that point. It is then stored in a 3D texture, which is then rendered on a box object. We employ raymarching to visualize the densities of each visible voxel (value on a three-dimensional grid) within the rendering volume. Raymarching is a rendering technique where a ray originates from the virtual camera and advances in uniform steps along its path. At each step, it samples density values

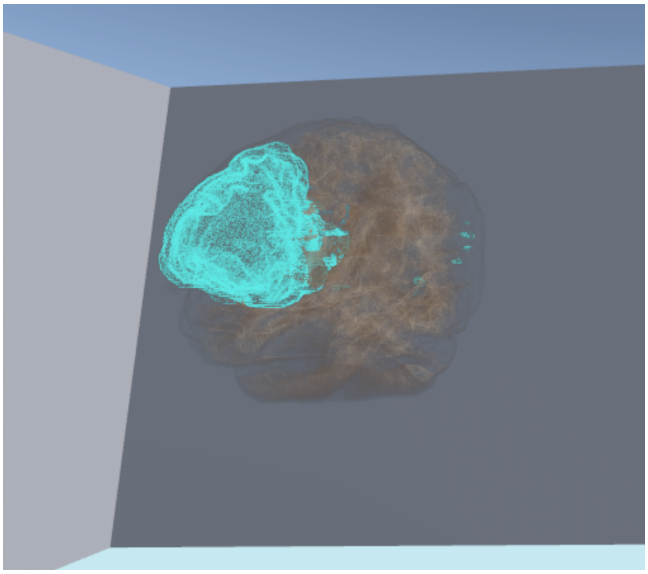


Figure 5: The volumetric render of an MRI Scan (grey volume) with highlighted segmentaiton area (teal volume).

and uses a predefined transfer function to render them. This process results in a real-time volumetric representation of the previously two-dimensional data.

Traditionally in 3D computer graphics, surface models are used to display 3D models, which are easier to render. However, surface models do not have any internal structure, thus being hollow. These surface models were therefore not suitable for our purposes, so we opted for volumetric renders, which while being more demanding to render, completely retain all detail from the original data.

We provide three slice renders (see Figure 6) to allow the user to view specific layers for all three planes of the brain (coronal, axial, and sagittal planes), which are often used for current 2D MRI viewing solutions. These slices update the view in real-time. Besides the slices in the brain, the individual current layers are also presented as images on the wall in the environment, for better viewing. Users can also simply "look into" the brain by moving closer to it for an increased understanding of the affected regions. They can scale the brain up and down in size, for either a blown-up view with higher detail or at a life-size scale for realism.

4 Results

While several models were able to produce output resembling the segmentation image associated with a given input, the VGG-16 implementation produced the strongest initial results and was carried forward for further refinement and validation. Results from the VGG-16 T1CE and FLAIR models support the use of FLAIR over T1CE for MRI segmentation as the recall score for the FLAIR model (76.16) was around 40% higher than the T1CE model (35.61). Additionally, while the FLAIR precision (98.28) was higher than T1CE (87.28), both were above 80%. Visual inspection of the resulting FLAIR VGG predictions shows noise around the brain, especially near the skull (see Figure 7).

Table 2: VGG precision and recall.

Model	Precision	Recall
VGG Flair	**98.28**	**76.16**
VGG T1CE	87.36	35.61

Table 3: Per-class recall with the final VGG16 model trained on FLAIR MRI data. Class 0 corresponds to empty pixels with no relevant segmentation information. Classes 1, 2, and 3 are all part of the segmentation.

Class	Recall
BG (Class 0)	**99.71**
Class 1	33.14
Class 2	88.75
Class 3	34.78
Differentiation	**98.44**

In addition to overall model precision and recall, the per-class and differentiation recall were calculated for the VGG FLAIR model. Differentiation is the ability of the model to classify a pixel as either background (both negative space and pixels with no segmentation data) or as part of the segmentation. For differentiation, the accuracy of classification for segmentation is not considered, only whether the model can differentiate between background and areas of interest (i.e. classified as 0 or as 1/2/3). As seen in Table 3, the VGG FLAIR model performs well at differentiating between background (class 0) and segmentation data (class 1, 2, or 3), however, the performance per class is not as strong with both class 1 classification and class 3 classification below 35%. Class 2 classification had the highest recall rate for a segmentation class at 88.75%. As reported in Table 1 2.56% of the data was comprised of Classes 1-3 and of that 2.56%, approximately 70% of those pixels were class 2, around 21% were class 3, and the remaining 9% were class 1.

5 Discussion

For our specific task, we found the VGG-16 model to be the most readily useful for the MRI image segmentation task. However, the lower performance from the other models may be a product of the high number of class 0 pixels present within a scan. Due to this, with simple accuracy measures such as pixel-by-pixel classification comparisons, a model could still achieve a high score simply by outputting a segmentation prediction of all 0s. Thus the use of more refined validation and comparison metrics was necessary. This issue could also be alleviated in future works by further manipulating and cleaning the dataset. While both FLAIR and T1CE data led to strong precision, meaning that when either of the VGG-16 models predicted a positive pixel that prediction was usually correct, the T1CE recall was far lower than the FLAIR model. This suggests that using FLAIR data for segmentation prediction allows for better segmentation output. Where the model suffered was in classifying class 1 and class 3 pixels. This is most likely due to the low number of these pixels represented in the data set with less than 1% of all pixels in a scan corresponding to class 1 or class 3. This may have inadvertently encouraged the model to predict class 2 more often.

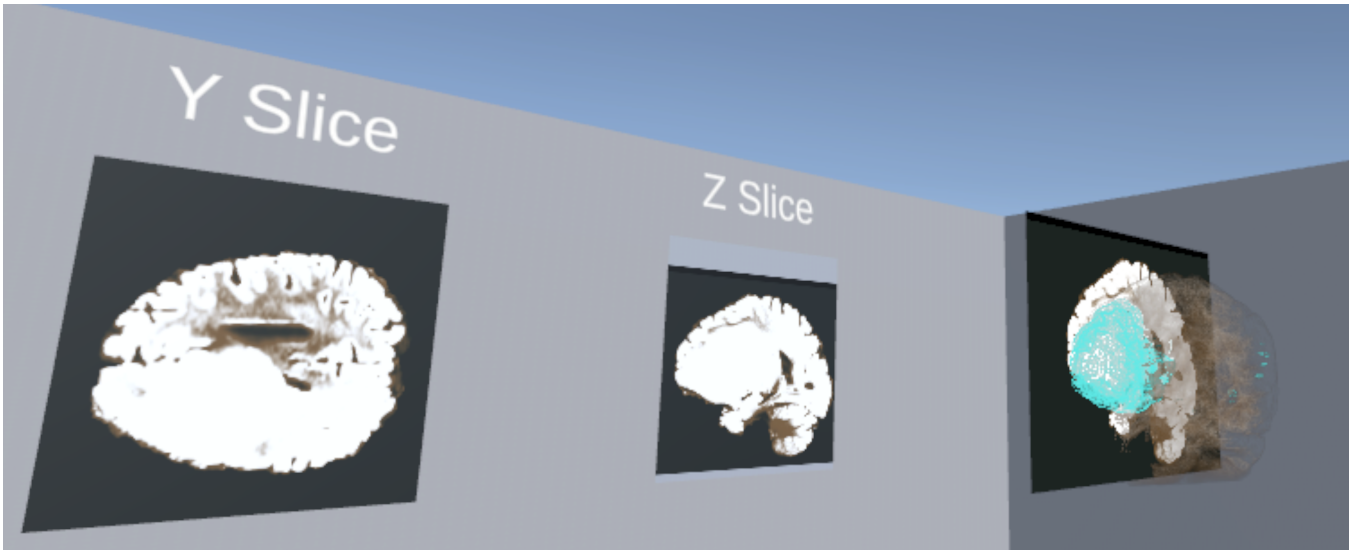


Figure 6: Brain render with slicing plane features.

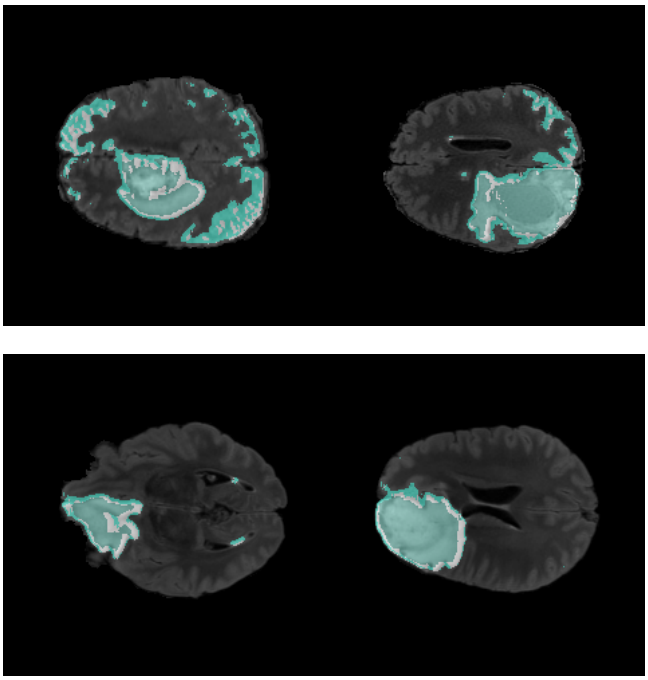


Figure 7: Results from the VGG flair model overlaid onto the original scan.

While the per-class recall was lacking, the trained model's ability to differentiate pixels with no relevant data for segmentation (class 0) and sections containing segmentation data (Classes 1-3) was strong, as evidenced by the differentiation recall value in Table 3. This means that the area of segmentation was highly accurate even if the specific regions denoted by classes 1-3 were not as sensitive. So, while there are potential improvements that would be necessary for better segmentation, for the purposes of our immersive visualization this does not significantly impact the rendering as the volumetric render step does not differentiate between segmentation classes (classes 1-3). Instead any pixel not classified as class 0 would be included in the volume without distinction. Further, data manipulation and hyper-parameter tuning may help to increase the Class 1 and Class 3 recall values.

5.1 Limitations and Future Work

Further measures would need to be taken to improve the overall segmentation accuracy of the VGG-16 model. This may include further hyper-parameter tuning and model refinement, or different data cleaning/augmentation processes (such as mirroring the scans or repeating scans with higher Class 1 and Class 3 presences). Additionally, there may be other models and modifications, either not fully validated or completely absent from this work, that may provide stronger results. Differentiating between the different classes of lesions may provide for a greater understanding of the immersive volumetric renderings, but would require a higher per-class recall for Class 1 and Class 3.

The visualization tool has not been evaluated from a medical perspective yet. This is important for future work to establish the effectiveness of the visualization tool in relevant contexts, such as in supporting diagnoses. With added networking capabilities, the tool could work better in a consultation setting, where the medical professional and the patient could be in the environment together.

Also, using a more powerful headset could allow the application to run on a standalone headset, without the need for a dedicated computer, allowing for more flexible use. A more powerful headset, for example, Apple's Vision Pro, would additionally be able to support a pass-through view, where the brain renders would be displayed in the real world as captured by outside cameras, instead of a fully virtual environment Lastly, adding support for marker-based

tracking could allow medical professionals to correctly overlay that render on a real patient, to get a better spatial understanding of the patient they are operating on.

Despite these limitations, this work has demonstrated the potential for combining deep learning convolutional neural networks with immersive virtual reality visualization for MRI segmentation.

6 Conclusion

In this work, we present a pipeline for combining deep-learning image segmentation techniques with immersive virtual reality visualization. We began by training several models using two forms of MRI data. From this initial training, we found that the implementation of a VGG-16 model best suited our needs. This model was then further refined with both FLAIR and T1CE data for comparison. While both forms of data allowed for high precision the recall from the VGG-16 FLAIR model was over 40% higher. While the per-class recall of this model was less than stellar, the ability for it to differentiate between background and relevant segmentation instances was high allowing us to produce useful predictions for immersive volumetric rendering in virtual reality. While there are opportunities to improve upon this work, the established pipeline demonstrates the potential for combing deep learning for image segmentation with virtual reality to provide new tools for medical consultation and visualization.

References

2024. Install, Uninstall, and Upgrade XR Plugin | Oculus Developers — developer.oculus.com. https://developer.oculus.com/documentation/unity/unity-xr-plugin/.

2024. XR Interaction Toolkit | 2.4.3. https://docs.unity3d.com/Packages/com.unity.xr.interaction.toolkit@2.4/manual/index.html.

Ikpe Justice Akpan and Murali Shanker. 2019. A comparative evaluation of the effectiveness of virtual reality, 3D visualization and 2D visual interactive simulation: an exploratory meta-analysis. *Simulation* 95, 2 (2019), 145–170.

Ikpe Justice Akpan, Murali Shanker, and Rouzbeh Razavi. 2020. Improving the success of simulation projects using 3D visualization and virtual reality. *Journal of the Operational Research Society* 71, 12 (2020), 1900–1926.

Bakas et al. 2018. Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation, Progression Assessment, and Overall Survival Prediction in the BRATS Challenge. (2018). https://doi.org/10.48550/ARXIV.1811.02629 Publisher: [object Object] Version Number: 3.

Spyridon Bakas, Hamed Akbari, Aristeidis Sotiras, Michel Bilello, Martin Rozycki, Justin Kirby, John Freymann, Keyvan Farahani, and Christos Davatzikos. 2017a. Segmentation Labels for the Pre-operative Scans of the TCGA-GBM collection. https://doi.org/10.7937/K9/TCIA.2017.KLXWJJ1Q

Spyridon Bakas, Hamed Akbari, Aristeidis Sotiras, Michel Bilello, Martin Rozycki, Justin Kirby, John Freymann, Keyvan Farahani, and Christos Davatzikos. 2017b. Segmentation Labels for the Pre-operative Scans of the TCGA-LGG collection. https://doi.org/10.7937/K9/TCIA.2017.GJQ7R0EF

Spyridon Bakas, Hamed Akbari, Aristeidis Sotiras, Michel Bilello, Martin Rozycki, Justin S. Kirby, John B. Freymann, Keyvan Farahani, and Christos Davatzikos. 2017c. Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features. *Scientific Data* 4, 1 (Sept. 2017), 170117. https://doi.org/10.1038/sdata.2017.117

Kaoutar Ben Ahmed, Lawrence O. Hall, Dmitry B. Goldgof, and Robert Gatenby. 2022. Ensembles of Convolutional Neural Networks for Survival Time Estimation of High-Grade Glioma Patients from Multimodal MRI. *Diagnostics* 12, 2 (Jan. 2022), 345. https://doi.org/10.3390/diagnostics12020345

A. Berger. 2002. How does it work?: Magnetic resonance imaging. *BMJ* 324, 7328 (Jan. 2002), 35–35. https://doi.org/10.1136/bmj.324.7328.35

Doug A Bowman and Ryan P McMahan. 2007. Virtual reality: how much immersion is enough? *Computer* 40, 7 (2007), 36–43.

Jan Brabec, Faris Durmo, Filip Szczepankiewicz, Patrik Brynolfsson, Björn Lampinen, Anna Rydelius, Linda Knutsson, Carl-Fredrik Westin, Pia C. Sundgren, and Markus Nilsson. 2022. Separating Glioma Hyperintensities From White Matter by Diffusion-Weighted Imaging With Spherical Tensor Encoding. *Frontiers in Neuroscience* 16 (April 2022), 842242. https://doi.org/10.3389/fnins.2022.842242

H.D. Cheng, X.H. Jiang, Y. Sun, and Jingli Wang. 2001. Color image segmentation: advances and prospects. *Pattern Recognition* 34, 12 (Dec. 2001), 2259–2281. https://doi.org/10.1016/S0031-3203(00)00149-7

Francois Chollet et al. 2015. *Keras*. https://github.com/fchollet/keras

Matthieu Coupet, Thierry Urruty, Teerapong Leelanupab, Mathieu Naudin, Pascal Bourdon, Christine Fernandez Maloigne, and Rémy Guillevin. 2022. A multi-sequences MRI deep framework study applied to glioma classfication. *Multimedia Tools and Applications* 81, 10 (April 2022), 13563–13591. https://doi.org/10.1007/s11042-022-12316-1

C. DeCarli and S.N. Lockhart. 2014. Cerebrovascular Disease. In *Encyclopedia of the Neurological Sciences*. Elsevier. https://doi.org/10.1016/B978-0-12-385157-4.00443-7

Andriy Fedorov, Reinhard Beichel, Jayashree Kalpathy-Cramer, Julien Finet, Jean-Christophe Fillion-Robin, Sonia Pujol, Christian Bauer, Dominique Jennings, Fiona M Fennessy, Milan Sonka, John Buatti, Stephen R Aylward, James V Miller, Steve Pieper, and Ron Kikinis. 2012. 3D Slicer as an Image Computing Platform for the Quantitative Imaging Network. *Magnetic Resonance Imaging* 30, 9 (Nov 2012), 1323–1341.

R. Douglas Fields. 2010. Change in the Brain's White Matter. *Science* 330, 6005 (Nov. 2010), 768–769. https://doi.org/10.1126/science.1199139

Divam Gupta. 2023. Image segmentation keras: Implementation of segnet, fcn, unet, pspnet and other models in keras. *arXiv preprint arXiv:2307.13215* (2023).

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2015. Deep Residual Learning for Image Recognition. (2015). https://doi.org/10.48550/ARXIV.1512.03385 Publisher: [object Object] Version Number: 1.

Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. 2017. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. (2017). https://doi.org/10.48550/ARXIV.1704.04861 Publisher: [object Object] Version Number: 1.

Aiyue Huang, Li Jiang, Jiangshan Zhang, and Qing Wang. 2022. Attention-VGG16-UNet: a novel deep learning approach for automatic segmentation of the median nerve in ultrasound images. *Quantitative Imaging in Medicine and Surgery* 12, 6 (June 2022), 3138–3150. https://doi.org/10.21037/qims-21-1074

Koto Ishida and Brett L. Cucchiara. 2021. White Matter Lesions on Magnetic Resonance Imaging. In *Decision-Making in Adult Neurology*. Elsevier, 238–241.e1. https://doi.org/10.1016/B978-0-323-63583-7.00116-8

R. Kates, D. Atkinson, and M. Brant-Zawadzki. 1996. Fluid-attenuated inversion recovery (FLAIR): clinical prospectus of current and future applications. *Topics in magnetic resonance imaging: TMRI* 8, 6 (Dec. 1996), 389–396.

Martin Lavik. 2024. *UnityVolumeRendering*. https://github.com/mlavik1/UnityVolumeRendering

Ying-Zong Liang, Bin Fang, Yi Wang, Lei Wu, and Peng Chen. 2011. Design of an interactive 3D medical visualization system. In *2011 International Conference on Wavelet Analysis and Pattern Recognition*. IEEE, 60–64.

Jessie Mann, Nicholas Polys, Rachel Diana, Manasa Ananth, Brad Herald, and Sweetuben Platel. 2017. Virginia tech's study hall: A virtual method of loci mnemotechnic study using a neurologically-based, mechanism-driven, approach to immersive learning research. In *2017 IEEE Virtual Reality (VR)*. IEEE, 383–384.

Menze et al. 2015. The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). *IEEE Transactions on Medical Imaging* 34, 10 (Oct. 2015), 1993–2024. https://doi.org/10.1109/TMI.2014.2377694

Shervin Minaee, Yuri Boykov, Fatih Porikli, Antonio Plaza, Nasser Kehtarnavaz, and Demetri Terzopoulos. 2020. Image Segmentation Using Deep Learning: A Survey. http://arxiv.org/abs/2001.05566 arXiv:2001.05566 [cs].

Mohamed Negm, Ahmed Mohamed Housseini, Mohamed Abdelfatah, and Alshimaa Asran. 2018. Relation between migraine pattern and white matter hyperintensities in brain magnetic resonance imaging. *The Egyptian Journal of Neurology, Psychiatry and Neurosurgery* 54, 1 (Dec. 2018), 24. https://doi.org/10.1186/s41983-018-0027-x

Eric D Ragan, Doug A Bowman, and Karl J Huber. 2012. Supporting cognitive processing with spatial information presentations in virtual environments. *Virtual Reality* 16 (2012), 301–314.

Sreelakshmi S, Malu G, Elizabeth Sherly, and Robert Mathew. 2023. M-Net: An encoder-decoder architecture for medical image analysis using ensemble learning. *Results in Engineering* 17 (March 2023), 100927. https://doi.org/10.1016/j.rineng.2023.100927

Alexander Sergeev and Mike Del Balso. 2018. Horovod: fast and easy distributed deep learning in TensorFlow. http://arxiv.org/abs/1802.05799 arXiv:1802.05799 [cs, stat].

Karen Simonyan and Andrew Zisserman. 2014. Very Deep Convolutional Networks for Large-Scale Image Recognition. (2014). https://doi.org/10.48550/ARXIV.1409.1556 Publisher: [object Object] Version Number: 6.

Youssef S Tanagho, Gerald L Andriole, Alethea G Paradis, Kerry M Madison, Gurdarshan S Sandhu, J Esteban Varela, and Brian M Benway. 2012. 2D versus 3D visualization: impact on laparoscopic proficiency using the fundamentals of laparoscopic surgery skill set. *Journal of Laparoendoscopic & Advanced Surgical Techniques* 22, 9 (2012), 865–870.

Unity Technologies. 2023. Unity 2022.3.7. https://unity.com/releases/editor/whats-new/2022.3.7.

Meral A. Tubi, Franklin W. Feingold, Deydeep Kothapalli, Evan T. Hare, Kevin S. King, Paul M. Thompson, and Meredith N. Braskie. 2020. White matter hyperintensities and their relationship to cognition: Effects of segmentation algorithm. *NeuroImage* 206 (Feb. 2020), 116327. https://doi.org/10.1016/j.neuroimage.2019.116327

Amin Ul Haq, Jianping Li, Muhammad Hammad Memon, Jalaluddin Khan, Salah Ud Din, Ijaz Ahad, Ruinan Sun, and Zhilong Lai. 2018. Comparative Analysis of the Classification Performance of Machine Learning Classifiers and Deep Neural Network Classifier for Prediction of Parkinson Disease. In *2018 15th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*. IEEE, Chengdu, China, 101–106. https://doi.org/10.1109/ICCWAMTIP.2018.8632613

A. Victor Ikechukwu, S. Murali, R. Deepu, and R.C. Shivamurthy. 2021. ResNet-50 vs VGG-19 vs training from scratch: A comparative analysis of the segmentation and classification of Pneumonia from chest X-ray images. *Global Transitions Proceedings* 2, 2 (Nov. 2021), 375–381. https://doi.org/10.1016/j.gltp.2021.08.027

Nan Wu, Jason Phang, Jungkyu Park, Yiqiu Shen, Zhe Huang, Masha Zorin, Stanislaw Jastrzebski, Thibault Fevry, Joe Katsnelson, Eric Kim, Stacey Wolfson, Ujas Parikh, Sushma Gaddam, Leng Leng Young Lin, Kara Ho, Joshua D. Weinstein, Beatriu Reig, Yiming Gao, Hildegard Toth, Kristine Pysarenko, Alana Lewin, Jiyon Lee, Krystal Airola, Eralda Mema, Stephanie Chung, Esther Hwang, Naziya Samreen, S. Gene Kim, Laura Heacock, Linda Moy, Kyunghyun Cho, and Krzysztof J. Geras. 2020. Deep Neural Networks Improve Radiologists' Performance in Breast Cancer Screening. *IEEE Transactions on Medical Imaging* 39, 4 (April 2020), 1184–1194. https://doi.org/10.1109/TMI.2019.2945514

Wanni Xu, You-Lei Fu, and Dongmei Zhu. 2023. ResNet and its application to medical image processing: Research progress and challenges. *Computer Methods and Programs in Biomedicine* 240 (Oct. 2023), 107660. https://doi.org/10.1016/j.cmpb.2023.107660

Matei Zaharia, Reynold S Xin, Patrick Wendell, Tathagata Das, Michael Armbrust, Ankur Dave, Xiangrui Meng, Josh Rosen, Shivaram Venkataraman, Michael J Franklin, et al. 2016. Apache spark: a unified engine for big data processing. *Commun. ACM* 59, 11 (2016), 56–65.