Interactive Visualization of Deep Learning Models in an Immersive Environment

Hikaru Nagasaka
Advanced Technology Lab, Recruit Co., Ltd.
Tokyo, Japan
hikaru_nagasaka@r.recruit.co.jp

Motoya Izuhara
Advanced Technology Lab, Recruit Co., Ltd.
Tokyo, Japan
izuhara@r.recruit.co.jp

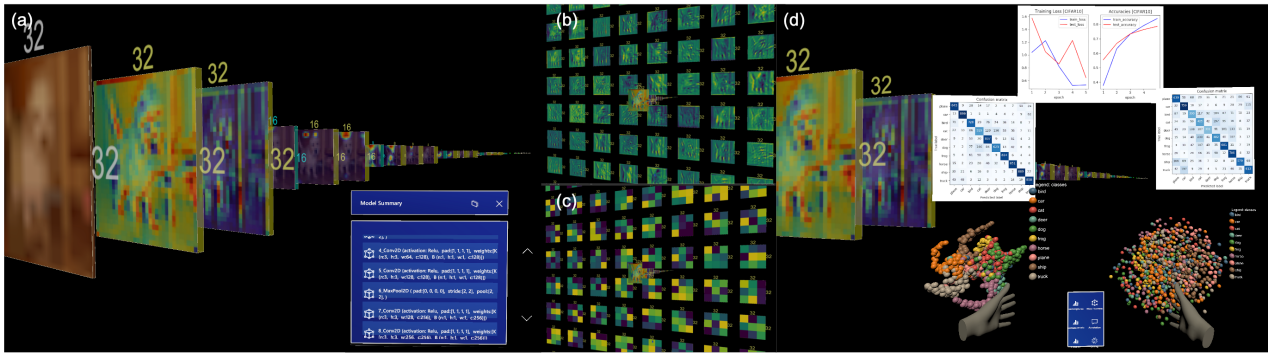

Figure 1: Overview of the developed system: (a) Automatic visualization of a deep learning model structure as a 3D computational graph. (b)–(c) Interactive visualization of propagation data (such as Grad-CAM, feature maps, and filters) for each layer. (d) Free placement of multiple 2D and 3D views in an immersive workspace to evaluate model performance data.

ABSTRACT

The development of deep learning (DL) models has been prevalent among software engineers. However, it is difficult for non-experts to analyze and understand their behavior. Hence, we propose an interactive visualization system of DL models in an immersive environment. Because an immersive environment offers unlimited displays and visualization of high-dimensional data, it enables a comprehensive analysis on data propagations through the layers, and compares the multiple performance metrics. In this research, we implemented a prototype system, demonstrated it to machine learning engineers, and discussed the future benefits of visualizing DL models in an immersive environment. Accordingly, our concept received positive feedback; however, we inferred that most of the engineers consider the visualization technology as a unique introduction to the immersive environment.

CCS CONCEPTS

• **Human-centered computing** → **Visualization systems and tools**; **Virtual reality**.

KEYWORDS

immersive analytics, deep learning

VRST '21, December 8–10, 2021, Osaka, Japan

ACM ISBN 978-1-4503-9092-7/21/12.
https://doi.org/10.1145/3489849.3489956

ACM Reference Format:
Hikaru Nagasaka and Motoya Izuhara. 2021. Interactive Visualization of Deep Learning Models in an Immersive Environment. In *27th ACM Symposium on Virtual Reality Software and Technology (VRST '21), December 8–10, 2021, Osaka, Japan.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3489849.3489956

1 INTRODUCTION

Recently, several software engineers are developing various DL models. In general, after training a DL model, they check the structure and performance of the model to optimize its parameters. For example, software engineers analyze the data propagation through each layer, and evaluate the performance (such as Accuracy, Precision, Recall) of the DL model updated after every epoch. As DL models become more complex, it becomes difficult to check all the layers and the test results in a desktop environment, where the display area and number of displays is limited. Therefore, experts address these by selecting a few essential data; however, this process is difficult for non-experts.

To address this challenge, we propose an interactive visualization system for DL models in an immersive environment, which can visualize high-dimensional data without being limited by the size or number of displays. It elucidates data propagation through each layer, and allows multiple DL models and performance metrics to be freely positioned and compared. Related studies have proposed easier methods for building and visualizing the components of a neural network in a VR workspace [5, 10]. However, they could only visualize a simple DL model, limiting the number of layers in the workspace. In this study, we focus on visualizing the structure of the developed complex DL models and comparing the performance data of each DL model without display limitations.

2 SYSTEM DESIGN

2.1 Visualization of Deep Learning Model

Figure 1 presents an overview of the proposed system. The system was implemented in Unity 2019.3 and deployed on Oculus Quest 2 as a virtual reality (VR) head-mounted display. First, the system loads the trained DL models in an Open Neural Network Exchange (ONNX)[1] format using Unity Barracuda[2]. Several famous DL model development frameworks (Keras, TensorFlow, PyTorch, etc.) support the ONNX export format. Next, the system converts the DL models into 2D and 3D visualization data in an immersive workspace. It builds on the existing 3D visualization methods provided by TensorSpace.js[3] and NeuralVis [11]. If the model is a convolutional neural network, it also offers conventional 2D visualization techniques for filter and feature map images such as [1] and Grad-CAM [6]. These visualizations provide an overview of the structural architecture for a trained model and a detailed analysis of the activation status on a specified layer (Figure 1 (a)–(c)).

The system also provides a multi-dimensional visualization for DL model performance metrics such as loss and accuracy, images, text, and audio data. We adopted Mixed Reality Toolkit[4] and DXR [7], which are toolkits for interactive 2D or 3D visualizations in an immersive workspace. Users can freely place and compare performance data, as illustrated Figure 1 (d).

2.2 Example Scenario

We illustrate the advantages of the proposed, citing a representative scenario. In this scenario, a user tries to develop a DL model for CIFAR-10 [3]. First, the user tunes hyper-parameters and trains on three prevalent DL models (MobileNet [2], VGG-16 [8], and PyTorch Tutorial[5]), using a PC. After every work, the user exports the trained DL models as ONNX format and training data (e.g., learning curve, confusion matrix as accuracy, and t-SNE [9] as clustering) as 2D and 3D data. Next, the user views and analyzes the inference results of the DL models and the data in the VR space using our system, as presented in Figure 2.

3 EXPLORATORY INTERVIEW

We conducted an exploratory interview on the developed system with five machine learning engineers (5 males) who know a lot about DL techniques, but not familiar with DL model visualization. We demonstrated the example scenario in Section 2.2 using existing DL model visualization techniques (TensorSpace and [1, 11]) and our developed system (Figure 2), and then discussed the future benefits of visualizing DL models in an immersive environment. Participants' feedbacks are below.

- P1: It is interesting to observe how the DL model works. In other cases, I believe it would be effective to arrange images in VR space and annotate them by tapping.
- P2: I like the Grad-CAM output. I also want to try this for the Transformer (a kind of DL model). I want to see the head is applying attention, and the direction of its attention.

[1]https://onnx.ai/
[2]http://docs.unity3d.com/Packages/com.unity.barracuda@1.3/manual/index.html
[3]https://tensorspace.org/
[4]https://github.com/microsoft/MixedRealityToolkit-Unity
[5]https://pytorch.org/tutorials/beginner/blitz/cifar10_tutorial.html

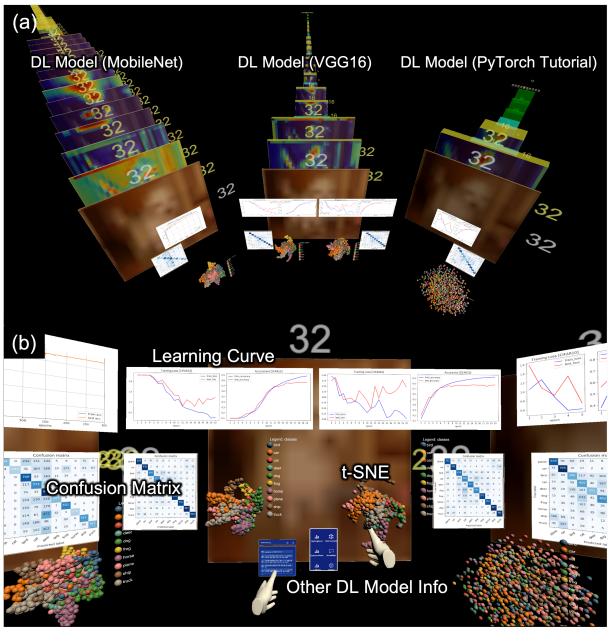

Figure 2: Visualization results of the considered scenario. (a) 2D and 3D views of the DL model structure and training data. (b) The user can position these data freely then ease to review the structure and metrics of each model.

- P3: If this system can interactively visualize 3D data in an immersive environment, 4D data visualization is also desirable.
- P4: I hope this system elucidates why the developed DL model mis-recognizes dogs and cats. However, it is crucial to observe the learning process; hence, I believe that the ability to visualize the difference data of each epoch is an advantage.
- P5: I hypothesize that you need to clarify the appropriate use cases for this system. For example, the design approach of the DL techniques required by the target users, including their visualization goals.

4 CONCLUSION AND FUTURE WORK

In this study, we proposed a DL model visualization system in an immersive environment, which can comprehensively analyze the data propagation through the layers, and compare the performance metrics of each model. We implemented a prototype system, designed an example scenario using the system, demonstrated it to machine learning engineers, and discussed the future benefits of visualizing DL models in an immersive environment. Consequently, our concept received positive feedbacks. However, we deduced that most of them consider the visualization technology a unique introduction to the immersive environment and experience design.

Future work will improve our system design and evaluate its usability by comparing its performance with an existing desktop-based system (e.g., TensorBoard [4] and TensorSpace). Furthermore, we will explore the benefits of an immersive visualization system in analyzing DL models. Finally, our concept is to develop hybrid systems in which our system does not replace existing tools, but attempts to complement them instead.

REFERENCES

[1] Adam W Harley. 2015. An interactive node-link visualization of convolutional neural networks. In *International Symposium on Visual Computing*. Springer, Cham, Switzerland, AG, 867–877. https://doi.org/10.1007/978-3-319-27857-5_77

[2] 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. *CoRR* abs/1704.04861 (2017). arXiv:1704.04861 http://arxiv.org/abs/1704.04861

[3] Alex Krizhevsky et al. 2009. Learning Multiple Layers of Features from Tiny Images. (2009). Retrieved October 8, 2021 from https://www.cs.toronto.edu/~kriz/cifar.html

[4] D Mané et al. 2015. TensorBoard: TensorFlow's visualization toolkit. Retrieved October 8, 2021 from https://tensorboard.dev/

[5] Nadine Meissler, Annika Wohlan, Nico Hochgeschwender, and Andreas Schreiber. 2019. Using Visualization of Convolutional Neural Networks in Virtual Reality for Machine Learning Newcomers. In *2019 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*. IEEE, 152–1526. https://doi.org/10.1109/AIVR46125.2019.00031

[6] Ramprasaath R. Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra. 2017. Grad-CAM: Visual Explanations From Deep Networks via Gradient-Based Localization. In *2017 IEEE International Conference on Computer Vision (ICCV)*. IEEE. https://doi.org/10.1109/ICCV.2017.74

[7] Ronell Sicat, Jiabao Li, JunYoung Choi, Maxime Cordeil, Won-Ki Jeong, Benjamin Bach, and Hanspeter Pfister. 2018. DXR: A Toolkit for Building Immersive Data Visualizations. *IEEE transactions on visualization and computer graphics* 25, 1 (2018), 715–725. https://doi.org/10.1109/TVCG.2018.2865152

[8] Karen Simonyan and Andrew Zisserman. 2015. Very Deep Convolutional Networks for Large-Scale Image Recognition. In *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, Yoshua Bengio and Yann LeCun (Eds.). http://arxiv.org/abs/1409.1556

[9] Laurens van der Maaten and Geoffrey Hinton. 2008. Visualizing Data using t-SNE. *Journal of Machine Learning Research* 9, 86 (2008), 2579–2605. http://jmlr.org/papers/v9/vandermaaten08a.html

[10] Kevin C. VanHorn, Meyer Zinn, and Murat Can Cobanoglu. 2019. Deep Learning Development Environment in Virtual Reality. *CoRR* abs/1906.05925 (2019). arXiv:1906.05925 http://arxiv.org/abs/1906.05925

[11] Xufan Zhang, Ziyue Yin, Yang Feng, Qingkai Shi, Jia Liu, and Zhenyu Chen. 2019. NeuralVis: Visualizing and Interpreting Deep Learning Models. In *2019 34th IEEE/ACM International Conference on Automated Software Engineering (ASE)*. IEEE, 1106–1109. https://doi.org/10.1109/ASE.2019.00113