The Journal of Open Source Software

Albumentations3D: A Python Package for 3D Medical Imaging Augmentation

J. McIntosh [1¶], Qian Cao [1], Berkman Sahiner [1], Nicholas Petrick [1], and M. Mehdi Farhangi [1]

1 Division of Imaging, Diagnostics, and Software Reliability, CDRH, U.S. Food and Drug Administration, Silver Spring, MD 20993, USA ¶ Corresponding author

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Editor: Open Journals

Reviewers:

- @openjournals

Submitted: 01 January 1970
Published: unpublished

Summary

Albumentations3D is a Python package based on the popular image augmentation library Albumentations (Buslaev et al., 2020), with specific enhancements for working with volumetric 3D images, such as CT scans and other volumetric imaging. This package provides a collection of powerful and efficient augmentation techniques that can be seamlessly integrated into a machine-learning pipeline to augment volumetric images. The package was designed to incorporate the metadata available in dicom headers, allowing users to create transformations that are consistent with an image's acquisition parameters. Albumentations3D extends the success of the Albumentations library for two-dimensional (2D) image augmentation to the realm of volumetric three-dimensional (3D) images, offering a comprehensive set of transformations and augmentations, ranging from pixel-level intensity transformations to spatial transformations, all designed and optimized for 3D data.

Statement of need

The recent advancements in machine learning have significantly improved the performance of AI models across various domains and applications. However, the success of machine learning models still largely relies on a significant amount of labeled and annotated training data. This is a particular limitation in medical imaging applications where imaging data annotation is expensive, time consuming and require clinician expertise to establish a reliable reference standard. Albumentations3D offers contribution in this regard by providing data augmentations that are consistent with pre-existing imaging data annotations available in form of object boundary masks, bounding boxes, or keypoints.

Besides different applications for volumetric images, Albumentation3D offers a unique advantage through the integration of physics-based augmentations that leverage metadata from DICOM files in CT imaging; during the reconstruction process of CT images, different manufactures employ different reconstruction kernels, leading to distinct noise textures in the resulting images, often characterized by the Noise Power Spectrum (Solomon et al., 2012). Albumentation3D utilizes the NPS profiles from different reconstruction kernels to generate noisy images that are consistent with the imaging acquisition and reconstruction parameters. To illustrate, we computed the Noise Power Spectrum (NPS) obtained from the noise insertion transformations offered by Albumentations3D and compared it against the outcome of white Gaussian noise, which is a common noise insertion method for data augmentation (e.g., MONAI (Cardoso et al., 2022) and Volumentations (Solovyev et al., 2022)).

For this experiment, we utilized a water phantom (*Phantom Testing: CT*, n.d.) scanned by SIEMENS using an imaging acquisition protocol with a B30f reconstruction kernel, 120 kVp, 0.48 mm pixel spacing, and an exposure of 240 mAs. The 2D NPS was estimated using the method proposed by Solomon et al. (Solomon et al., 2012); from the water phantom, we extracted a uniform volumetric region of 128*128*70 dimensions and computed the 2D NPS of each slice by taking the square root of the Fourier transform:

$$NPS(u,v) = \frac{d_x d_y}{N_x N_y}. \mid F[I(x,y) - P(x,y)] \mid^2 \tag{1}$$

In equation Equation 1, u and v represent spatial frequency (mm^{-1}) in the x and y directions, respectively, d_x and d_y are pixel size in millimeter space, N_x and N_y are the number of pixels in the x and y directions within the selected ROI (i.e., 128*128 px), F denotes the 2D Fourier transform, $I(x,y)$ is the pixel density in Hounsfield Unit at position (x,y) and P is the second order polynomial fit of I. Each estimated 2D NPS within the volumetric ROI are normalized by its integral across all frequencies and converted into a one-dimensional radial representation, $r^2 = \sqrt{u^2 + v^2}$. The final nNPS(r) was obtained by averaging 2D radial NPS across 70 slices within the volumetric image.

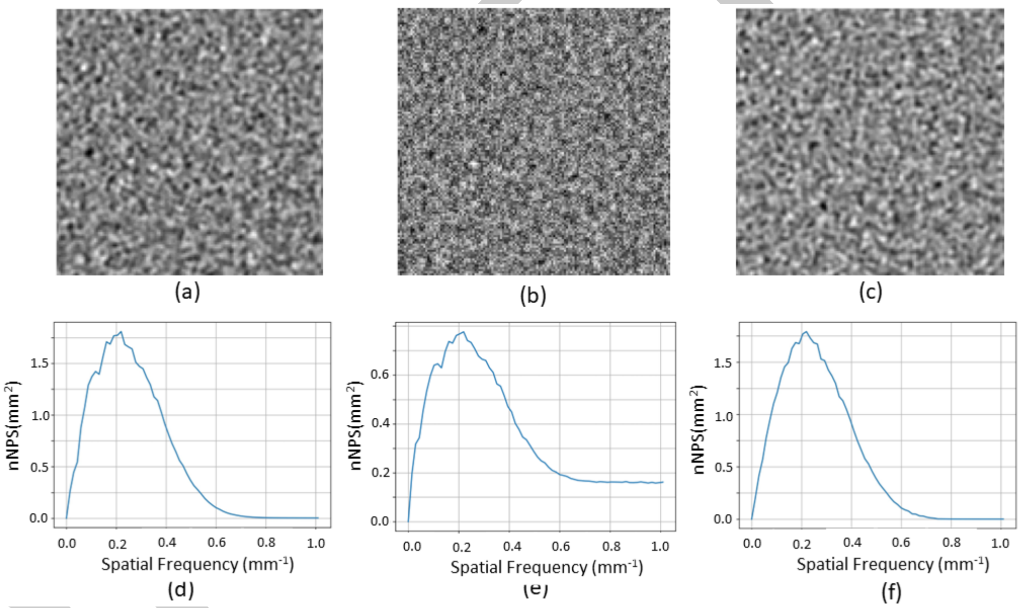

Figure 1: (a) Water phantom scanned by SIEMENS with imaging parameters set at 120 kVp, 0.48 mm pixel spacing, and an exposure of 240 mAs, and reconstructed using the B30f kernel. The noise magnitude is computed as 9.46. (b) Increasing the noise magnitude by adding white Gaussian noise, resulting in a noisy image with a magnitude of 16.01. (c) Increasing the noise magnitude using Albumentations3D, resulting in a noise image with a magnitude of 16.08. Figures (d), (e), and (f) illustrate the benefits of Albumentations3D in offering augmentations consistent with the imaging parameters, demonstrated in terms of the Noise Power Spectrum (NPS) of the augmented image.

Figure 1(a) illustrates a slice of the noise region extracted from the uniform water phantom along with the corresponding estimation of the normalized Noise Power Spectrum (nNPS). Figure 2(b) illustrates the outcome of noise insertion using white Gaussian noise, a common approach for noise insertion (Cardoso et al., 2022), demonstrating the correlation between adjacent pixel values affected by the inserted Gaussian noise. On the right, the outcome of noise insertion by Albumentation3D is illustrated. Despite the increased magnitude of the noise, the correlation among adjacent pixels remained mostly unchanged, as evident in the normalized Noise Power Spectrum (nNPS) shown at the bottom of this figure.

Acknowledgements

Jacob McIntosh was supported by an appointment to the Research Participation Program at the Center for Devices and Radiological Health administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and the U.S. Food and Drug Administration. The mention of softwares, commercial products, their sources, or their use in connection with material reported herein is not to be construed as either an actual or implied endorsement of such products by the Department of Health and Human Services.

References

Buslaev, A., Iglovikov, V. I., Khvedchenya, E., Parinov, A., Druzhinin, M., & Kalinin, A. A. (2020). Albumentations: Fast and flexible image augmentations. *Information*, *11*(2). https://doi.org/10.3390/info11020125

Cardoso, M. J., Li, W., Brown, R., Ma, N., Kerfoot, E., Wang, Y., Murrey, B., Myronenko, A., Zhao, C., Yang, D., & others. (2022). Monai: An open-source framework for deep learning in healthcare. *arXiv Preprint arXiv:2211.02701*.

Phantom Testing: CT. (n.d.). https://accreditationsupport.acr.org/support/solutions/articles/11000056197-phantom-testing-ct-revised-11-9-2022-.

Solomon, J. B., Christianson, O., & Samei, E. (2012). Quantitative comparison of noise texture across CT scanners from different manufacturers. *Med. Phys.*, *39*(10), 6048–6055.

Solovyev, R., Kalinin, A. A., & Gabruseva, T. (2022). 3D convolutional neural networks for stalled brain capillary detection. *Computers in Biology and Medicine*, *141*, 105089. https://doi.org/10.1016/j.compbiomed.2021.105089