JOSS
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

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 medical images, such as CT scans. This package provides a collection of powerful and efficient augmentation techniques that can be seamlessly integrated into a machine-learning pipeline to augment volumetric medical 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 for specific CT systems and reconstruction kernels. Albumentations3D extends the success of the Albumentations library for two-dimensional (2D) image augmentation to the realm of 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 deep learning models across various domains and applications. However, the success of these models still largely rely on a significant amount of labeled and annotated training data. This is a particular limitation in medical imaging applications where imaging data annotation and the establishment of a reliable reference standard is expensive, time consuming, and typically requires clinician expertise. Albumentations3D supports the efficient use of available labeled and annotated data by providing augmentations that are consistent with preexisting imaging data annotations available in the form of masks, bounding boxes, and keypoints.

Albumentations3D provides a variety of transformations designed to enhance the diversity of 3D image data. A subset of these transformations includes blurring techniques, such as median blur and Gaussian blur, that must utilize convolution to achieve the desired output. Notably, Albumentations3D offers a unique capability by allowing users to choose between a 2D by-slice convolution or a full 3D convolution when applying these blurring augmentations in a pipeline. This flexibility caters to a variety of use cases and ensures that the blurring effects can be consistent across an entire 3D volume, preserving the spatial relationships between slices.

Additionally, Albumentations3D enables users to construct custom augmentation pipelines with control over each individual transformation. Mimicking the Albumentations (Buslaev et al., 2020) package, Albumentations3D utilizes the concept of augmentation probabilities. This feature allows users to assign a probability of occurrence to each transformation within a pipeline, enabling selective and stochastic augmentation of an image dataset. By specifying different probabilities for various transformations, users can create diverse and balanced augmented datasets that reflect real-world variability. This level of control ensures that specific

⁴² augmentations, whether it be intensity adjustments or geometric transformations, are applied
⁴³ with user-defined frequencies, therefore tailoring the augmentation process to the characteristics
⁴⁴ and use case of the dataset.

⁴⁵ Besides different incorporations for volumetric images, Albumentation3D offers a unique
⁴⁶ advantage through the addition of physics-based augmentations that leverage metadata from
⁴⁷ DICOM files in CT imaging. During the reconstruction process of CT images, different
⁴⁸ manufacturers employ different reconstruction kernels, leading to distinct noise textures in the
⁴⁹ resulting images often characterized by the Noise Power Spectrum (Solomon et al., 2012).
⁵⁰ Albumentation3D utilizes the NPS profiles from different reconstruction kernels to generate
⁵¹ noisy images that are consistent with the imaging acquisition and reconstruction parameters
⁵² (Tward & Siewerdsen, 2008). To illustrate, we computed the Noise Power Spectrum (NPS)
⁵³ obtained from the noise insertion transformations offered by Albumentations3D and compared
⁵⁴ it against simple white Gaussian noise, which is a common noise insertion method for data
⁵⁵ augmentation (e.g., MONAI (Cardoso et al., 2022) and Volumentations (Solovyev et al.,
⁵⁶ 2022)).

⁵⁷ For this experiment, we utilized a water phantom (*Phantom Testing: CT*, n.d.) scanned by
⁵⁸ SIEMENS with an exposure of 240 mAs, 120 kVp, and reconstructed with a B30f reconstruction
⁵⁹ kernel at 0.48 mm in-plane pixel spacing. The 2D NPS was estimated using the method
⁶⁰ proposed by Solomon et al. (Solomon et al., 2012); from the water phantom, we extracted a
⁶¹ uniform volumetric region of $128 \times 128 \times 70$ dimensions and computed the 2D NPS of each
⁶² slice by taking the square of the Fourier transform:

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

⁶³ In Equation 1, u and v represent spatial frequency (mm^{-1}) in the x and y directions,
⁶⁴ respectively, d_x and d_y are pixel size in millimeter space, N_x and N_y are the number of pixels
⁶⁵ in the x and y directions within the selected ROI (i.e. 128×128 px), F denotes the 2D
⁶⁶ Fourier transform, $I(x,y)$ is the pixel density of the uniform region in Hounsfield Unit at
⁶⁷ position (x,y) and P is the second order polynomial fit of I. Each estimated 2D NPS within
⁶⁸ the volumetric ROI is normalized by its integral across all frequencies and converted into a
⁶⁹ one-dimensional radial representation, $r^2 = \sqrt{u^2 + v^2}$. The final normalized NPS (nNPS) was
⁷⁰ obtained by averaging the radial NPS curves across 70 slices within the volumetric image.
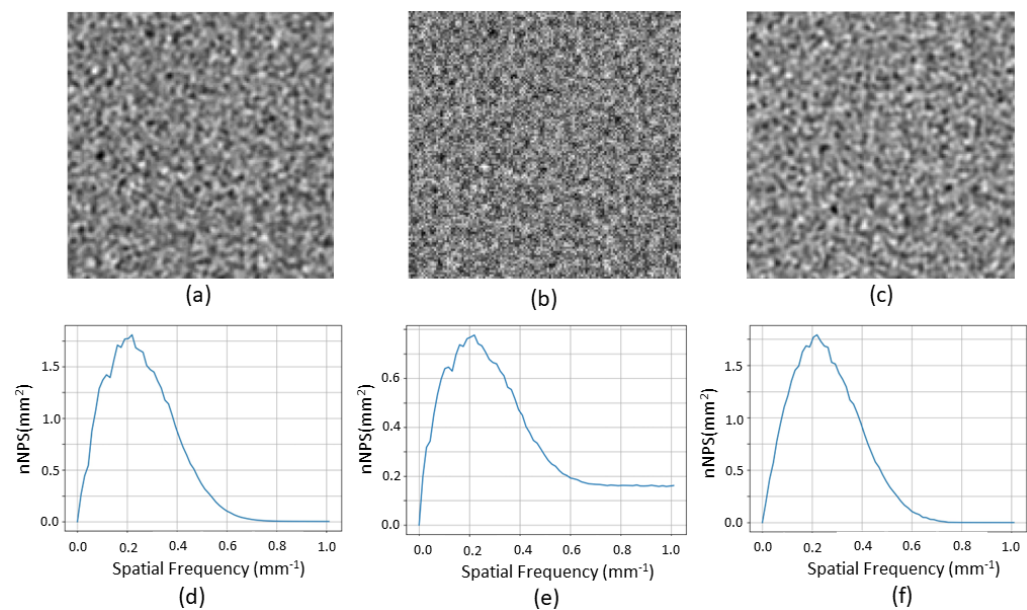
Figure 1: (a) Water phantom scanned by SIEMENS with imaging parameters set at 120 kVp, 0.48 mm pixel spacing, and an exposure of 240 mAs, and reconstructed using the B30f kernel. The noise magnitude is computed as 9.46. (b) Increasing the noise magnitude by adding white Gaussian noise, resulting in a 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 lack of correlation between adjacent pixels. 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 essentially unchanged, as evident in the normalized Noise Power Spectrum (nNPS) shown at the bottom of this figure. This shows that an increased CT exposure can be simulated through our augmentation process for specific CT imaging systems and kernels.

Acknowledgements

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

Tward, D. J., & Siewerdsen, J. H. (2008). Cascaded systems analysis of the 3D noise transfer characteristics of flat-panel cone-beam CT. *Med. Phys.*, *35*(12), 5510–5529.