UDC 004.89

NEW ERA IN IMAGE AUGMENTATION – NEURAL RADIANCE FIELDS

Shashko Oleksii
Scientific supervisor – Assoc. Prof., Ph. D.,
Assoc. Prof. AI Dep. Deineko A.
Kharkiv National University of Radio Electronics, AI Department,
Kharkiv, Ukraine
tel. +38(095) 080-69-50, e-mail: oleksii.shashko@nure.ua

In this article, a new method of image augmentation using Neural Radiance Fields is being considered. NeRF has significant advantages in generating photo-realistic images with high resolution and fidelity, and it can be used for a wide range of applications, including 3D rendering, virtual or augmented reality, and robotics. Despite its computational cost and limitations in some types of scenes, NeRFs are becoming more and more popular in many fields of computer vision, with ongoing research to improve its limitations. Also, this work describes a practical example of how it can preprocess and augment an existing dataset to use it for a specific aim.

Image augmentation is an essential technique in the field of computer vision, used to generate new training examples by applying various transformations to existing images. The goal of image augmentation is to improve the robustness and generalization capabilities of deep learning models, especially in cases where training data is limited. Recently, a new technique called Neural Radiance Fields (NeRF) has emerged, which promises to revolutionize the field of image augmentation. NeRF is a deep learning-based method that can represent the 3D structure of a scene by learning a continuous function that maps 3D coordinates to radiance values, which determine the color and intensity of light passing through a particular point in space.

The key idea behind NeRF is to represent the scene as a collection of small volume elements called voxels, each of which contains a radiance value. By training a deep neural network to predict these radiance values from the 3D coordinates of the voxel, NeRF can generate high-quality synthetic images of the scene from any viewpoint. The key innovation in NeRF is that it uses a positional encoding scheme to represent the 3D coordinates of each voxel as a high-dimensional vector. This allows the neural network to learn a continuous function that can map any 3D coordinate to its corresponding radiance value, without the need for explicit 3D reconstruction or surface representation.

One of the most significant advantages of NeRF is its ability to generate photo-realistic images with high resolution and fidelity. This is because NeRF can capture the complex interactions between light and surfaces, including reflections, refractions, and occlusions, which are often challenging to model with traditional methods.

Moreover, NeRF can be used for a wide range of applications, including 3D rendering, virtual or augmented reality, and robotics. For example, NeRF can be used to create realistic 3D models of objects and scenes from a small set of images, which can be used for augmented reality applications. NeRF can also be used to train robots to navigate in complex environments by generating synthetic images that mimic real-world scenarios.

One good example of how NeRF can significantly increase the amount of training data for the dataset is task generation depth from stereo (DFS). Models in this field should predict depth maps using stereo pairs. But the main problem here is the high-quality dataset with needed camera parameters (such as focal length, baseline between cameras, rectification, etc.). Here, NeRF can help resolve obstacles – from simple indoor image datasets, this technology can create neural volume representations and generate a large variety of novel views with needed camera parameters and even create ground truth depth maps.

Despite its numerous advantages, NeRF has some limitations that need to be addressed. One of the main limitations of NeRF is its computational cost, as it requires a large amount of training data and computational resources to train the neural network. Additionally, NeRF may not be suitable for all types of scenes, especially those with complex geometries or textures.

Despite these limitations, NeRFs are becoming more and more popular in many fields of computer vision. Modern research shows how we can improve this technology to avoid its limitations, such as training NeRF in sparse training data dimensionality that can give good results on a small number of images, or using hashing to increase training and inference speed.

In conclusion, Neural Radiance Fields represent a new era in image augmentation, offering a powerful and versatile tool for generating high-quality synthetic images of scenes and objects. As research in this area continues, we can expect to see even more exciting applications of NeRF in computer vision, robotics, and other fields.

References:

1. Li, T.-M., Aittala, M., Durand, F., & Lehtinen, J. (2019). Differentiable Monte Carlo ray tracing through edge sampling. ACM Transactions on Graphics, 37(6), 1–11. https://doi.org/10.1145/3272127.3275109

2. Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., & Ng, R. (2020). NeRF: representing scenes as neural radiance fields for view synthesis. У Computer vision – ECCV 2020 (с. 405–421). Springer International Publishing. https://doi.org/10.1007/978-3-030-58452-8_24

3. Curless, B., & Levoy, M. (1996). A volumetric method for building complex models from range images. У The 23rd annual conference. ACM Press. https://doi.org/10.1145/237170.237269