DiagnoNeRF: Neural Radiance Fields for Reconstructing 3D-Aware CT-Projections from 2D X-ray

by

MEHADI HASAN

999

Dissertation in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

IN

COMPUTER SCIENCE AND ENGINEERING

Faculty of Mathematical and Physical Sciences

Jahangirnagar University

Dhaka, Bangladesh

October, 2023

Abstract

The diagnosis of various illnesses is frequently performed in the field of clinical medicine using computed tomography (CT), a reliable medical imaging technique. Multidetector CT imaging Technological advancements have made it possible for additional functions, such as the creation of thin-slice multiplanar body cross-section imaging and 3D reconstructions. However, Patients will be exposed to a large dose in this situation of ionizing radiation. Ionizing radiation in excess can result in impacts on the body that are predictable and detrimental. This document proposes a Deep Learning model that may be trained to reconstruct CT projections from a few or simply a single X-ray. This is based on an innovative architecture that develops from Neural Radiance Field (**NeRF**) and learns a continuous representation of CT scans by separating the volumetric depth and shape of surface and internal anatomical structures from 2D images. Our model will be trained on data sets for the chest and arm, and we show examples of qualitative and quantitative high-fidelity renderings and contrast our method with other recent radiance field-based techniques.

Clinical significance: Our model will be able to deduce the anatomical 3D structure from a few or a single-view X-ray, suggesting possibilities for less ionizing radiation exposure during the imaging process in the future.

Chapter 1

Introduction

A deep learning technique called a Neural Radiance Field (NeRF) is used to rebuild a scene's three dimensions from its two dimensions. It is a graphics primitive that can be enhanced to create a 3D scene from a collection of 2D photos. The NeRF model can learn the geometry of the scene, the camera postures, and the reflectance characteristics of the objects in the scene, enabling it to display fresh perspectives on the picture from different angles. A group from UC Berkeley, Google Research, and UC San Diego first presented the technique in 2020.

Neural Radiance Fields (NeRF) are able to deconvolute the form and volumetric depth of external and internal anatomical structures from 2D pictures to learn a continuous representation of CT scans, which may then be used to rebuild 3D images from 2D x-ray images. This is based on an innovative architecture that develops from neuronal radiance fields and learns a continuous representation of CT scans by separating the volumetric form and depth of surface and internal anatomical structures from 2D images.

Chapter 2

Literature Review

Specifying values for the location of the patient, the imaging source, and the detector are part of the workflow for 3D medical imaging, which frequently entails merging numerous 2D slices from **CT** or Magnetic Resonance Imaging (**MRI**). Hundreds of X-ray projections with a narrow slice thickness are necessary for a **CT** 3D depiction to be accurate and of high quality [1]. Additionally, depending on the test, this procedure subjects patients to higher levels of ionizing radiation than standard X-rays and necessitates that they stay immobile for up to an hour [2]. Radiologists would have access to continuous 3D representations that would provide views of every point in the imaged interior anatomy. Even though these representations are helpful, using them in CT poses some practical difficulties because of the increased radiation dose, angle-dependent structures, and time requirements [3].

Analytical and iterative methods on pre-existing input data were employed in earlier efforts to medical image reconstruction. However, they frequently run across discrepancies between the mathematical model and the physical characteristics of the imaging system. A more recent reformulation for estimating a 3D volumetric representation from photographs is the Neural Radiance Fields (**NeRF**) model [4]. Such representations incorporate the scene's light field and density into the neural network's parameters. By using volume rendering to create new views from point samples along cast rays, the neural network learns to synthesis new ones. However, these depictions are frequently recorded in predetermined environments [5]. The scene is first captured quickly by a group of fixed cameras. Second, the entire scene's material is static, whereas real visuals frequently require masking. Due to these limitations, **NeRF** cannot be directly used to the medical field, where the imaging system is very different from that of traditional cameras and the images are taken over a long period of time, making it difficult for the patient to remain still. In order to overcome these difficulties, we suggest **DiagnoNeRF**, a model that uses Generative Radiance Fields (**GRAF**) [6] in the medical field to produce CT projections given a few or even a single-view X-ray. Our method, which does not require 3D supervision, not only generates realistic images but also collects data in a variety of ways, allowing us to show continuously how the attenuation and volumetric depth of anatomical structures change with angle.

Chapter 3

Proposed System

3.1 Data Collection

Instead of acquiring paired X-rays and accompanying CT reconstructions, which would subject patients to further radiation exposure, we will create DRRs to train our algorithms. Aside from erasing patient data, DRR production also gives users choice over the capture ranges and resolutions. Through the use of 10 CT knee images and 25 CT chest scans, we will create DRRs. A wide range of patients will be covered by these scans at various contrast levels, displaying both normal and pathological anatomy. Assuming that the imaging panel and radiation source revolve about the vertical axis, 72 DRRs will be produced for each item, with a DRR of 128128 resolution produced every five degrees. We will use the whole set of 72 DRRs (or one-fifth of all views during a full 360° vertical rotation) for training purposes and let the model display the remaining views. Institutional Review Board permission won't be needed for our work because it won't include using either human or animal subjects in experiments.

3.2 Model

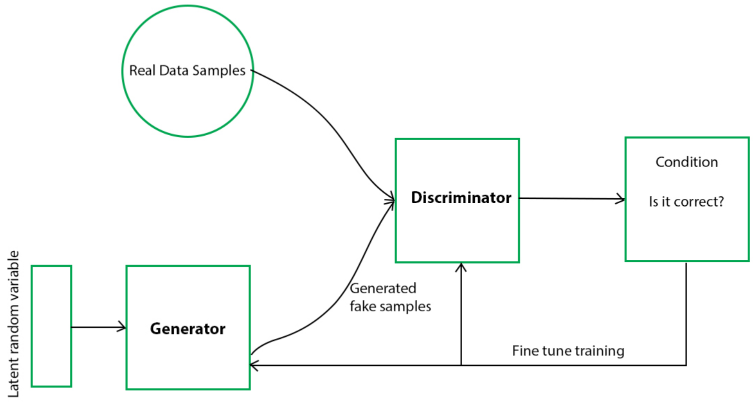

Figure 3.1: Image generation using Generative Adversarial Networks (**GAN**) diagram

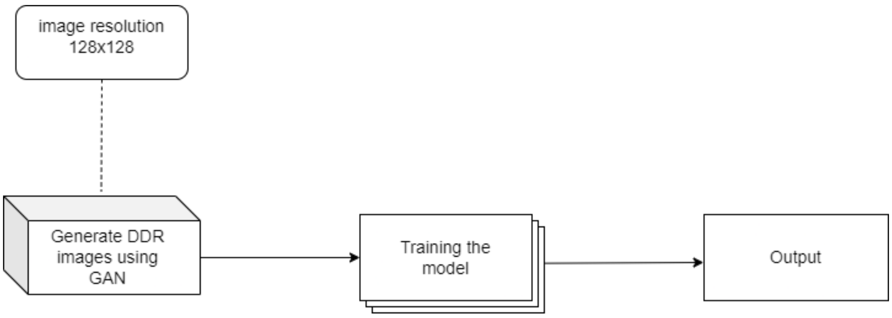

Figure 3.2: Proposed model for **DiagnoNeRF**

Chapter 4

Expected Results

Given a single-view X-ray as input, we will assess the representation of our model for 3D-aware DRR synthesis. Although the implicit linear network has a limited capacity, we will discover that our model can separate the 3D anatomical identity and attenuation response of various medical cases, which will be retrieved using the reconstruction reformulation given in II-C.3. Our model will make it easier to differentiate between bone and tissue using a contrast transformation since it will produce brighter pixel values for denser components. The quality of the reconstructed signals and human subjective similarity will be measured by the peak signal-to-noise ratio (**PSNR**) and structural similarity (**SSIM**), respectively, and these metrics will be used to summarize our findings in Table. We shall discover that, when compared to the truth, our generative loss may produce renderings with a fair perception-distortion curve and exhibit consistency with regard to the placement and volumetric depth of anatomical structures from continuous viewpoints.

References

[1] P. Suetens, *Visualization for diagnosis and therapy*, 2nd ed. Cambridge University Press, 2009, pp. 190–218.

[2] P. Lo, B. van Ginneken, J. M. Reinhardt, T. Yavarna, P. A. de Jong, B. Irving, C. Fetita, M. Ortner, R. Pinho, J. Sijbers, M. Feuerstein, A. Fabijańska, C. Bauer, R. Beichel, C. S. Mendoza, R. Wiemker, J. Lee, A. P. Reeves, S. Born, O. Weinheimer, E. M. van Rikxoort, J. Tschirren, K. Mori, B. Odry, D. P. Naidich, I. Hartmann, E. A. Hoffman, M. Prokop, J. H. Pedersen, and M. de Bruijne, "Extraction of airways from CT (EXACT'09)," *IEEE Trans. Med. Imaging*, vol. 31, no. 11, pp. 2093–2107, Nov. 2012.

[3] M. Coffey and A. Vaandering, "Patient setup for PET/CT acquisition in radiotherapy planning," *Radiother. Oncol.*, vol. 96, no. 3, pp. 298–301, Sep. 2010.

[4] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, "Nerf: Representing scenes as neural radiance fields for view synthesis," *CoRR*, vol. abs/2003.08934, 2020.

[5] R. Martin-Brualla, N. Radwan, M. S. M. Sajjadi, J. T. Barron, A. Dosovitskiy, and D. Duckworth, "Nerf in the wild: Neural radiance fields for unconstrained photo collections," in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021, pp. 7206–7215.

[6] K. Schwarz, Y. Liao, M. Niemeyer, and A. Geiger, "Graf: Generative radiance fields for 3d-aware image synthesis," in *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, Eds., vol. 33. Curran Associates, Inc., 2020, pp. 20 154–20 166. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2020/file/e92e1b476bb5262d793fd40931e0ed53-Paper.pdf