

Multimodal Geometric Deep Learning in Medical Image Analysis

SONATA 20

Popular Science Abstract

Artificial intelligence (AI) has recently entered the lives of most people with great momentum, often without them even realizing it. Achievements in the field of large language models have caught the world's attention, and technology companies responsible for AI development have recorded great results on global stock exchanges.

However, it is worth recalling that the precursor to the development of large language models was the initiation of the development of deep learning models in computer vision. Thanks to these advancements, the largest technology companies became interested in the subject at the beginning of the second decade of the current century and began investing in it. The development of methods based on convolutional neural networks led to the increased popularity of AI and the adaptation of proposed methods to other fields. Consequently, this led to the creation of the transformer architecture dedicated to text processing, which became the foundation for large language models.

Nevertheless, computer vision and natural language processing — despite their undeniable popularity and societal awareness — are not the only areas where AI has found applications. The proposed project concerns another, yet equally important area of knowledge for society — the automatic analysis of medical images.

Almost every person at some point in their life benefits from the achievements of medical imaging techniques and their automatic analysis. Medical imaging includes applications such as ultrasound, computed tomography, magnetic resonance imaging, and other ways of acquiring images in radiology or computational pathology. Currently, the primary problem with effectively conducting diagnostic imaging tests is not the availability of equipment but the lack of a sufficient number of physicians and specialists who can interpret the results and prepare the associated reports. Unfortunately, the results of some tests can even take several weeks, which may be crucial for conducting effective and minimally invasive therapy.

Artificial intelligence is also helpful in this regard. Currently, research is being conducted worldwide on methods for effective and automatic medical image analysis. Researchers from leading scientific institutes worldwide, as well as private research and development departments, spend whole days developing and testing increasingly accurate solutions. AI algorithms do not get tired, can work continuously, are objective when properly trained, and importantly — they work very quickly. Thus, they are an excellent tool that can support doctors in their work, accelerating and improving the quality of diagnostics and therapy.

The proposed project addresses one of the issues in the use of artificial intelligence in the automatic analysis of medical images. Most current algorithms are based on a strictly defined representation, called a voxel grid, which directly results from the way imaging data is reconstructed by imaging devices. Unfortunately, this representation, despite many advantages, also has significant limitations. The number of voxels increases cubically with the reduction of their physical size, allowing for more precise diagnostics but leading to significant difficulties in developing effective methods for their processing and analysis. Consequently, this results in compromises that negatively impact the performance of the aforementioned diagnostic systems.

There are also other methods for representing imaging data, such as point clouds, surface meshes, or implicit representations. Each of them has its own characteristic advantages and disadvantages. Currently, researchers attempting to use other geometric representations in their studies typically focus on one single selected representation, e.g., point clouds for surface reconstruction, surface meshes for the classification of surgical tools, or indirect representations to support cancer diagnostics. Nevertheless, the most desirable methods would be those that combine the advantages of different geometric representations while minimizing their drawbacks. This area of knowledge is currently unexplored.

Therefore, the project's goal is to answer the following questions:

How can various geometric representations be used simultaneously in the automatic analysis of medical images? Will this approach improve the quality of diagnostics? In which downstream tasks can the proposed approach find practical application?