BraintankAI ASIA

DISPATCH · MEDICAL IMAGE SYNTHESIS · 3 MIN

ASTANA · Nazarbayev University

Chest X-ray generator lets users choose where a disease appears

LatAlignCXR steers an image generator with anatomy and pathology location, and blinded radiologists rated its images more spatially consistent than a rival's.

11 November 2026 · based on a paper in Scientific Reports

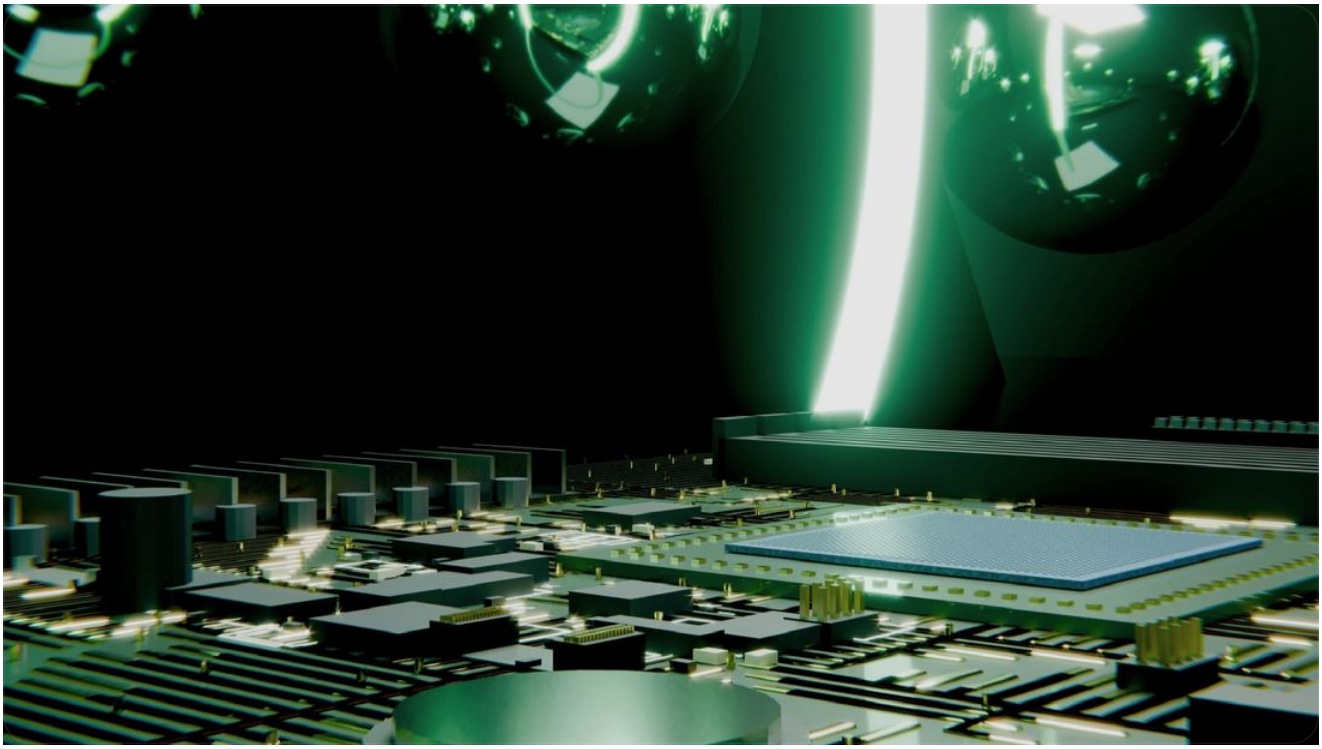


Photo: Brecht Corbeel

Image generators that work from text can be asked for a chest X-ray showing a particular disease, but they often give users little say over where in the chest that disease should sit or how far it should spread. A paper published in Scientific Reports in 2026 by Isabella Schlattner, Vadim Atlassov, Azimzhan Abdrakhmanov, Dong-Ok Won and Min-Ho Lee describes LatAlignCXR, a system built to give that spatial control. The full text was not available to braintankAI, so this report works from the paper's abstract.

The central idea is to put a structured intermediate step between the request and the picture. LatAlignCXR first translates any request into a latent representation, a compact internal description used by the model, that is structured to represent three things: the layout of the chest's anatomy, the location of any disease, and a remainder that covers the rest of the image's

appearance. A request can arrive as a written prompt, as an existing radiograph used as a reference, or as coordinates defined by the user. All three routes lead into the same representation, so one system handles text-, image- and coordinate-driven generation.

From that representation the system decodes three guiding signals, which the authors call priors: one for anatomy, one for pathology and one for the structure of the chest. These steer the image generator itself, a scalable interpolant transformer (SiT). That is a transformer-based generative model of the family that learns to turn random noise into a realistic image through a gradual series of steps.

The authors tested the system on MIMIC-CXR, a collection of chest radiographs. On an anatomical version of the Dice score, an overlap measure on which higher values mean closer agreement, LatAlignCXR scored highest of the systems compared, at 0.6564. It also had the lowest Fréchet distance in three of the four feature spaces used. Fréchet distance compares the statistics of generated and real images as seen through an image-analysis network, with lower values meaning the synthetic set looks more like the real one; the abstract does not name the four feature spaces.

The spatial test reported in the abstract centres on pneumonia. LatAlignCXR showed the highest pneumonia-localisation consistency of the systems compared, with a Dice score of 0.1385 and an IoU of 0.0741. IoU, or intersection over union, is a related overlap measure on which higher values likewise mean closer agreement. Both figures sit far from perfect overlap, and because the abstract does not give the competing systems' scores, the size of the margin cannot be judged from it.

In a blinded evaluation, radiologists rated images without knowing which system had produced them. LatAlignCXR received mean ratings of 4.63 for pathology consistency and 4.50 for spatial consistency, against 3.63 and 3.37 for a comparison system named Cheff. The abstract does not say how many radiologists took part or what rating scale they used.

The authors also asked whether the synthetic images help train an image classifier. With 1000 real images available, adding synthetic ones raised accuracy from 69.0% to 76.0%. As more real data became available, the effect of adding synthetic images was smaller and non-monotonic, meaning it did not rise or fall steadily with the amount of real data. The abstract does not specify which conditions the classifier was trained to recognise.

The authors conclude that the results support structured latent conditioning for spatially controllable chest X-ray synthesis, particularly for pathology representation and spatial consistency. The localisation figures in the abstract concern pneumonia alone, and the training effect is smaller and uneven once more real data is available. The study is an evaluation of generated images and a training experiment; the abstract describes no use of the system in clinical care.

PAPER	*Latent-aligned scalable interpolant transformer for controllable chest X-ray synthesis*
AUTHORS	Isabella Schlattner, Vadim Atlassov, Azimzhan Abdrakhmanov, Dong-Ok Won, Min-Ho Lee
VENUE	Scientific Reports · 30 October 2026
LINK	doi:10.1038/s41598-026-75231-7
INDEX	stacks.braintank.ai