

Final Project: Auto_Stereo

Aidan Mokalla

CSCI 378

May 2025

My final project, “auto_stereo,” is a program for generation and interpretation of autostereograms, i.e. “magic eye” images. Autostereograms are single-image stereograms that create visual illusions of 3D scenes from repeated 2D *textures*, using *stereopsis* and *disparity* to create the visual illusion of a 3D scene.

I initially considered training a CNN directly to convert stereograms into depth maps. This would’ve involved learning to parse the complex relationship between autostereogram texture and its underlying depth. After experimenting with different architectures and training methods I was unsuccessful here, and had a hard time learning the specific spatial relationship in autostereograms except in the case with especially binary depth maps (i.e. with few intermediate depth intensities).

So instead, I recognized the complexity and effort that has gone into existing work in this domain, and the project evolved into a novel combination of existing work. Two external models were composed to create new functionalities.

The first is *Neural Magic Eye*. This model showed that a deep CNN equipped with a special layer can recover the depth behind an autostereogram especially well. Neural Magic Eye had a novel convolutional layer called *disparity* convolution, which simulates stereopsis and encodes disparity from quasi-periodic textures. The model also implemented a standard convolutional autoencoder architecture and learned, under self-supervision, from a large 3D object dataset. Neural Magic Eye had state-of-the-art performance in stereogram-to-depth recovery compared to other models that existed at the time that did not have this layer. In the four years since, it has remained the state of the art.

The second model, *Depth Anything*, estimates depth from regular images. Its development was supported by TikTok, inter alia. Depth Anything V2 is the specific model version I used in this project. Depth Anything had an interesting approach to training: because there's a lack of high quality image datasets with precise depth labels, it used a teacher-student training strategy where a teacher model, trained on synthetic data for which the researchers had depth information, then labeled real-world images. These real-world images and labels then trained a student model. I used pre-trained versions of both of these models, since I had no need to re-train them myself.

The main outcome is a tool that integrates components from these models into a configurable and flexible pipeline for working with autostereograms and depth maps. Auto_stereo can perform the following conversions:

1. Using Depth Anything V2, it can take a standard 2D image and generate a corresponding depth map.
2. Specifically using the trained *decoding* part of the model from the Neural Magic Eye project, auto_stereo can extract depth information embedded within an existing autostereogram.
3. Auto_stereo can generate a new autostereogram by applying a specific *texture*, in addition to the five predefined textures in the program (random blocks, noise, stripes, dots, and gradients). This is done with depth-dependent horizontal pixel shifting based on the depth map and a repeating texture pattern. Parameters like effect strength (controlling the disparity shift based on depth) can also be adjusted, but are set to defaults that work in most cases.

As a result, this tool can convert directly from images to stereograms in a single step. It can also convert stereograms from one texture to another. Neither functionality existed previously.

Looking ahead, potential areas for further development or research include improving packaging (e.g. creating a fully functional website), implementing automatic input type detection within the main script itself (instead of specifying whether the input is an image, depth map, or

autostereogram), exploring watermarking techniques within stereograms (potentially enabling authentication or extended capacity), researching depth map to image conversion (i.e. depth map assisted image generation), and further tuning the model and generation parameters.

The auto_stereo project is a functional tool that integrates the existing models Neural Magic Eye and Depth Anything. This approach taught me about several interesting concepts, such as intuition for how stereopsis influences visual depth cognition in the first place, the utility of synthetic distributions in training, the role of priors and inductive biases in architecture design, and the effectiveness of building upon previous efforts.