

A vertical strip of 12 small, square images showing the progression of a plant's growth from a seedling to a mature plant. The images are arranged in a single column, with the seedling at the top and the mature plant at the bottom. The plants are shown in various stages of development, from a small seedling to a mature plant with many leaves. The images are arranged in a single column, with the seedling at the top and the mature plant at the bottom. The plants are shown in various stages of development, from a small seedling to a mature plant with many leaves.

- Now imagine a prompt just for a persons face at first the pixel intensities are chosen randomly so the Image/video looks like pure noise, from here this pure noise video is passed into a transformer (like LLM one) but insted of outputting text it outputs another Image/video that has a littel bit of structure this new Img/video is added to the pure noise Img/video and then its passed in the transformer again this new img/vid has more structure and is getting closer to the prompt and so we repeat this steps over and over until the model thinks its enough at which point the Img/vid matches the prompt (this can be Ex 50 iterations)

