

1. How does a Conditional GAN differ from a vanilla GAN?

Extra input:

- Vanilla GAN: Generator takes only random noise; Discriminator sees only “real vs. fake” samples.
- Conditional GAN: Both G and D also receive a condition vector (e.g. class label, text embedding, segmentation map).

Controlled outputs:

- By feeding in a label or other side-information, you steer the generator toward a specific mode (e.g. digit “7” instead of any digit).

Real-world application:

- **Medical imaging:** Generate MRI scans conditioned on pathology labels (tumor vs. healthy), enabling data augmentation for rare disease classes.
- **Text-to-image:** Systems like DALL-E use text prompts to condition image synthesis on semantic content.

2. What does the discriminator learn in an image-to-image GAN?

Joint realism: It doesn’t just check “is this output realistic?” but “is this output realistic *given* this input?”

Paired consistency: By training on (input, ground-truth) pairs, the D learns the correspondence between structure in the input (e.g. a sketch or low-res image) and the desired output (e.g. a photorealistic image).

Why pairing matters: Without explicit pairing, the model might produce realistic images that bear no relation to the input. Pairing enforces pixel- or feature-level alignment so that the generated image both looks real and matches the conditioning input.