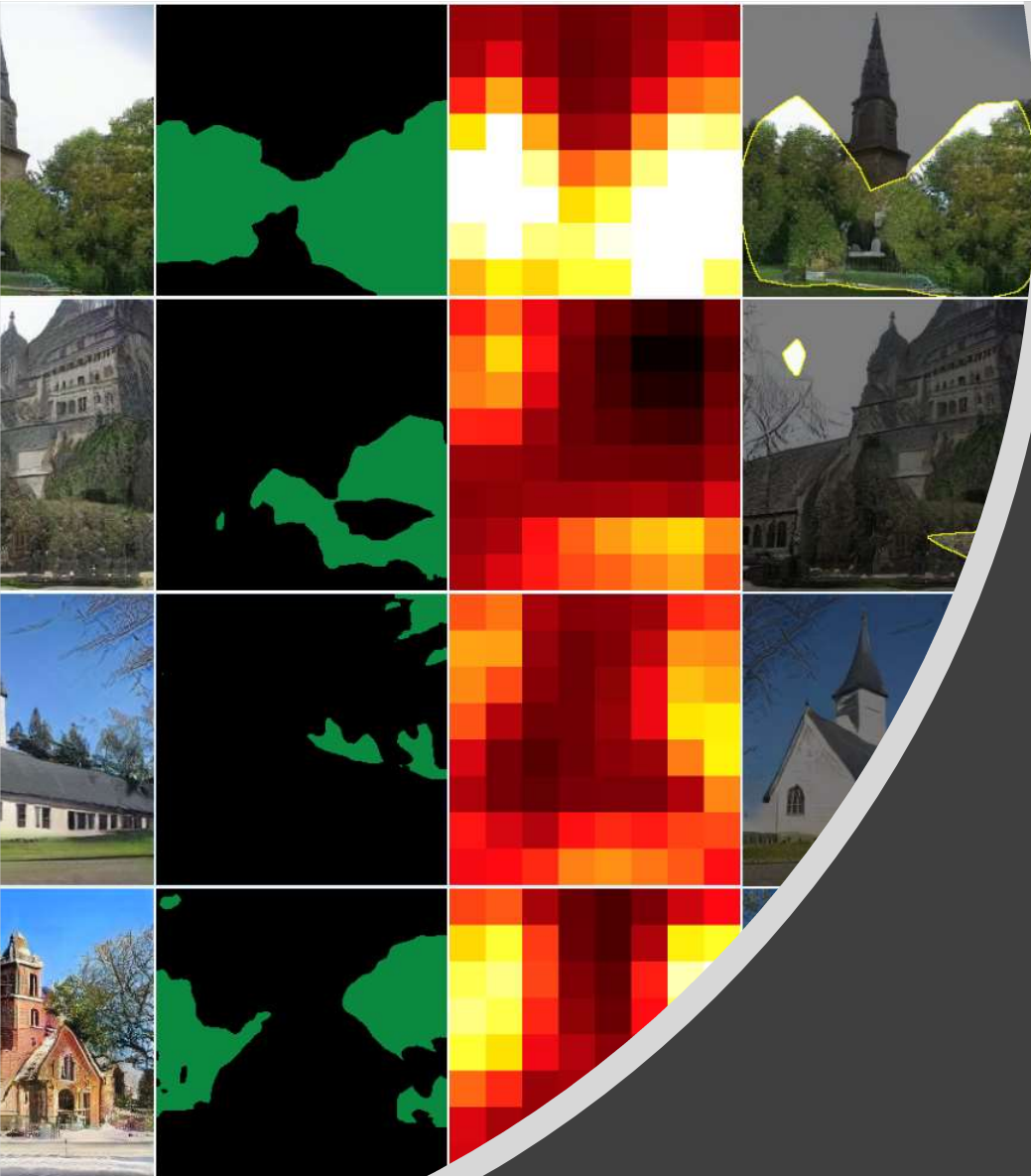# Tutorial: GAN Dissection
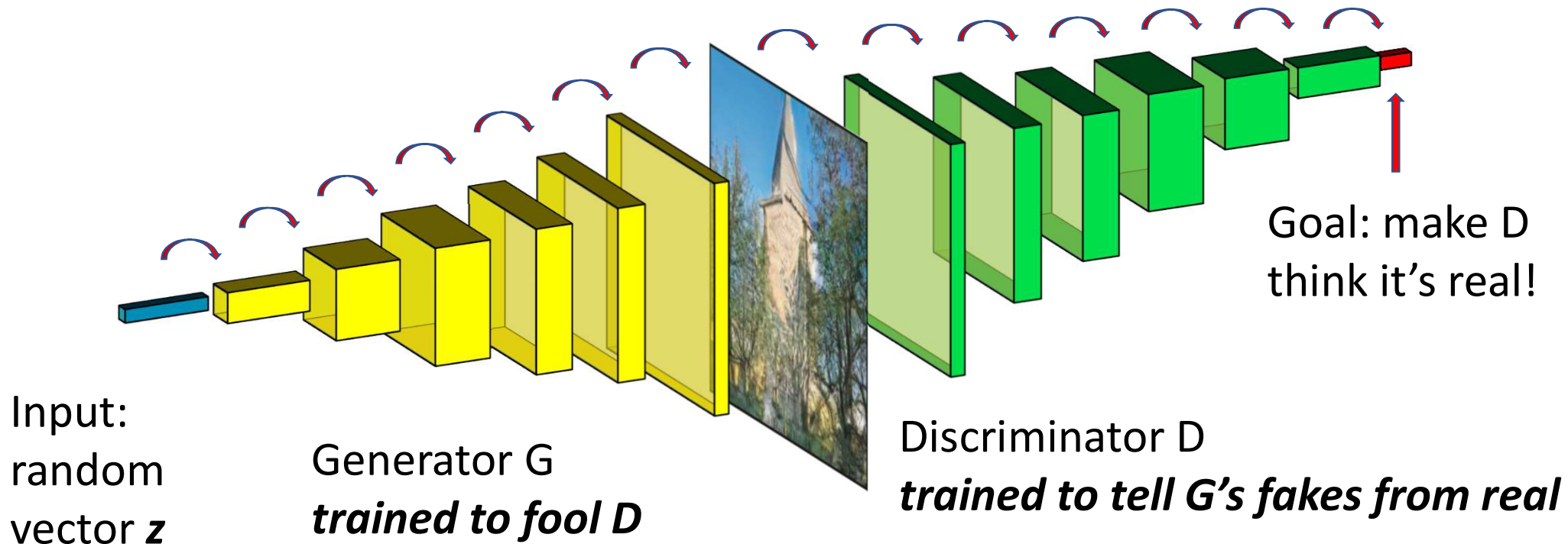
What is learned inside a GAN?

David Bau

To follow along: http://bit.ly/gandtut
We will go through a Jupyter notebook.
Ideal machine has git, conda, and a GPU.

Generative Adversarial Networks

The generator learns to beat a discriminator….

Goal: make D
think it's real!

Input:
random
vector z

Generator G
trained to fool D

Discriminator D
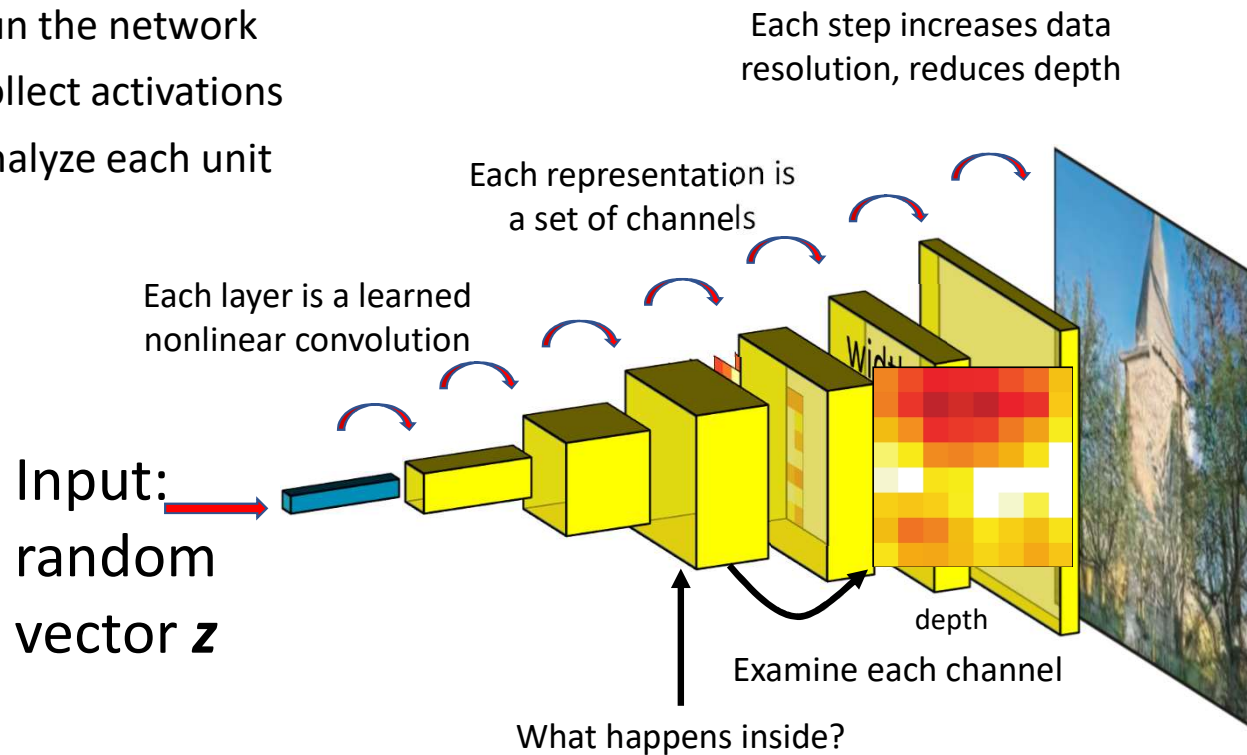**trained to tell G's fakes from real**

What do we learn?

In the end the generator can synthesize terrific images.

But how does it work? What does it actually learn?

Our Plan: GAN Dissection

1. Run the network

2. Collect activations

3. Analyze each unit

Each step increases data resolution, reduces depth

Output: random realistic image **x**

Each representation is a set of channels

Each layer is a learned nonlinear convolution

Input: random vector **z**

Width

depth

What happens inside?

Examine each channel

You can play along. Tutorial setup…

Instructions on **http://bit.ly/gandtut**

```
git clone --branch tutorial https://github.com/CSAILVision/gandissect.git
cd gandissect
script/setup_env.sh        # Create a conda environment with dependencies
script/make_dirs.sh        # Create the dataset and dissect directories
source activate netd       # Enter the conda environment
pip install -v -e .        # Link the local netdissect package into the env
cd notebooks
jupyter notebook &         # Run jupyter
```

Then run the notebook called: **dissect_progan.ipynb**