

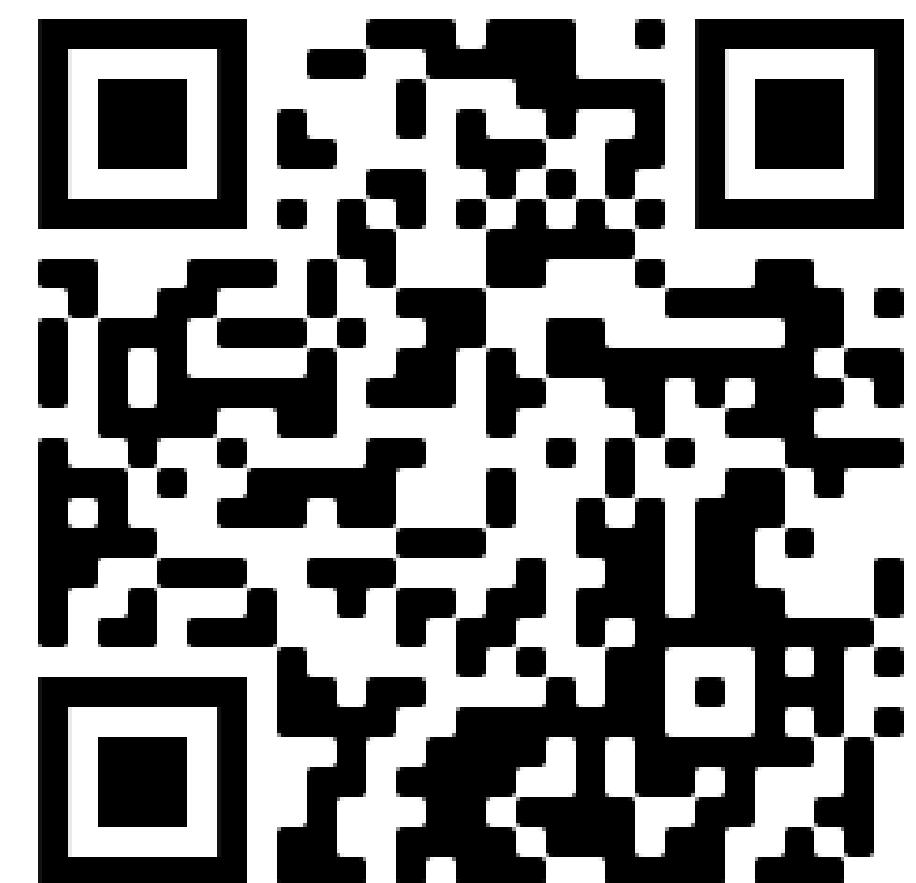
Generative AI Innovative Applications

NVIDIA Generative AI with Diffusion Models Part II

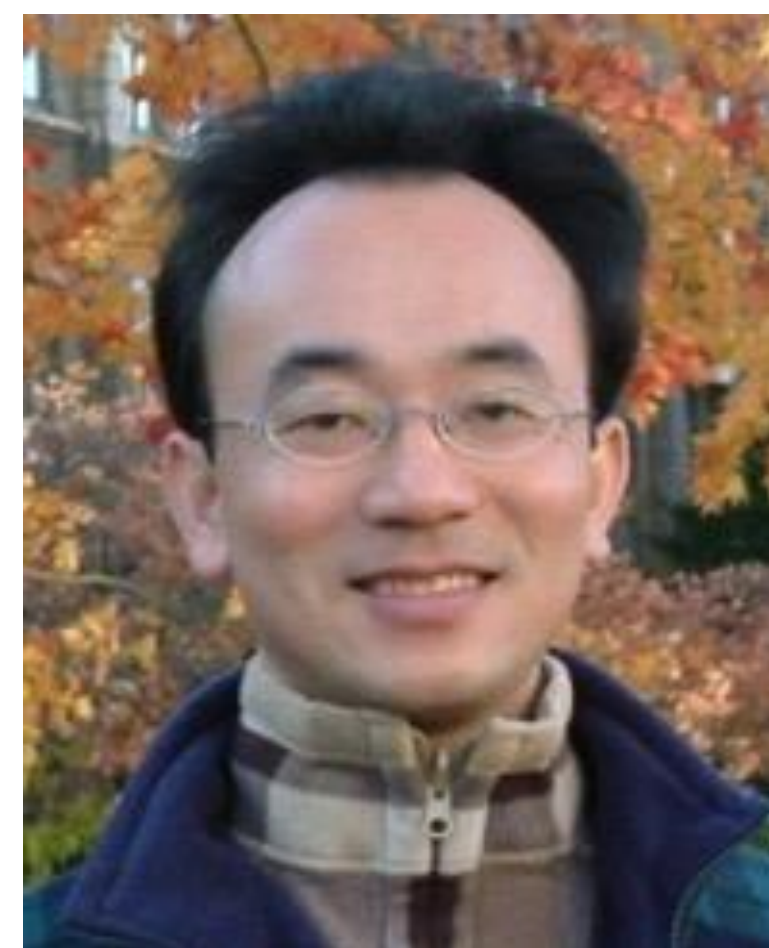
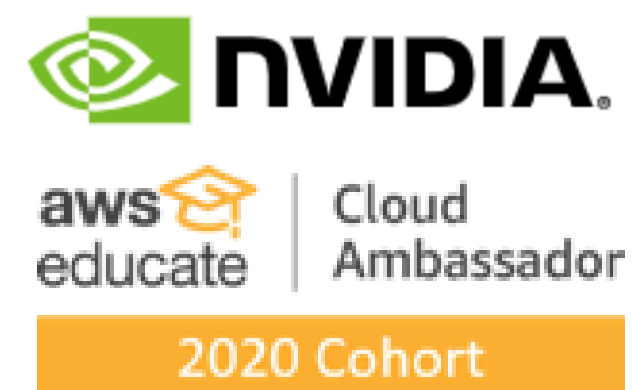
1132GAIIA08

MBA, IM, NTPU (M6031) (Spring 2025)

Tue 2, 3, 4 (9:10-12:00) (B3F17)



<https://meet.google.com/pai-zhhj-mya>



Min-Yuh Day, Ph.D.,
Professor

Institute of Information Management, National Taipei University

<https://web.ntpu.edu.tw/~myday>

2025-04-29



Syllabus

Week Date Subject/Topics

1 2025/02/18 Introduction to Generative AI Innovative Applications

**2 2025/02/25 Transformers for Natural Language Processing and
Computer Vision**

**3 2025/03/04 Large Language Models (LLMs),
NVIDIA Building RAG Agents with LLMs Part I**

4 2025/03/11 Case Study on Generative AI Innovative Applications I

5 2025/03/18 NVIDIA Building RAG Agents with LLMs Part II

6 2025/03/25 NVIDIA Building RAG Agents with LLMs Part III

Syllabus

Week Date Subject/Topics

7 2025/04/01 Self-Learning

8 2025/04/08 Midterm Project Report

9 2025/04/15 Generative AI for Multimodal Information Generation

10 2025/04/22 NVIDIA Generative AI with Diffusion Models Part I

11 2025/04/29 NVIDIA Generative AI with Diffusion Models Part II

12 2025/05/06 Case Study on Generative AI Innovative Applications II

Syllabus

Week Date Subject/Topics

**13 2025/05/13 NVIDIA DLI and NVIDIA-Certified Associate (NCA)
Generative AI LLMs Certification [Invited Talk]**

**14 2025/05/20 NVIDIA Generative AI with Diffusion Models Part III
AI Agents and Large Multimodal Agents (LMAs)**

15 2025/05/27 Final Project Report I

16 2025/06/03 Final Project Report II

NVIDIA

Generative AI with Diffusion Models

Part II

[0,1,2] [3, 4] [5, 6]

Generative AI with Diffusion Models

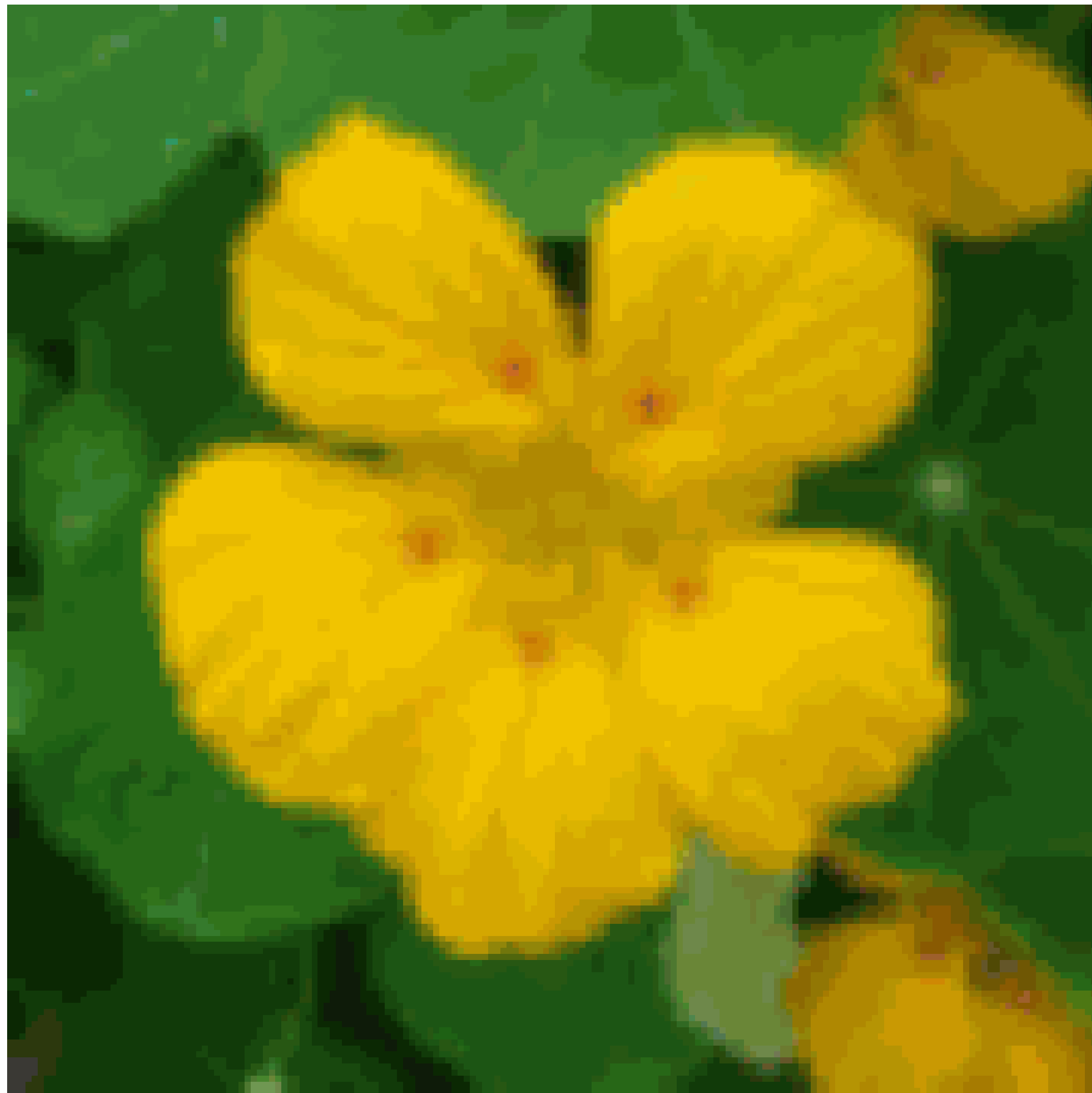
- **Part 1: From U-Nets to Diffusion**
- **Part 2: Denoising Diffusion Probabilistic Models**
- **Part 3: Optimizations**
- **Part 4: Classifier-Free Diffusion Guidance (CFG)**
- **Part 5: CLIP (Contrastive Language-Image Pre-Training)**
- **Part 6: Wrap-up & Assessment**

Diffusion Models

diffusion models are trained to denoise noisy images,
and can generate images by iteratively denoising pure noise.

Forward diffusion

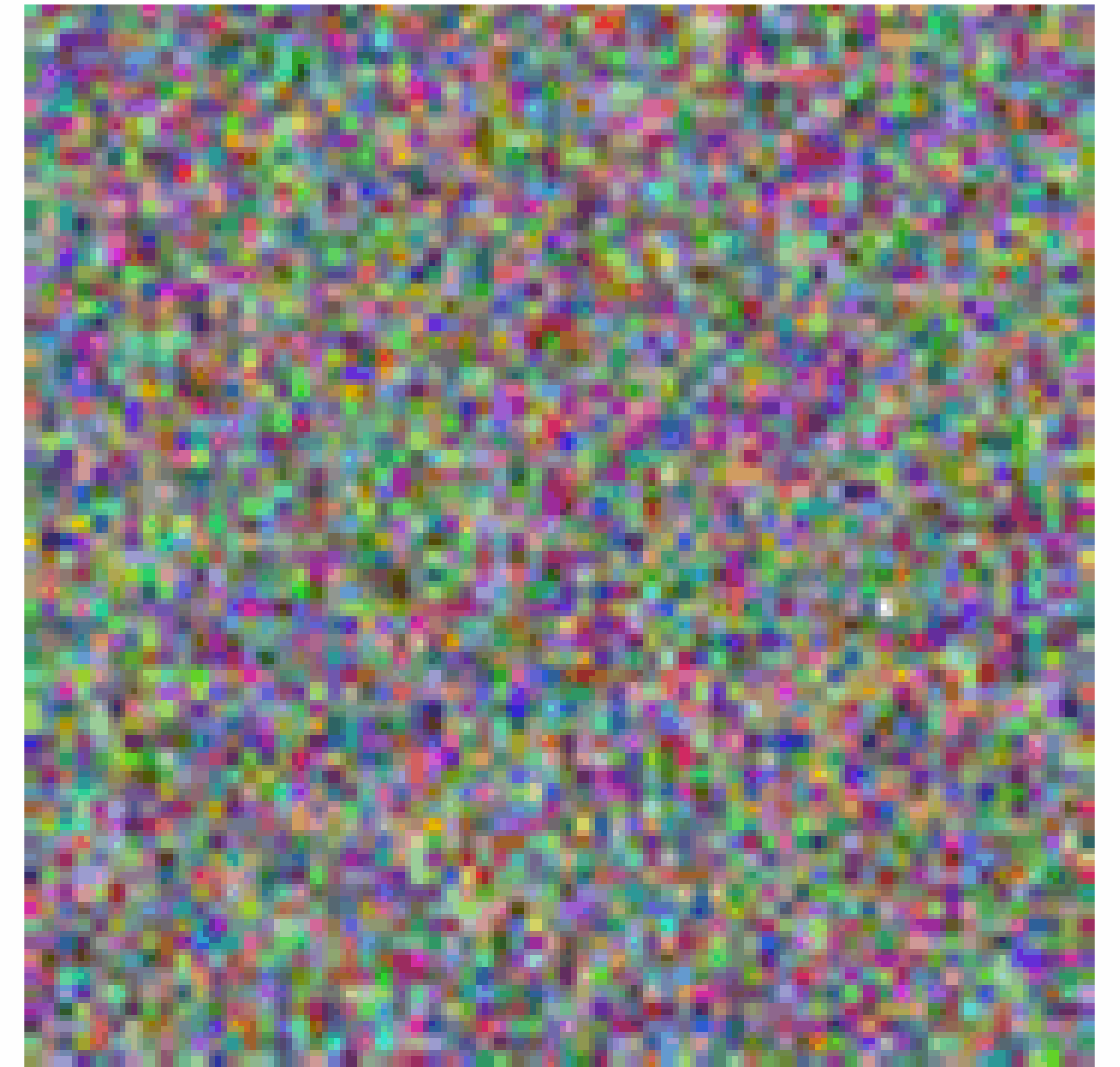
image



noisy image



noise



Source: <https://keras.io/examples/generative/ddim/>

build.nvidia.com Models: Image Generation


[Explore](#)
[Models](#)
[Blueprints](#)
[GPUs](#)

 Search




Models

Deploy and scale models on your GPU infrastructure of choice with NVIDIA NIM inference microservices

Filter by text

Sort By: Most Recent

Publisher

Use Case (1)

NIM Type

Use Case: Image Generation

Clear Filters



black-forest-labs
FLUX.1-dev

FLUX.1 is a state-of-the-art suite of image generation...

image generation +3



nvidia
consistory

Generates consistent characters across a series o...

image generation +2



briaai
BRIA-2.3

An enterprise-grade text-to-image model trained on a...

image generation +2



stabilityai
stable-diffusion-3-...

Advanced text-to-image model for generating high...

image generation +2



nvidia
visual-changenet

Visual Changenet detects pixel-level change maps...

image +9



stabilityai
stable-diffusion-xl

Generate images and stunning visuals with...

image generation +2



stabilityai
stable-video-diffusi...

Stable Video Diffusion (SVD) is a generative diffusion...

image generation +2



stabilityai
sdxl-turbo

A fast generative text-to-image model that can...

image generation +2

https://build.nvidia.com/models?filters=usecase%3Ausecase_image_gen

NVIDIA

ExploreModelsBlueprintsGPUs

?

M

stable-diffusion-xlPREVIEW

Generate images and stunning visuals with realistic aesthetics.

image generationtext-to-image

Get API Key

ExperienceProjectsModel Card

API ReferenceAccelerated by DGX Cloud

AI models generate responses and outputs based on complex algorithms and machine learning techniques, and those responses or outputs may be inaccurate, harmful, biased or indecent. By testing this model, you assume the risk of any harm caused by any response or output of the model. Please do not upload any confidential information or personal data unless expressly permitted.

InputTryShellNodePython

Prompt * ⓘ

underwater world, plants, shells, creatures, high detail, sharp focus, 4k

GOVERNING TERMS:

This trial service is governed by the [NVIDIA API Trial Terms of Service](#). Use of this model is governed by the [Stable Diffusion XL License](#).

ResetRun

25 / 25 requests remaining

stabilityai
stable-diffusion-xlPREVIEW

Generate images and stunning visuals with realistic aesthetics.

image generationtext-to-image

Get API Key

ExperienceProjectsModel Card

API ReferenceAccelerated by DGX Cloud

AI models generate responses and outputs based on complex algorithms and machine learning techniques, and those responses or outputs may be inaccurate, harmful, biased or indecent. By testing this model, you assume the risk of any harm caused by any response or output of the model. Please do not upload any confidential information or personal data unless expressly permitted.

InputTryShellNodePython

Prompt ⓘ

Cute Yellow Cat, Cherry Blossom, high detail, sharp focus, 4k

Hide Parameters

Negative Prompt ⓘ

Sampler ⓘ

DDIM

K_EULER_ANCESTRALK_LMS

City Scale ⓘ

11

0.850.91.01.1

5

Steps ⓘ

7

10152025

25

ResetRun

1 / 25 requests remaining

OutputPreviewJSON



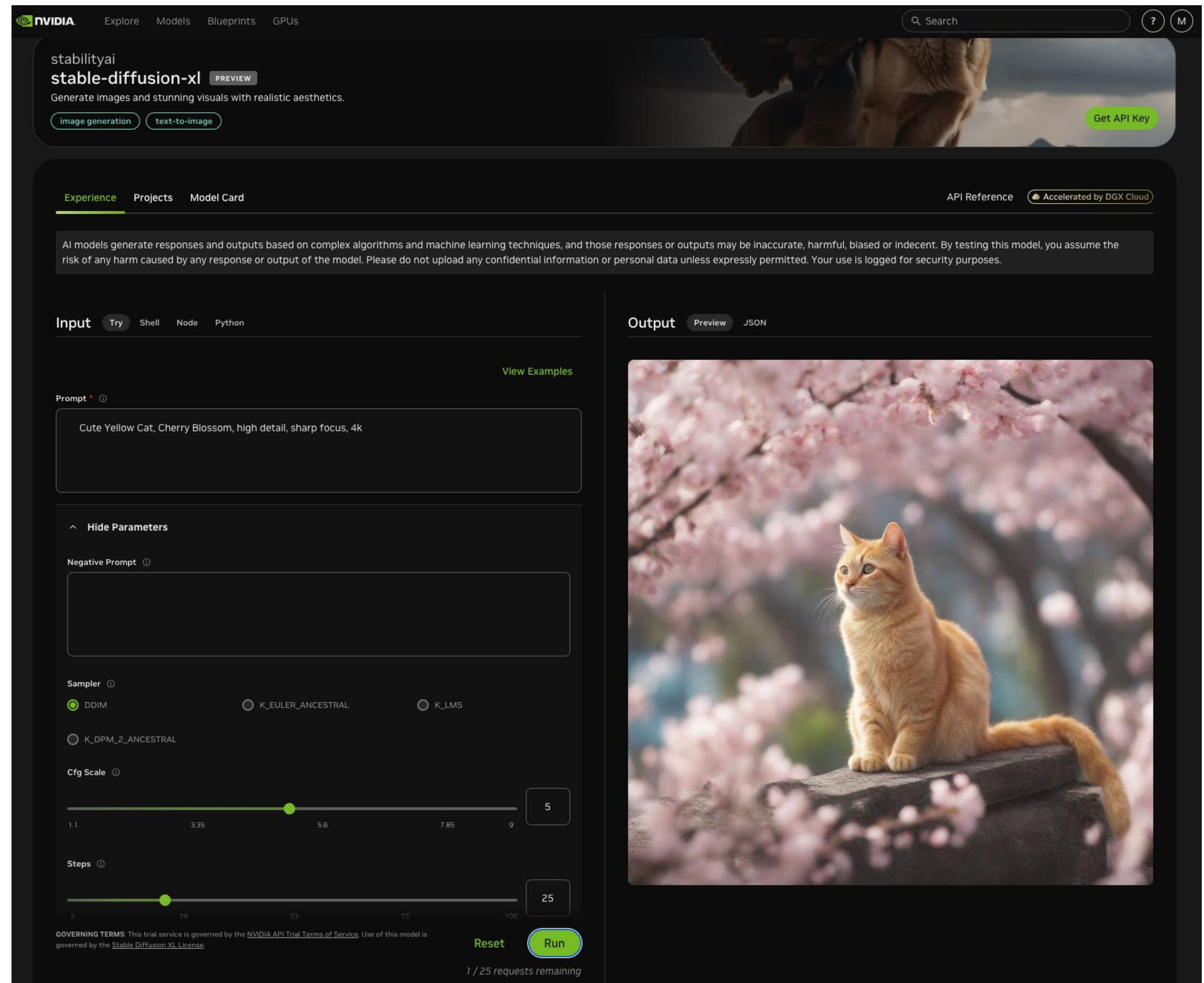
Techniques, and those responses or outputs may be inaccurate, harmful, biased or indecent. By testing this model, you assume the risk of any harm caused by any response or output of the model. Please do not upload any confidential information or personal data unless expressly permitted.

OutputPreviewJSON



Input Prompt

Cute Yellow Cat, Cherry Blossom, high detail, sharp focus, 4k



Spring 2025

Generative AI Innovative Applications



University Ambassador



This certificate acknowledges that

Min-Yuh Day

has been certified to deliver NVIDIA instructor-led workshop for
academia

A handwritten signature in black ink, appearing to read "Greg A. Estes", written over a horizontal line.

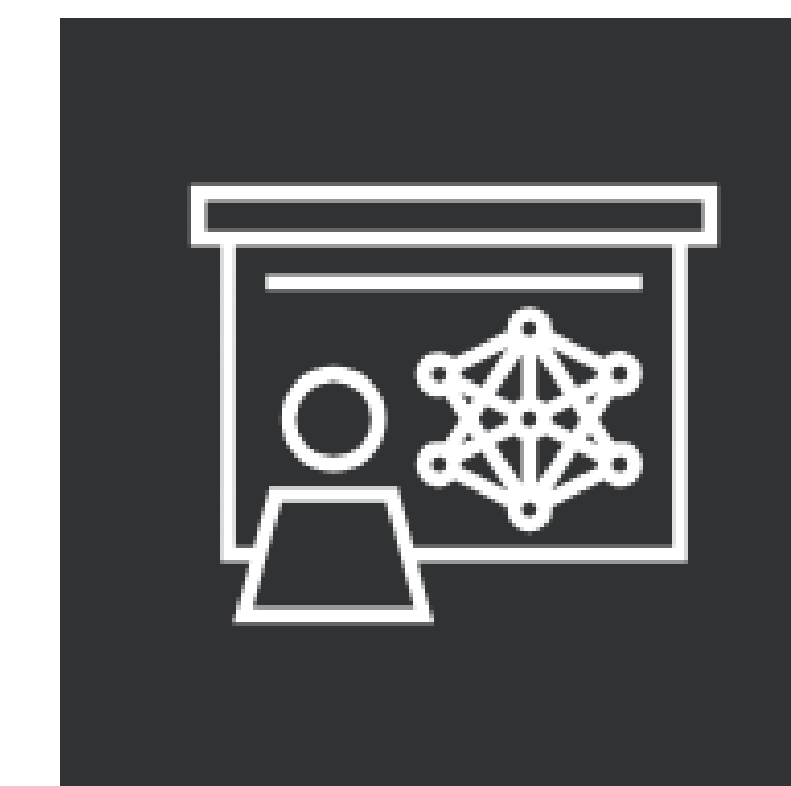
Greg Estes

Vice President, NVIDIA

Issue Date: : March 7, 2025

Ambassador Certification ID: cCFh1ZWWTvqKTq7dcKkEWw

<https://learn.nvidia.com/certificates?id=cCFh1ZWWTvqKTq7dcKkEWw>



Certified Instructor

This certificate acknowledges that

Min-Yuh Day

has been certified to deliver the instructor-led workshop

Building RAG Agents with LLMs

A handwritten signature in black ink, appearing to read "Greg A. Estes", written over a horizontal line.

Greg Estes

Vice President, NVIDIA

Issue Date: : March 7, 2025

Certification ID: OVmqY4cSSya0BdMQBWHxzw

<https://learn.nvidia.com/certificates?id=OVmqY4cSSya0BdMQBWHxzw>

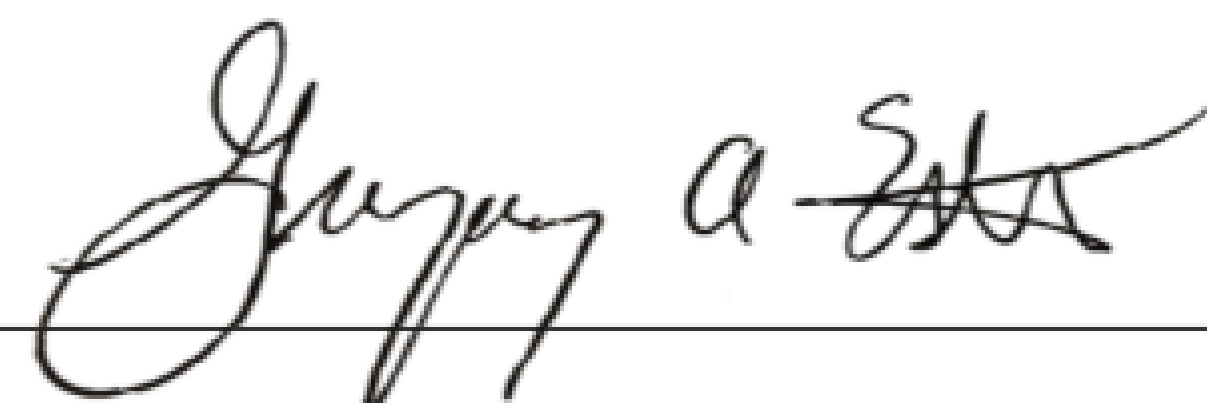
Certificate of Completion

This certificate is awarded to

Min-Yuh Day

for successfully completing

Building RAG Agents with LLMs



Greg Estes

Vice President, NVIDIA

Issue Date: : December 8, 2024

Certification ID: ed-qOCIMQaatzU8SNUNxgw |

https://learn.nvidia.com/certificates?id=ed-qOCIMQaatzU8SNUNxgw/courses/course?course_id=course-v1:DLI+S-FX-15+V1

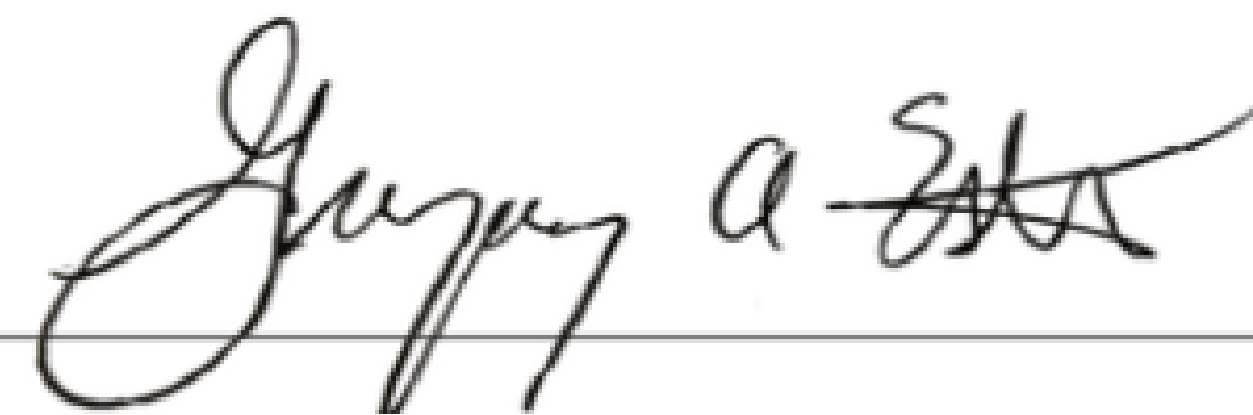
<https://learn.nvidia.com/certificates?id=ed-qOCIMQaatzU8SNUNxgw>

Certificate of Competency

This certificate is awarded to

Min-Yuh Day

for demonstrating competence in the completion of
Generative AI with Diffusion Models



Greg Estes

Vice President, NVIDIA

Issue Date: : February 28, 2025

Certification ID: q3oo-oBhTQKtyCCote2E-Q

NVIDIA Certified Instructors (12 from Taiwan)

Shop Drivers Support 

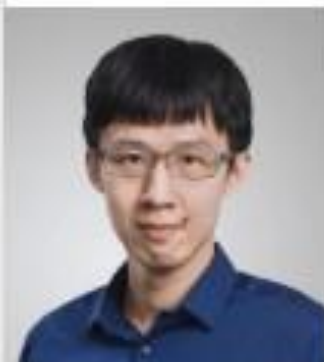
Don't miss NVIDIA CEO Jensen Huang's GTC keynote March 18. Join early for the pregame event. 01d:04h:15m:12s [Add to Calendar](#)

Deep Learning Institute Find Training Self Paced Courses Instructor-Led Workshops Educator Programs ...

Filters

Workshop Certification Location [1] Organization Specialization Reset filters Search for instructors by name. Minimum

Sort by: Name A-Z

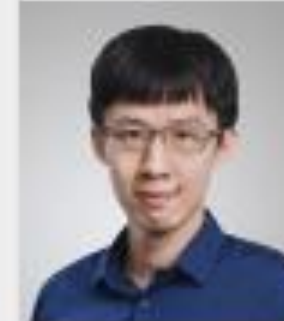
Chi-Hung Chuang  Chung Yuan Christian University in Featured Workshop Fundamentals of Deep Learning, Taiwanese	Chia Yu Hsu  National Taiwan University of Science and Technology(NTUST) Featured Workshop Applications of AI for Predictive Maintenance, Taiwanese	Chien-Yu Chen  National Taiwan University (NTU) Featured Workshop Fundamentals of Deep Learning, Taiwanese
Chun-Yi Lee  National Taiwan University (NTU) Platinum Instructor Featured Workshop Fundamentals of Deep Learning, Taiwanese	David Tseng  Cavedu Platinum Instructor Featured Workshop Getting Started with AI on Jetson Nano, Taiwanese	Hsinmin Lu  National Taiwan University (NTU) Featured Workshop Fundamentals of Deep Learning, English
Min-Yuh Day  National Taipei University (NTPU) in Featured Workshop Building RAG Agents with LLMs, English	Ming-Che Chen  Southern Taiwan University of Science and Technology(STUST) Featured Workshop Getting Started with AI on Jetson Nano, Taiwanese	MingChe Hu  Chung Yuan Christian University Featured Workshop Fundamentals of Deep Learning, Taiwanese
Ping-Chun Hsieh  National Yang Ming Chiao Tung University (NYCU) Featured Workshop Fundamentals of Deep Learning, English	Po-Chih Kuo  National Tsing Hua University (NTHU) in Featured Workshop Fundamentals of Deep Learning, English	Shu-Kai Hsieh  National Taiwan University (NTU) in Featured Workshop Fundamentals of Deep Learning, Taiwanese

NVIDIA Certified Instructors (14 from Taiwan) (2 from NTPU)(April 2025)

Filters

Workshop Certification Location [1] Organization Specialization Reset filters Search for instructors by name. Minisur

Sort by: Name A-Z

Chi-Hung Chuang  Chung Yuan Christian University in Featured Workshop Fundamentals of Deep Learning, Taiwanese	Chia Yu Hsu  National Taiwan University of Science and Technology (NTUST) Featured Workshop Applications of AI for Predictive Maintenance, Taiwanese	Chien-Yu Chen  National Taiwan University (NTU) Featured Workshop Fundamentals of Deep Learning, Taiwanese
Chun-Yi Lee  National Taiwan University (NTU) Platinum Instructor Featured Workshop Fundamentals of Deep Learning, Taiwanese	David Tseng  Cavedu Platinum Instructor Featured Workshop Building Transformer-Based Natural Language Processing Applications, Taiwanese	Hsinmin Lu  National Taiwan University (NTU) Featured Workshop Fundamentals of Deep Learning, English
Hung-Wen Chen  National Tsing Hua University (NTHU) in Featured Workshop Fundamentals of Deep Learning, Taiwanese	Ko-Chia Yu  National Taipei University (NTPU) in Featured Workshop Building RAG Agents with LLMs, English	Min-Yuh Day  National Taipei University (NTPU) in Featured Workshop Building RAG Agents with LLMs, English
Ming-Che Chen  Southern Taiwan University of Science and Technology (STUST) Featured Workshop Getting Started with AI on Jetson Nano, Taiwanese	MingChe Hu  Chung Yuan Christian University Featured Workshop Fundamentals of Deep Learning, Taiwanese	Ping-Chun Hsieh  National Yang Ming Chiao Tung University (NYCU) Featured Workshop Fundamentals of Deep Learning, English
Po-Chih Kuo  National Tsing Hua University (NTHU) in Featured Workshop Fundamentals of Deep Learning, English	Shu-Kai Hsieh  National Taiwan University (NTU) in Featured Workshop Fundamentals of Deep Learning, Taiwanese	

NVIDIA Developer Program

<https://developer.nvidia.com/join-nvidia-developer-program>

NVIDIA

Deep Learning Institute (DLI)

<https://learn.nvidia.com/>

Get NVIDIA DLI Certificate

- **Step 1. Join NVIDIA Developer Program (Free)**
<https://developer.nvidia.com/join-nvidia-developer-program>
- **Step 2. Visit NVIDIA Deep Learning Institute (DLI)**
<https://learn.nvidia.com/>
- **Step 3. Enroll "Generative AI with Diffusion Models"**
Self-Paced Course (\$90)
https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Join the NVIDIA Developer Program

take one of the
complimentary
technical self-
paced courses
(worth up to \$90)

Generative AI and LLMs Graphics and Simulation Accelerated Computing Data Science Deep Learning

8 hours
Getting Started With Deep Learning
Explore the fundamentals of deep learning by training neural networks and using results to improve performance and capabilities.

2 hours
Modeling Time-Series Data With Recurrent Neural Networks in Keras
Explore how to classify and forecast time-series data using recurrent neural networks (RNNs), such as modeling a patient's health over time.

4 hours
Deploying a Model for Inference at Production Scale
Learn how to deploy your own machine learning models on a GPU server.

8 hours
Building Real-Time Video AI Applications
Gain the knowledge and skills needed to enable the real-time transformation of raw video data from widely deployed camera sensors into deep learning-based insights.

2 hours
Introduction to Graph Neural Networks
Learn the basic concepts, models, and applications of graph neural networks.

4 hours
Introduction to Physics-Informed Machine Learning With Modulus
Learn the various building blocks of NVIDIA Modulus, which turbocharges use cases by building physics-based deep learning models that are 100,000X faster than traditional methods and offers high-fidelity simulation results.

2 hours
Get Started With Highly Accurate Custom ASR for Speech AI
Learn to build, train, fine-tune, and deploy a GPU-accelerated automatic speech recognition (ASR) service with NVIDIA® Riva that includes customized features.

2 hours
Integrating Sensors With NVIDIA DRIVE
Find out how to integrate automotive sensors into your applications using NVIDIA DRIVE®.

<https://developer.nvidia.com/join-nvidia-developer-program>

NVIDIA Deep Learning Institute (DLI)

Self-Paced Course

Generative AI Explained

Free
2 hours

Self-Paced Course

Getting Started With Deep Learning

Certificate available
\$90
8 hours

Instructor-Led Workshop

Fundamentals of Deep Learning

Certificate available
\$500
8 hours

Self-Paced Course

Introduction to Transformer-Based Natural Language Processing

Certificate available
\$30
6 hours

Self-Paced Course

Building RAG Agents With LLMs

Certificate available
Free
8 hours

Instructor-Led Workshop

Building RAG Agents With LLMs

Certificate available
\$500
8 hours

Self-Paced Course

Generative AI with Diffusion Models

Certificate available
\$90
8 hours

Instructor-Led Workshop

Generative AI with Diffusion Models

Certificate available
\$500
8 hours

What do you want to learn today?

Filters

Level +

Format +

Topics -

- ☐ Deep Learning
- ☐ Accelerated Computing
- ☐ Generative AI/LLM
- ☐ Graphics and Simulation
- ☐ OpenUSD
- ☐ Data Science
- ☐ NIMS
- ☐ NIM
- ☐ RAPIDS

Free / Paid +

Language +

Generative AI



Sort by: --

Showing 19 results

Generative AI x

Generative AI

All Courses

Self-paced

Generative AI Explained

Free
02:00

Self-paced

Generative AI with Diffusion Models

\$90
08:00

Instructor-Led

Generative AI with Diffusion Models

08:00

Self-paced

Augment your LLM Using

Self-paced

Introduction to Transformer-

Instructor-Led

Rapid Application

Generative AI with Diffusion Models

Self-paced Course

Generative AI with Diffusion Models

Take a deeper dive into denoising diffusion models, which are a popular choice for text-to-image pipelines, with applications in creative content generation, data augmentation, simulation and planning, anomaly detection, drug discovery, personalized recommendations, and more.

About Course

Objectives

Topics Covered

Course Outline

Stay Informed

Contact Us

Continue Learning

About this Course

Thanks to improvements in computing power and scientific theory, generative AI is more accessible than ever before. Generative AI plays a significant role across industries due to its numerous applications, such as creative content generation, data augmentation, simulation and planning, anomaly detection, drug discovery, personalized recommendations, and more. In this course, learners will take a deeper dive into denoising diffusion models, which are a popular choice for text-to-image pipelines.

Learning Objectives

Course Details

Duration: 08:00

Price: \$90

Subject: Generative AI/LLM

Language: English

Course Prerequisites:

A basic understanding of [Deep Learning Concepts](#).

https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Building RAG Agents with LLMs

Self-paced Course

Building RAG Agents with LLMs

Agents powered by large language models (LLMs) have shown great retrieval capability for using tools, looking at documents, and plan their approaches. This course will show you how to deploy an agent system in practice with the flexibility to scale up your system to meet the demands of users and customers.



About this Course

This course is free for a limited time.

The evolution and adoption of large language models (LLMs) have been nothing short of revolutionary, with retrieval-based systems at the forefront of this technological leap. These models are not just tools for automation; they are partners in enhancing productivity, capable of holding informed conversations by interacting with a vast array of tools and documents. This course is designed for those eager to explore the potential of these systems, focusing on practical deployment and the efficient implementation required to manage the considerable demands of both users and deep learning models. As we delve into the intricacies of LLMs, participants will gain insights into advanced orchestration techniques that include internal reasoning, dialog management, and effective tooling strategies.

Course Details

Duration: 08:00

Price: Free

Level: Technical - Intermediate

Subject: Generative AI/LLM

Language: English

Course Prerequisites:

Introductory deep learning knowledge, with comfort

https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-15+V1

Generative AI Explained

Deep Learning Institute

[Find Training](#)

[Self Paced Courses](#)

[Instructor-Led Workshops](#)

[Educator Programs](#)

[Enterprise Solutions](#)

[Certification](#)

[Resources](#)

Self-paced Course

Generative AI Explained

In this no-coding course, learn Generative AI concepts and applications, as well as the challenges and opportunities in this exciting field.

[About Course](#)

[Objectives](#)

[Topics Covered](#)

[Course Outline](#)

[Stay Informed](#)

[Contact Us](#)

[Continue Learning](#)

About this Course

Generative AI describes technologies that are used to generate new content based on a variety of inputs. In recent time, Generative AI involves the use of neural networks to identify patterns and structures within existing data to generate new content. In this course, you will learn Generative AI concepts, applications, as well as the challenges and opportunities in this exciting field.

Learning Objectives

Upon completion, you will have a basic understanding of Generative AI and be able to more effectively use the various tools built on this

Course Details

Duration: 02:00

Price: Free

Level: Technical - Beginner

Subject: Generative AI/LLM

Language: English

https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-15+V1

Introduction to Transformer-Based Natural Language Processing

Self-paced Course

Introduction to Transformer-Based Natural Language Processing

Learn how Transformers are used as the building blocks of modern large language models (LLMs). You'll then use these models for various NLP tasks, including text classification, named-entity recognition (NER), author attribution, and question answering.

About Course

[Objectives](#)

[Topics Covered](#)

[Course Outline](#)

[Stay Informed](#)

[Contact Us](#)

Continue Learning

About this Course

Large Language Models (LLMs), or Transformers, have revolutionized the field of natural language processing (NLP). Driven by recent advancements, applications of NLP and generative AI have exploded in the past decade. With the proliferation of applications like chatbots and intelligent virtual assistants, organizations are infusing their businesses with more interactive human-machine experiences. Understanding how Transformer-based large language models (LLMs) can be used to manipulate, analyze, and generate text-based data is essential. Modern pre-trained LLMs can encapsulate the nuance, context, and sophistication of language, just as humans do. When fine-tuned and deployed correctly, developers can use these LLMs to build powerful NLP applications that provide natural and seamless human-computer interactions within chatbots, AI voice agents, and more. In this course, you'll learn how Transformers are used as the building blocks of modern large language models (LLMs). You'll then use these models for various NLP

Course Details

Duration: 06:00

Price: \$30

Level: Technical - Beginner

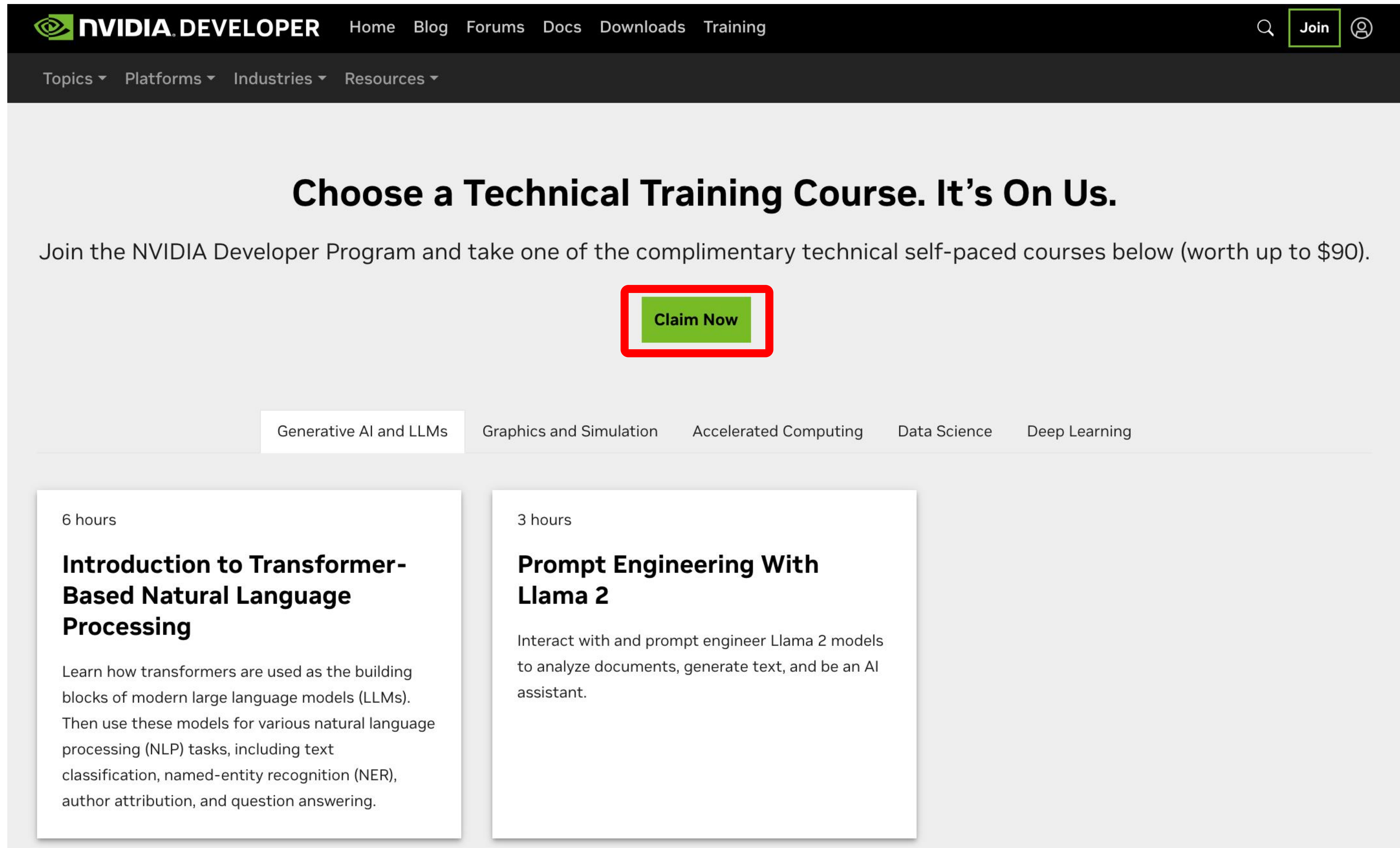
Subject: Generative AI/LLM

Language: English

https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-08+V1

Join the NVIDIA Developer Program

take one of the complimentary technical self-paced courses (worth up to \$90)



The screenshot shows the NVIDIA Developer Program website. At the top is a black navigation bar with the NVIDIA Developer logo on the left and links for Home, Blog, Forums, Docs, Downloads, and Training on the right. A search icon and a 'Join' button are also present. Below the navigation bar is a dark grey bar with dropdown menus for Topics, Platforms, Industries, and Resources. The main content area has a light grey background. It features a large heading 'Choose a Technical Training Course. It's On Us.' followed by a paragraph: 'Join the NVIDIA Developer Program and take one of the complimentary technical self-paced courses below (worth up to \$90).' Below this is a green 'Claim Now' button with a red border. Underneath the button is a horizontal menu with five categories: 'Generative AI and LLMs' (which is highlighted with a white background), 'Graphics and Simulation', 'Accelerated Computing', 'Data Science', and 'Deep Learning'. Below the menu are two course cards. The first card is titled 'Introduction to Transformer-Based Natural Language Processing', has a duration of '6 hours', and describes learning about transformers and their applications in NLP tasks like text classification and question answering. The second card is titled 'Prompt Engineering With Llama 2', has a duration of '3 hours', and describes interacting with and prompting Llama 2 models for tasks like document analysis and AI assistance.

Choose a Technical Training Course. It's On Us.

Join the NVIDIA Developer Program and take one of the complimentary technical self-paced courses below (worth up to \$90).

Claim Now

Generative AI and LLMs Graphics and Simulation Accelerated Computing Data Science Deep Learning

6 hours

Introduction to Transformer-Based Natural Language Processing

Learn how transformers are used as the building blocks of modern large language models (LLMs). Then use these models for various natural language processing (NLP) tasks, including text classification, named-entity recognition (NER), author attribution, and question answering.

3 hours

Prompt Engineering With Llama 2

Interact with and prompt engineer Llama 2 models to analyze documents, generate text, and be an AI assistant.

<https://developer.nvidia.com/join-nvidia-developer-program>



Join the NVIDIA Developer Program

take one of the complimentary technical self-paced courses (worth up to \$90)

NVIDIA

NVIDIA

Sign in or sign up with your email address

Email Address

user@domain.com

Continue

[Privacy Policy](#) | [Legal Info](#) | [Contact Us](#)

<https://developer.nvidia.com/join-nvidia-developer-program>

Join the NVIDIA Developer Program

take one of the complimentary technical self-paced courses (worth up to \$90)

Date of birth
1970-00-20

Password
.....

Confirm password
.....

☒ Stay logged in

☐ By proceeding, I agree to the [Privacy Policy](#) and [Legal Info](#)

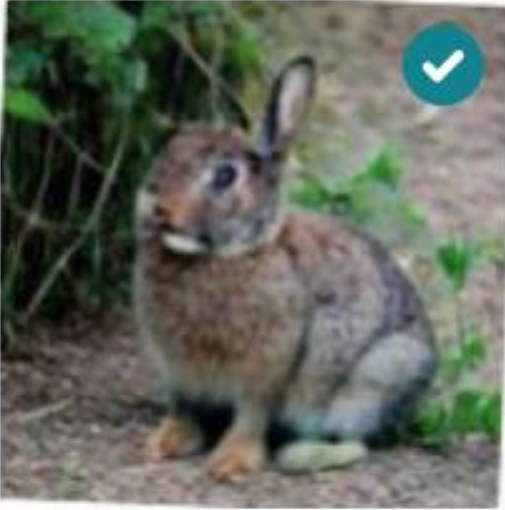
EN

Privacy Policy | Legal Info | Contact Us

Copyright © 2024 NVIDIA Corporation

Select the images matching the toy in the example image.
If there are None, click Skip



EN

⋮ ↺  Verify

English (US) ▾

<https://developer.nvidia.com/join-nvidia-developer-program>

Join the NVIDIA Developer Program

take one of the complimentary technical self-paced courses (worth up to \$90)

 NVIDIA DEVELOPER DEEP LEARNING INSTITUTE PROGRAM BENEFITS

SELECT YOUR FREE COURSE.

Thank you for your participation in the NVIDIA Developer Program. Please select your free DLI course below.

English ▾

Integrating Sensors with NVIDIA DRIVE®

Getting Started with Deep Learning

Deploying a Model for Inference at Production Scale

Get Started with Highly Accurate Custom ASR for Speech AI

Introduction to Graph Neural Networks

Introduction to Transformer-Based Natural Language Processing

Prompt Engineering with LLaMA-2 (Access Expires Dec. 5th 2025)

Generative AI with Diffusion Models

Building Real-Time Video AI Applications

Introduction to Robotic Simulations in Isaac Sim

Generative AI with Diffusion Models

Take a deeper dive into denoising diffusion models, which are a popular choice for text-to-image pipelines, with applications in creative content generation, data augmentation, simulation and planning, anomaly detection, drug discovery, personalized recommendations, and more.

 Certificate Available

 Duration: 08:00

Continue >



Join the NVIDIA Developer Program

take one of the complimentary technical self-paced courses (worth up to \$90)

 **NVIDIA. DEVELOPER** DEEP LEARNING INSTITUTE PROGRAM BENEFITS

SELECT YOUR FREE COURSE.

Thank you for your participation in the NVIDIA Developer Program. Please select your free DLI course below.

English ▾

Modeling Time Series Data with Recurrent Neural Networks in Keras (Access ends 10/16/2024)

Optimizing CUDA Machine Learning Codes With Nsight Profiling Tools

Getting Started with Accelerated Computing in CUDA C/C++

Fundamentals of Accelerated Computing with CUDA Python

Fundamentals of Accelerated Computing with OpenACC

Integrating Sensors with NVIDIA DRIVE®

Getting Started with Deep Learning

Deploying a Model for Inference at Production Scale

Get Started with Highly Accurate Custom ASR for Speech AI

Getting Started with Deep Learning

Learn how deep learning works through hands-on exercises in computer vision and natural language processing.

 Certificate Available

 Duration: 08:00

Continue >

<https://developer.nvidia.com/join-nvidia-developer-program>



NVIDIA Deep Learning Institute (DLI)

Deep Learning Institute

Find Training

Self Paced Courses

Instructor-Led Workshops

Educator Programs

Enterprise Solutions

Certification

Resources

Search



Monthly Activity

Skill Points 0

Time Spent

Courses in Progress 1

Courses Completed 0

Watched Videos

Assessments

Skills

Certificates



No Certificates

You don't have any certificates yet.

Courses in Progress

Self-paced

Generative AI with Diffusion Models

0% Completed

08:00

<https://learn.nvidia.com/my-learning>

Generative AI with Diffusion Models

Course Progress Bookmarks Updates

Generative AI with Diffusion Models > Start Here > 0: Server Access

Generative AI with Diffusion Models

Start Here

Next Steps

Feedback

[Previous](#)[Next](#)

0: Server Access

 [Bookmark this page](#)

Welcome to Generative AI with Diffusion Models. Please click "Next" below to get started.

Underneath each video is a link to start your own private server for hands-on coding practice. Click the "Start" button to boot up the server. In a few minutes after the server is done loading, click "Launch" to access the code labs.

1: From U-Nets to Diffusion

 [Bookmark this page](#)

Theory

<https://learn.nvidia.com/my-learning>



NVIDIA Deep Learning Institute (DLI)



Products Solutions Industries For You

Shop Drivers Support Q

Min-Yuh Day

Deep Learning Institute

Find Training

Self Paced Courses

Instructor-Led Workshops

Educator Programs

Enterprise Solutions

Certification

Training Home

My Learning

NVIDIA Account

Logout

Generative AI with Diffusion Models

Course

Progress

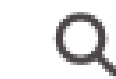
Bookmarks

Updates

<https://learn.nvidia.com/my-learning>

Deep Learning Institute (DLI)

Search



Monthly Activity

Skill Points	0
Time Spent	
Courses in Progress	16
Courses Completed	12
Watched Videos	
Assessments	

Skills

Certificates



Introduction to
Transformer-
Based Natural
Language
Processing



Building RAG
Agents with
LLMs



Building RAG
Agents with
LLMs



Accelerating
End-to-End Data
Science
Workflows



Generative AI
with Diffusion
Models



Building Agentic
AI Applications
with LLMs

Completed Courses

View more < >

Accelerating End-to-End
Data Science Workflows

100% Completed

Rapid Application
Development with Large
Language Models (LLMs)

100% Completed

Sizing LLM Inference Systems

100% Completed

Building RAG Agents with
LLMs

100% Completed

Generative AI with Diffusion
Models

100% Completed

Certificates



Introduction to
Transformer-
Based Natural
Language
Processing



Building RAG
Agents with
LLMs



Building RAG
Agents with
LLMs



Accelerating
End-to-End Data
Science
Workflows



Generative AI
with Diffusion
Models



Building Agentic
AI Applications
with LLMs



Building Agentic
AI Applications
with LLMs



Rapid
Application
Development
with Large
Language
Models (LLMs)



Getting Started
with Deep
Learning

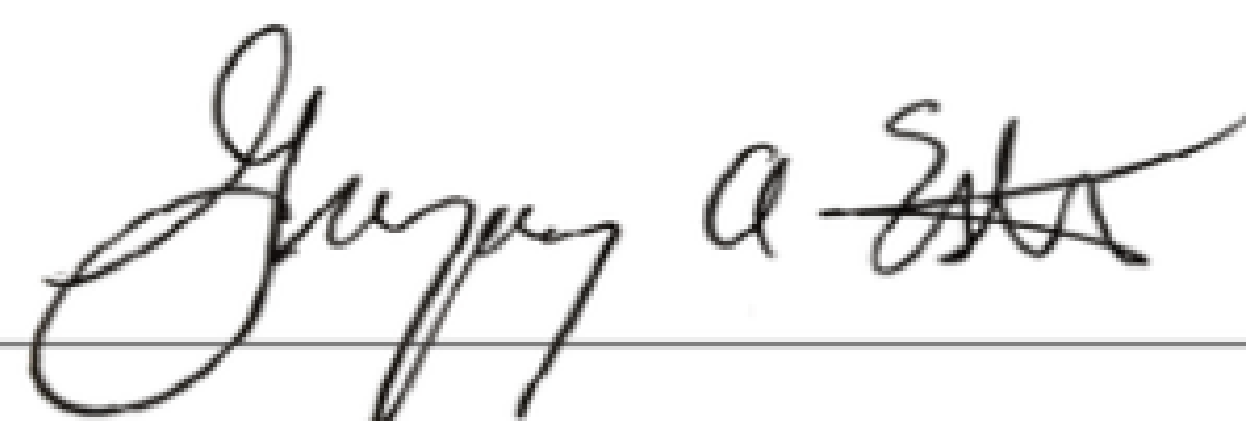
Certificate of Competency

This certificate is awarded to

Min-Yuh Day

for demonstrating competence in the completion of

Generative AI with Diffusion Models



Greg Estes

Vice President, NVIDIA

Issue Date: : February 28, 2025

Certification ID: q3oo-oBhTQKtyCCote2E-Q

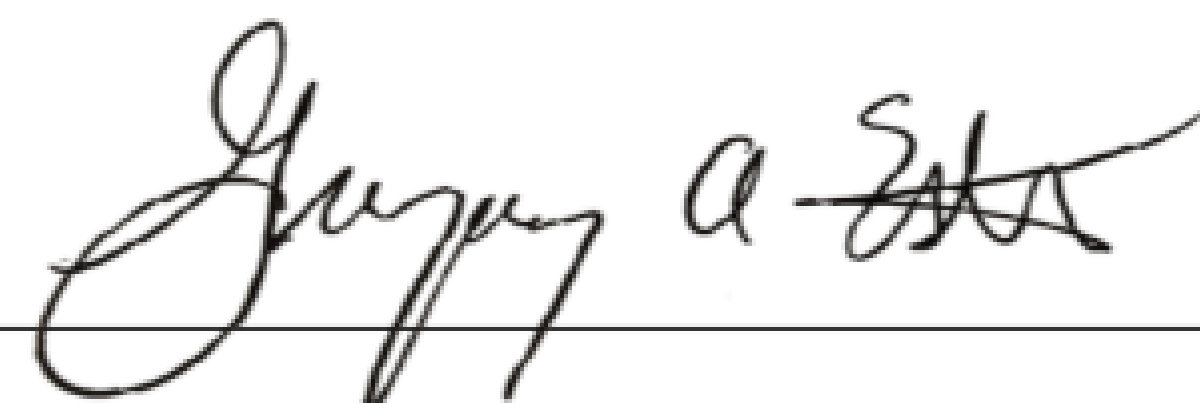
Certificate of Completion

This certificate is awarded to

Min-Yuh Day

for successfully completing

Building RAG Agents with LLMs



Greg Estes

Vice President, NVIDIA

Issue Date: : December 8, 2024

Certification ID: ed-qOCIMQaatzU8SNUNxgw |

https://learn.nvidia.com/certificates?id=ed-qOCIMQaatzU8SNUNxgw/courses/course?course_id=course-v1:DLI+S-FX-15+V1

<https://learn.nvidia.com/certificates?id=ed-qOCIMQaatzU8SNUNxgw>

All Self-Paced Courses

[Accelerated Computing](#)
[Data Science](#)
[Deep Learning](#)
[Generative AI/LLM](#)
[Graphics and Simulation](#)
[Infrastructure](#)
[Share Generative AI/LLM Courses](#)

Self-paced

Generative AI Explained

Free
02:00

Self-paced

**Introduction to NVIDIA NIM™
Microservices**

Free
02:00

Self-paced

**Introduction to Deploying
RAG Pipelines for Production
at Scale**

\$90
03:00

Self-paced

**Generative AI with Diffusion
Models**

\$90
08:00

Self-paced

**Techniques for Improving the
Effectiveness of RAG Systems**

\$30
03:00

Self-paced

**Introduction to Transformer-
Based Natural Language
Processing**

\$30
06:00

Self-paced

**Building LLM Applications
With Prompt Engineering**

\$90
08:00

Self-paced

**Synthetic Tabular Data
Generation Using
Transformers**

\$30
04:00

Self-paced

Sizing LLM Inference Systems

Free
03:00

Self-paced

**Building RAG Agents with
LLMs**

Free
08:00

Self-paced

**Augment your LLM Using
Retrieval Augmented
Generation**

Free
01:00

NVIDIA Deep Learning Institute (DLI)

Generative AI with Diffusion Models

Self-paced Course

Generative AI with Diffusion Models

Take a deeper dive into denoising diffusion models, which are a popular choice for text-to-image pipelines, with applications in creative content generation, data augmentation, simulation and planning, anomaly detection, drug discovery, personalized recommendations, and more.

About Course

Objectives

Topics Covered

Course Outline

Stay Informed

Contact Us

Continue Learning

About this Course

Thanks to improvements in computing power and scientific theory, generative AI is more accessible than ever before. Generative AI plays a significant role across industries due to its numerous applications, such as creative content generation, data augmentation, simulation and planning, anomaly detection, drug discovery, personalized recommendations, and more. In this course, learners will take a deeper dive into denoising diffusion models, which are a popular choice for text-to-image pipelines.

Learning Objectives

Course Details

Duration: 08:00

Price: \$90

Subject: Generative AI/LLM

Language: English

Course Prerequisites:

A basic understanding of [Deep Learning Concepts](#).

https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1

Generative AI with Diffusion Models

Deep Learning Institute

[Find Training](#)

[Self Paced Courses](#)

[Instructor-Led Workshops](#)

[Educator Programs](#)

[Enterprise Solutions](#)

[Certification](#)

[Resources](#)

Course

[Progress](#)

[Bookmarks](#)

[Updates](#)

Generative AI with Diffusion Models

[Start Here](#)

1: From U-Nets to Diffusion

Generative AI with Diffusion Models

[Start Here](#)

0: Server Access

1: From U-Nets to Diffusion

2: Denoising Diffusion Probabilistic Models

3: Optimizations

4: Classifier-Free Diffusion Guidance

5: CLIP

6: Wrap-up and Assessment

[Next Steps](#)

[Feedback](#)

[Previous](#)

[Next](#)



▶ 17:54 / 17:55

▶ Speed 1.0x



DEEP
LEARNING
INSTITUTE



START

Please click the start button to the top-right. The course will take about 5 minutes to load. The course lecture slide decks will appear here when you launch the course. The numbers on the lectures correspond to the numbers of the interactive notebooks in the course.

1: From U-Nets to Diffusion

Generative AI with Diffusion Models

 Products Solutions Industries For You

Shop Drivers Support    Min-Yuh Day

Deep Learning Institute Find Training Self Paced Courses Instructor-Led Workshops Educator Programs Enterprise Solutions Certification Resources

Generative AI with Diffusion Models

Course Progress Bookmarks Updates

Generative AI with Diffusion Models Start Here 1: From U-Nets to Diffusion

Generative AI with Diffusion Models

Start Here

0: Server Access

1: From U-Nets to Diffusion

2: Denoising Diffusion Probabilistic Models

3: Optimizations

4: Classifier-Free Diffusion Guidance

5: CLIP

6: Wrap-up and Assessment

Next Steps

Feedback



 DEEP LEARNING INSTITUTE

This Lab 0 : 10 : 54 / 1 : 30 : 00

Course 8 : 18 : 49 / 24 : 00 : 00

 LAUNCH

 STOP TASK

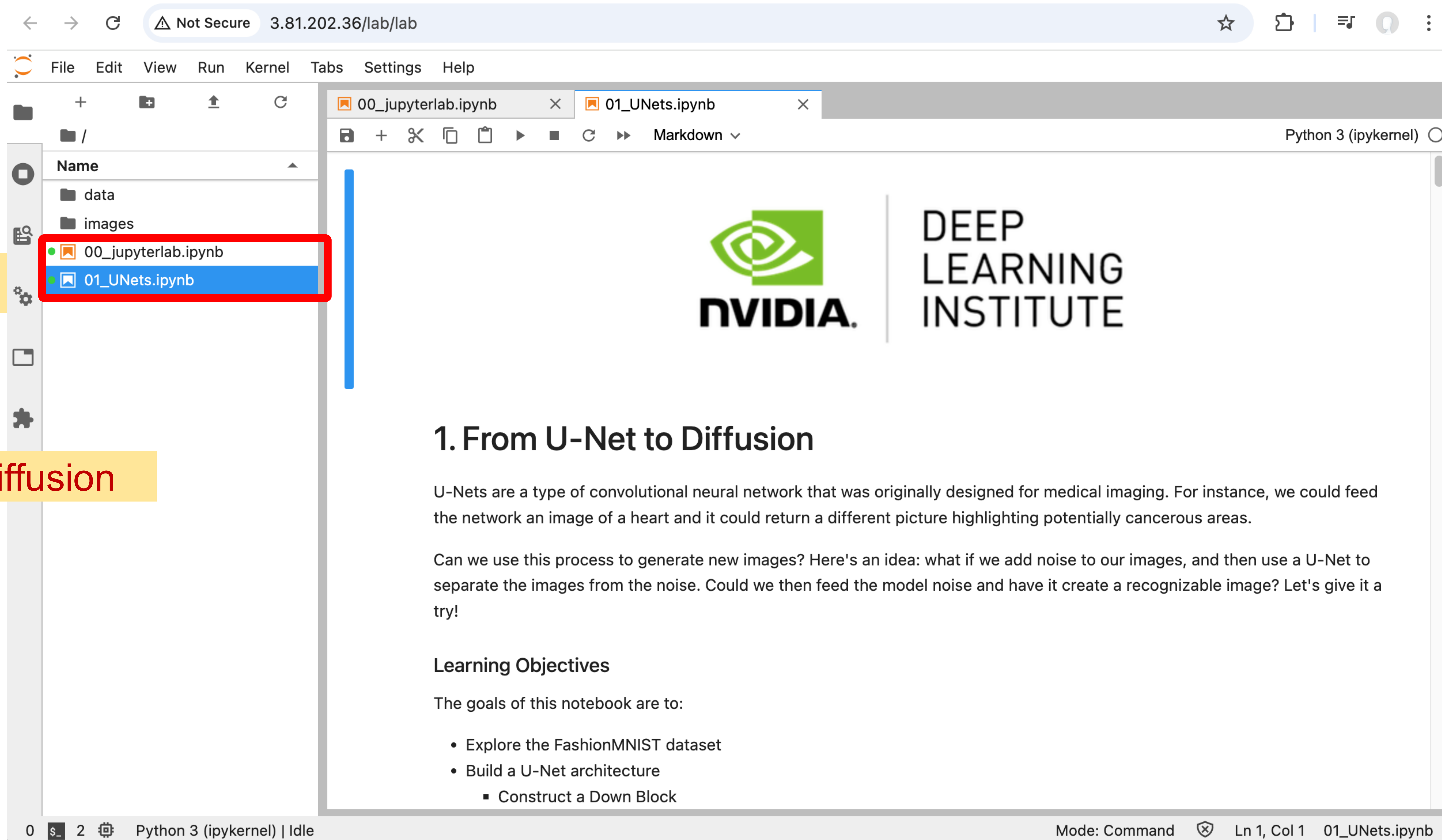
**Generative AI with Diffusion Models**
Part 1: From U-Nets to Diffusion

1: From U-Nets to Diffusion

Generative AI with Diffusion Models

01_UNets.ipynb

1: From U-Nets to Diffusion



00_jupyterlab.ipynb 01_UNets.ipynb

Python 3 (ipykernel)

DEEP LEARNING INSTITUTE

1. From U-Net to Diffusion

U-Nets are a type of convolutional neural network that was originally designed for medical imaging. For instance, we could feed the network an image of a heart and it could return a different picture highlighting potentially cancerous areas.

Can we use this process to generate new images? Here's an idea: what if we add noise to our images, and then use a U-Net to separate the images from the noise. Could we then feed the model noise and have it create a recognizable image? Let's give it a try!

Learning Objectives

The goals of this notebook are to:

- Explore the FashionMNIST dataset
- Build a U-Net architecture
 - Construct a Down Block

0 \$ 2 Python 3 (ipykernel) | Idle Mode: Command Ln 1, Col 1 01_UNets.ipynb



Generative AI with Diffusion Models

Part 1: From U-Nets to Diffusion

Agenda

- Part 1: From U-Nets to Diffusion
- Part 2: Denoising Diffusion Probabilistic Models
- Part 3: Optimizations
- Part 4: Classifier Free Diffusion
- Part 5: CLIP
- Part 6: Wrap-up & Assessment

Prerequisites

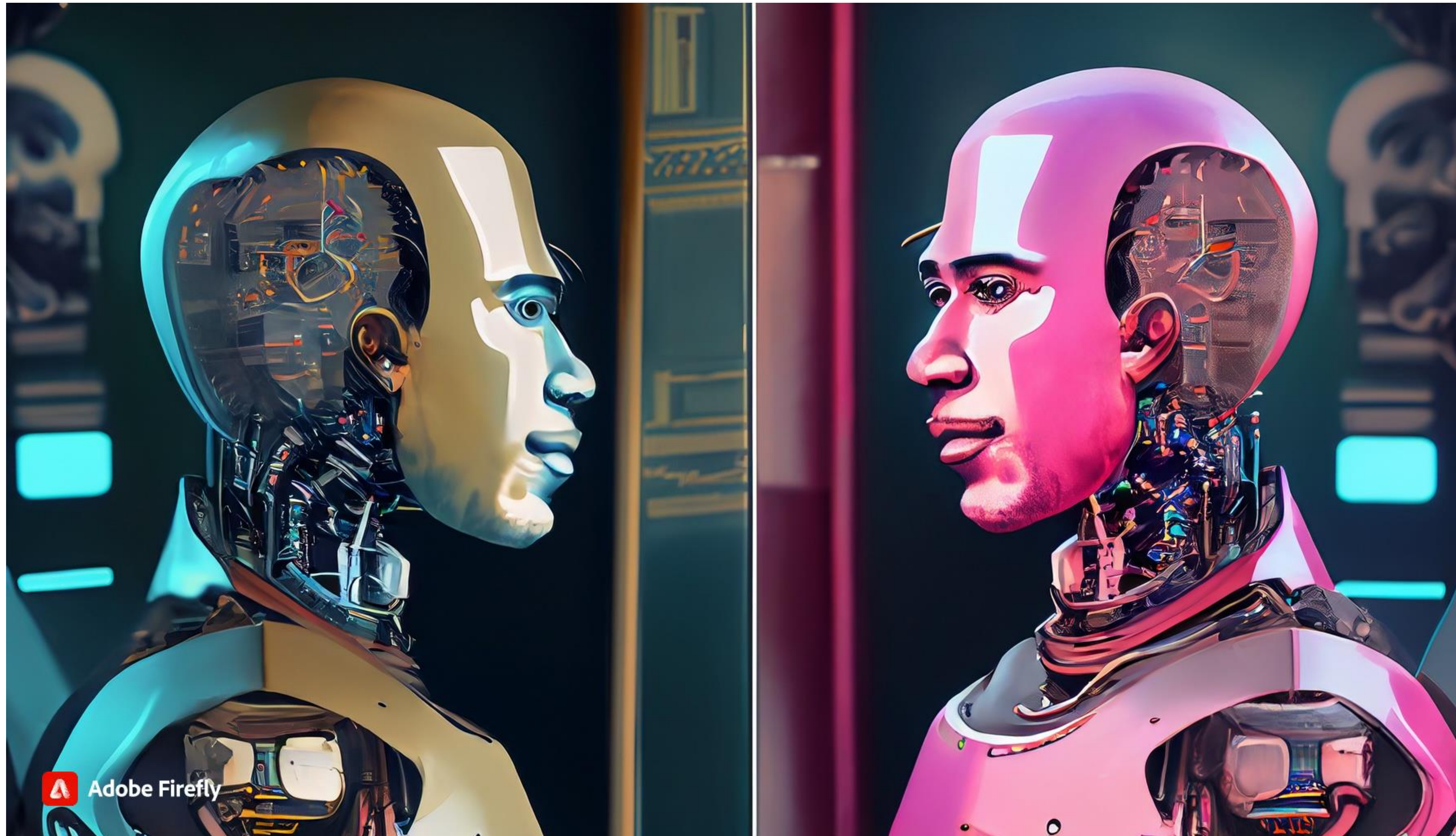
- Basic familiarity with convolutional neural networks (CNNs)
- Basic familiarity with a deep learning framework such as:
 - PyTorch
 - TensorFlow



A Brief History of Generative AI

The Imitation Game

A.K.A The Turing Test



A robot looks in a mirror and the reflection is human, cyberpunk

IBM 704

The First Singing Computer



Eliza

The First Gen AI Chatbot?

Talk to Eliza

> Hello, I am Eliza.
* Nice to meet you
> Oh... ?
* How are you
> Would you prefer if I were not ?
* were not what?
> Have you asked anyone else?

Input:



Generative AI of the 70's, 80's and 90s?

- Electronic music
- Video games graphics
- Video game AI
- Computer animation
- Instant messaging chatbots

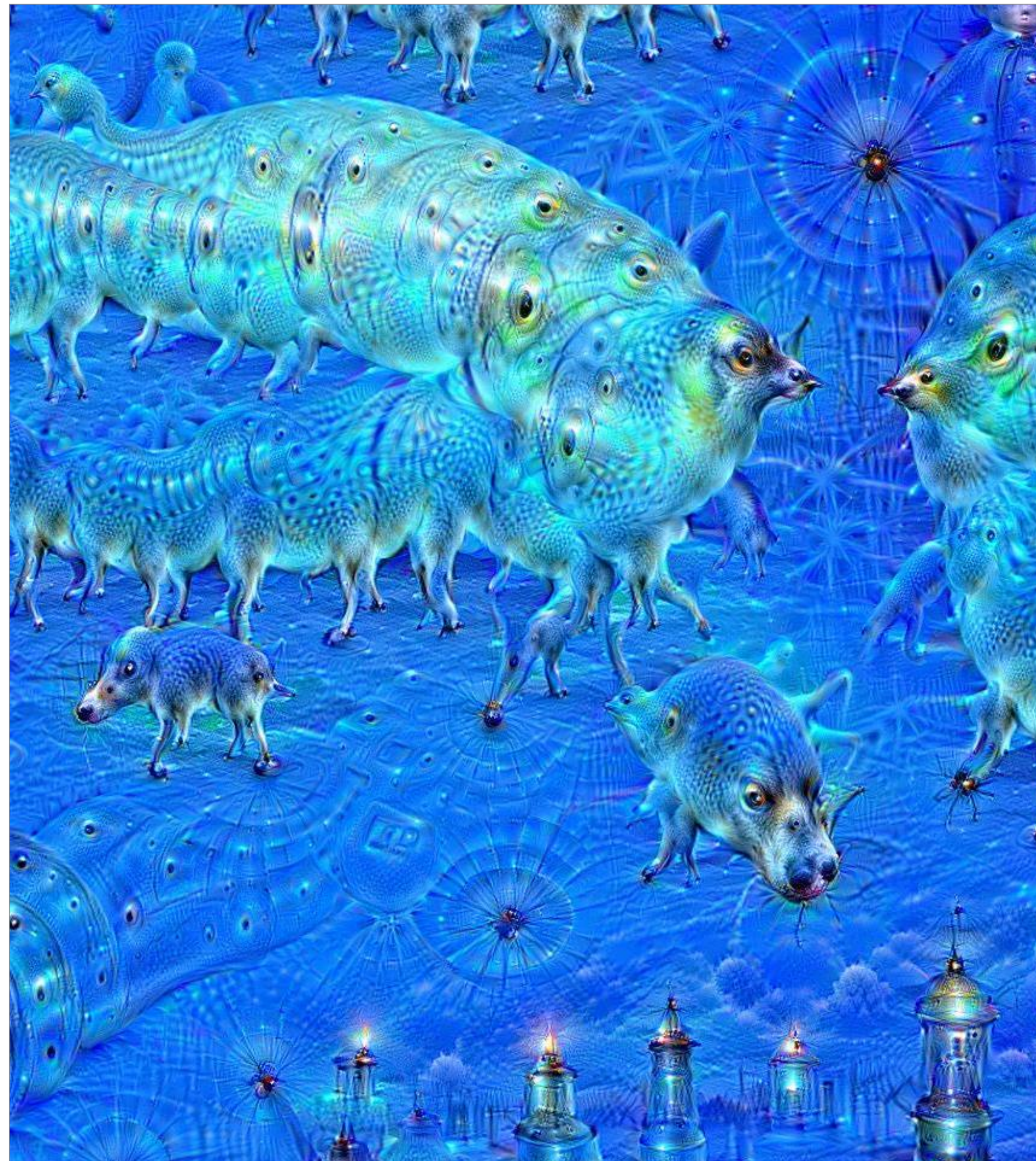
The Rise of Neural Networks

Deep Dreaming

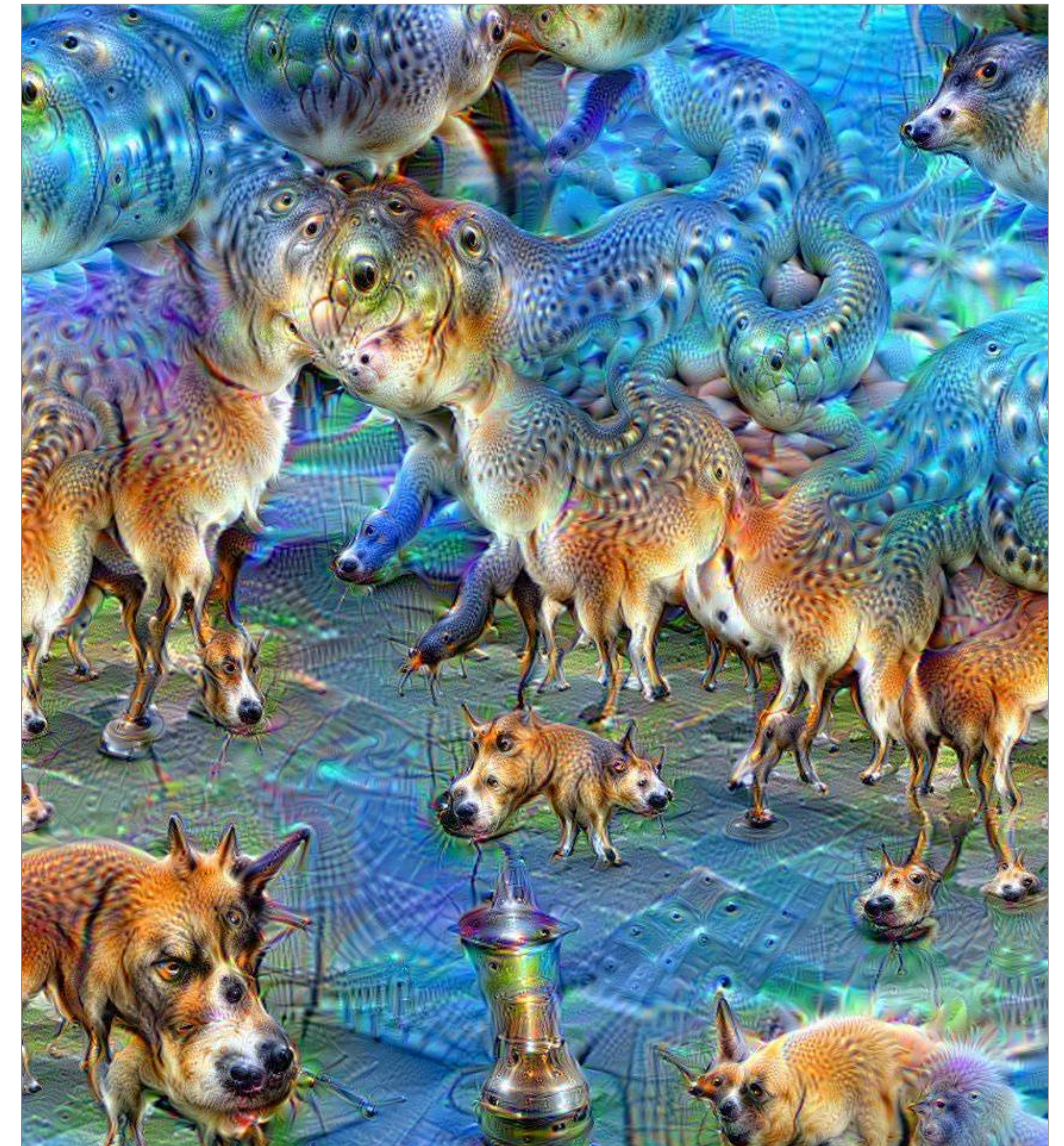
Images by Martin Thoma



Original



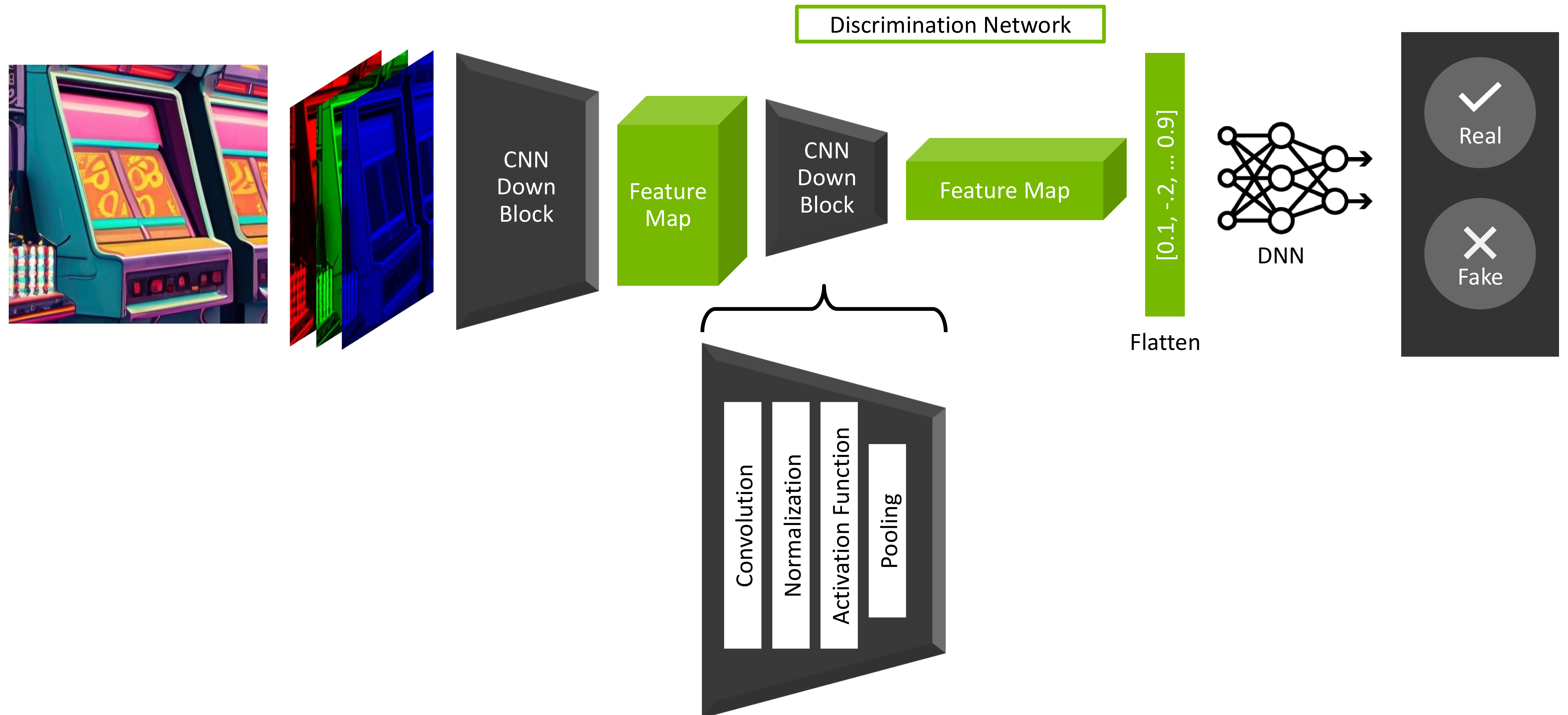
10 Iterations



50 Iterations

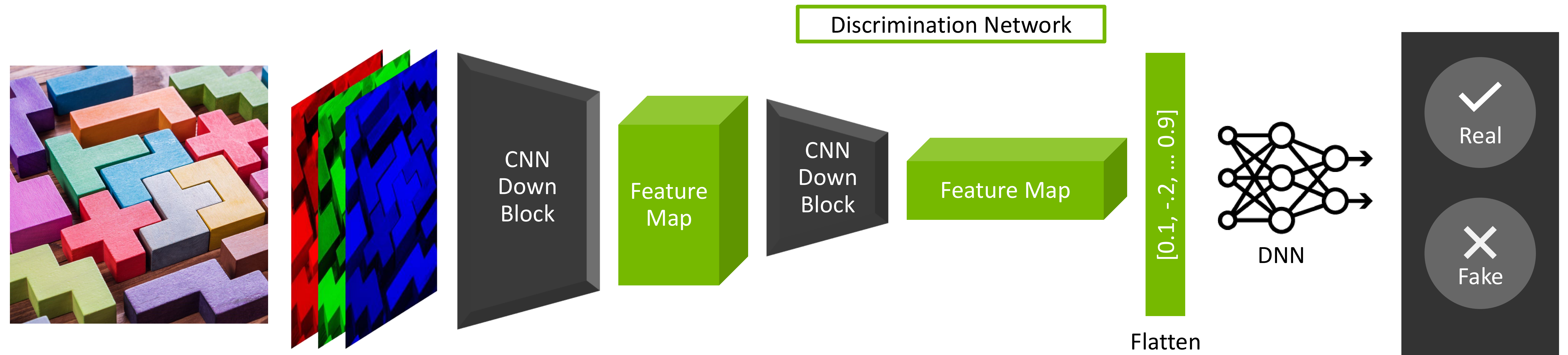
GANs

Generative Adversarial Networks



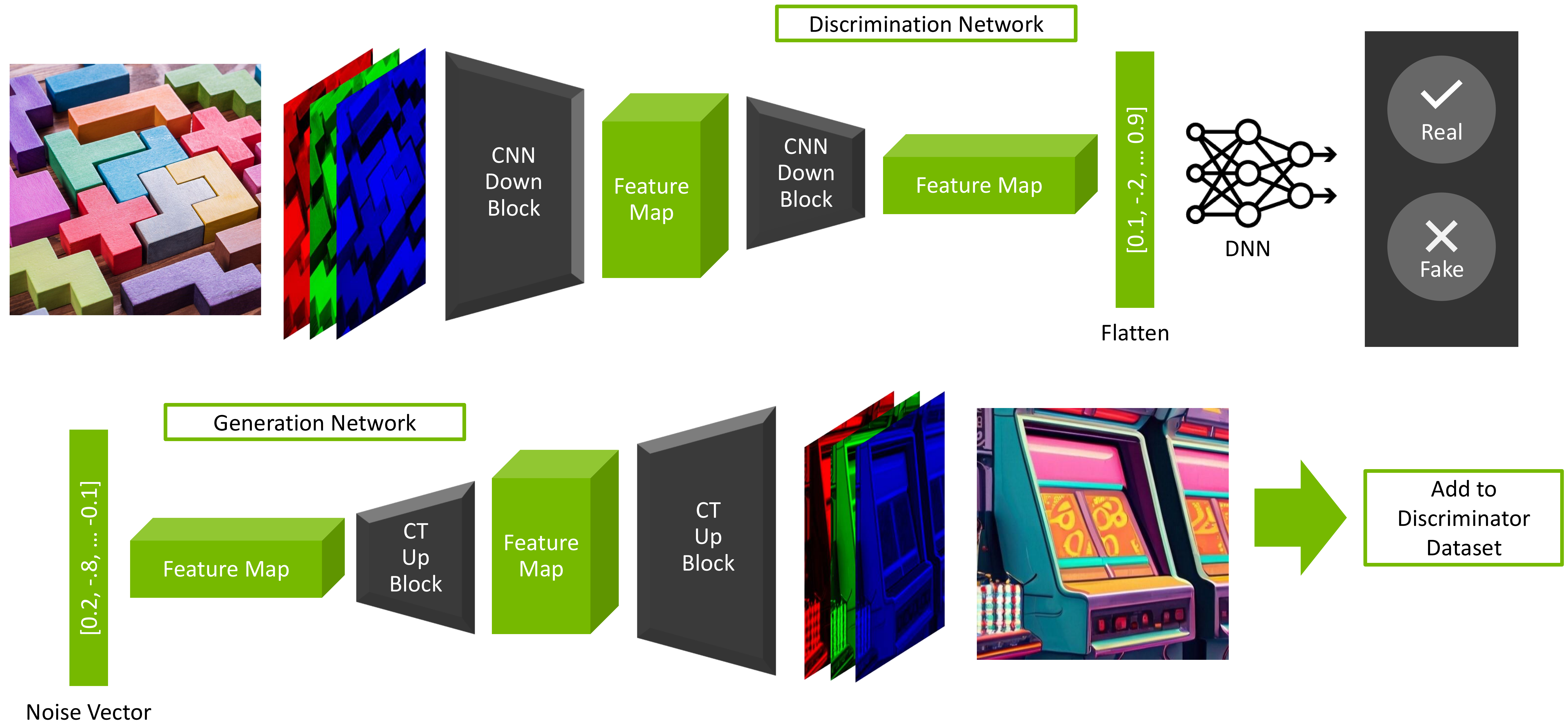
GANs

Generative Adversarial Networks



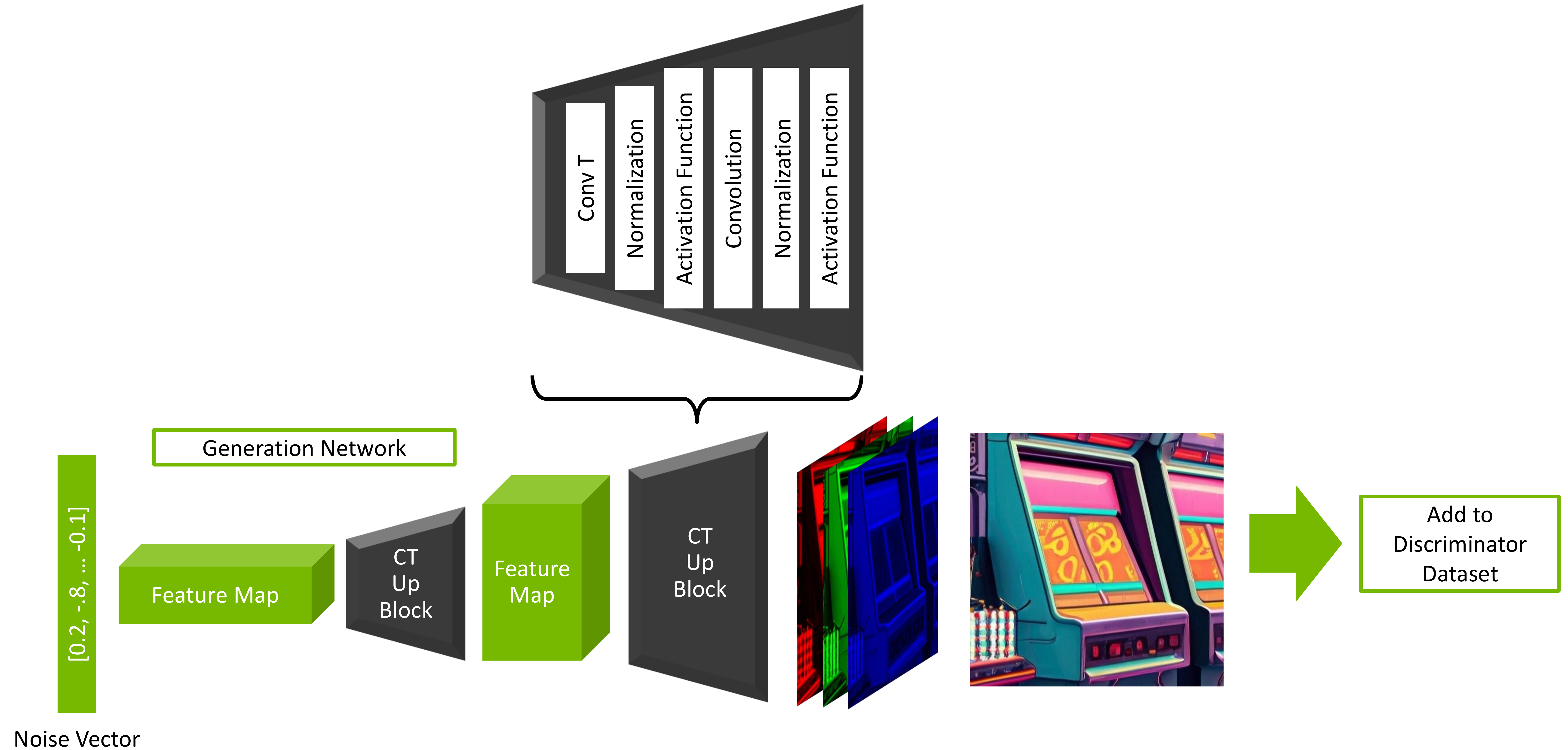
GANs

Generative Adversarial Networks



GANs

Generative Adversarial Networks



GANs

Generative Adversarial Networks

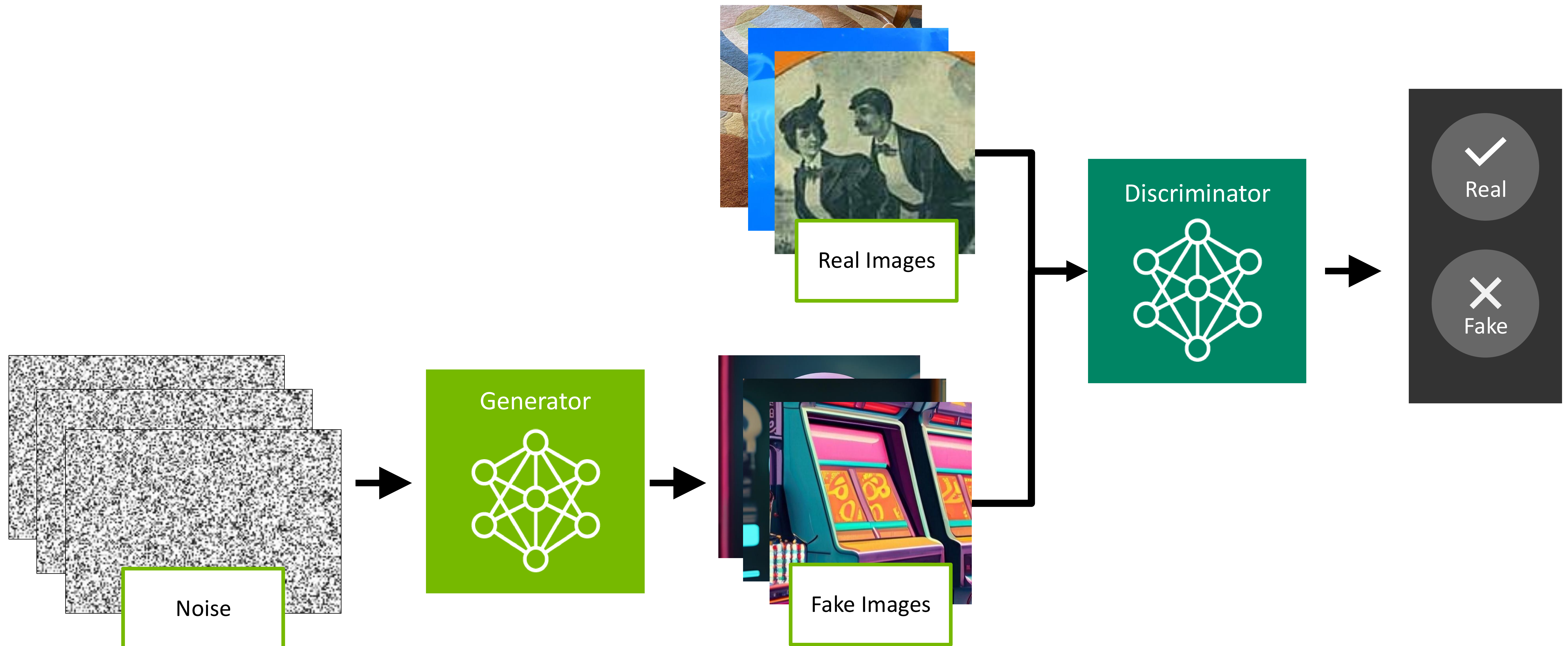


Image Segmentation

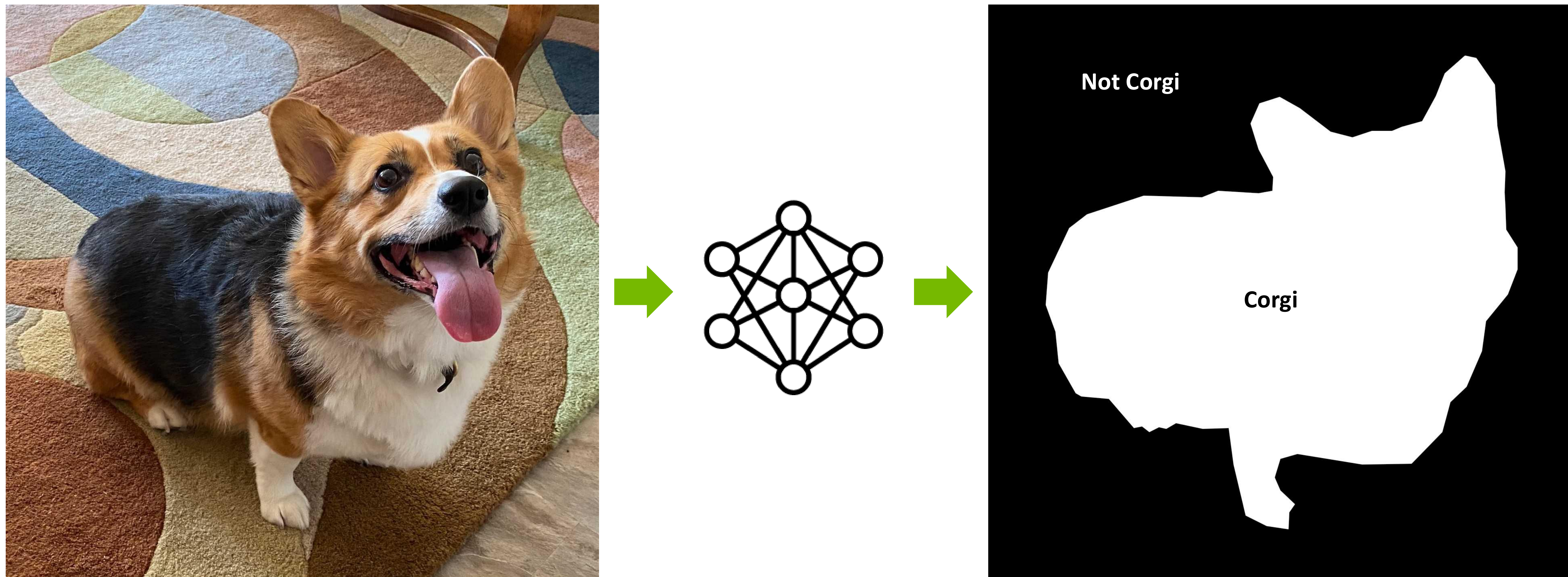


Image Segmentation + GANs

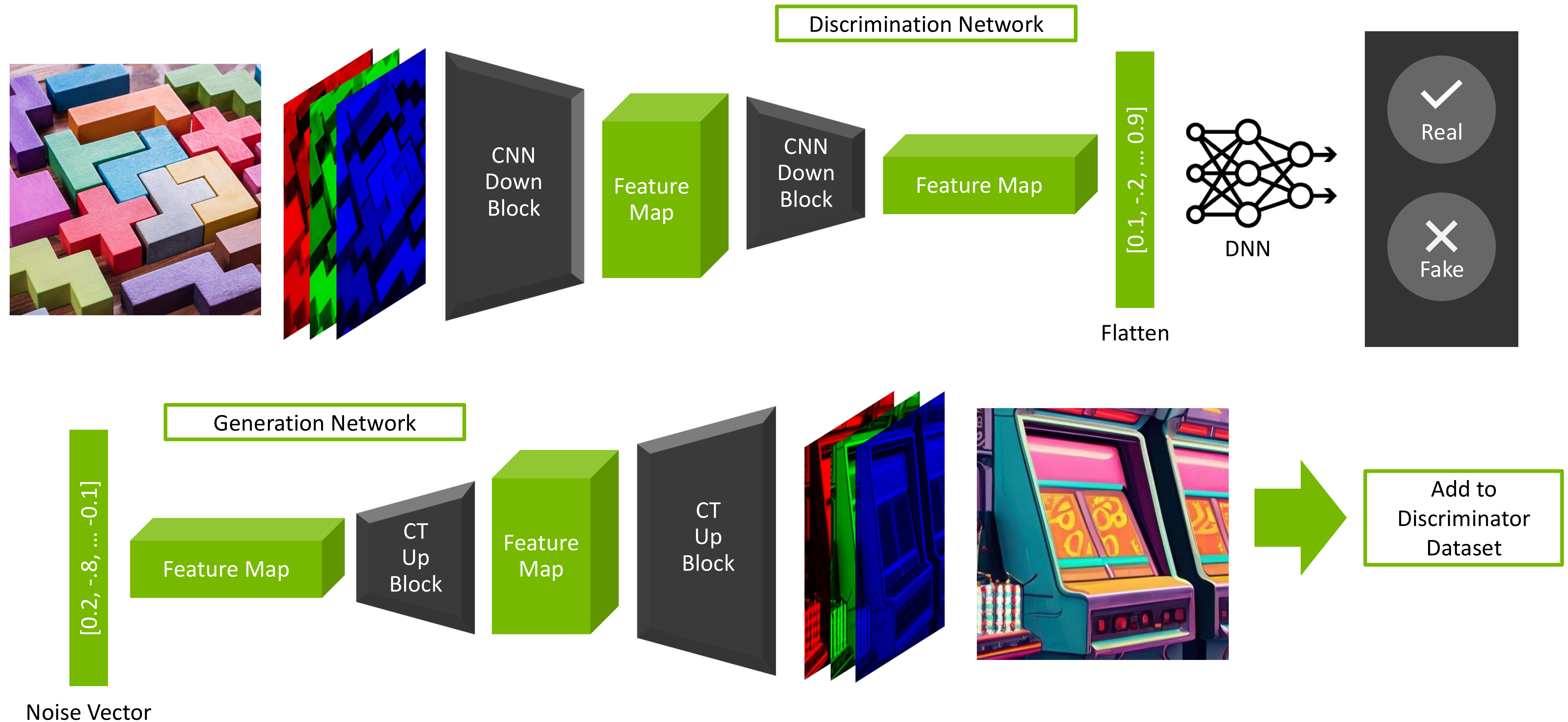
NVIDIA Spade



U-Nets

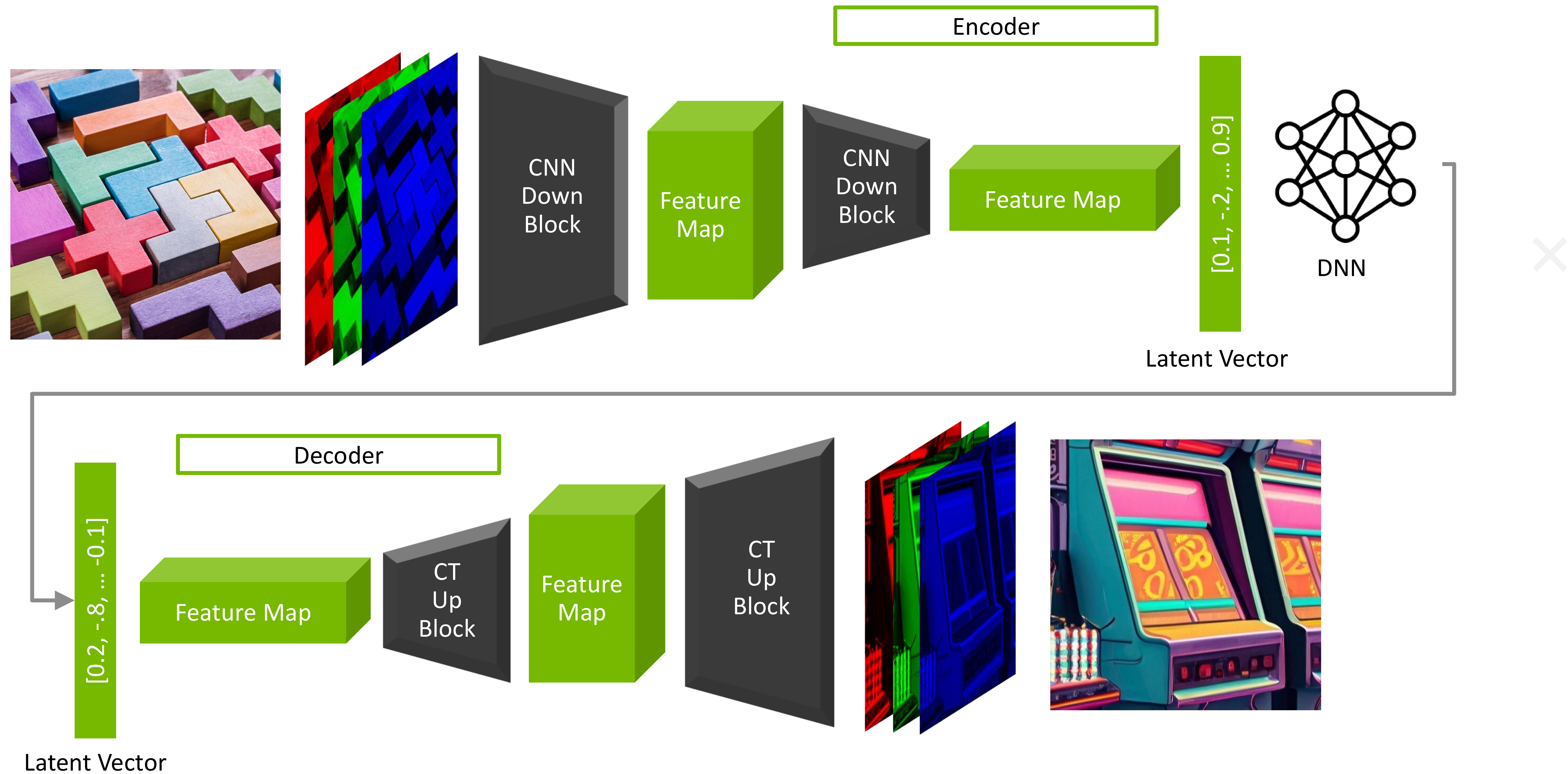
GANs

Generative Adversarial Networks



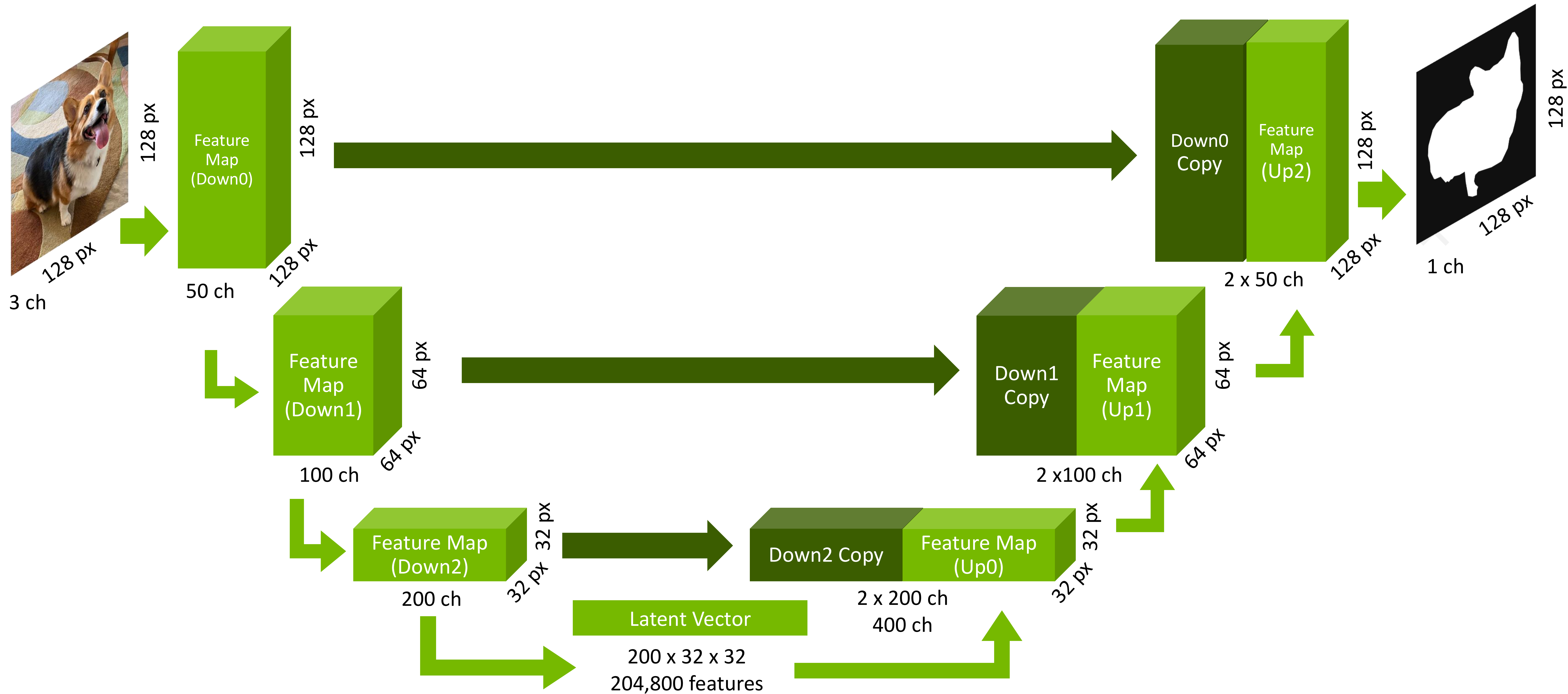
GANs U-Nets

The U shaped Autoencoder



U-Nets

The U shaped Autoencoder



Transposed Convolution

Transposed Convolution

Convolution Review

Kernel

.25	.25
.25	.25

Image

1	0	1
0	1	0
1	0	1

Output

Transposed Convolution

Convolution Review

Kernel

.25	.25
.25	.25

Image

$1 \bullet .25$	$0 \bullet .25$	1
$0 \bullet .25$	$1 \bullet .25$	0
1	0	1

Output

.5	

Transposed Convolution

Convolution Review

Kernel

.25	.25
.25	.25

Image

1	$0 \bullet .25$	$1 \bullet .25$
0	$1 \bullet .25$	$0 \bullet .25$
1	0	1

Output

.5	.5

Transposed Convolution

Convolution Review

Kernel

.25	.25
.25	.25

Image

1	0	1
$0 \bullet .25$	$1 \bullet .25$	0
$1 \bullet .25$	$0 \bullet .25$	1

Output

.5	.5
.5	

Transposed Convolution

Convolution Review

Kernel

.25	.25
.25	.25

Image

1	0	1
0	$1 \bullet .25$	$0 \bullet .25$
1	$0 \bullet .25$	$1 \bullet .25$

Output

.5	.5
.5	.5

Transposed Convolution

Image Upscaling

Kernel

.25	.25
.25	.25

Image

1	0	1
0	1	0
1	0	1

Output

Transposed Convolution

Image Upscaling

Kernel

.25	.25
.25	.25

Image

Stride = 2

1	0	0	0	1
0	0	0	0	0
0	0	1	0	0
0	0	0	0	0
1	0	0	0	1

Output

Transposed Convolution

Image Upscaling

Kernel

.25	.25
.25	.25

Image

Stride = 2

1 • .25	0 • .25	0	0	1
0 • .25	0 • .25	0	0	0
0	0	1	0	0
0	0	0	0	0
1	0	0	0	1

Output

.25			

Transposed Convolution

Image Upscaling

Kernel

.25	.25
.25	.25

Image

Stride = 2

1	0 • .25	0 • .25	0	1
0	0 • .25	0 • .25	0	0
0	0	1	0	0
0	0	0	0	0
1	0	0	0	1

Output

.25	0		

Transposed Convolution

Image Upscaling

Kernel

.25	.25
.25	.25

Image

Stride = 2

1	0	0	0	1
0	0	0	0	0
0	0	1	0	0
0	0	0	0	0
1	0	0	0	1

Output

.25	0	0	.25
0	.25	.25	0
0	.25	.25	0
.25	0	0	.25

Transposed Convolution

Stride

Image

Stride = 2

1	0	0	0	1
0	0	0	0	0
0	0	1	0	0
0	0	0	0	0
1	0	0	0	1

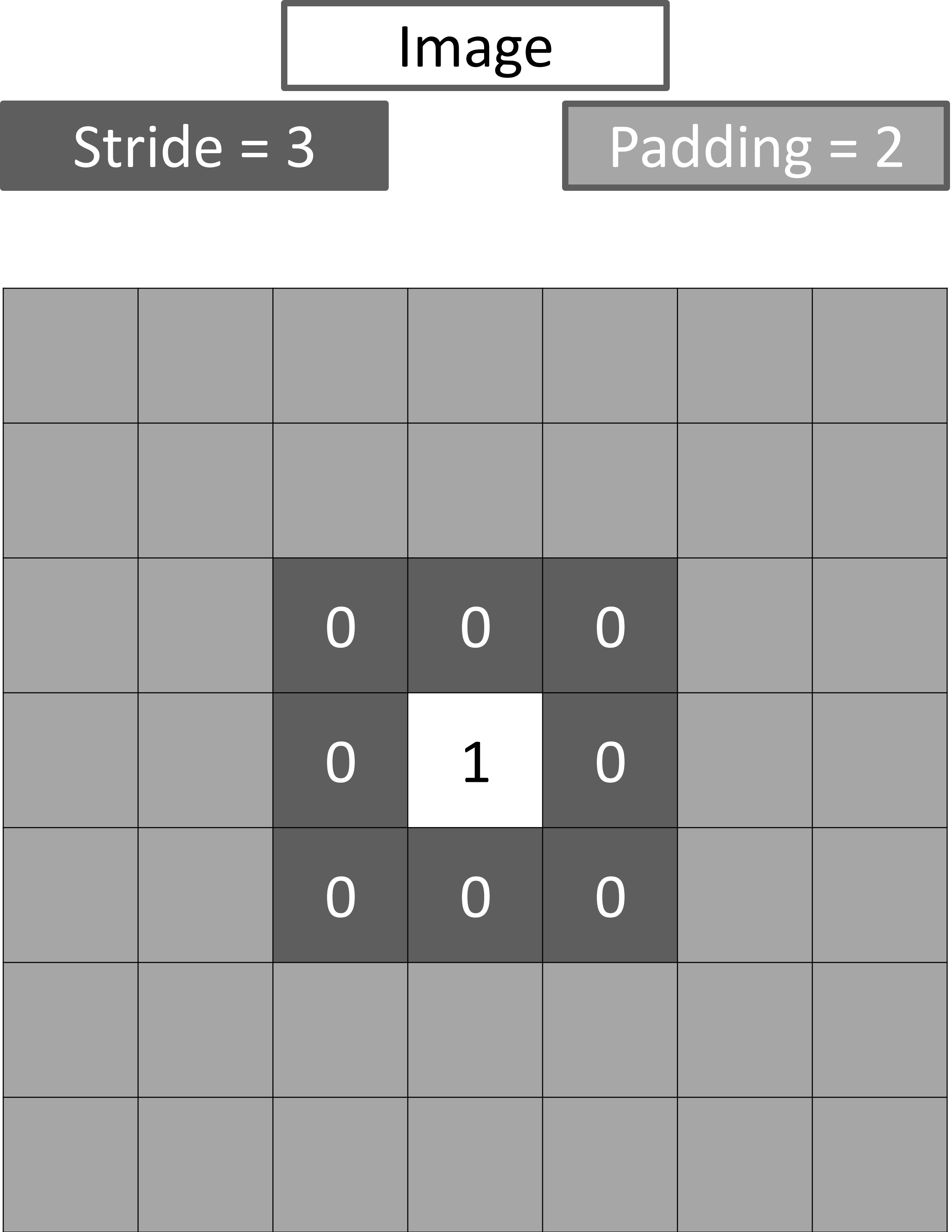
Image

Stride = 3

1	0	0	0	0	0	1
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	1	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
1	0	0	0	0	0	1

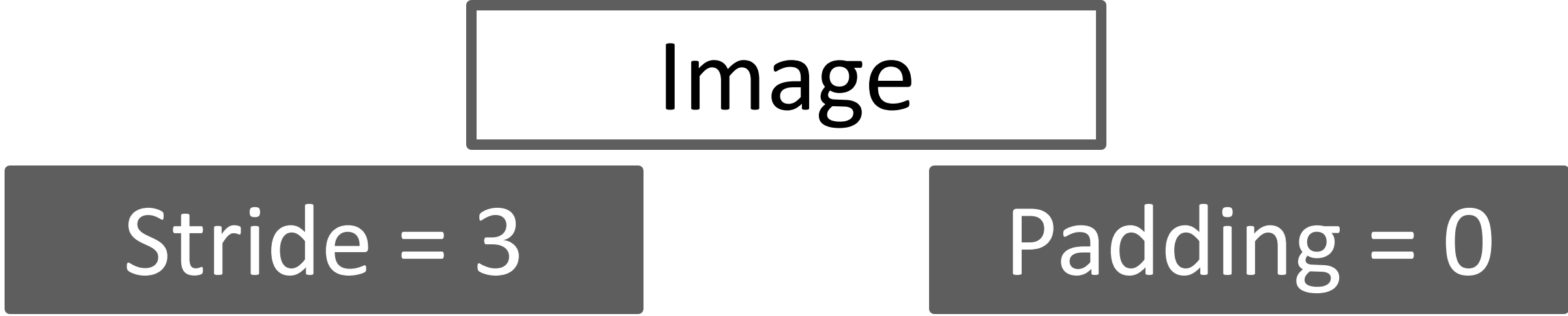
Transposed Convolution

Padding

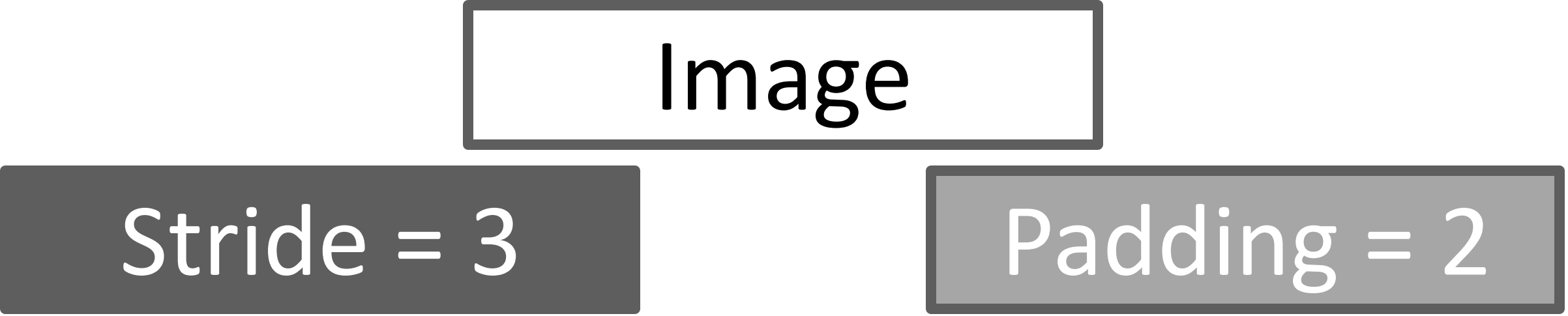


Transposed Convolution

Padding



1	0	0	0	0	0	1
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	1	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
1	0	0	0	0	0	1



0	0	0
0	1	0
0	0	0

Transposed Convolution

Out Padding

Image

Out Padding =
0

1	0	1
0	1	0
1	0	1

Image

Out Padding =
1

1	0	1	0
0	1	0	0
1	0	1	0
0	0	0	0

Image

Out Padding =
2

1	0	1	0	0
0	1	0	0	0
1	0	1	0	0
0	0	0	0	0
0	0	0	0	0

Image Resizing

Upsampling



64 px

64 px



128 px

128 px

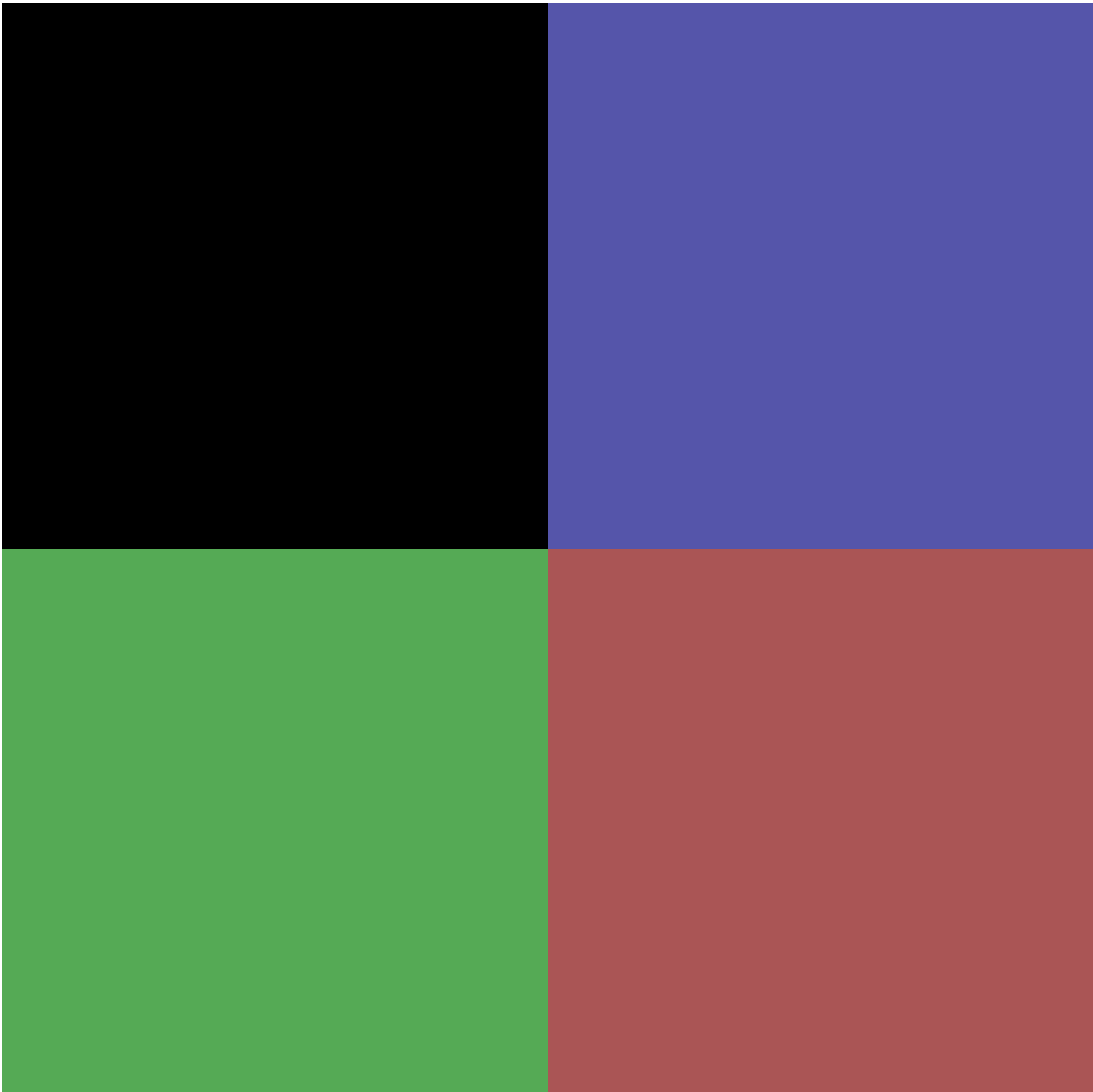


192 px

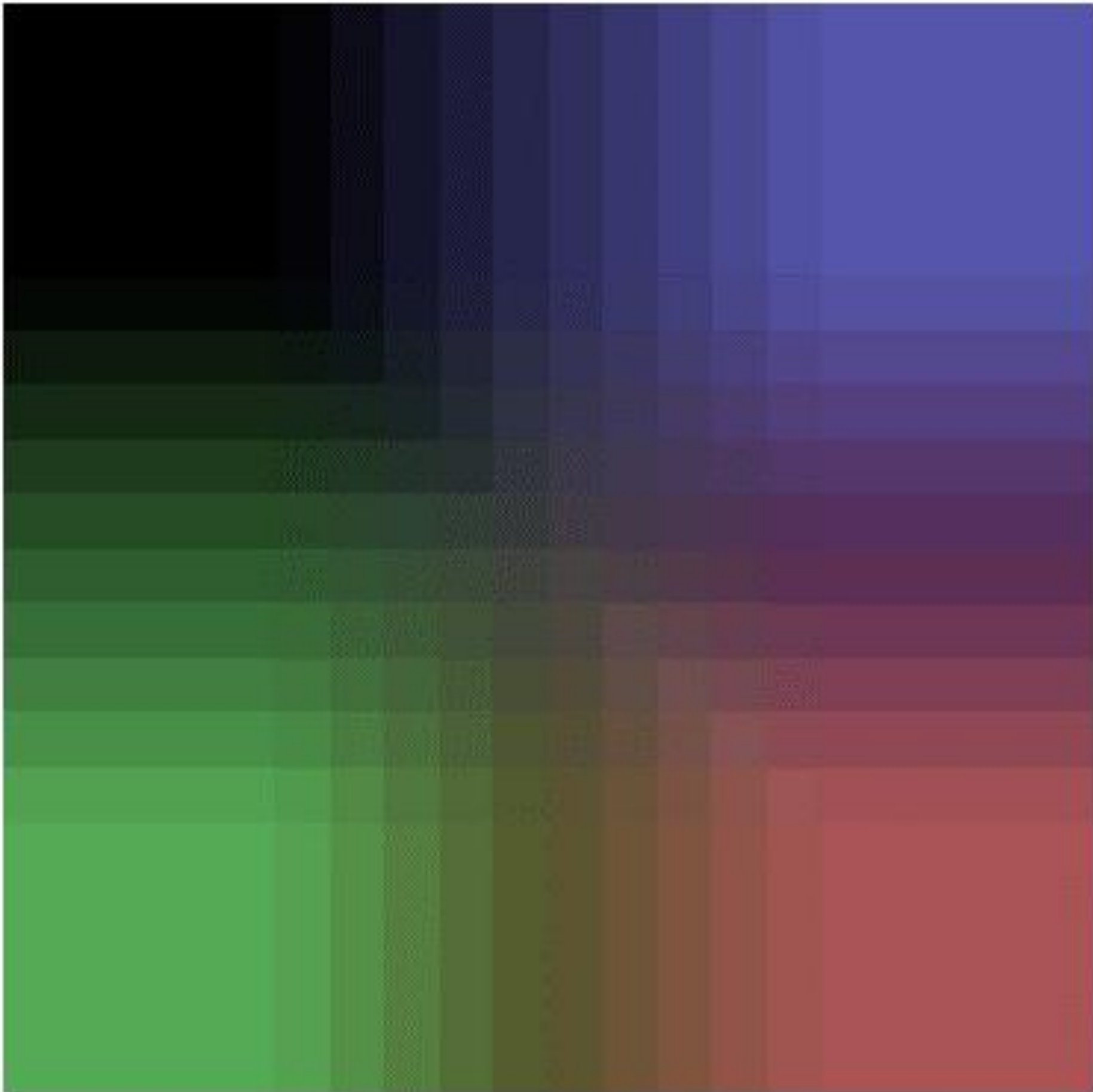
192 px

Image Resizing

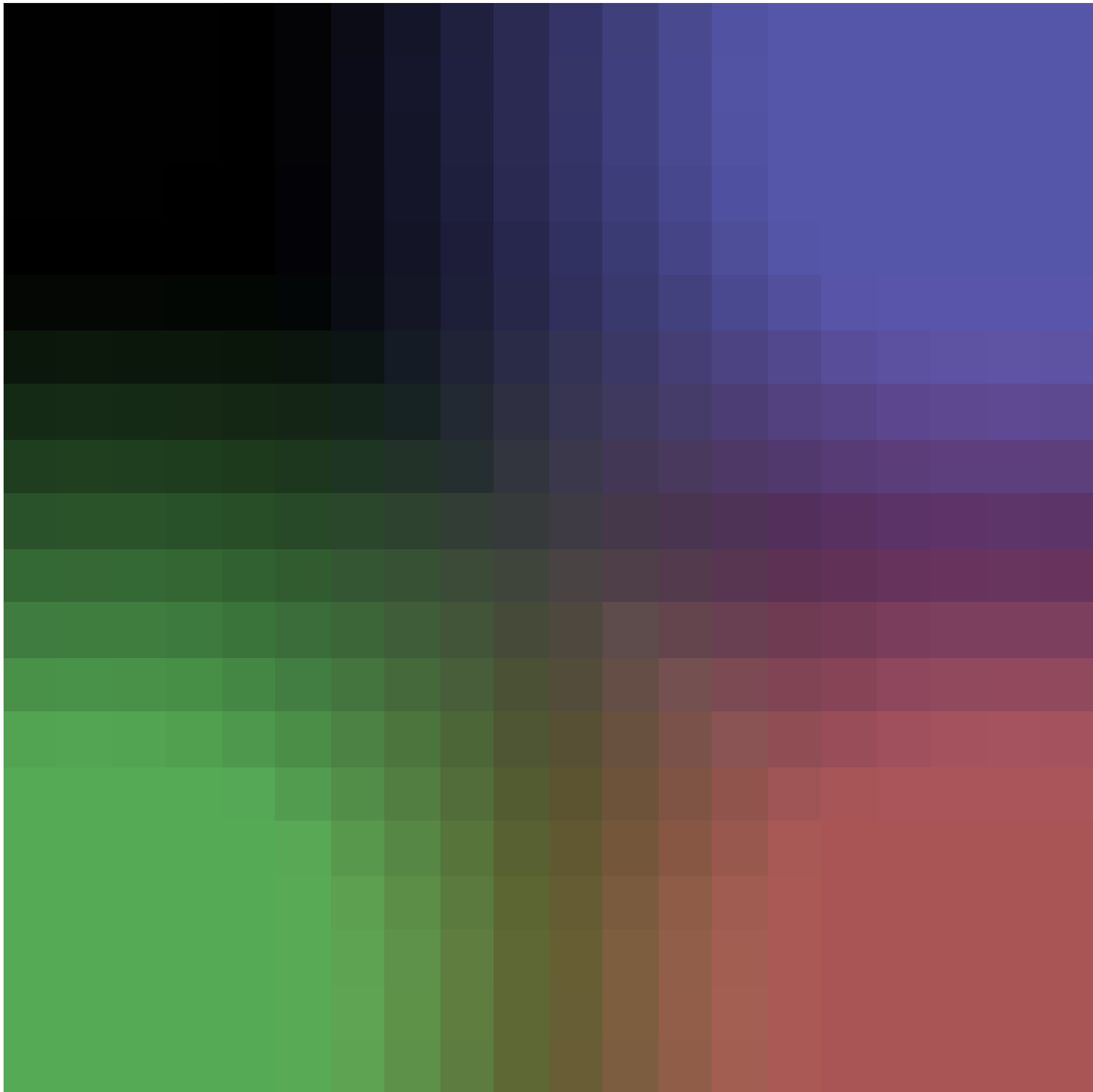
Upsampling



Nearest



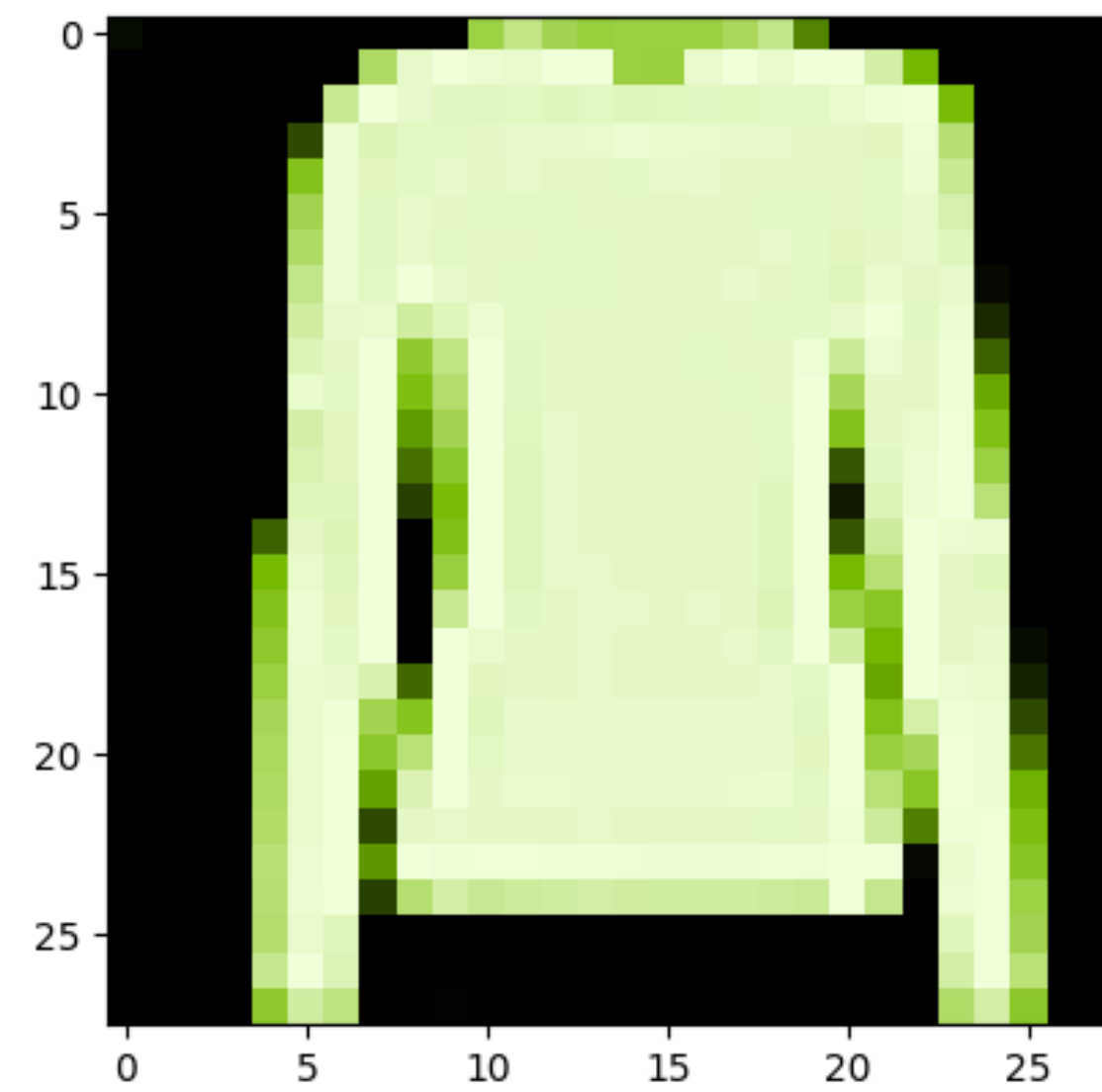
Bilinear



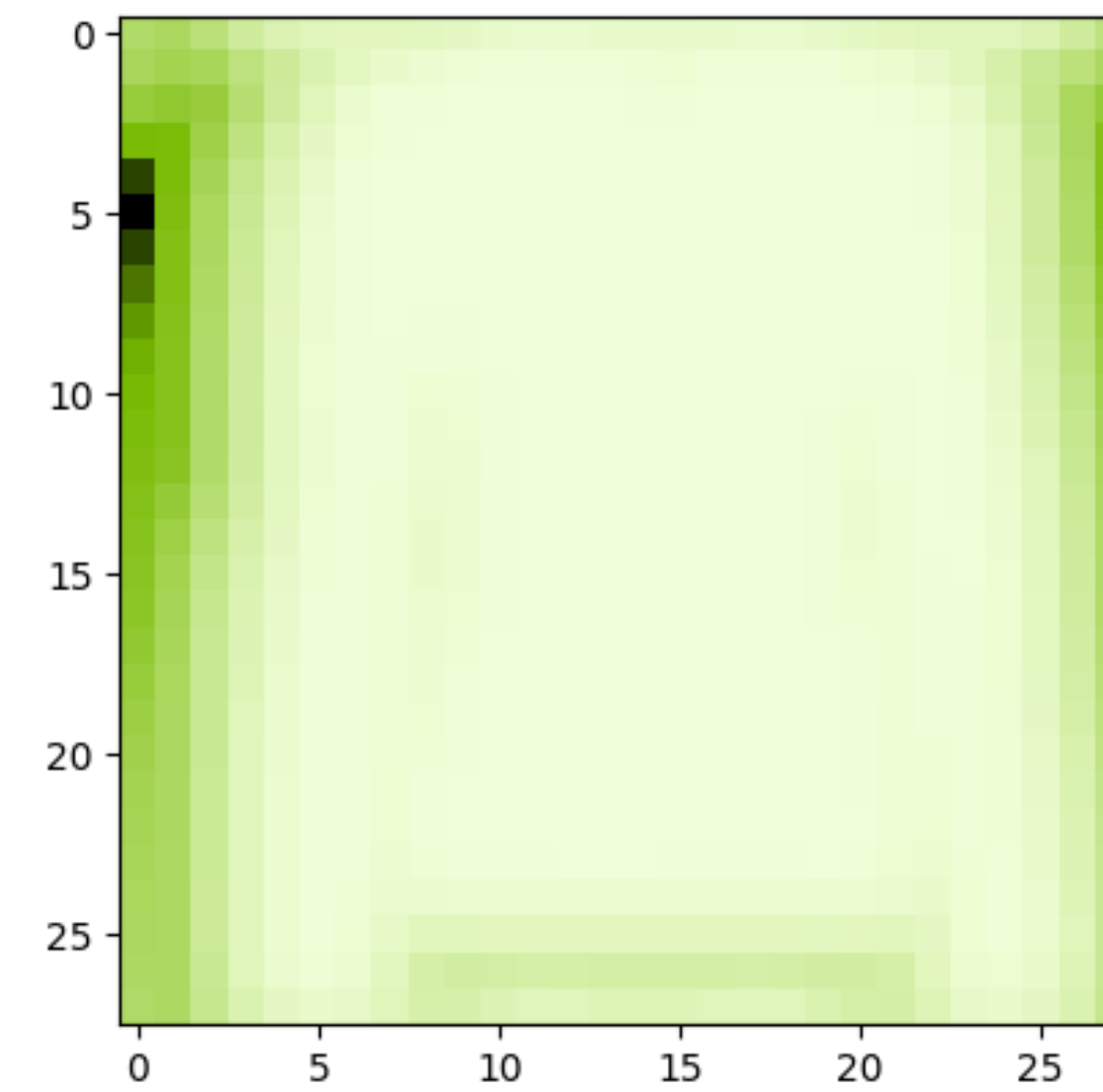
Bicubic

Deconvolution?

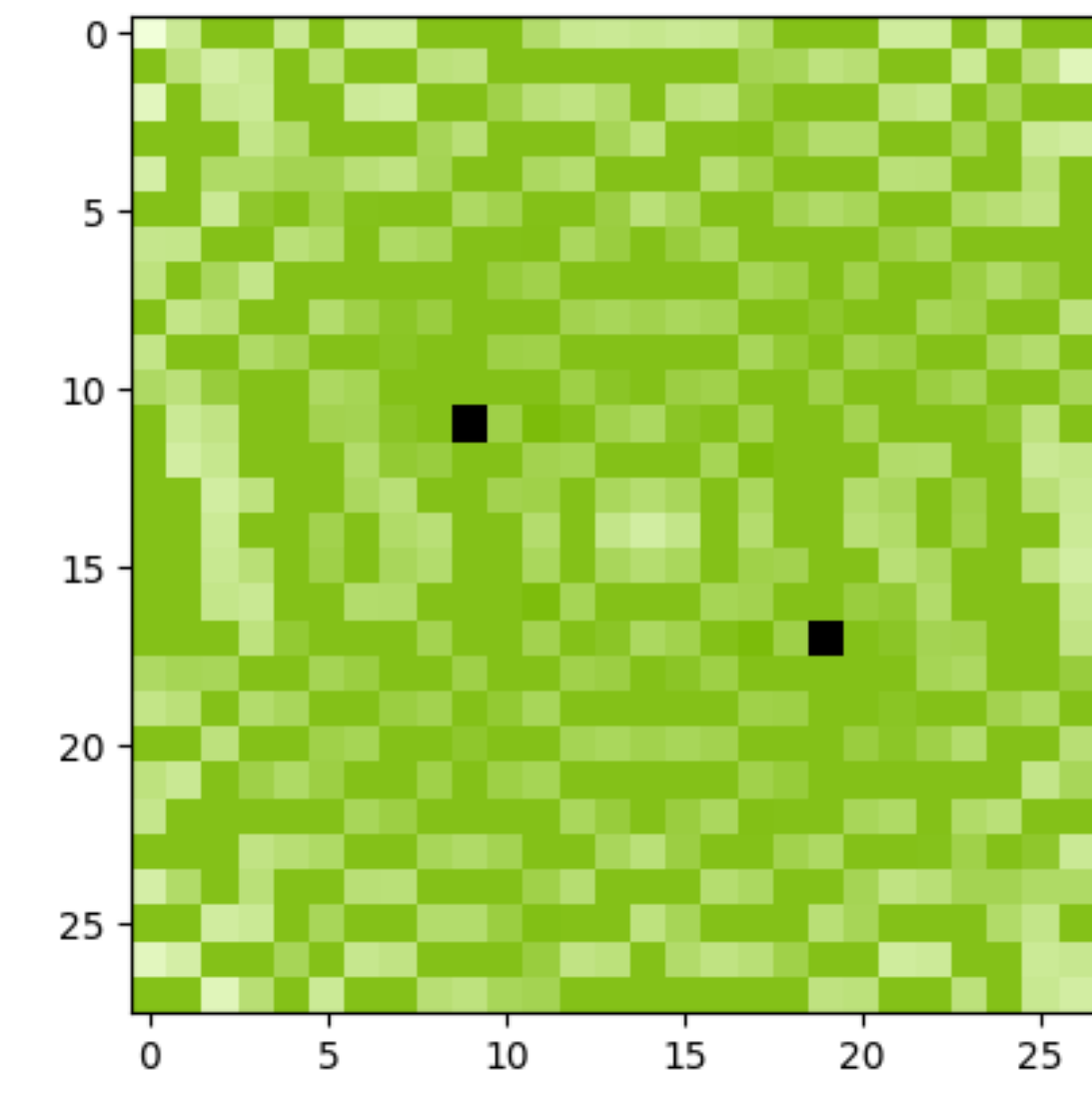
Same as Transposed Convolution?



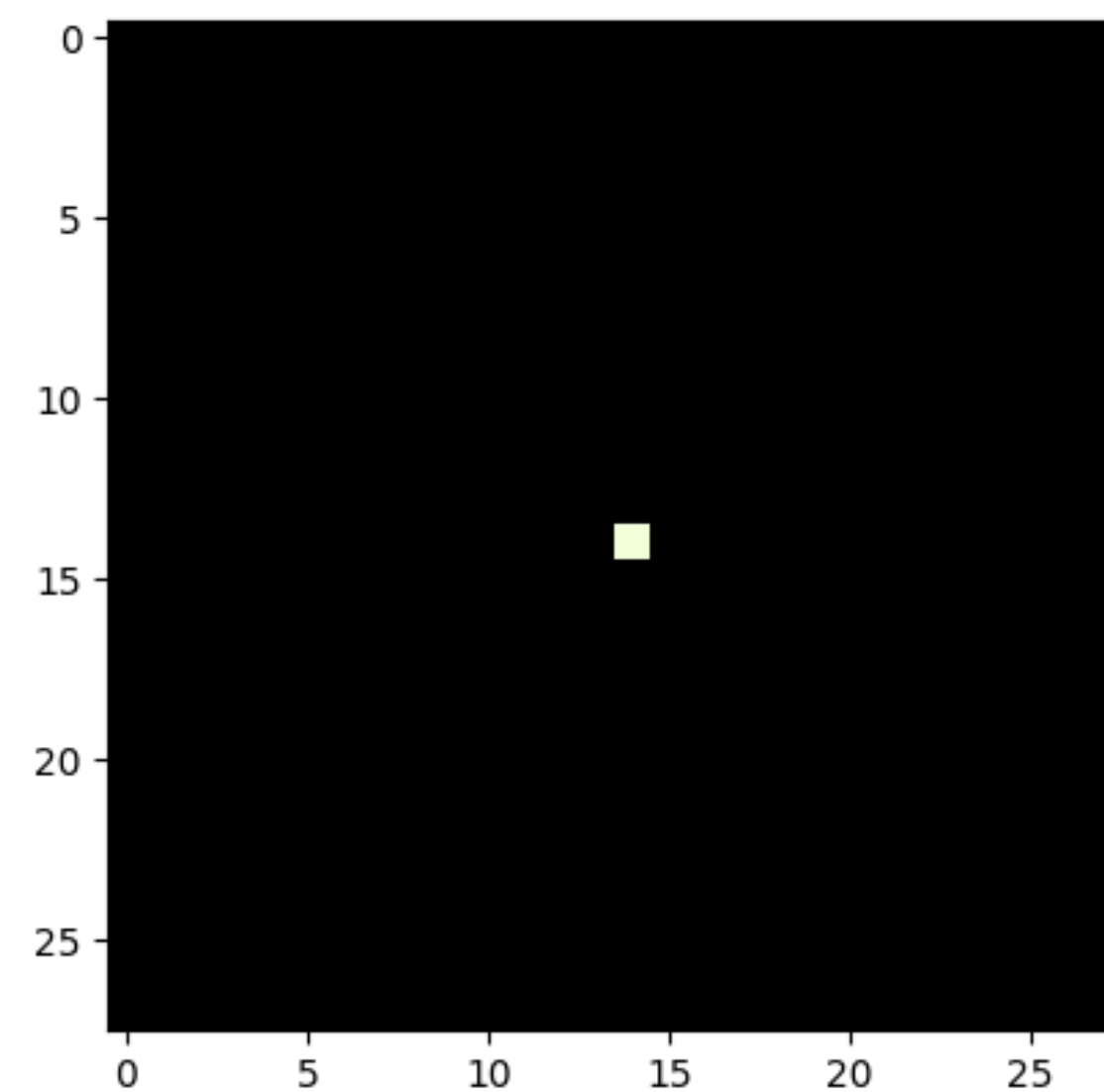
Image



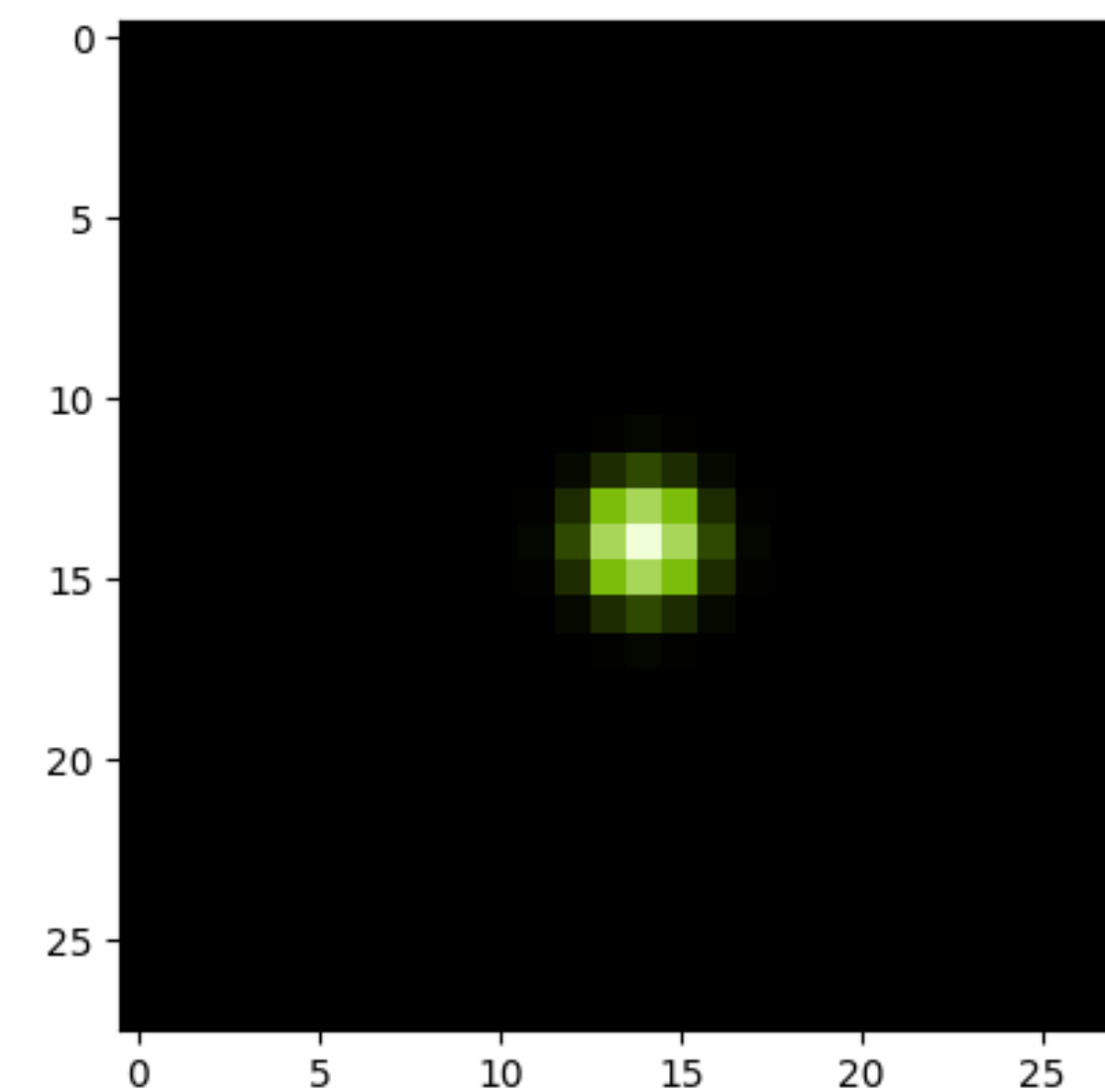
Convolved Image



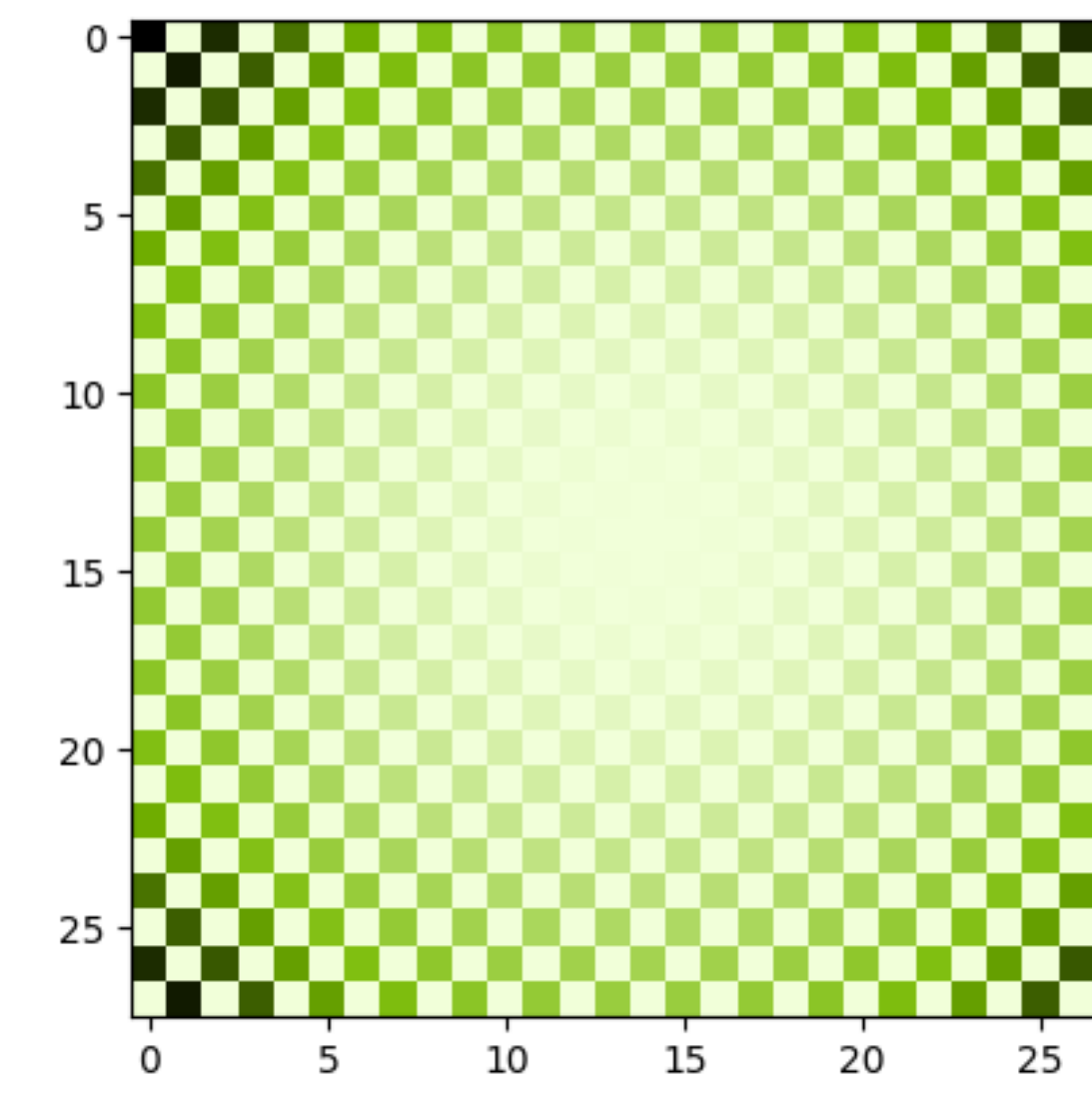
FFT Convolved



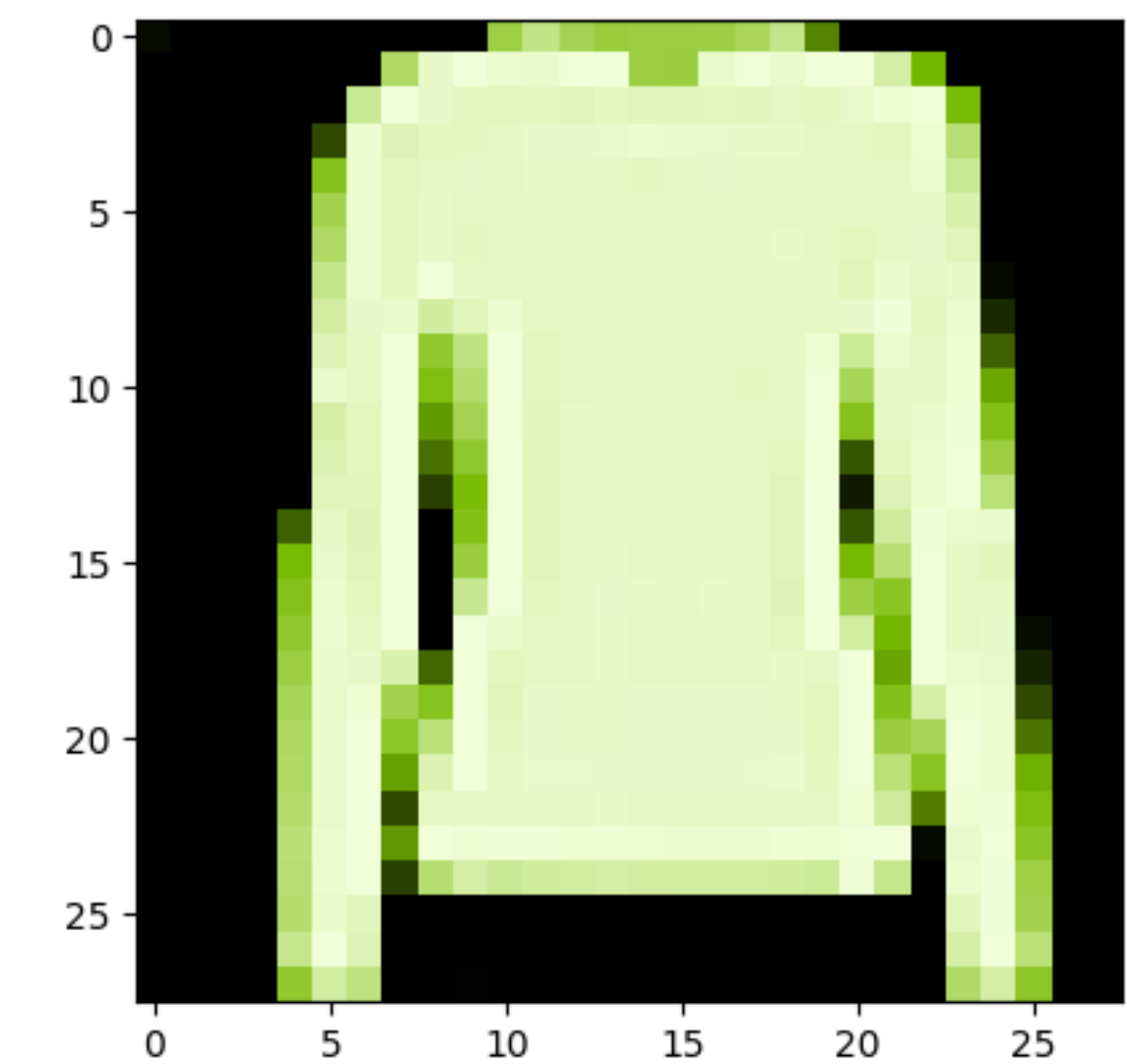
Point



Point Spread
Function



FFT Point Spread
Function

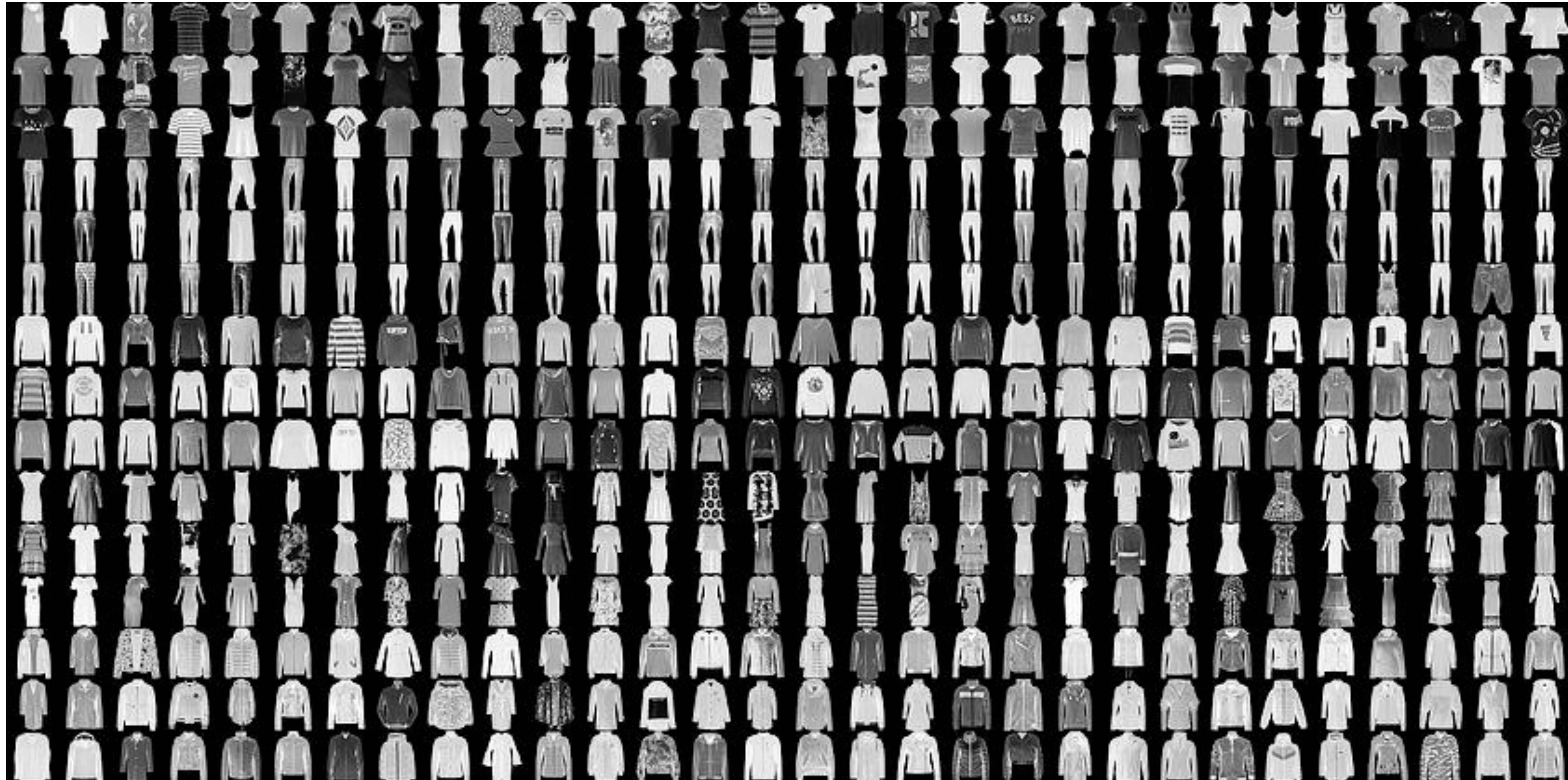


Deconvolved
Convolved Image

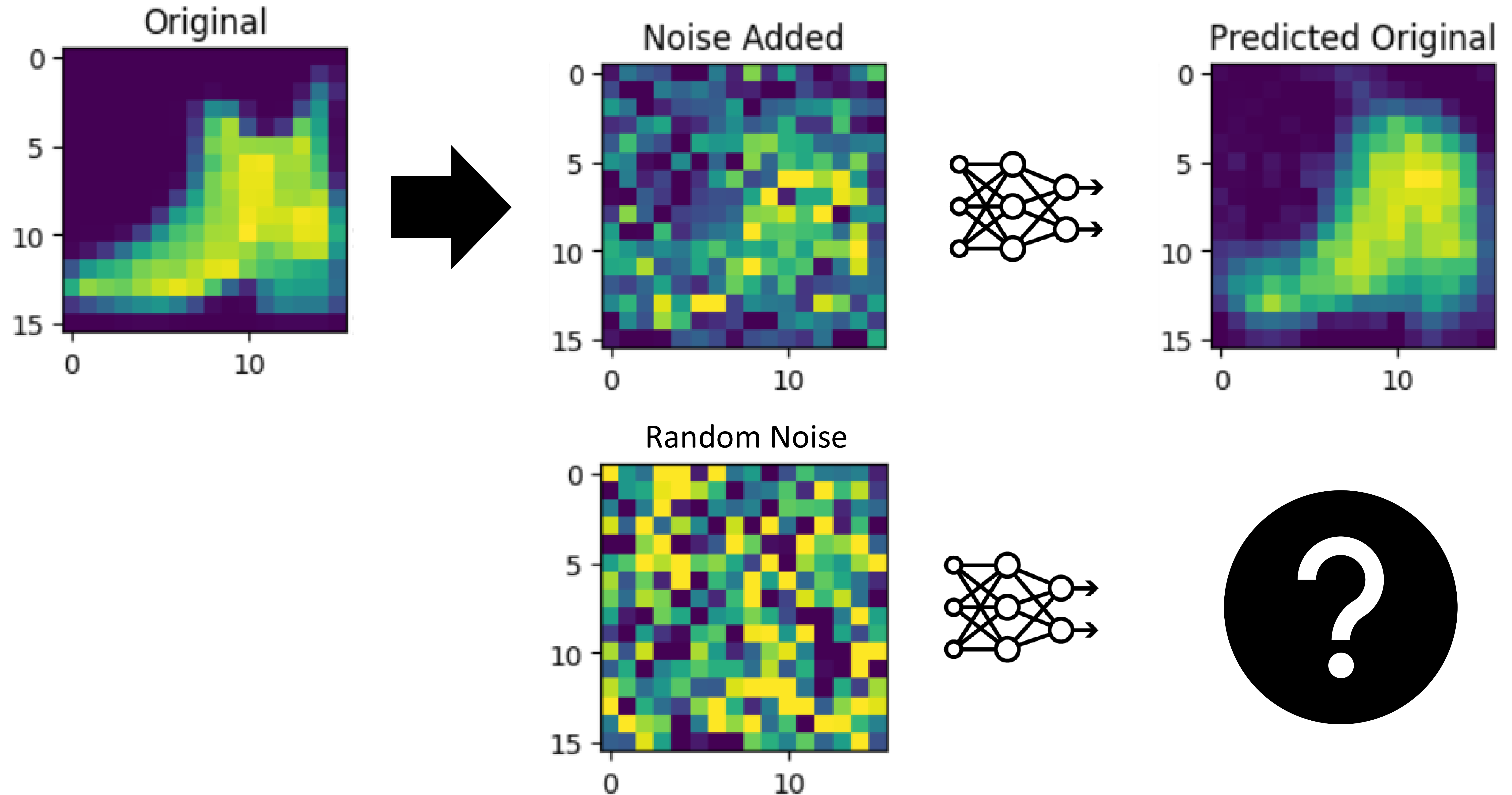
Lab

FashionMNIST

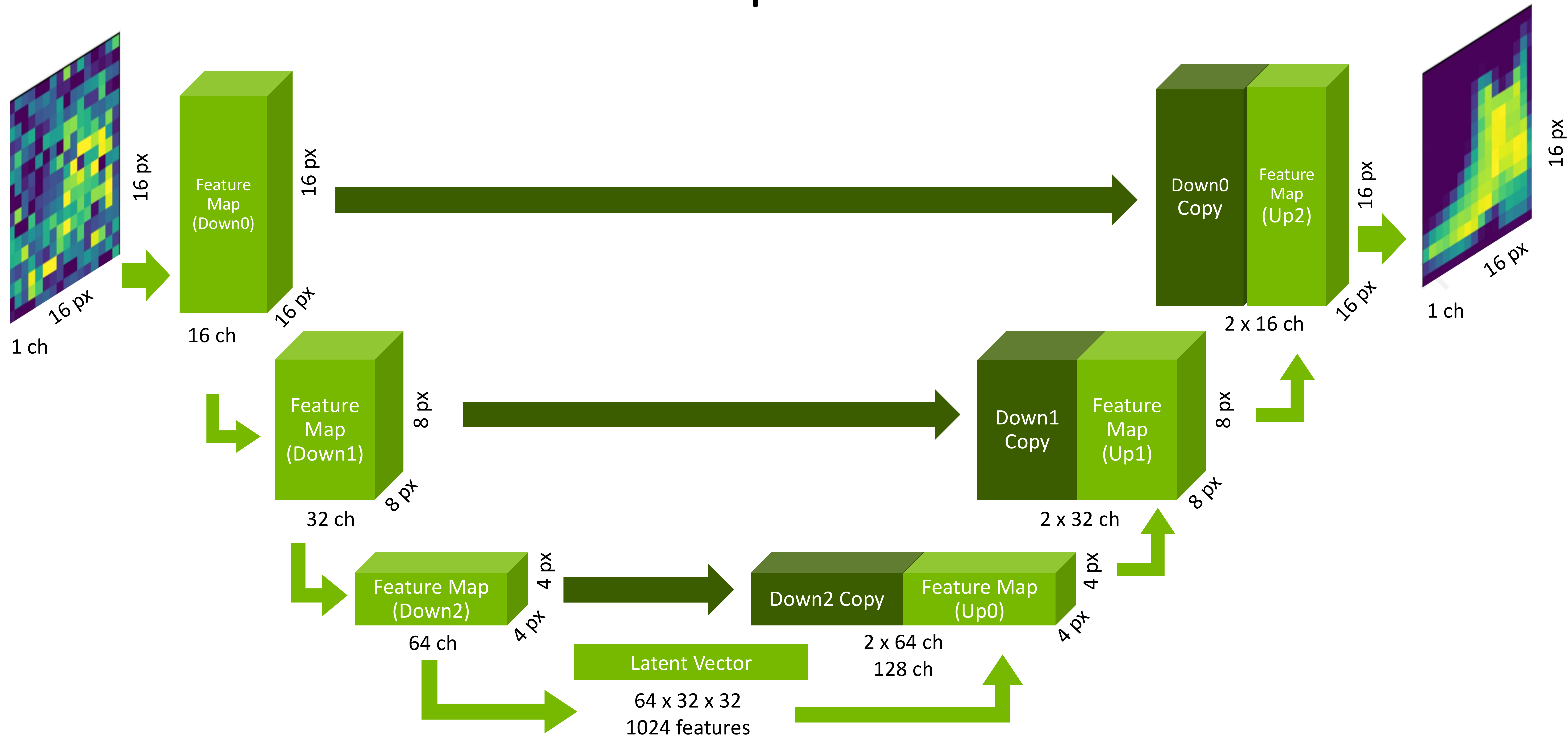
Convolutional Neural Network “Hello World”



Hypothesis: Generate an image from Noise



The Experiment



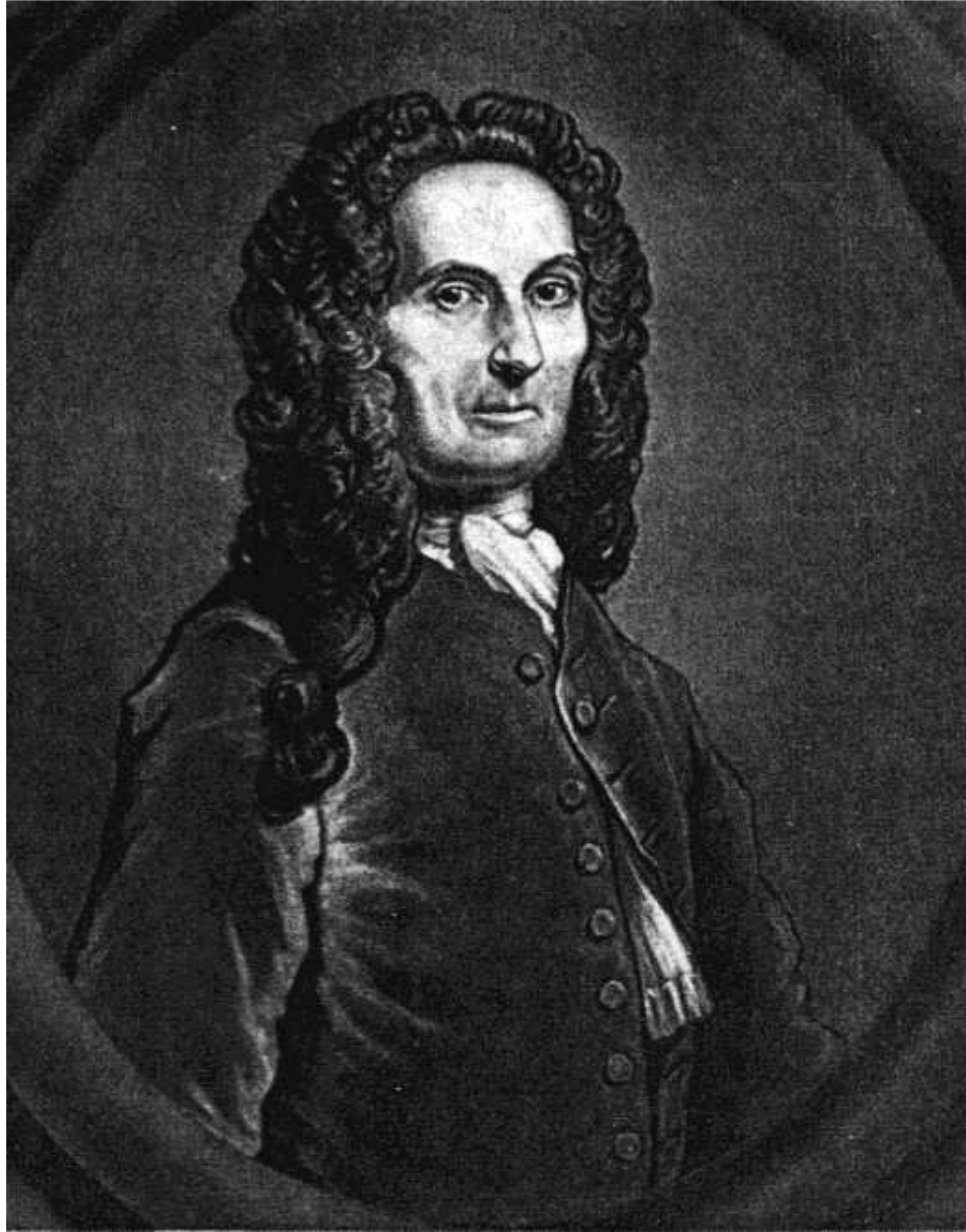
Let's get started!

1: From U-Nets to Diffusion

Appendix: The Normal Distribution

De Moivre

From Coin Flips to Bells



$$\Pr(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

$$\Pr(X = k) = \frac{n!}{k! (n - k)!} p^k (1 - p)^{n-k}$$

$$\Pr(X = 2) = \frac{4!}{2! (4 - 2)!} \left(\frac{1}{2}\right)^2 \left(1 - \frac{1}{2}\right)^{4-2}$$

$$\Pr(X = 2) = \frac{4 \cdot 3 \cdot 2 \cdot 1}{2 \cdot 1 \cdot 2 \cdot 1} \left(\frac{1}{4}\right) \left(\frac{1}{4}\right)$$

$$\Pr(X = 2) = \frac{6}{16}$$

A weighted coin flipping through the air
like a cartoon

$$p = \frac{1}{2}$$

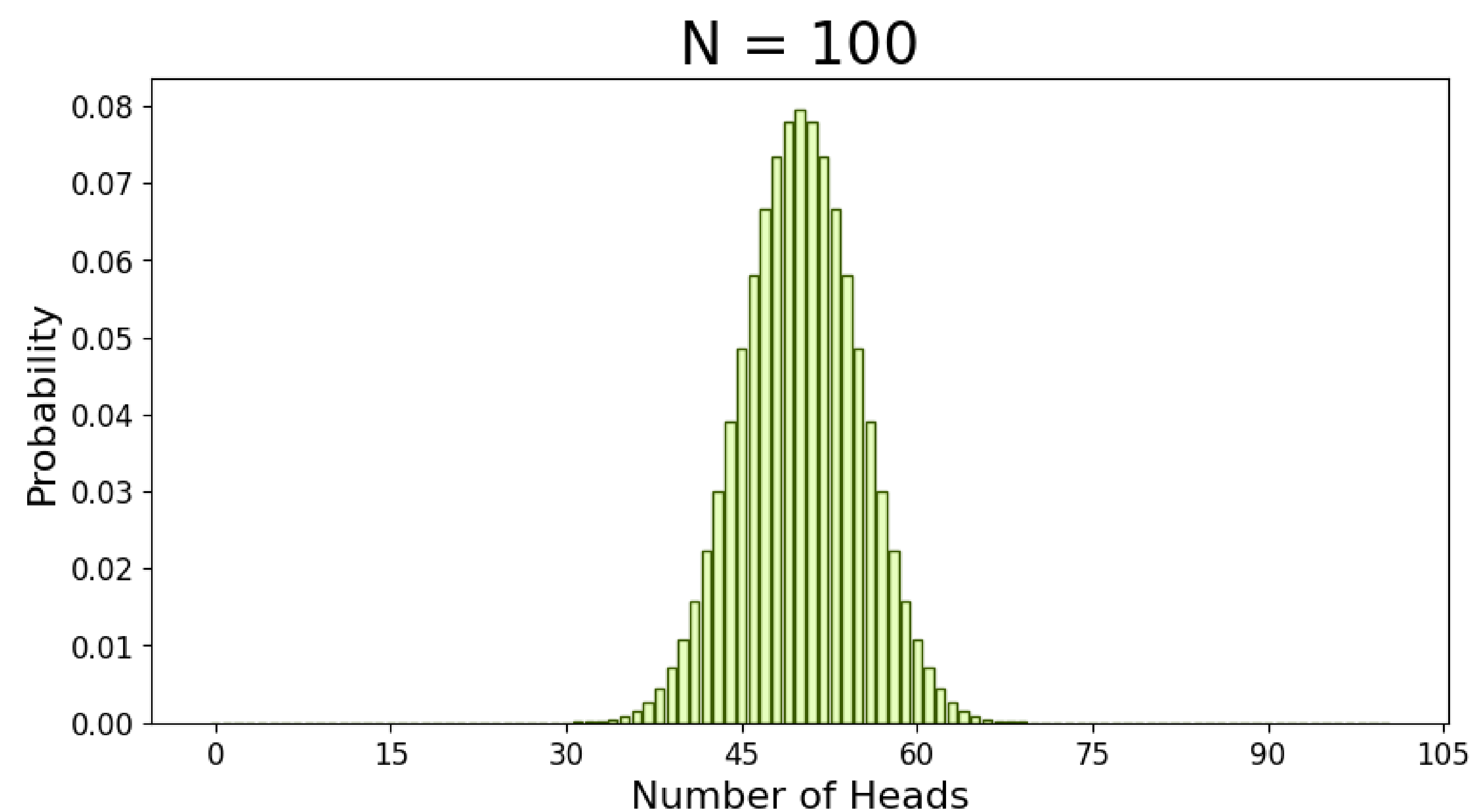
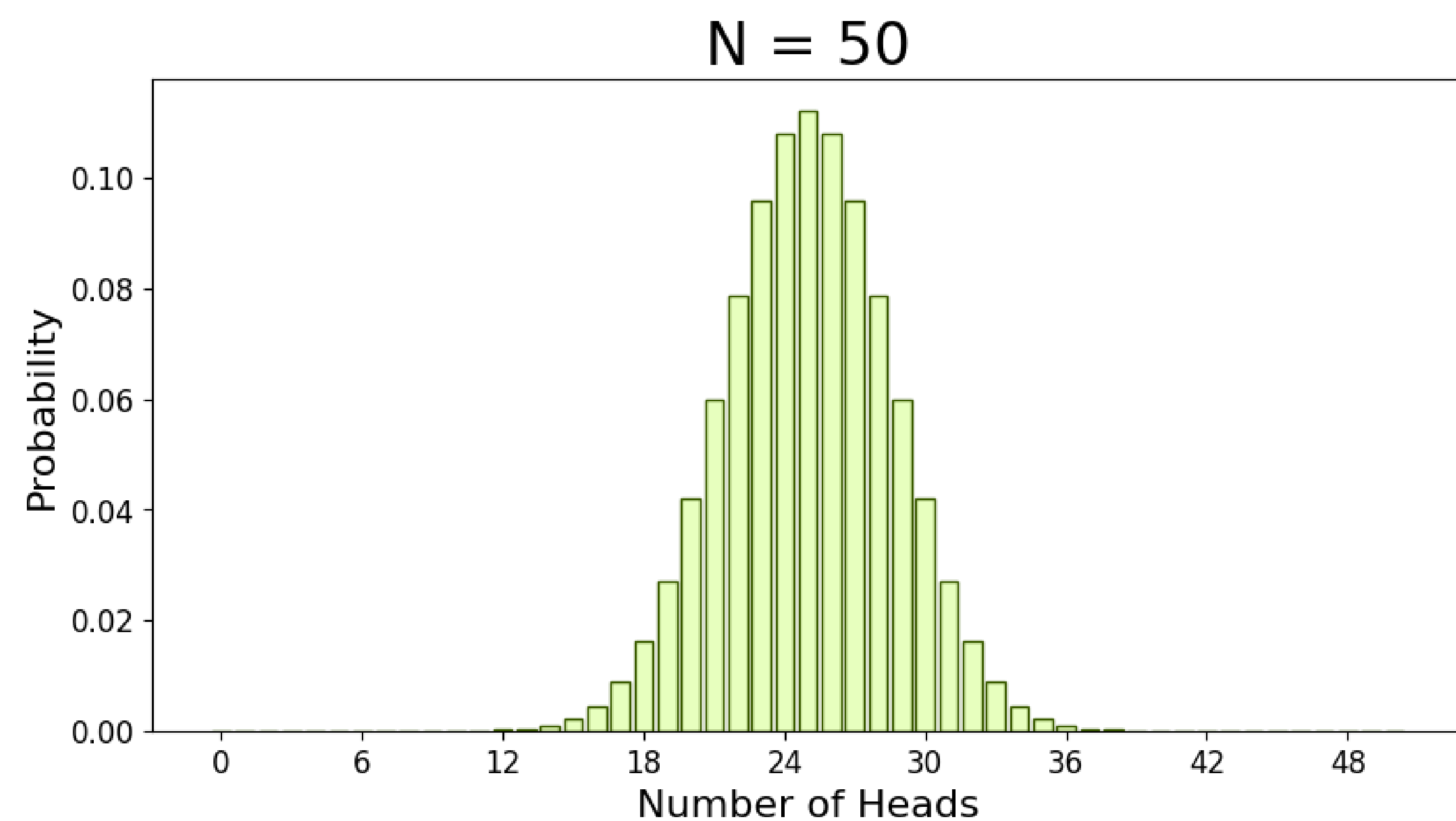
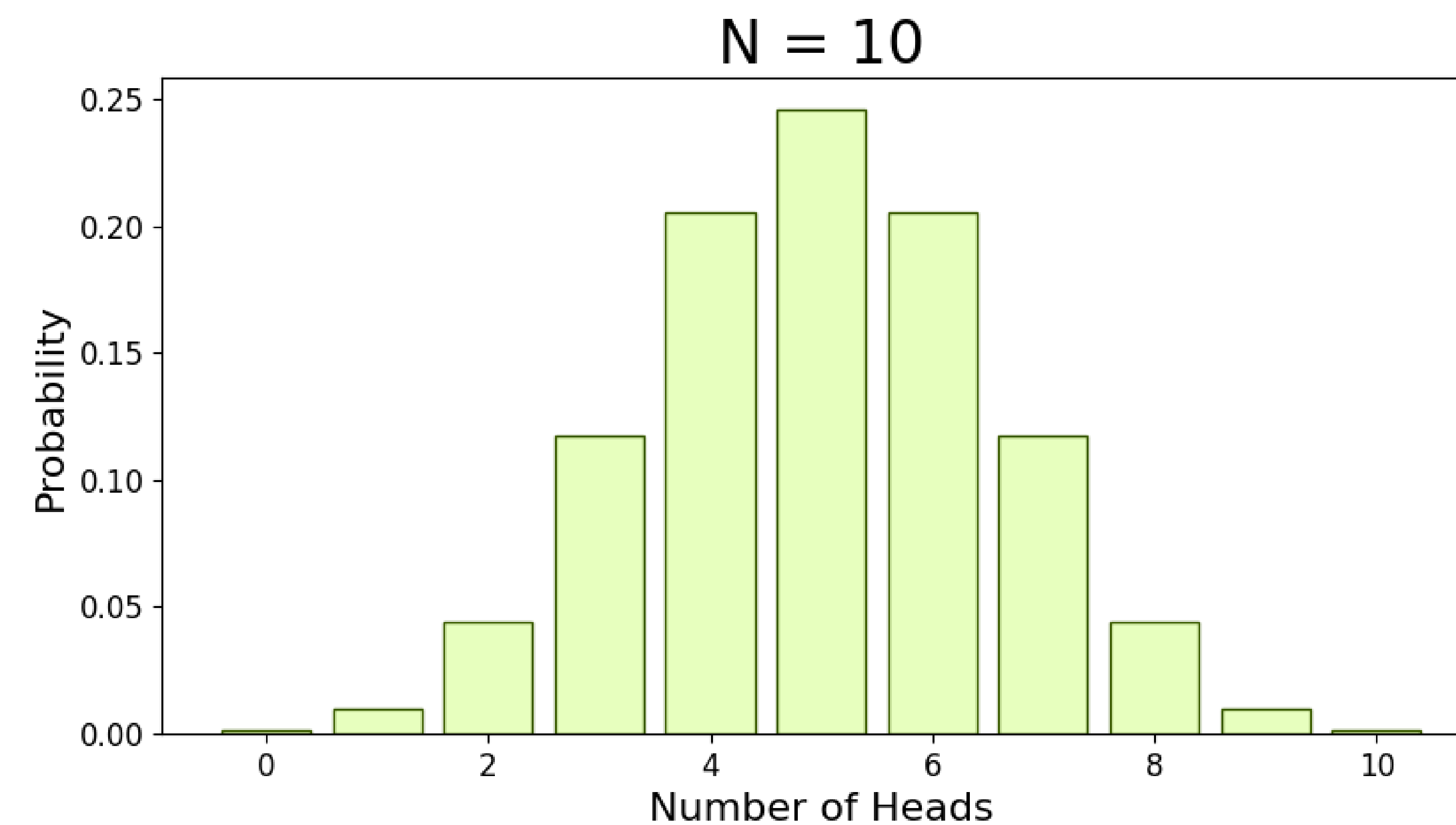
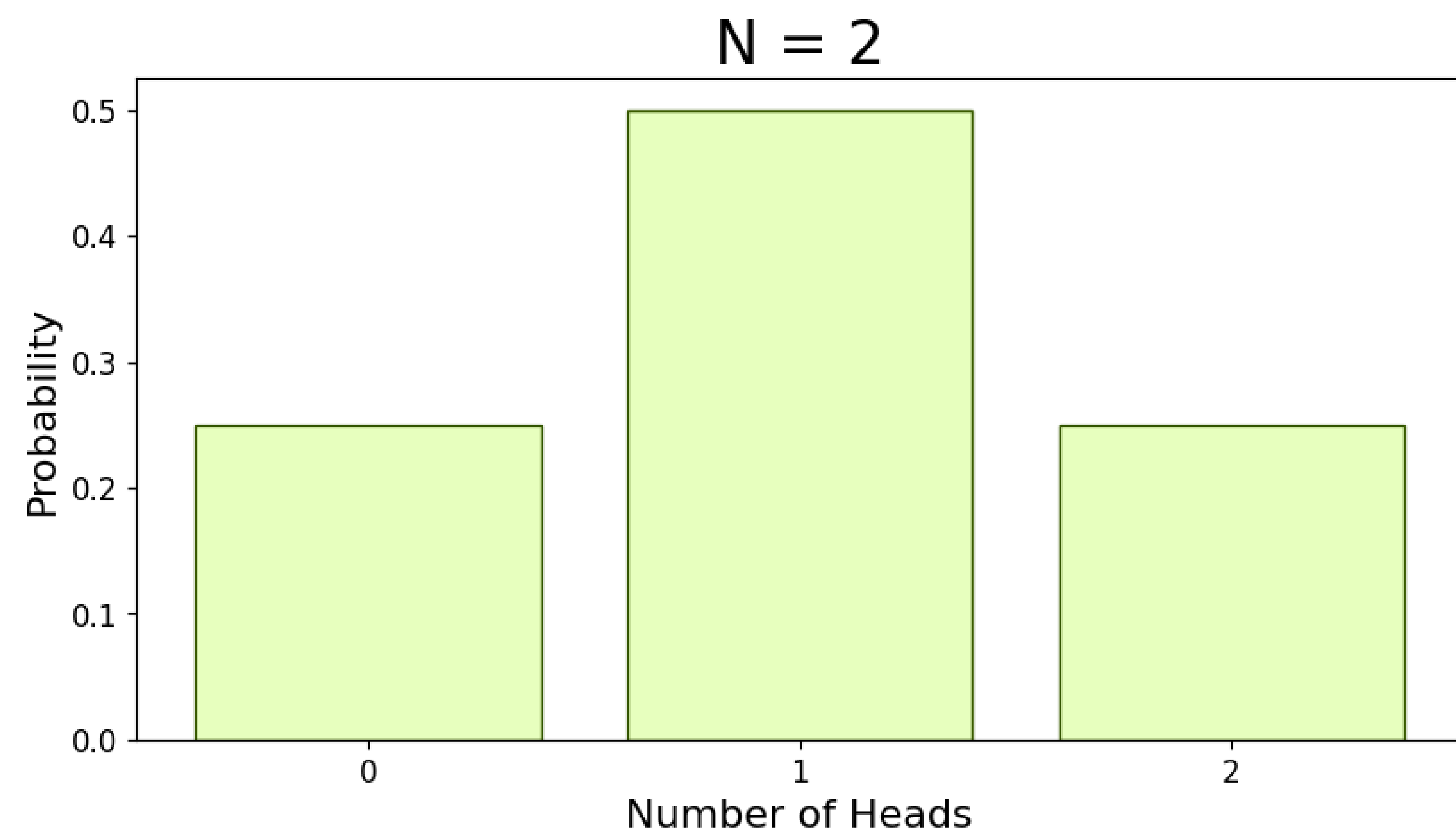
$$n = 4$$

$$k = 2$$



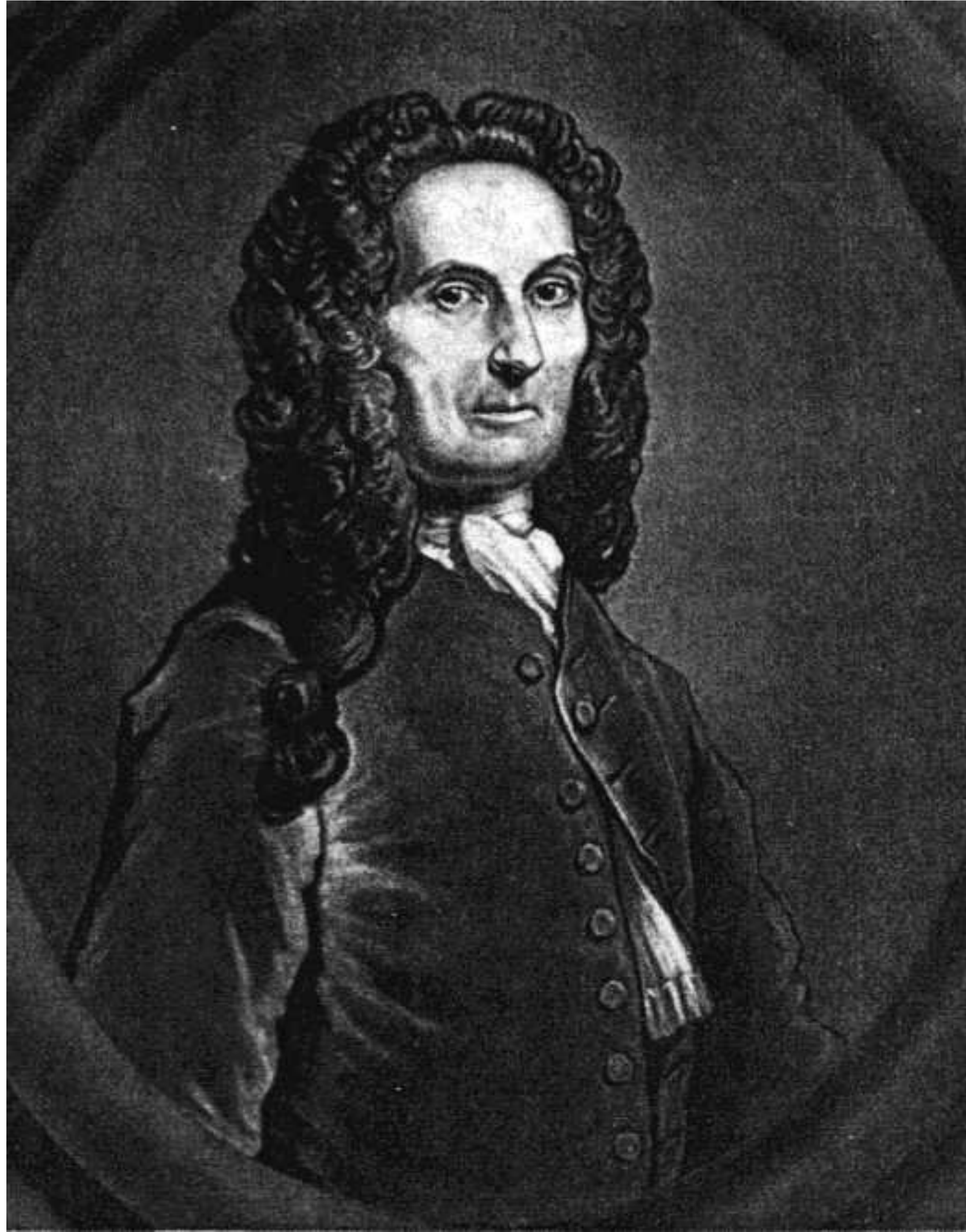
De Moivre

From Coin Flips to Bells



De Moivre

From Coin Flips to Bells



$$\Pr(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

$$\Pr(X = k) = \frac{n!}{k! (n - k)!} p^k (1 - p)^{n-k}$$

$$n! \approx \sqrt{2\pi} \left(\frac{n}{e}\right)^n$$

$$\binom{n}{k} p^k q^{n-k} \simeq \frac{1}{\sqrt{2\pi n p q}} e^{-\frac{(k-np)^2}{2npq}}$$

$$N(x; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

A weighted coin flipping through the air
like a cartoon



De Moivre

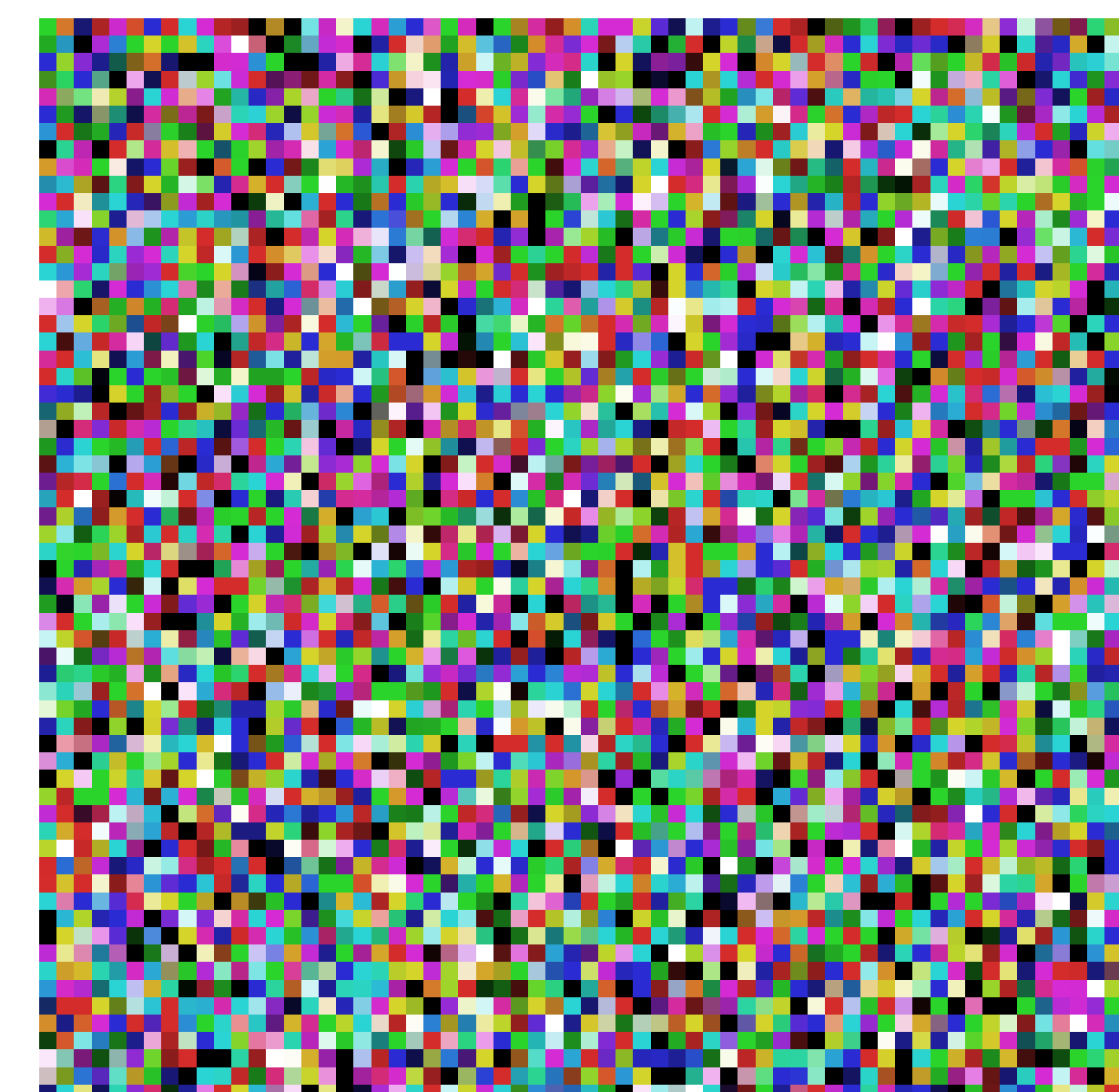
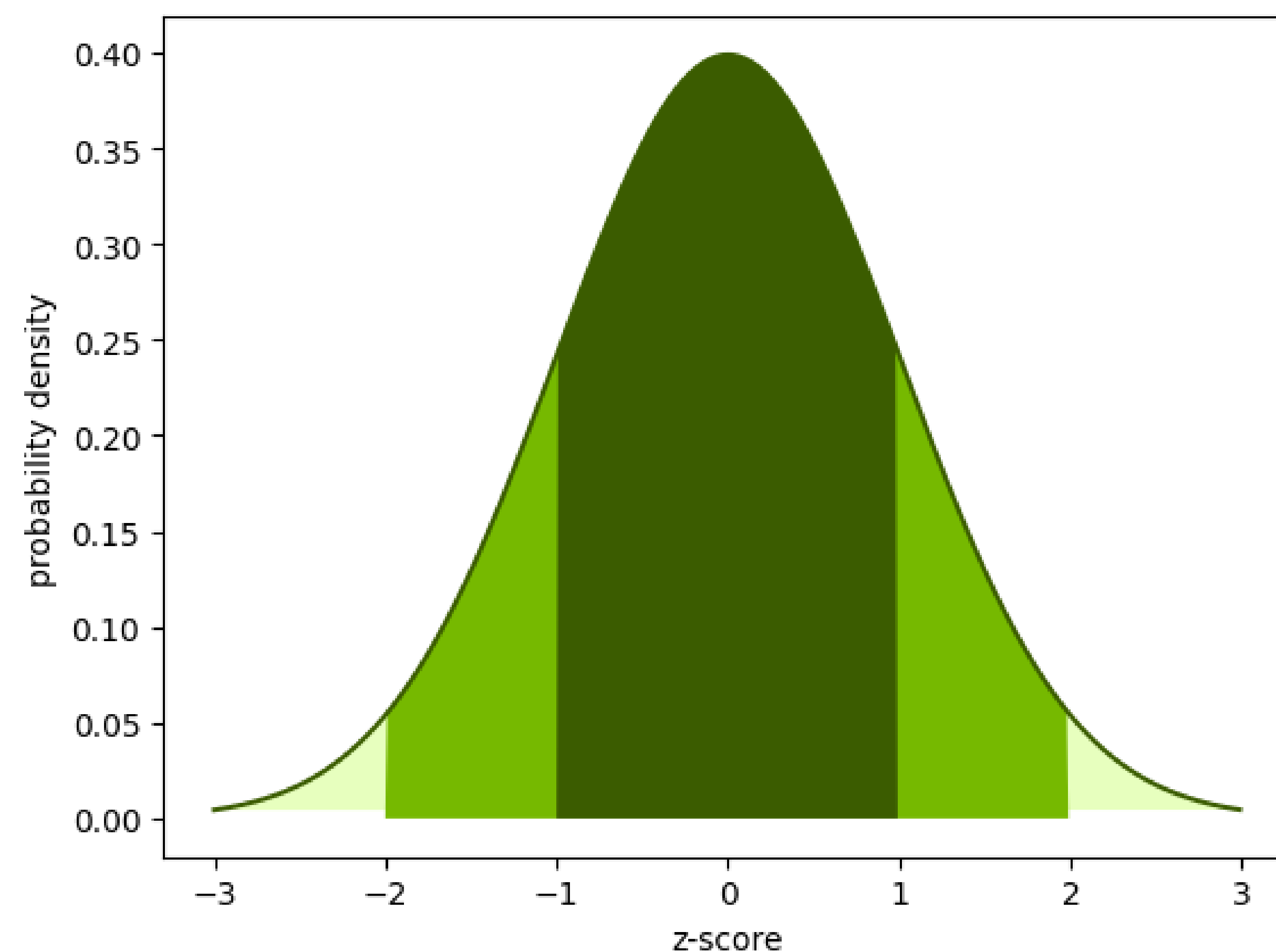
From Coin Flips to Bells

$$N(x; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$\mu = \text{mean, a.k.a. average}$

$\sigma = \text{standard deviation, a.k.a. spread}$

$$z = \frac{x - \mu}{\sigma}$$



Generative AI with Diffusion Models

Generative AI with Diffusion Models

Course Progress Bookmarks Updates

Generative AI with Diffusion Models Start Here 1: From U-Nets to Diffusion

Generative AI with Diffusion Models

Start Here

0: Server Access

1: From U-Nets to Diffusion

2: Denoising Diffusion Probabilistic Models

3: Optimizations

4: Classifier-Free Diffusion Guidance

5: CLIP

6: Wrap-up and Assessment

Next Steps

Feedback

Previous

Next

▶ 17:54 / 17:55

▶ Speed 1.0x



DEEP
LEARNING
INSTITUTE

This Lab 0 : 10 : 54 / 1 : 30 : 00

Course 8 : 18 : 49 / 24 : 00 : 00



LAUNCH



STOP TASK



**Generative AI with
Diffusion Models**

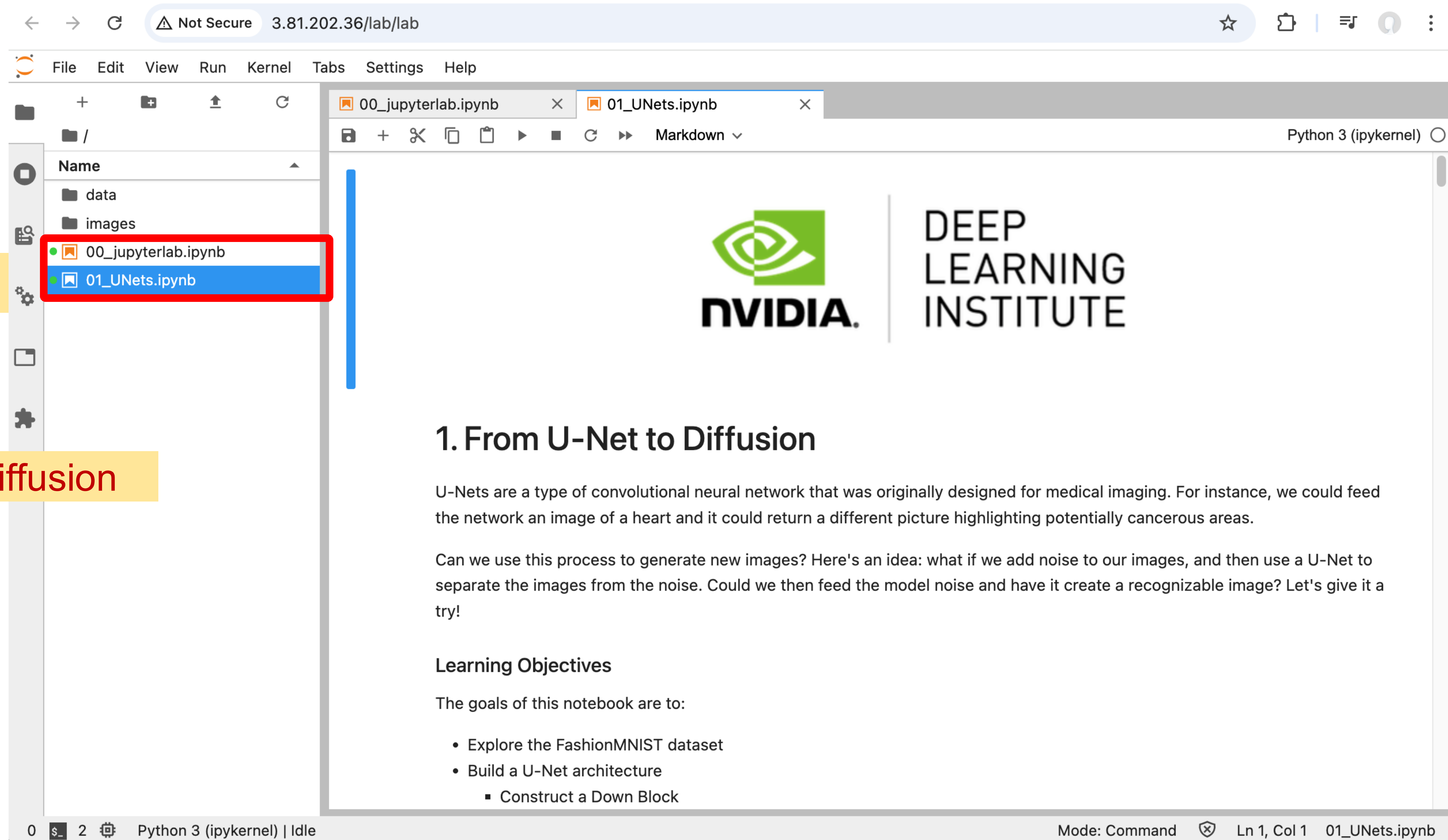
Part 1: From U-Nets to Diffusion

1: From U-Nets to Diffusion

Generative AI with Diffusion Models

01_UNets.ipynb

1: From U-Nets to Diffusion



00_jupyterlab.ipynb 01_UNets.ipynb

Python 3 (ipykernel)

DEEP LEARNING INSTITUTE

1. From U-Net to Diffusion

U-Nets are a type of convolutional neural network that was originally designed for medical imaging. For instance, we could feed the network an image of a heart and it could return a different picture highlighting potentially cancerous areas.

Can we use this process to generate new images? Here's an idea: what if we add noise to our images, and then use a U-Net to separate the images from the noise. Could we then feed the model noise and have it create a recognizable image? Let's give it a try!

Learning Objectives

The goals of this notebook are to:

- Explore the FashionMNIST dataset
- Build a U-Net architecture
 - Construct a Down Block

0 \$ 2 Python 3 (ipykernel) | Idle Mode: Command Ln 1, Col 1 01_UNets.ipynb



Generative AI with Diffusion Models

Part 2: Denoising Diffusion Probabilistic Models

Agenda

- Part 1: From U-Nets to Diffusion
- Part 2: Denoising Diffusion Probabilistic Models
- Part 3: Optimizations
- Part 4: Classifier-Free Diffusion Guidance
- Part 5: CLIP
- Part 6: Wrap-up & Assessment

Diffusion Inspiration

Thermodynamic Origins



Deep Unsupervised Learning using Nonequilibrium Thermodynamics

Jascha Sohl-Dickstein Stanford University	JASCHA@STANFORD.EDU
Eric A. Weiss University of California, Berkeley	EAWISS@BERKELEY.EDU
Niru Maheswaranathan Stanford University	NIRUM@STANFORD.EDU
Surya Ganguli Stanford University	SGANGULI@STANFORD.EDU

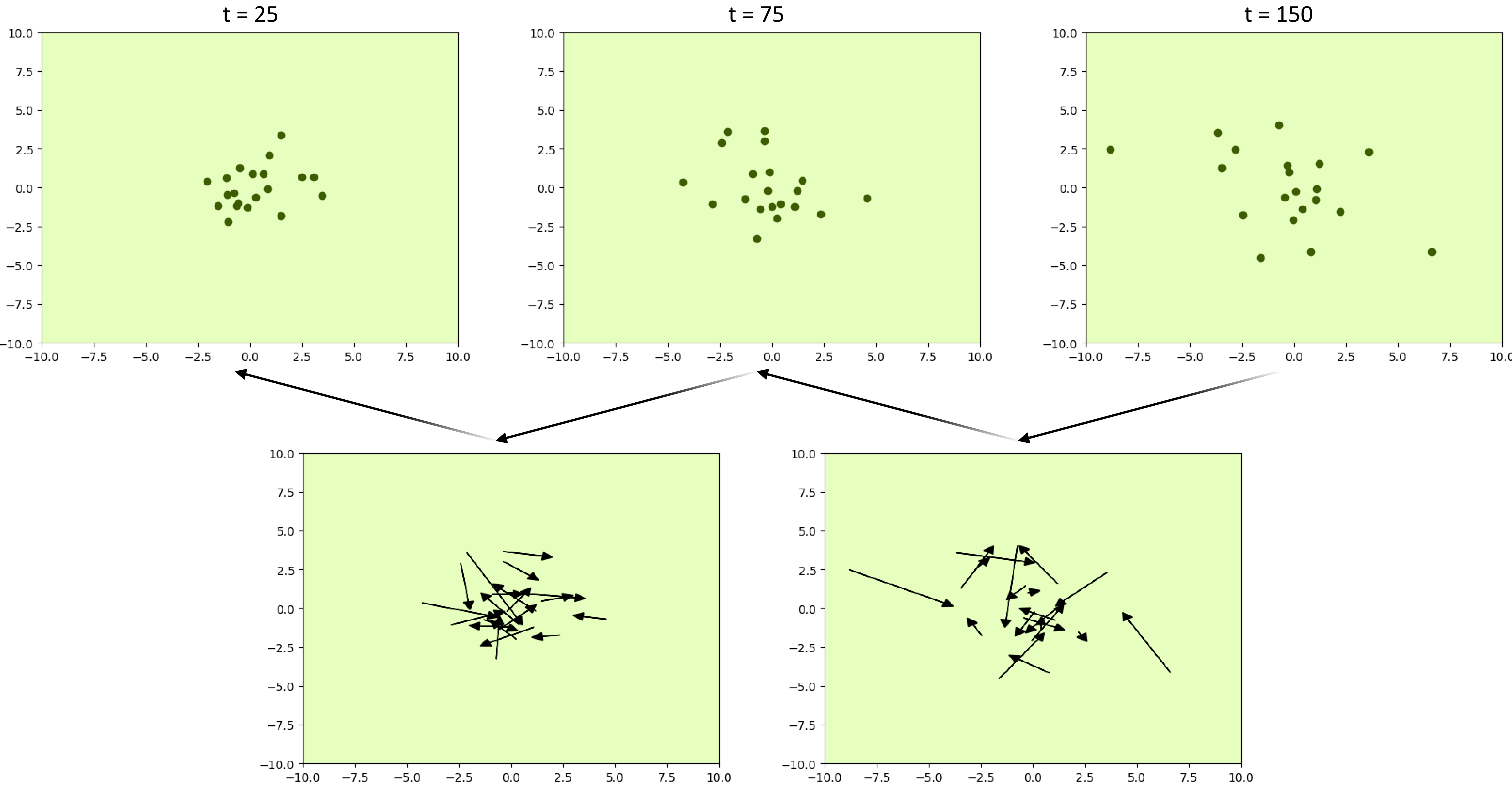
Abstract

A central problem in machine learning involves modeling complex data-sets using highly flexible families of probability distributions in which learning, sampling, inference, and evaluation

these models are unable to aptly describe structure in rich datasets. On the other hand, models that are *flexible* can be molded to fit structure in arbitrary data. For example, we can define models in terms of any (non-negative) function $\phi(\mathbf{x})$ yielding the flexible distribution $p(\mathbf{x}) = \frac{\phi(\mathbf{x})}{Z}$, where Z is a normalization constant. However, computing this

Green food coloring dissolving in a glass of water

Thermodynamic Origins



Denoising Diffusion Probabilistic Models

Getting Lost in the Noise

Denoising Diffusion Probabilistic Models

Jonathan Ho

UC Berkeley

jonathanho@berkeley.edu

Ajay Jain

UC Berkeley

ajayj@berkeley.edu

Pieter Abbeel

UC Berkeley

pabbeel@cs.berkeley.edu

Abstract

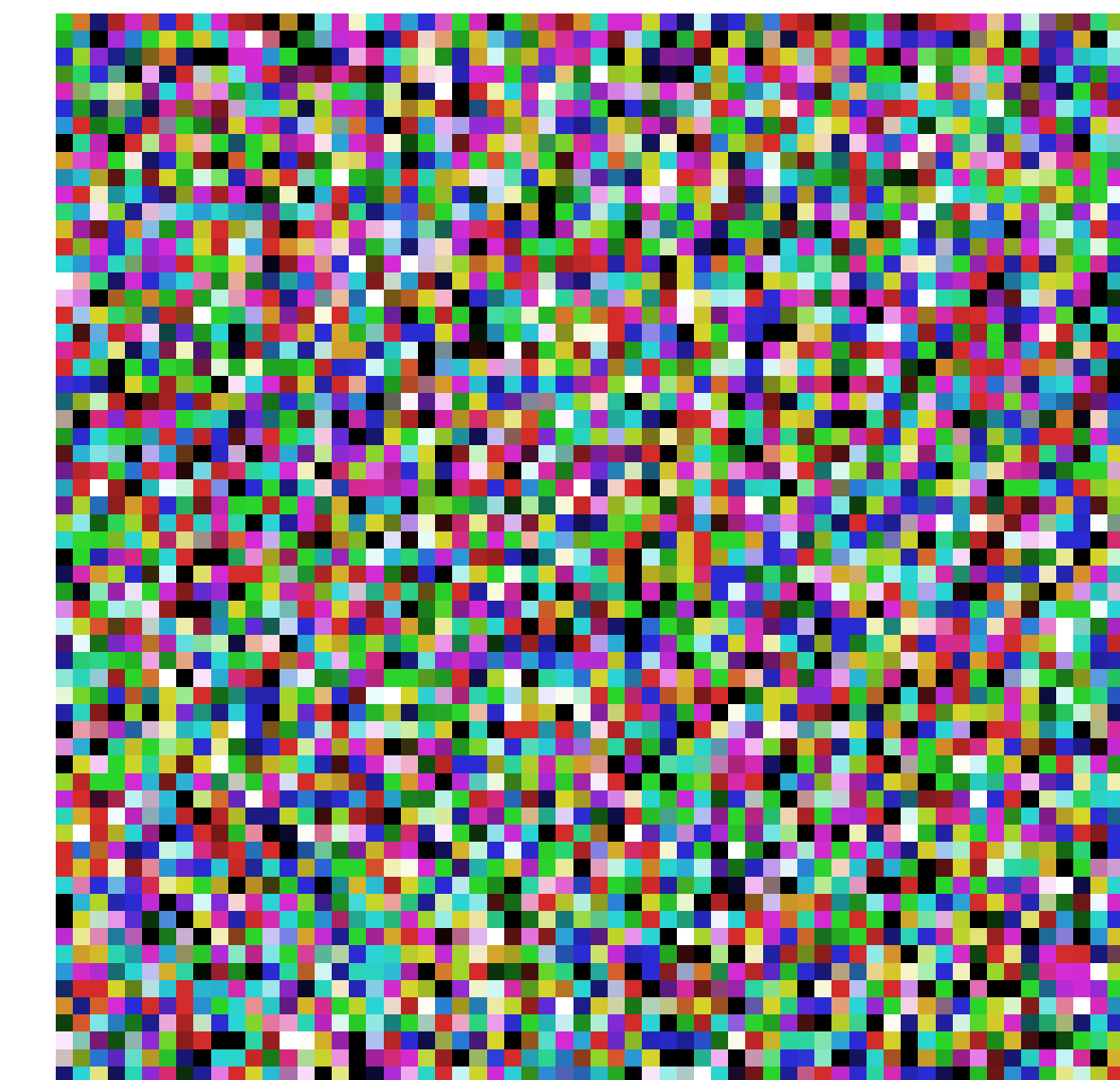
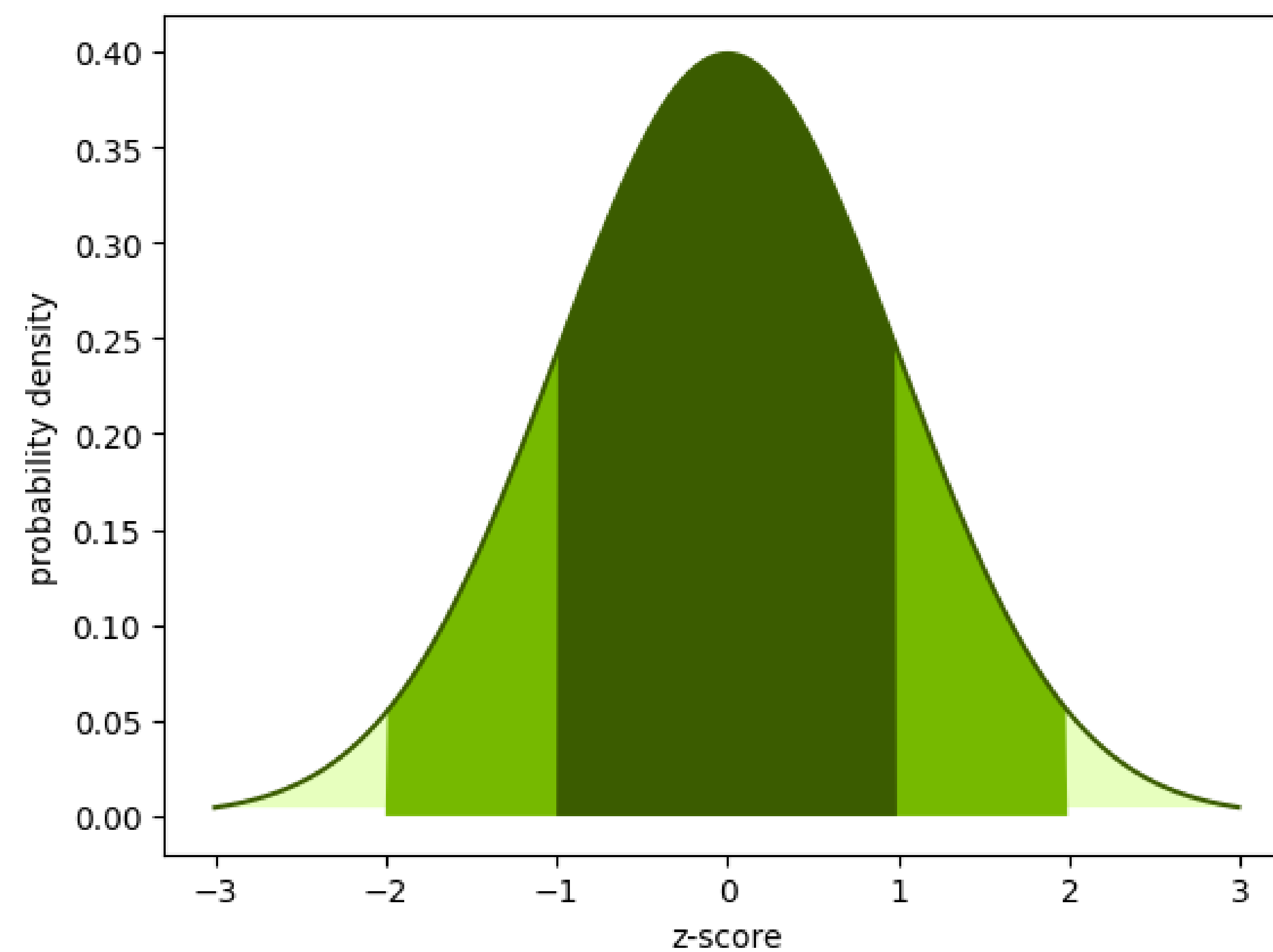
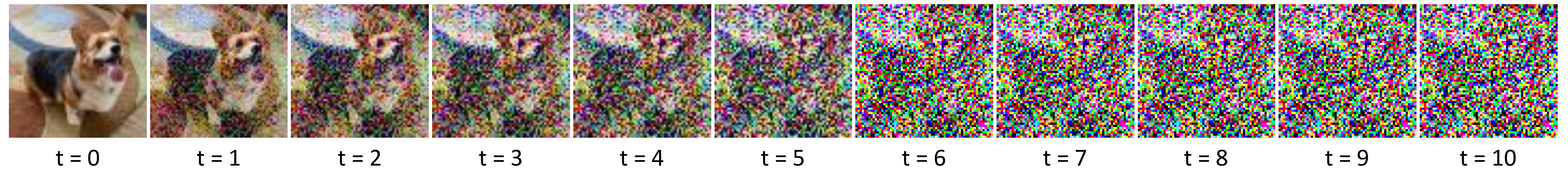
We present high quality image synthesis results using diffusion probabilistic models, a class of latent variable models inspired by considerations from nonequilibrium thermodynamics. Our best results are obtained by training on a weighted variational bound designed according to a novel connection between diffusion probabilistic models and denoising score matching with Langevin dynamics, and our models naturally admit a progressive lossy decompression scheme that can be interpreted as a



Forward Diffusion

Getting Lost in the Noise

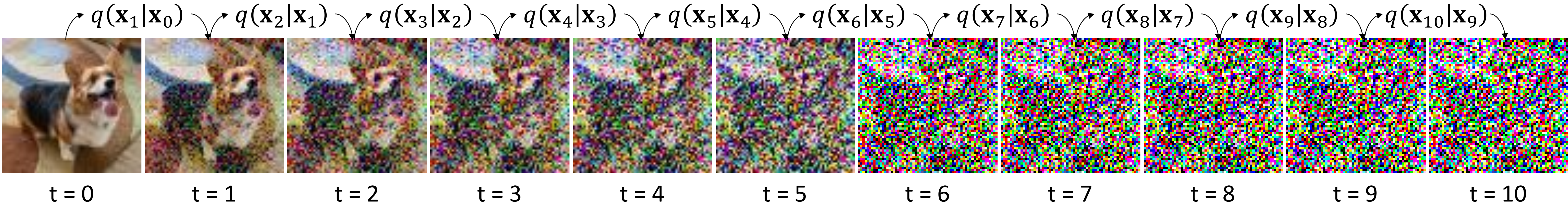
Forward Diffusion



$$N(x; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Getting Lost in the Noise

Forward Diffusion



$$\beta = [0.0001, 0.0023, 0.0045, 0.0067, \dots, 0.0200]$$

$t = 0 \quad t = 1 \quad t = 3 \quad t = 4 \quad t = 10$

$$T = 10$$

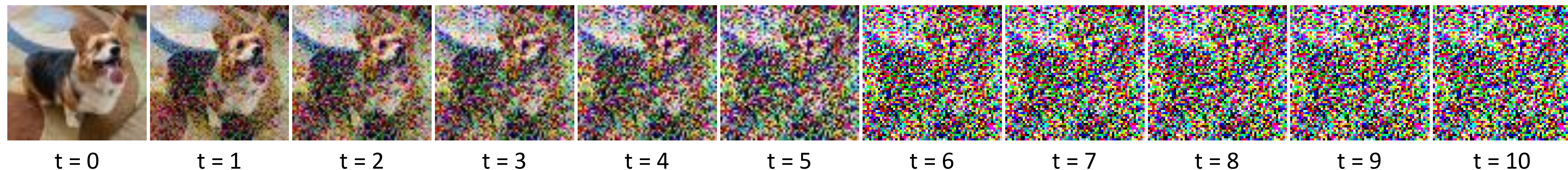
Total number of timesteps

$$\alpha = 1 - \beta$$

Convenient shorthand

Getting Lost in the Noise

Forward Diffusion



$$q(\mathbf{x}_t | \mathbf{x}_{t-1}) = N(\mathbf{x}_t; \sqrt{1 - \beta_t} \cdot \mathbf{x}_{t-1}, \beta_t \cdot \mathbf{I})$$

Code:

```
noise = torch.randn_like(x_t)
x_t = torch.sqrt(1 - B[t]) * x_t
      + torch.sqrt(B[t]) * noise
```

Step 2:

- Multiply image at previous step by $\sqrt{1 - \beta_t}$
- Add it to the result from step 1

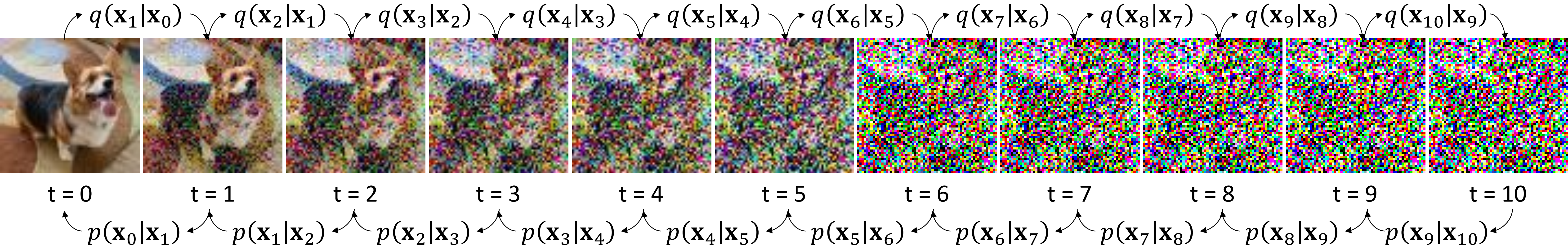
Step 1:

- Generate noise from a standard normal distribution
- Multiply result by $\sqrt{\beta_t}$

Reverse Diffusion

Getting Lost in the Noise

Reverse Diffusion

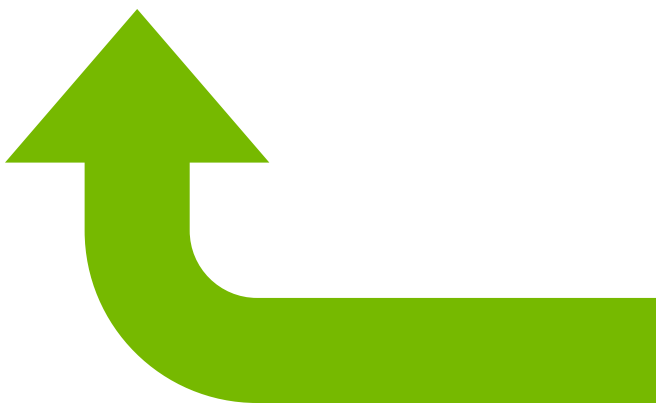


$$p(\mathbf{x}_{t-1}|\mathbf{x}_t) = N(\mathbf{x}_{t-1}; \boldsymbol{\mu}_\theta(\mathbf{x}_t, t), \Sigma_\theta(\mathbf{x}_t, t))$$

Difficult to calculate

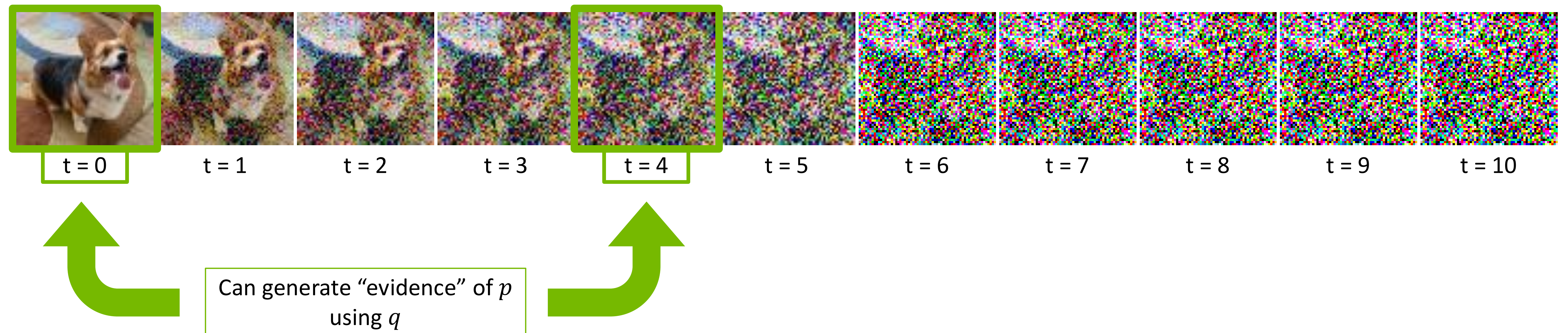


Approximate average used to
create $\mathbf{x}_t$



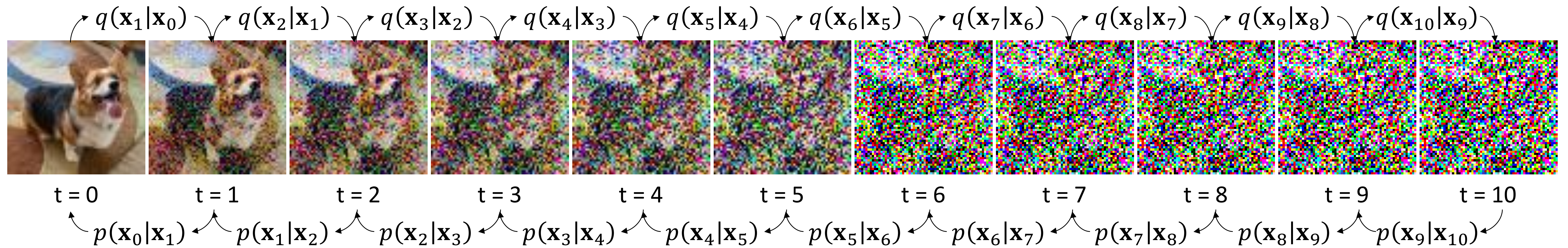
Getting Lost in the Noise

Reverse Diffusion



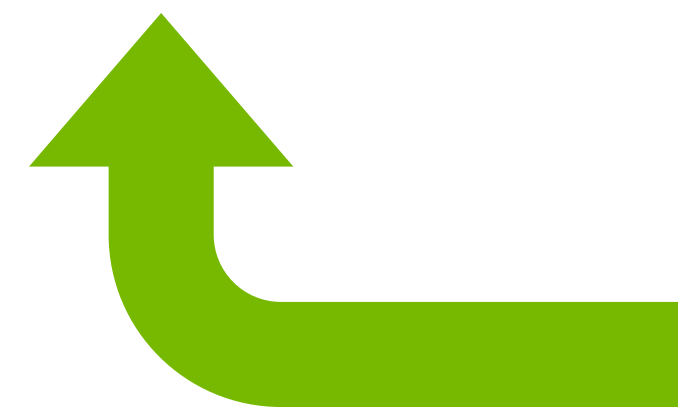
Getting Lost in the Noise

Reverse Diffusion



$$p(\mathbf{x}_{t-1} | \mathbf{x}_t) = N(\mathbf{x}_{t-1}; \boldsymbol{\mu}_\theta(\mathbf{x}_t, t), \Sigma_\theta(\mathbf{x}_t, t))$$

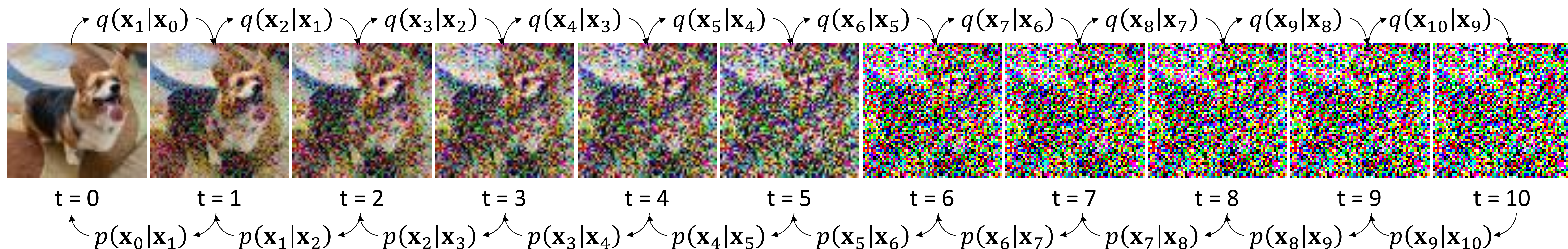
$$q(\mathbf{x}_{t-1} | \mathbf{x}_t, \mathbf{x}_0) = N(\mathbf{x}_{t-1}; \tilde{\boldsymbol{\mu}}(\mathbf{x}_t, \mathbf{x}_0), \tilde{\boldsymbol{\beta}}_t \cdot \mathbf{I}) \quad \tilde{\boldsymbol{\beta}}_t = \frac{1 - \bar{\alpha}_{t-1}}{1 - \bar{\alpha}_t} \boldsymbol{\beta}_t$$



Use Bayes' Rule and ...

Getting Lost in the Noise

Reverse Diffusion



$$p(\mathbf{x}_{t-1}|\mathbf{x}_t) = N(\mathbf{x}_{t-1}; \boldsymbol{\mu}_\theta(\mathbf{x}_t, t), \Sigma_\theta(\mathbf{x}_t, t))$$

$$q(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) = N(\mathbf{x}_{t-1}; \tilde{\boldsymbol{\mu}}(\mathbf{x}_t, \mathbf{x}_0), \tilde{\boldsymbol{\beta}}_t \cdot \mathbf{I})$$

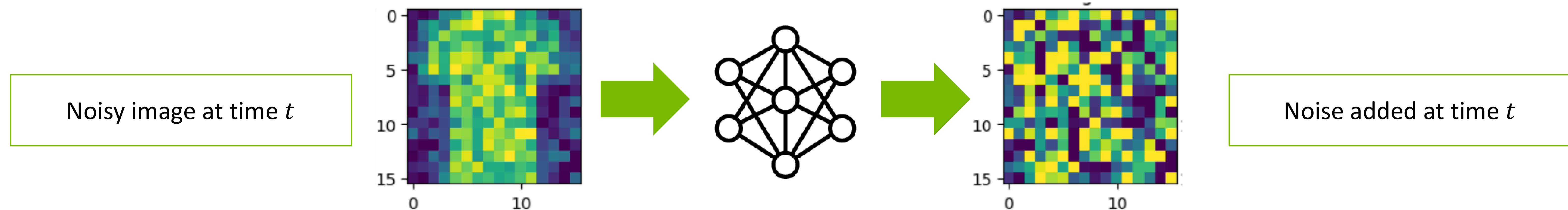
$$\tilde{\boldsymbol{\mu}}_t = \frac{1}{\sqrt{\alpha_t}} \left(x_t - \frac{1 - \alpha_t}{\sqrt{1 - \alpha_t}} \epsilon_t \right)$$

The noise added at time t
This is what the neural network will estimate

Model Training

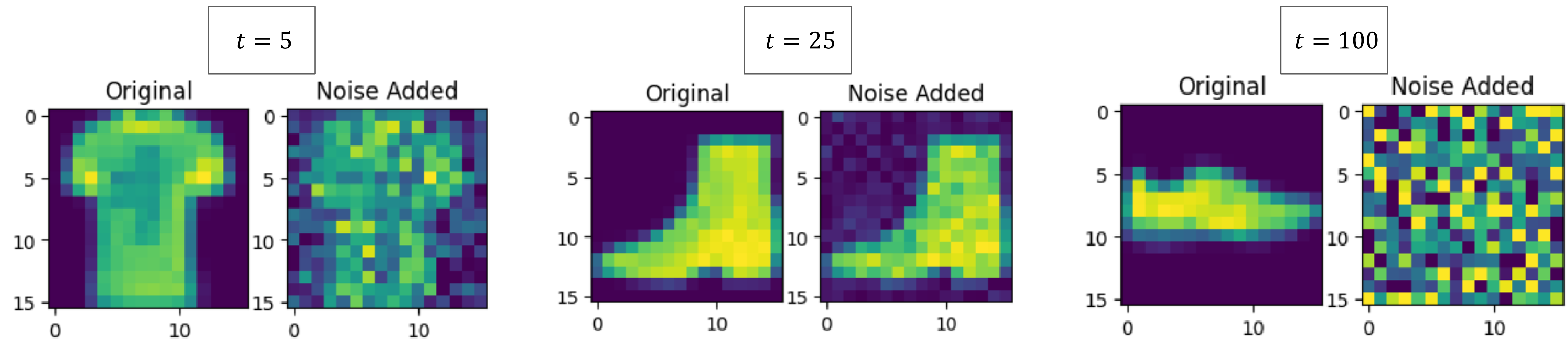
Model Training

Training Data



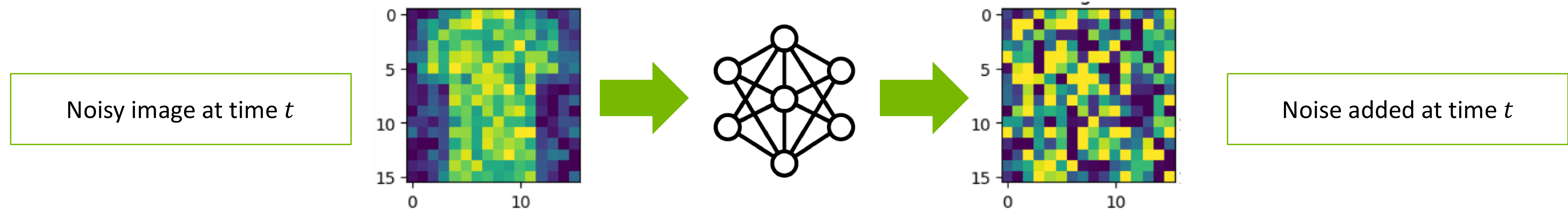
$$q(\mathbf{x}_t | \mathbf{x}_0) = N(\mathbf{x}_t; \sqrt{\bar{\alpha}_t} \cdot \mathbf{x}_0, (1 - \bar{\alpha}_t) \cdot \mathbf{I})$$

“Skip ahead” function



Model Training

ELBO Loss



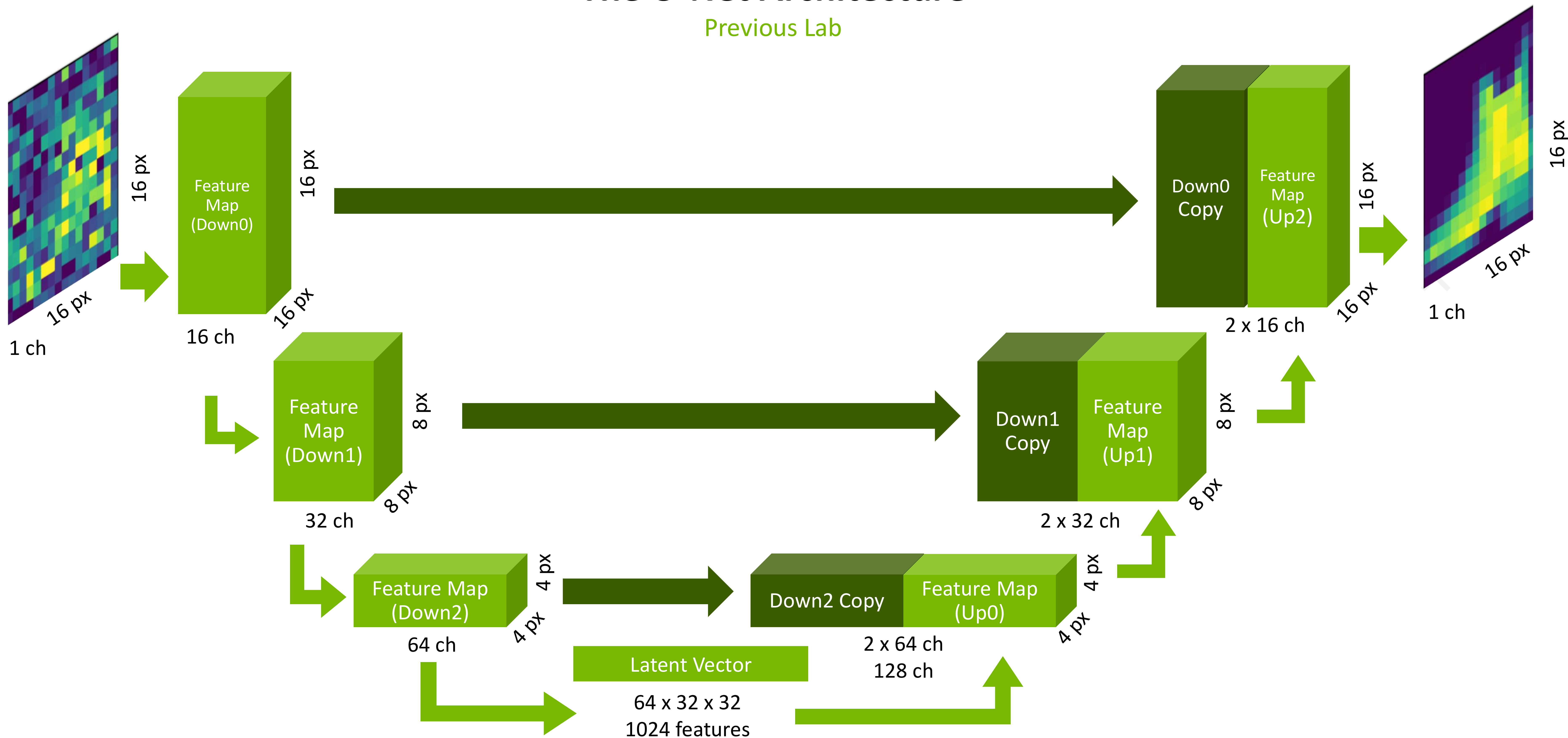
~~$ELBO = \mathbb{E}_{q(\phi)} \left[\log p(y_i | \phi) \right] - \mathbb{E}_{q(\phi)} \left[\log p(\phi) \right]$~~

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Model Architecture

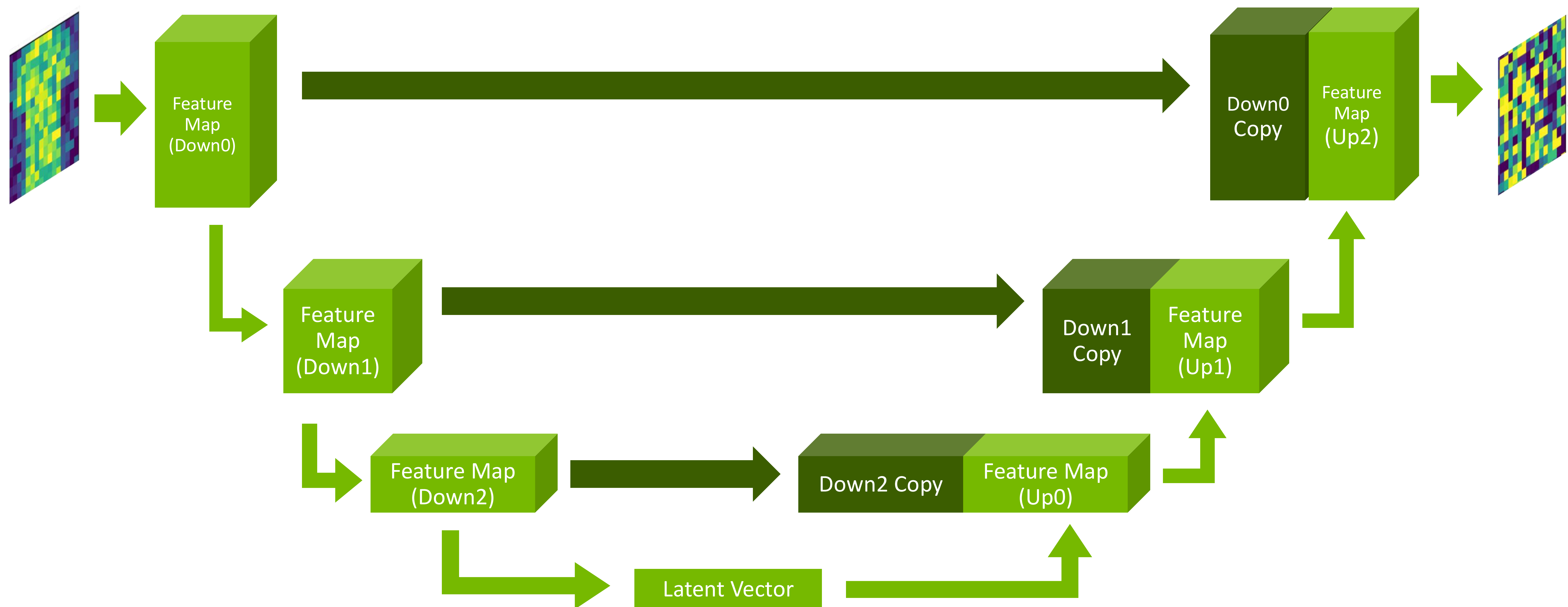
The U-Net Architecture

Previous Lab



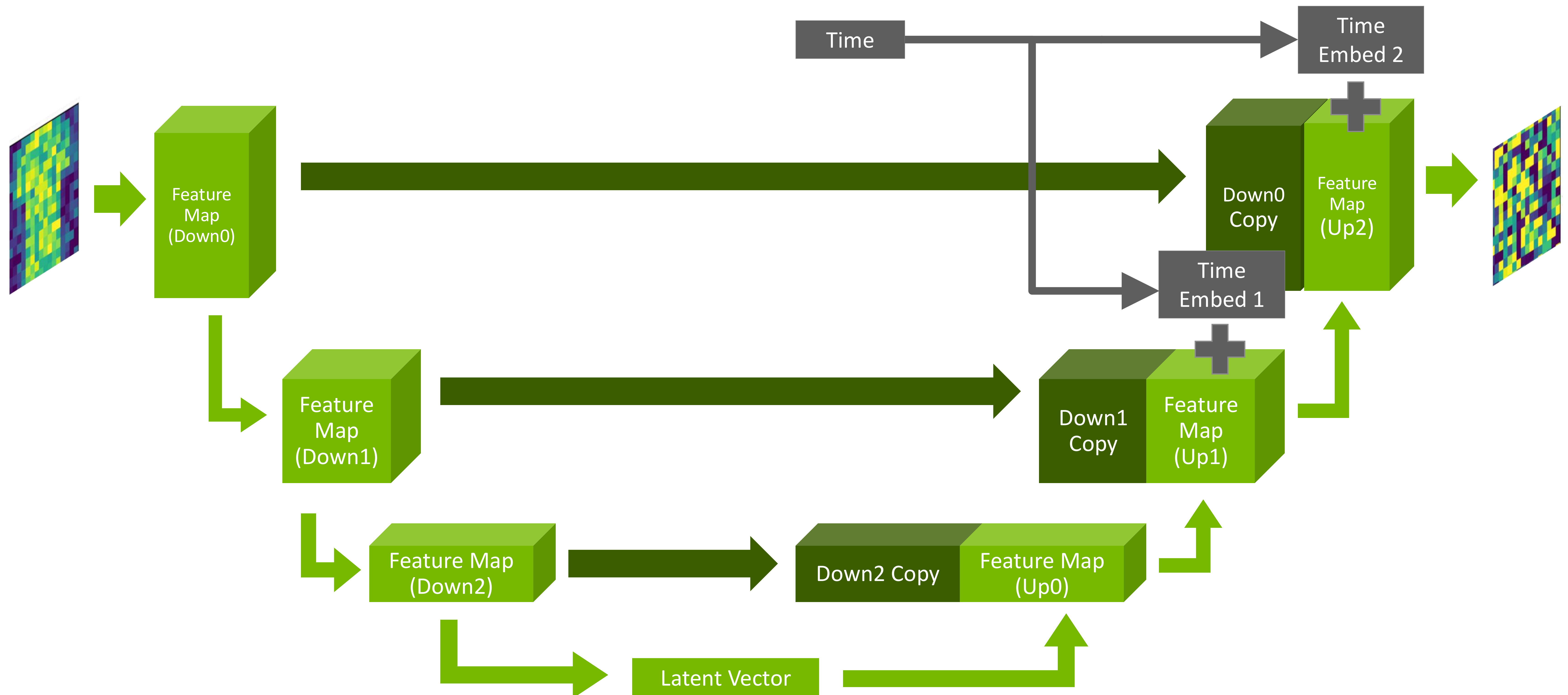
The U-Net Architecture

Next Lab



It's About Time

Adding Time



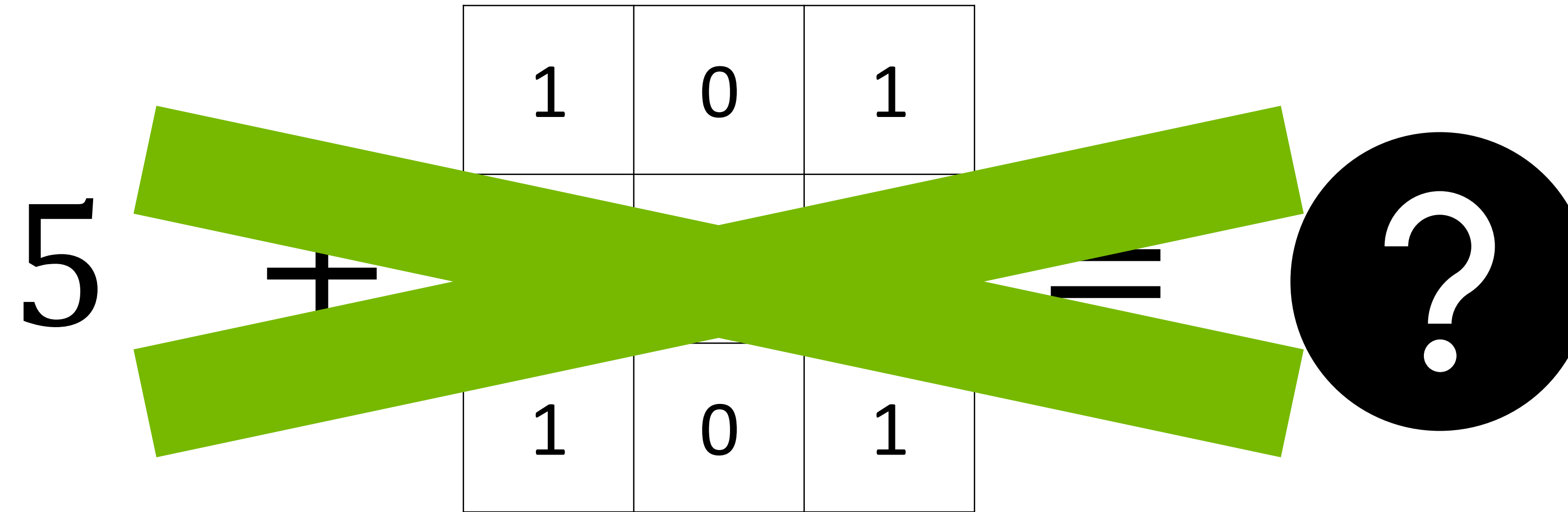
It's About Time

Broadcasting

5 +

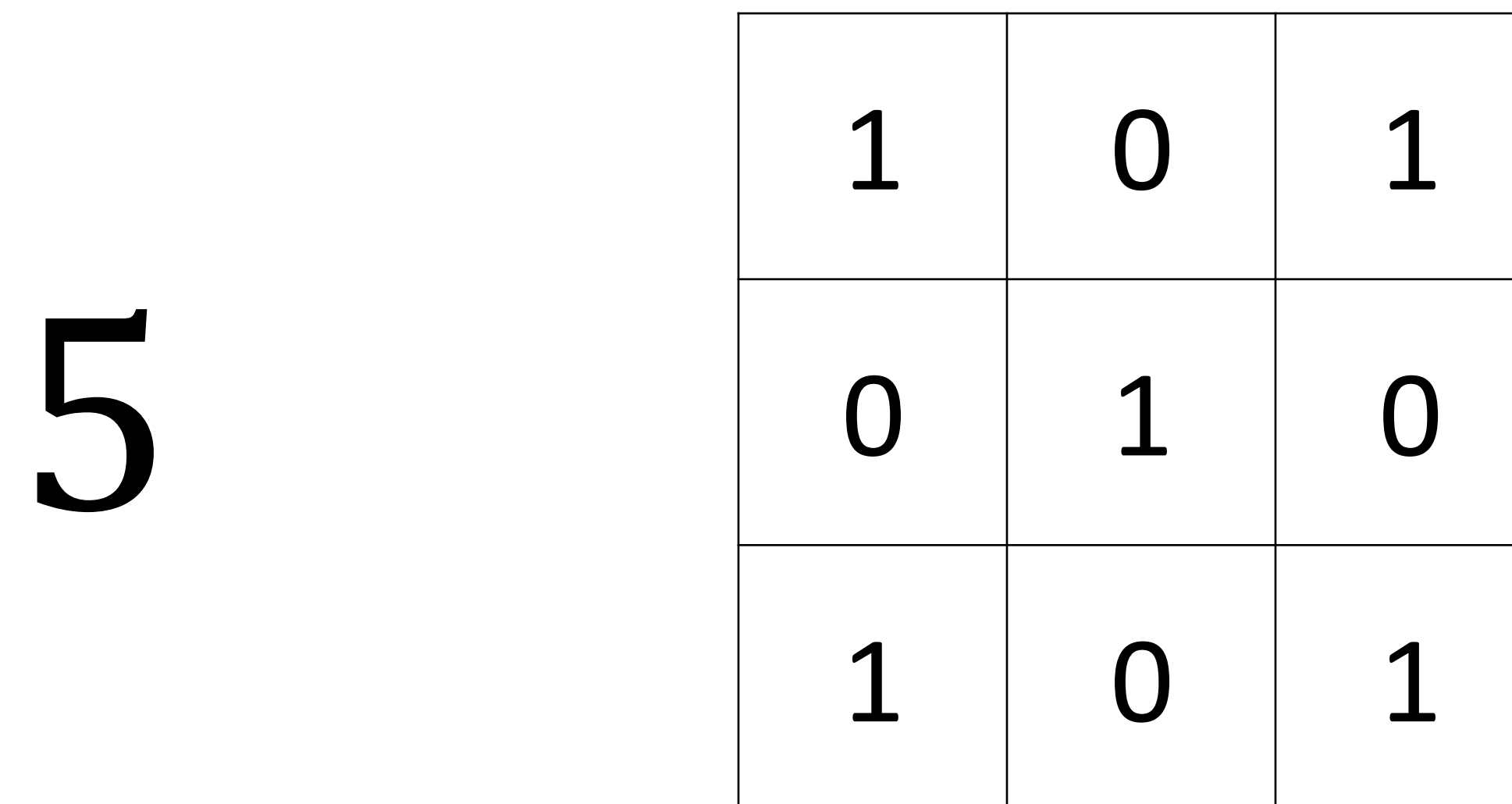
1	0	1
1	0	1

 = ?



5

1	0	1
0	1	0
1	0	1



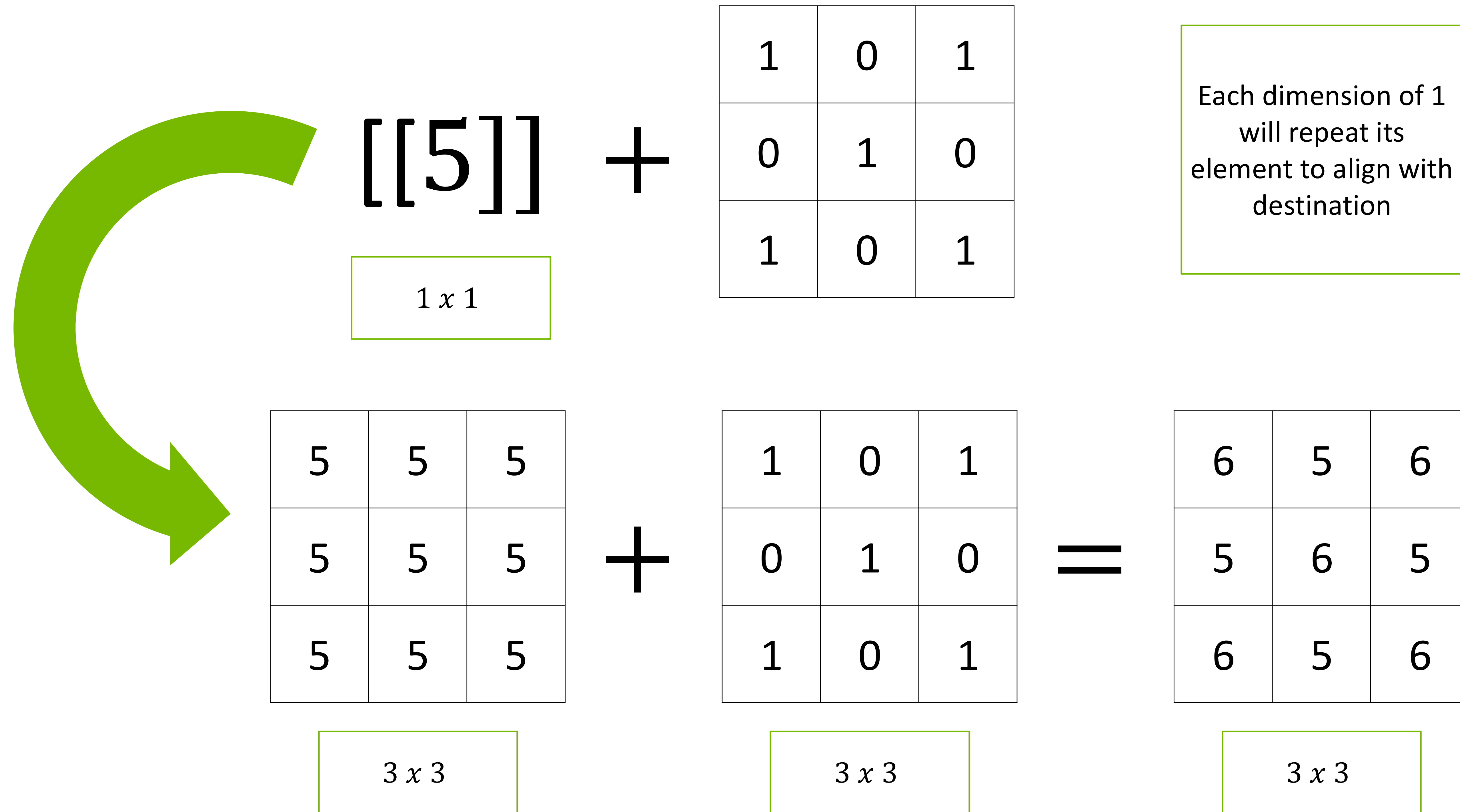
0 or 1

3 x 3

Dimensions

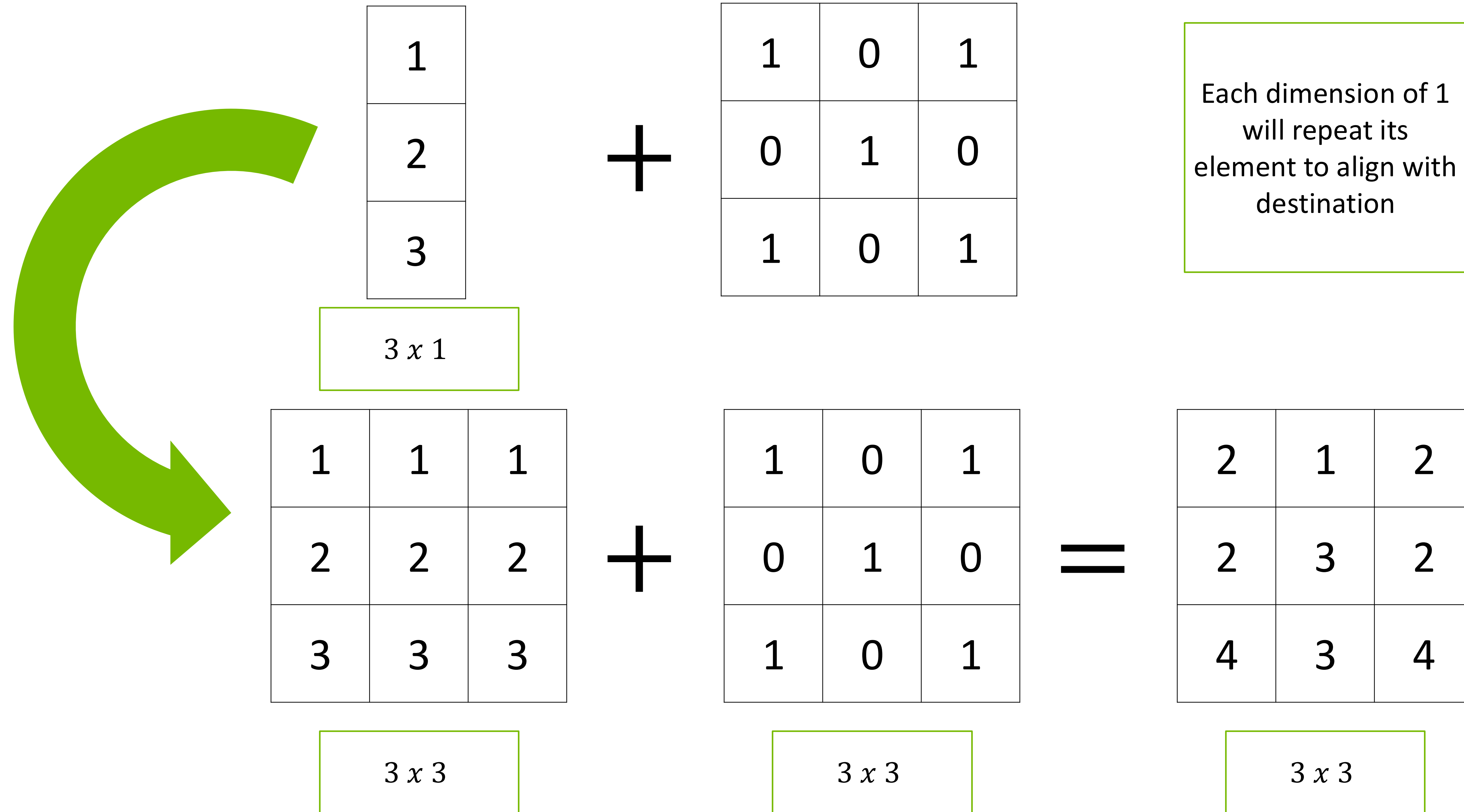
It's About Time

Broadcasting

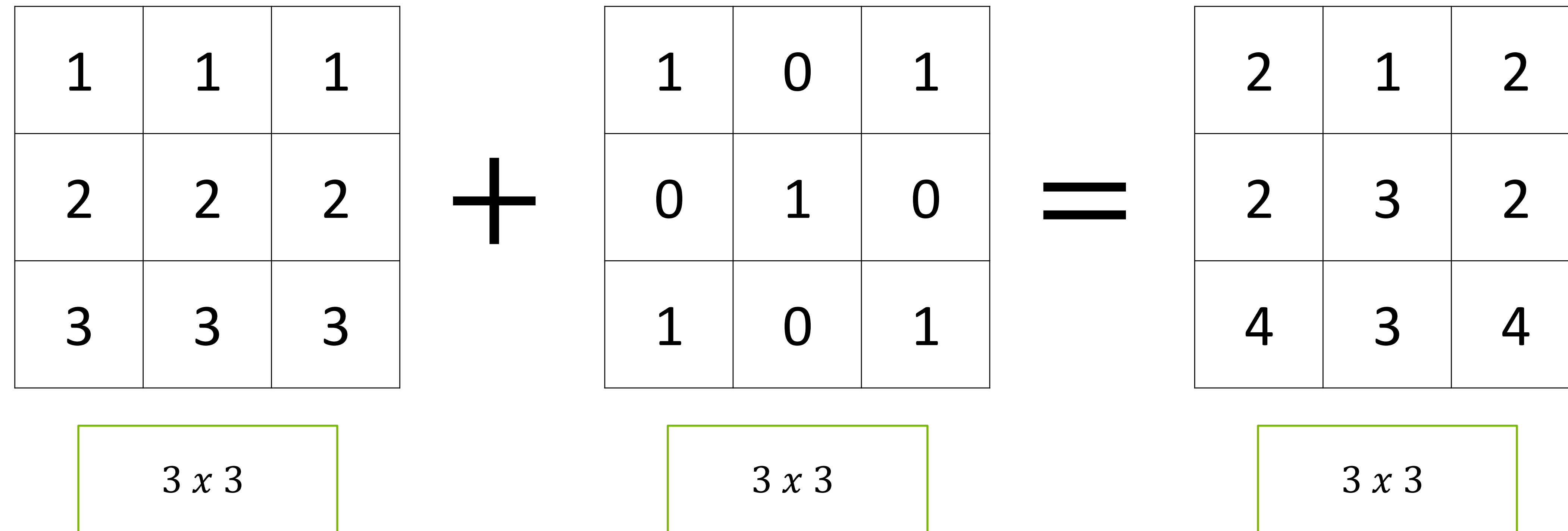


It's About Time

Broadcasting



Each dimension of 1
will repeat its
element to align with
destination



Let's get started!

2: Denoising Diffusion Probabilistic Models

Generative AI with Diffusion Models

Generative AI with Diffusion Models

Course Progress Bookmarks Updates

Generative AI with Diffusion Models Start Here 2: Denoising Diffusion Probabilistic Models

Generative AI with Diffusion Models

Start Here

0: Server Access

1: From U-Nets to Diffusion

2: Denoising Diffusion Probabilistic Models

3: Optimizations

4: Classifier-Free Diffusion Guidance

5: CLIP

6: Wrap-up and Assessment

Previous

Next



DEEP
LEARNING
INSTITUTE



START

Please click the start button to the top-right. The course will take about 5 minutes to load. The course lecture slide decks will appear here when you launch the course. The numbers on the lectures correspond to the numbers of the interactive notebooks in the course.

2: Denoising Diffusion Probabilistic Models

Generative AI with Diffusion Models

Generative AI with Diffusion Models

Course Progress Bookmarks Updates

Generative AI with Diffusion Models Start Here 2: Denoising Diffusion Probabilistic Models

Generative AI with Diffusion Models

Start Here

0: Server Access

1: From U-Nets to Diffusion

2: Denoising Diffusion Probabilistic Models

3: Optimizations

4: Classifier-Free Diffusion Guidance

5: CLIP

6: Wrap-up and Assessment

Next Steps

Feedback

Previous

Next



LOADING



STOP TASK



Generative AI with Diffusion Models

Part 2: Denoising Diffusion Probabilistic Models

2: Denoising Diffusion Probabilistic Models

Generative AI with Diffusion Models


[Products](#)
[Solutions](#)
[Industries](#)
[For You](#)

[Shop](#)
[Drivers](#)
[Support](#)



Min-Yuh Day

Deep Learning Institute
[Find Training](#)
[Self Paced Courses](#)
[Instructor-Led Workshops](#)
[Educator Programs](#)
[Enterprise Solutions](#)
[Certification](#)
[Resources](#)

Generative AI with Diffusion Models

Course
Progress
Bookmarks
Updates

Generative AI with Diffusion Models
Start Here
2: Denoising Diffusion Probabilistic Models

Generative AI with Diffusion Models

Start Here

0: Server Access

1: From U-Nets to Diffusion

2: Denoising Diffusion Probabilistic Models

3: Optimizations

4: Classifier-Free Diffusion Guidance

5: CLIP

6: Wrap-up and Assessment

Next Steps

Feedback

Previous
Next


10:52 / 10:54
Speed 1.0x


DEEP LEARNING INSTITUTE

This Lab 0 : 23 : 10 / 1 : 30 : 00
Course -68 : 43 : 45 / 24 : 00 : 00

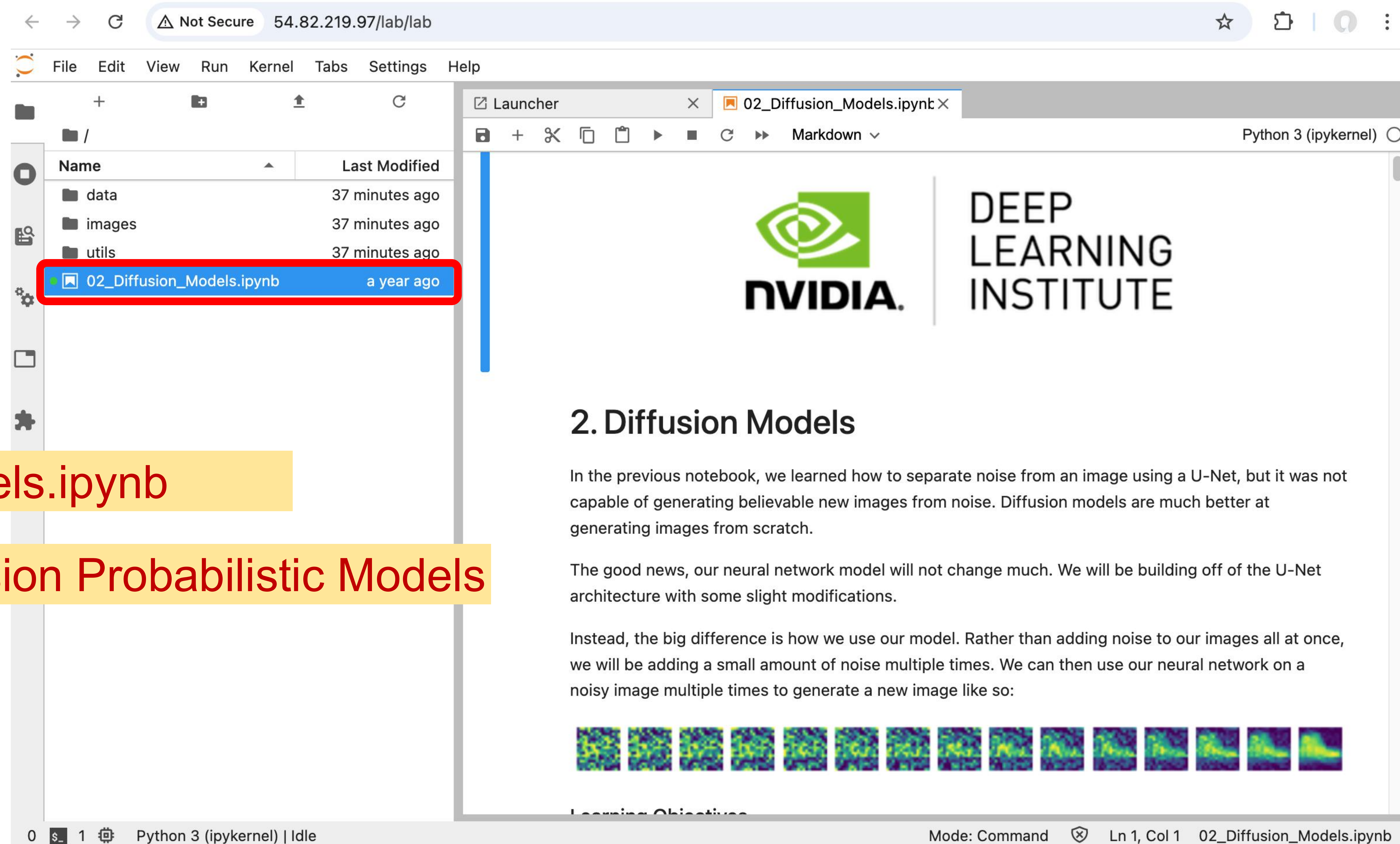

LAUNCH


STOP TASK


Generative AI with Diffusion Models
Part 2: Denoising Diffusion Probabilistic Models

2: Denoising Diffusion Probabilistic Models

Generative AI with Diffusion Models



← → ↻ ⚠ Not Secure 54.82.219.97/lab/lab ☆ 📁 🔍 ⋮

File Edit View Run Kernel Tabs Settings Help

📁 /

Name	Last Modified
data	37 minutes ago
images	37 minutes ago
utils	37 minutes ago
02_Diffusion_Models.ipynb	a year ago

📄 02_Diffusion_Models.ipynk

Python 3 (ipykernel)

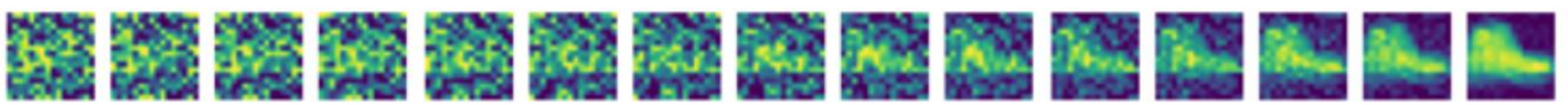
DEEP LEARNING INSTITUTE

2. Diffusion Models

In the previous notebook, we learned how to separate noise from an image using a U-Net, but it was not capable of generating believable new images from noise. Diffusion models are much better at generating images from scratch.

The good news, our neural network model will not change much. We will be building off of the U-Net architecture with some slight modifications.

Instead, the big difference is how we use our model. Rather than adding noise to our images all at once, we will be adding a small amount of noise multiple times. We can then use our neural network on a noisy image multiple times to generate a new image like so:



Learning Objectives

0 \$ 1 Python 3 (ipykernel) | Idle Mode: Command Ln 1, Col 1 02_Diffusion_Models.ipynb

02_Diffusion_Models.ipynb

2: Denoising Diffusion Probabilistic Models

NVIDIA

Generative AI with

Diffusion Models

Part II

[0,1,2] [3, 4] [5, 6]

Generative AI with Diffusion Models

- **Part 1: From U-Nets to Diffusion**
- **Part 2: Denoising Diffusion Probabilistic Models**
- **Part 3: Optimizations**
- **Part 4: Classifier-Free Diffusion Guidance (CFG)**
- **Part 5: CLIP (Contrastive Language-Image Pre-Training)**
- **Part 6: Wrap-up & Assessment**

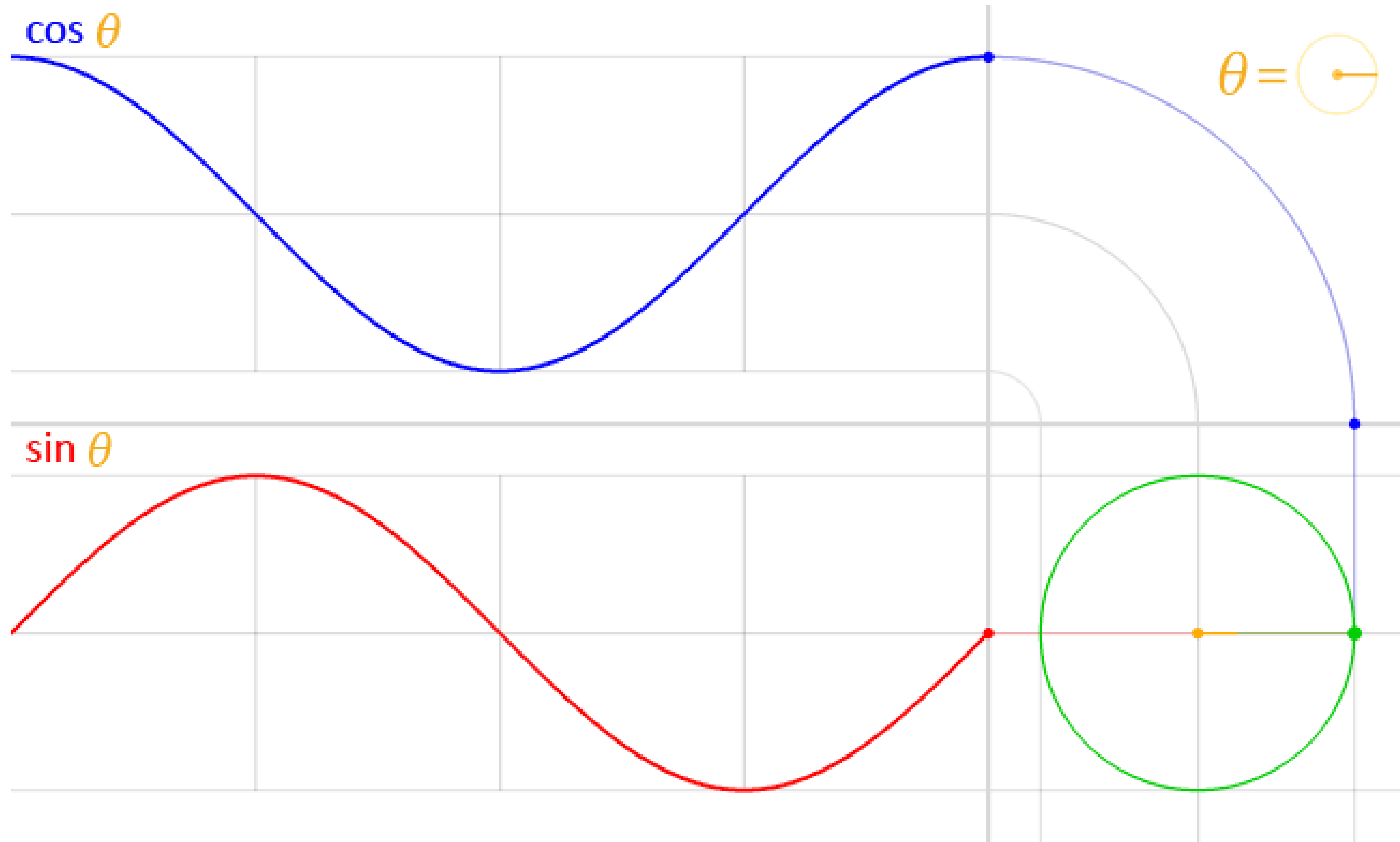
Generative AI with Diffusion Models

Part 3: Optimizations

Technique	Key Benefit
Group Normalization	Stable training
Gaussian Error Linear Unit (GELU) Activation	Smoother learning
Rearrange Pooling	Smooth upsampling
Sinusoidal Position Embeddings (Sine Wave)	Better timestep understanding
Deeper Networks	Finer image generation

Sinusoidal Position Embeddings for Diffusion Model

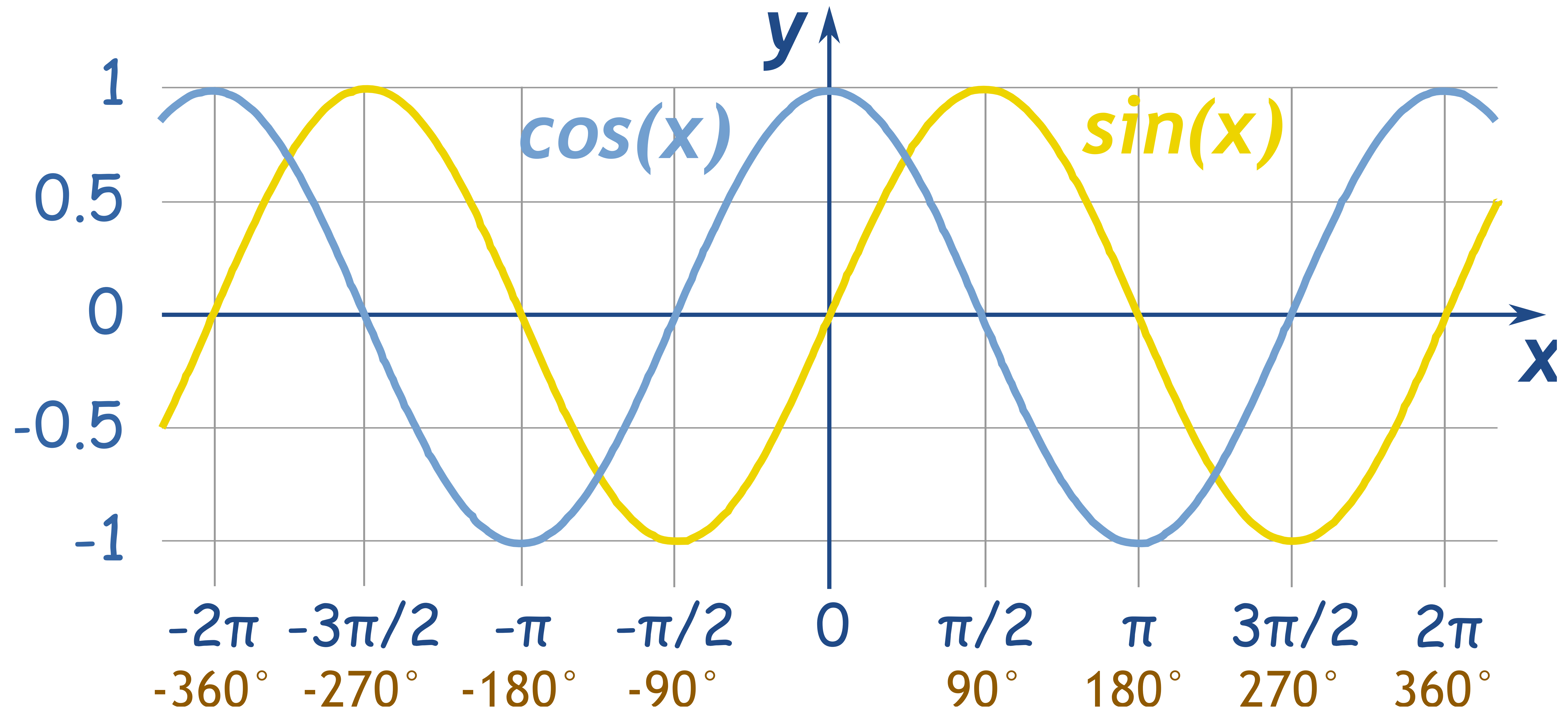
Sine and Cosine Waves to Represent Time as a Sequence



Source: https://en.m.wikipedia.org/wiki/File:Circle_cos_sin.gif

Sinusoidal Position Embeddings for Diffusion Model

Sine and Cosine Waves to Represent Time as a Sequence



Generative AI with Diffusion Models

 Products Solutions Industries For You

Shop Drivers Support    Min-Yuh Day

Deep Learning Institute Find Training Self Paced Courses Instructor-Led Workshops Educator Programs Enterprise Solutions Certification Resources

Course Progress Bookmarks Updates

Generative AI with Diffusion Models Start Here 3: Optimizations

Generative AI with Diffusion Models

Start Here

0: Server Access

1: From U-Nets to Diffusion

2: Denoising Diffusion Probabilistic Models

3: Optimizations

4: Classifier-Free Diffusion Guidance

5: CLIP

6: Wrap-up and Assessment

Next Steps

Feedback

Previous

Next

▶ 8:25 / 8:28

▶ Speed 1.0x  

 DEEP LEARNING INSTITUTE

**Generative AI with Diffusion Models**
Part 3: Optimizations

LOADING  STOP TASK 

START

3: Optimizations



Generative AI with Diffusion Models Part 3: Optimizations

Agenda

- Part 1: From U-Nets to Diffusion

- Part 2: Denoising Diffusion Probabilistic Models

- Part 3: Optimizations

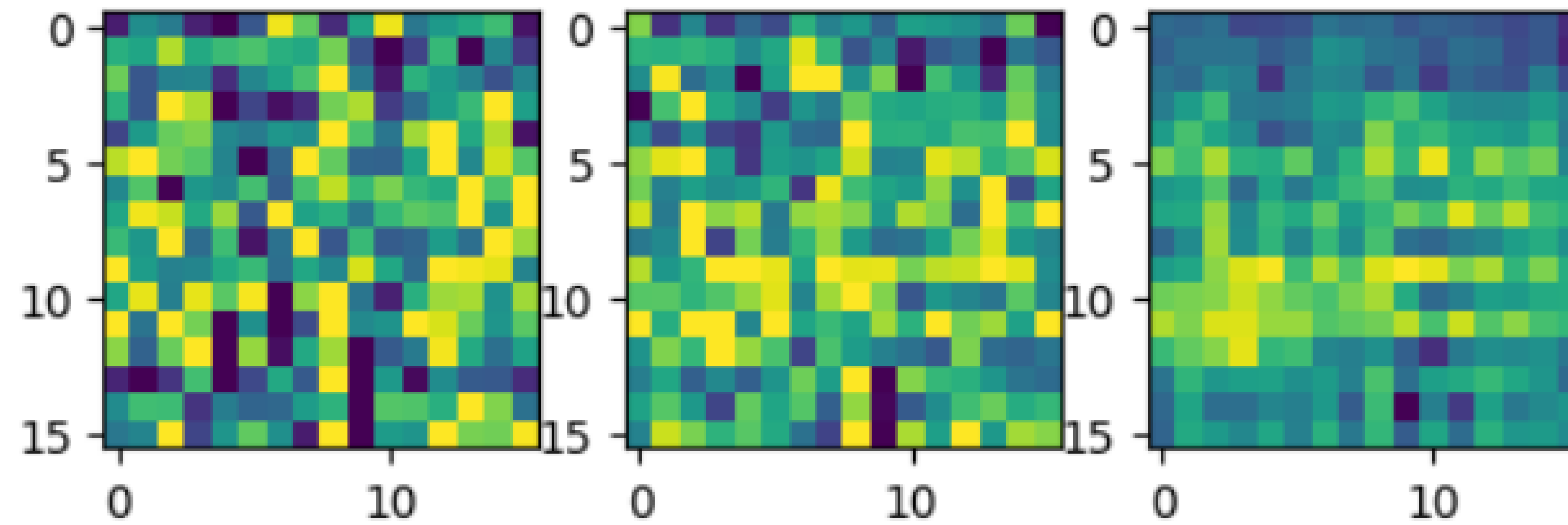
- Part 4: Classifier-Free Diffusion Guidance

- Part 5: CLIP

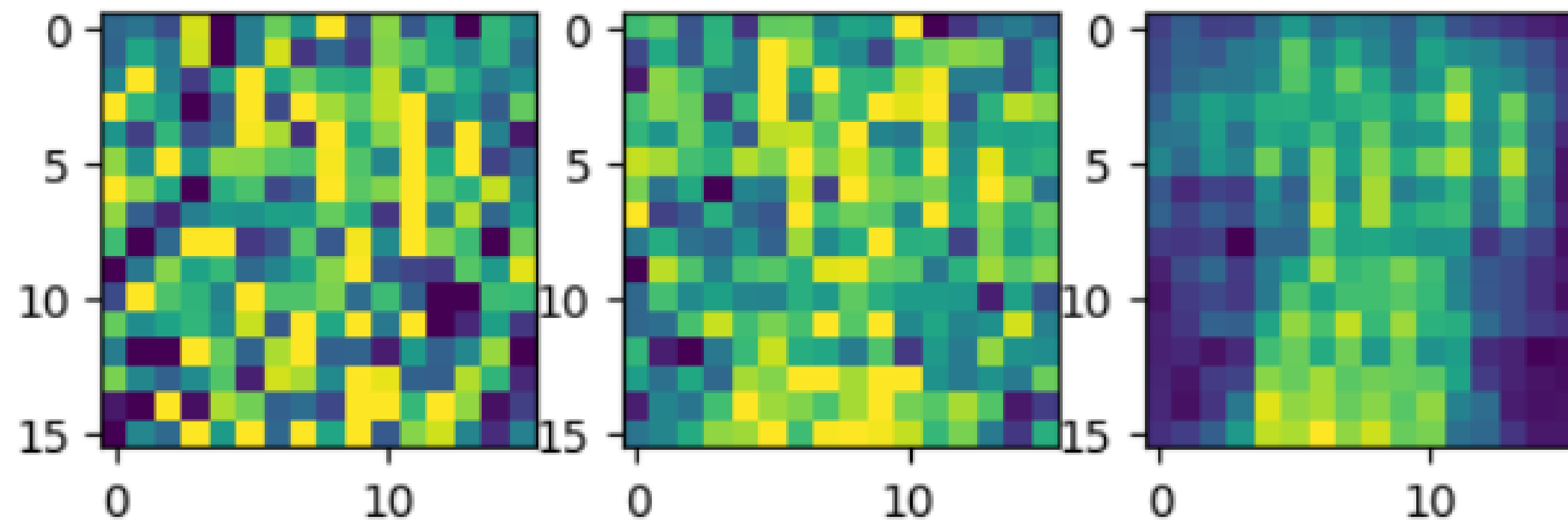
- Part 6: Wrap-up & Assessment

The Checkerboard Problem

The Checkerboard Problem



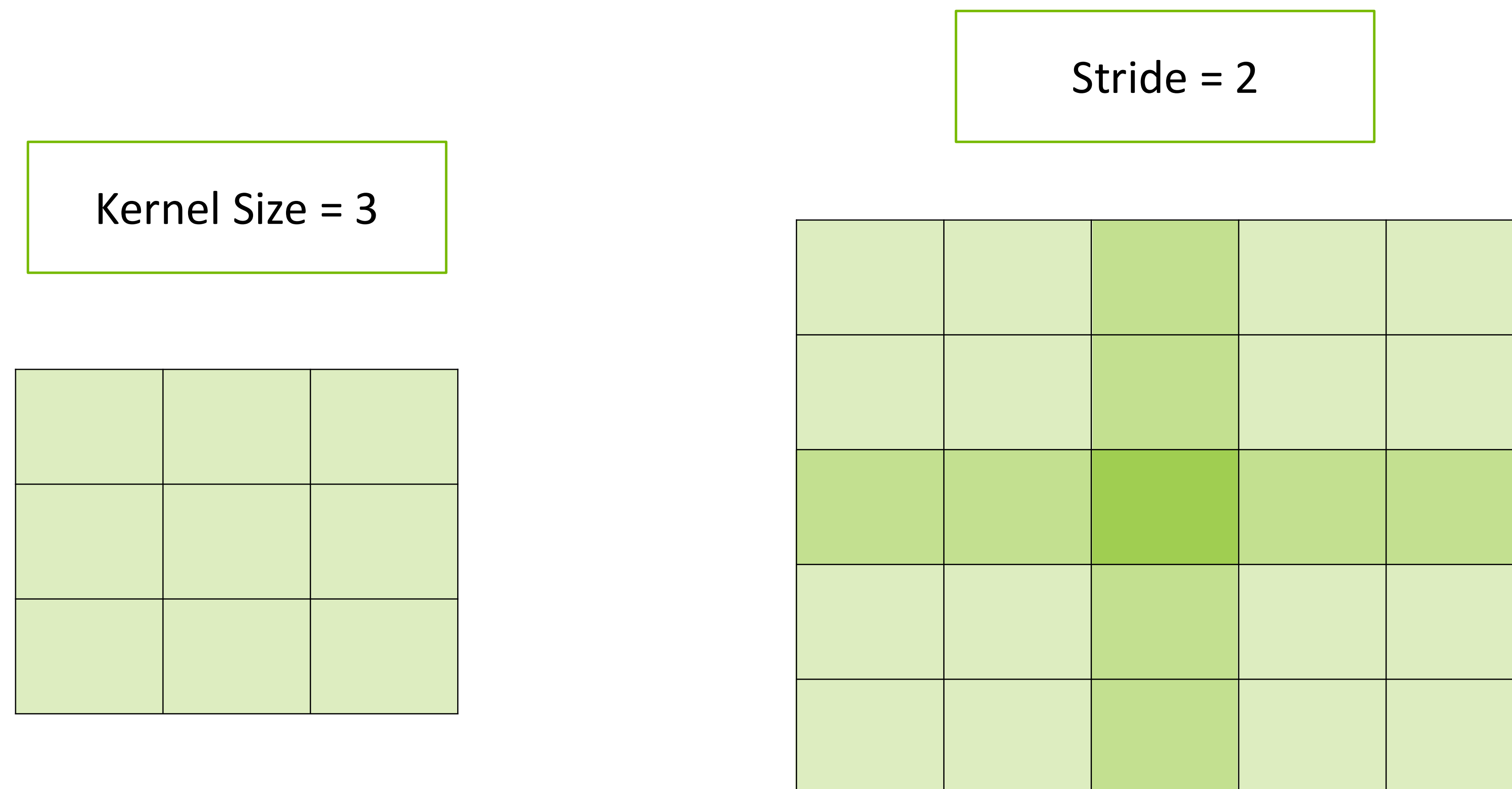
Shoe?



Shirt?

The Checkerboard Problem

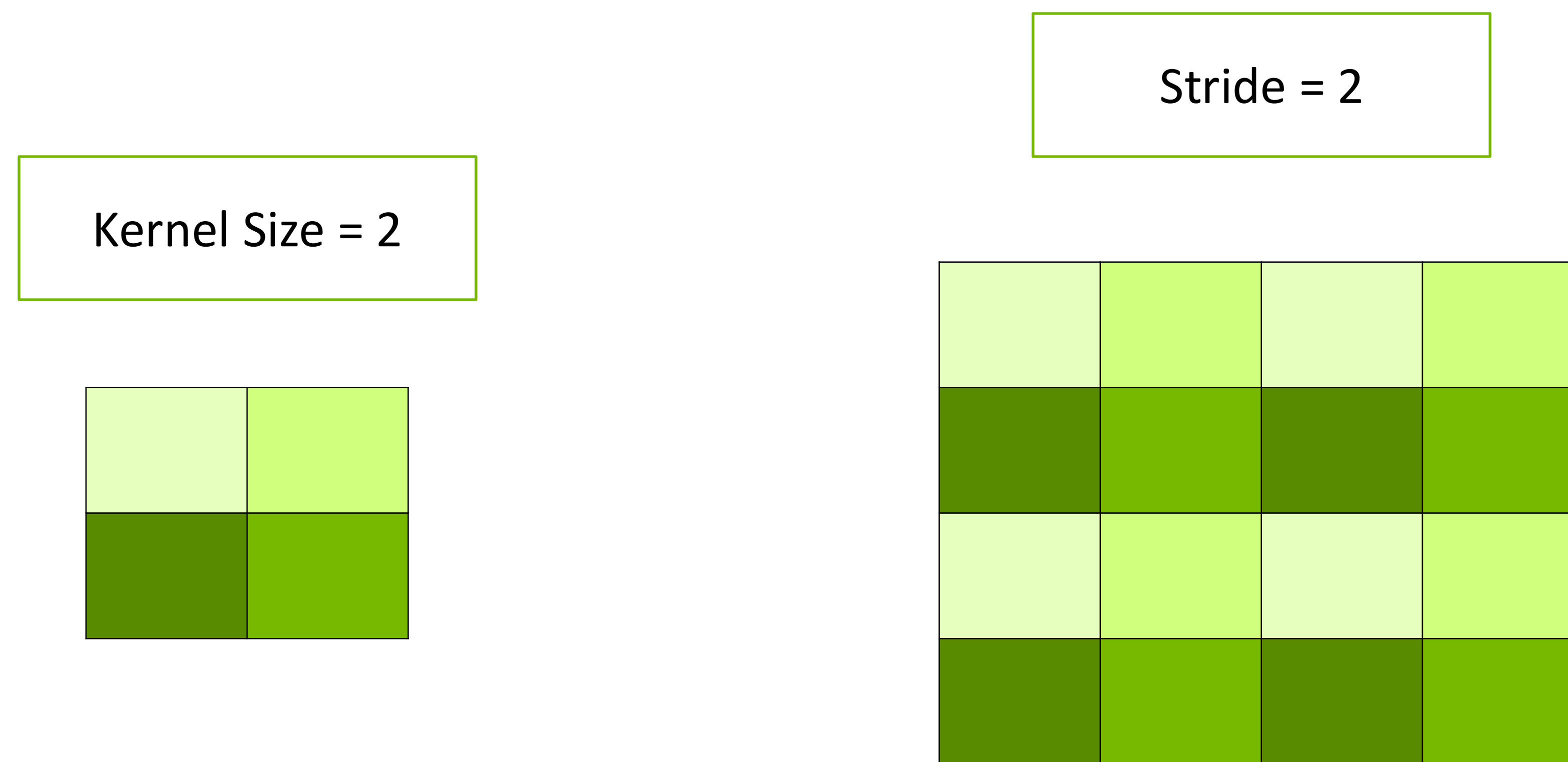
The Impact of Stride



A cozy green plaid blanket, fairy tale drawing

The Checkerboard Problem

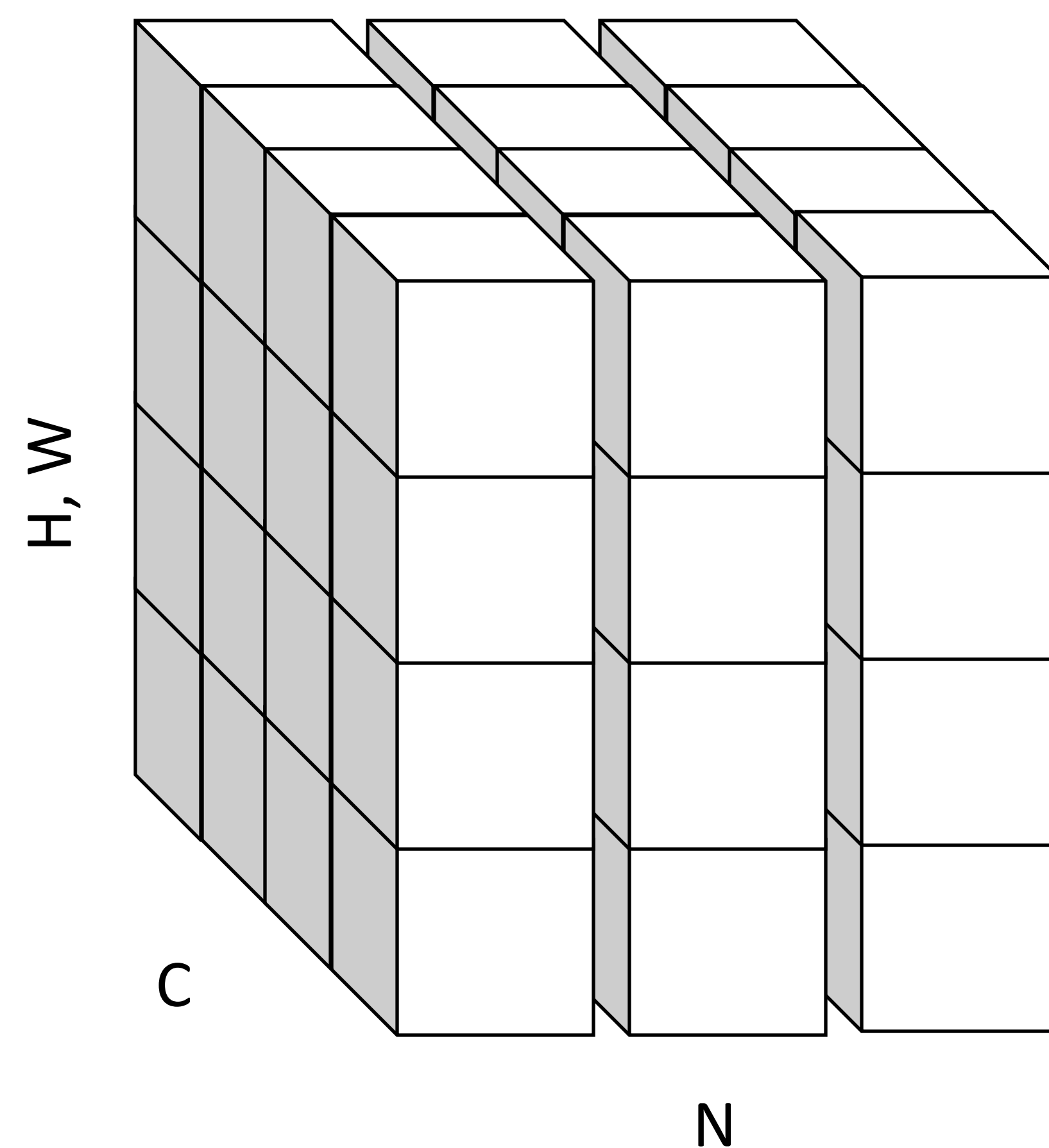
The Impact of Stride



Group Normalization

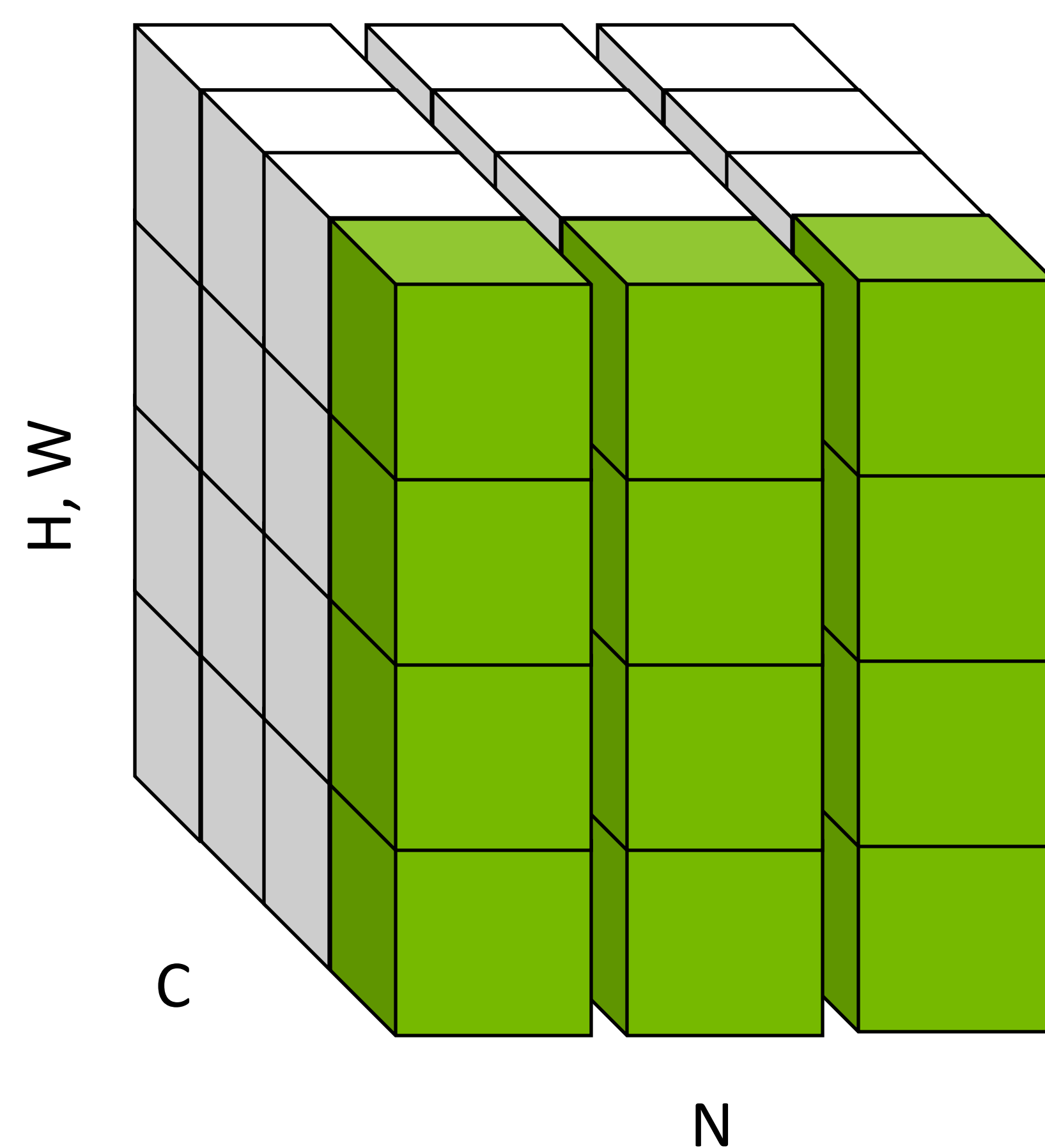
Group Normalization

Batch Normalization Review



Group Normalization

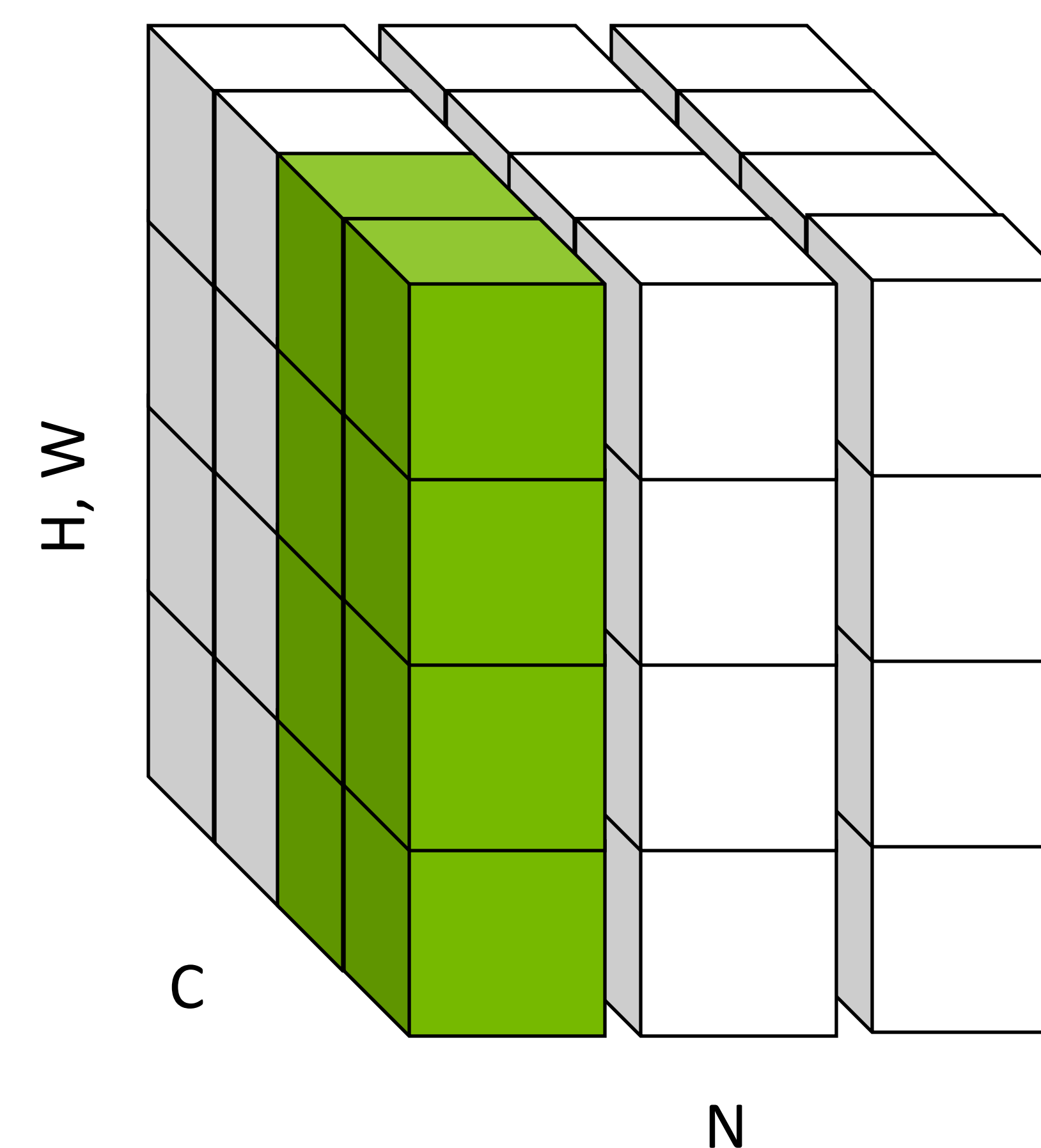
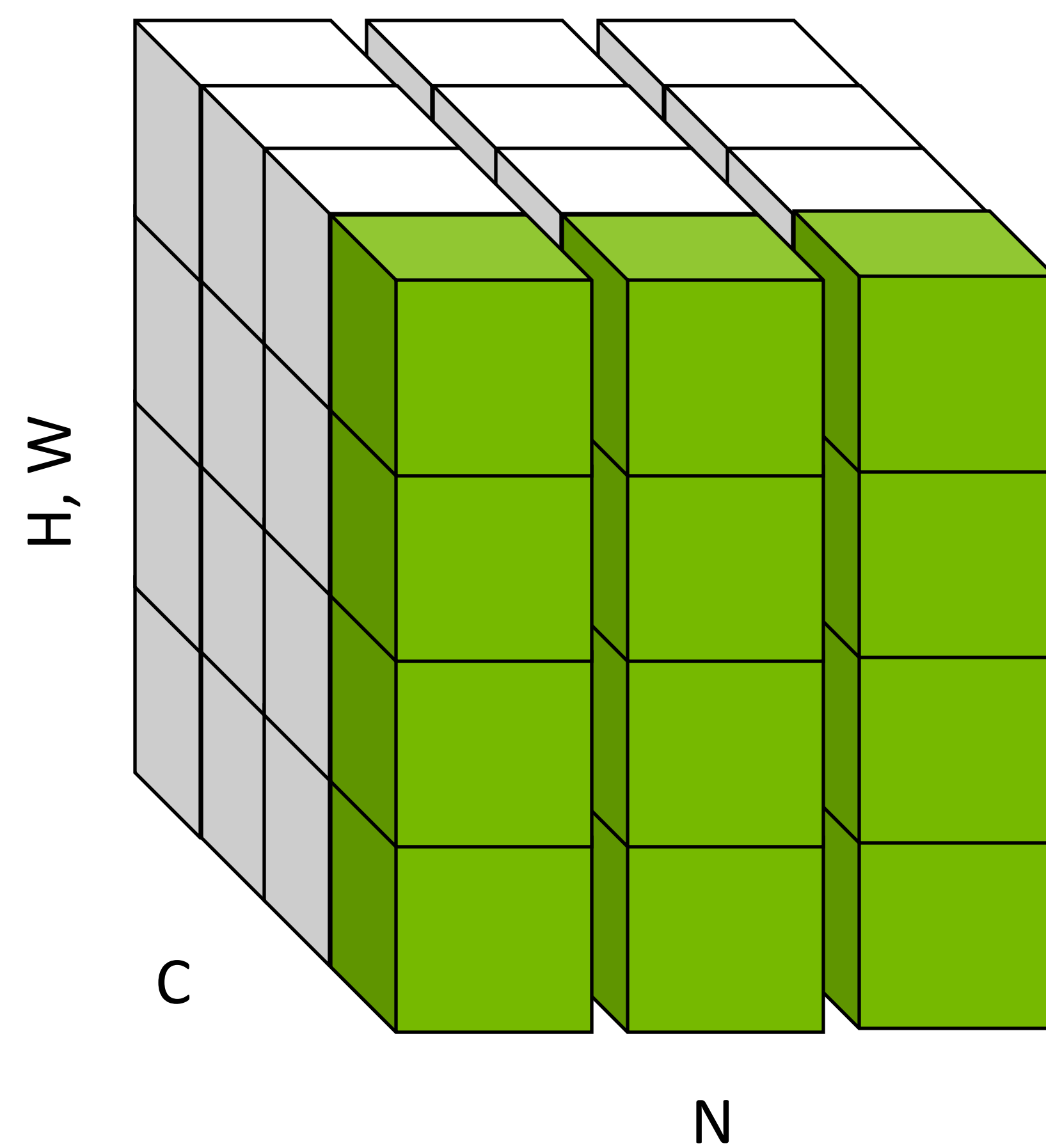
Batch Normalization Review



Batch Normalization

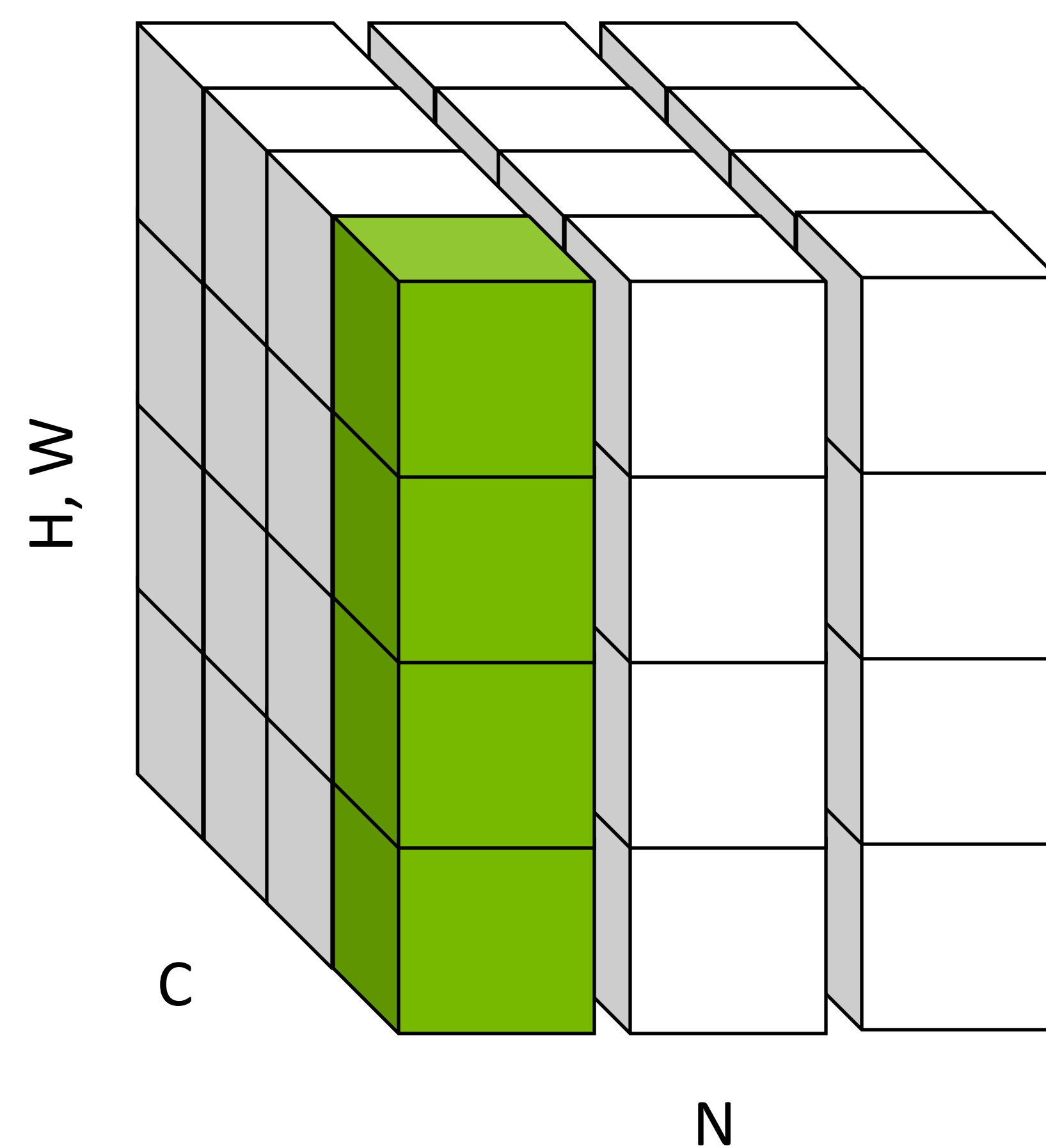
Group Normalization

Types of Group Normalization



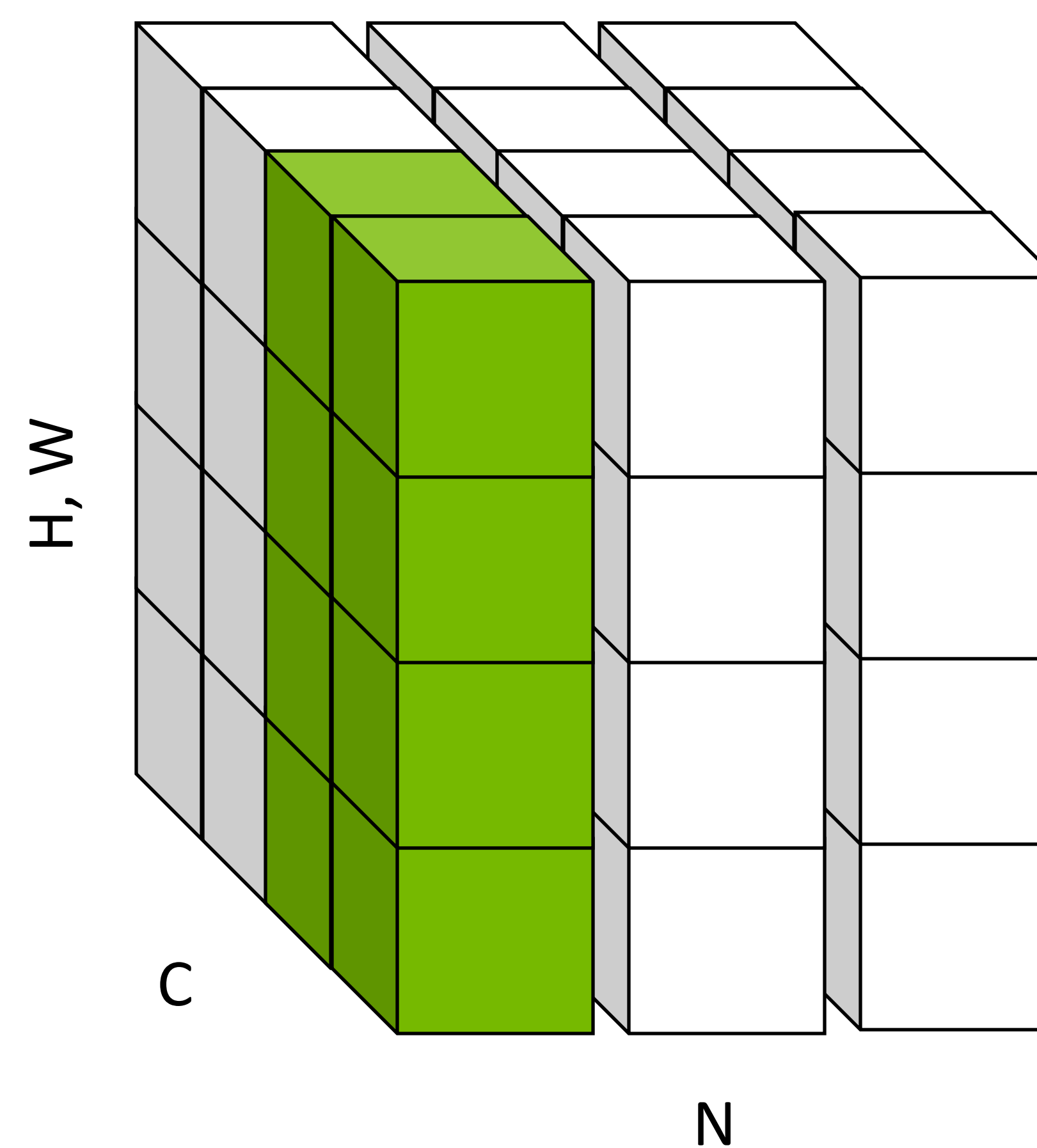
Group Normalization

Types of Group Normalization



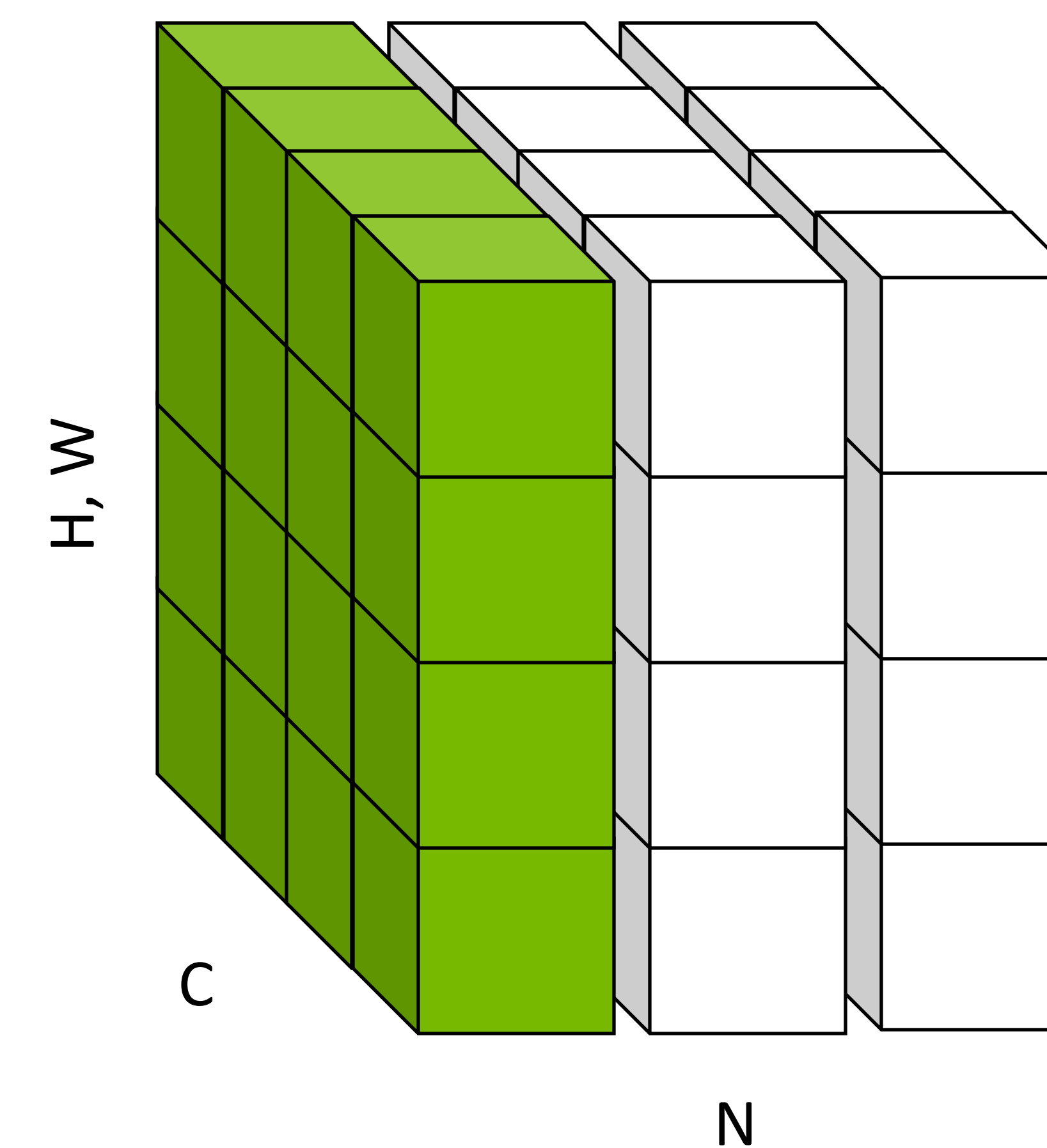
Instance Normalization

Group Size = 1



Group Normalization

Group Size = 2

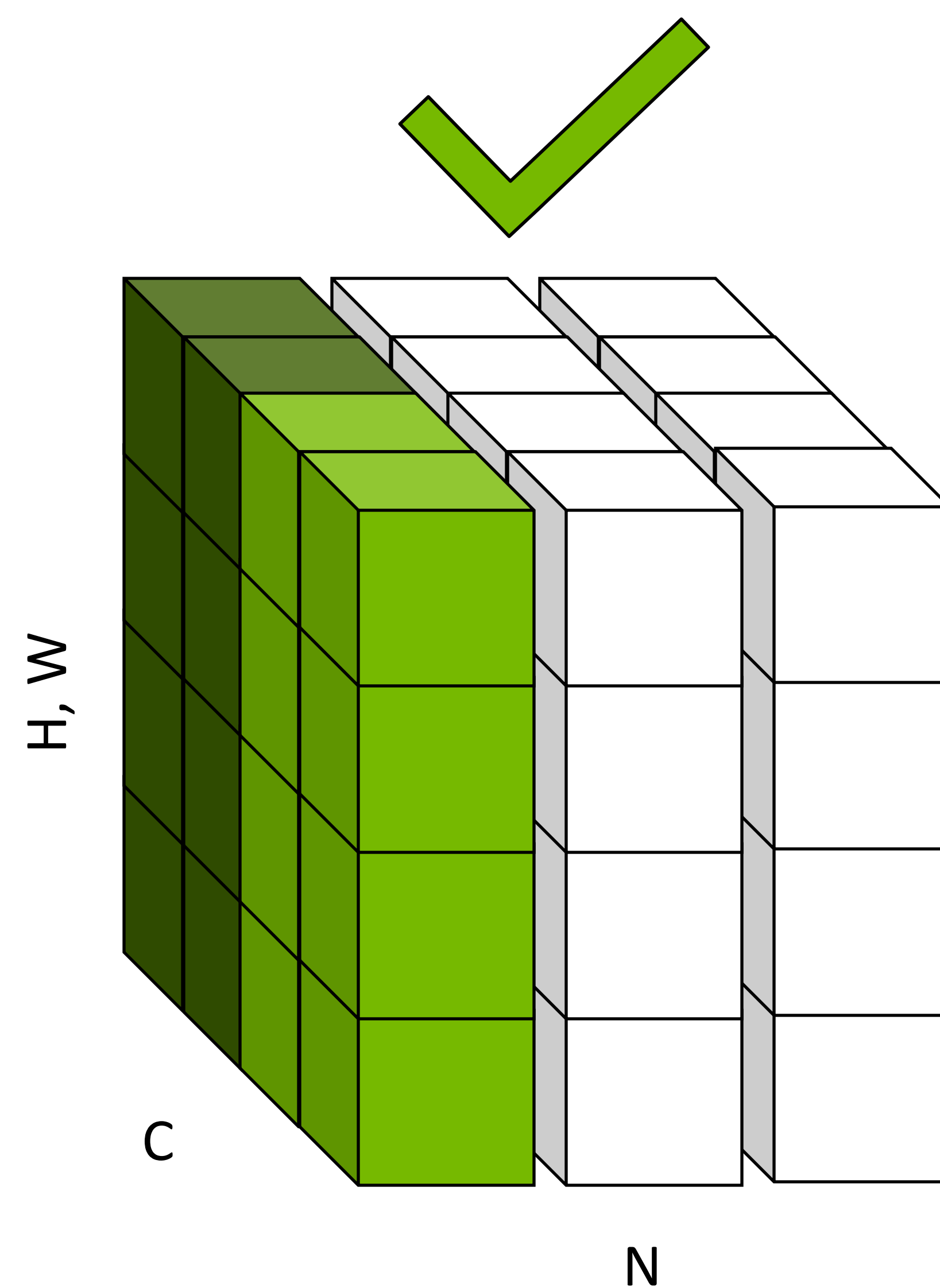


Layer Normalization

Group Size = All

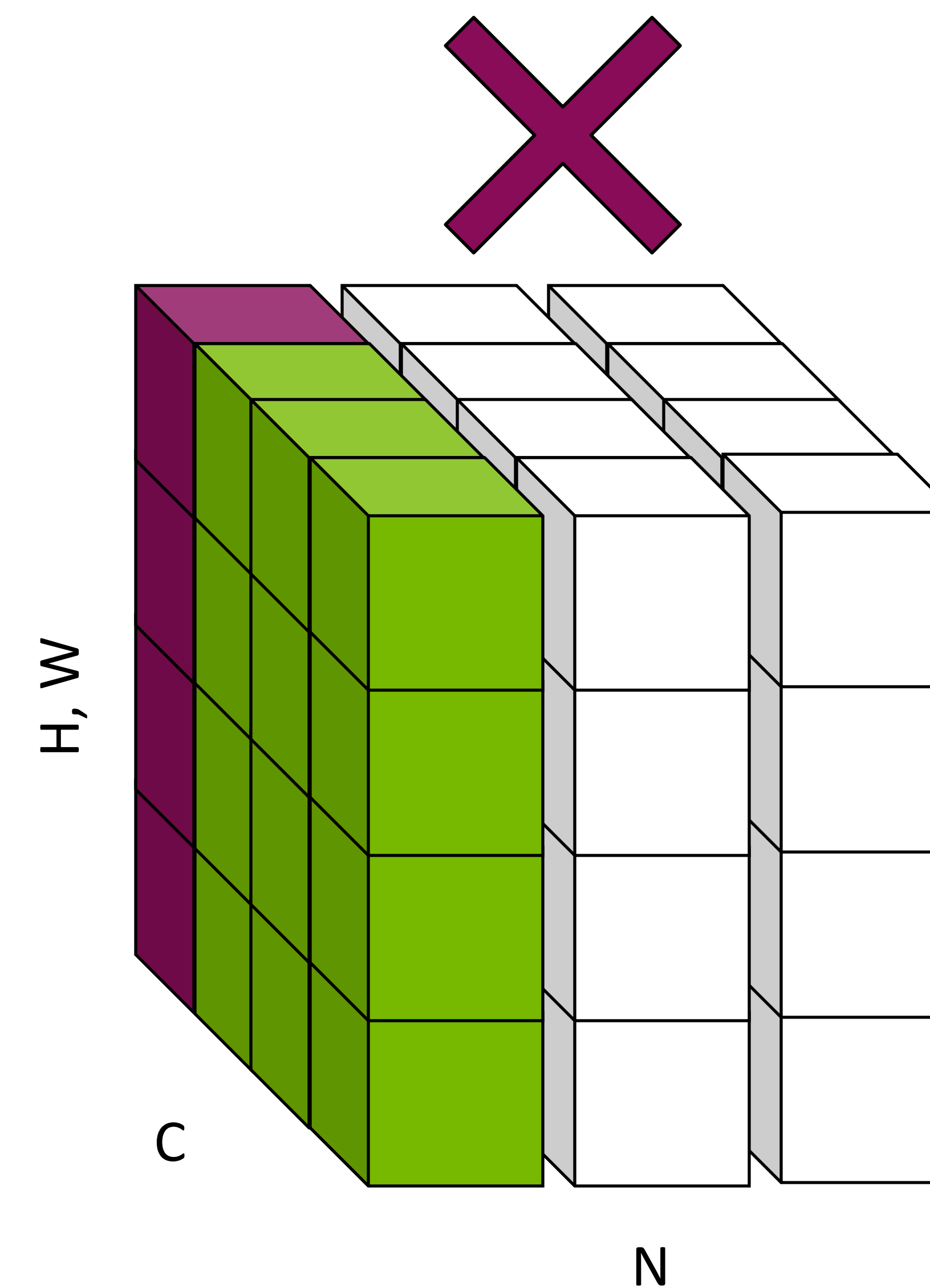
Group Normalization

Types of Group Normalization



Group Normalization

Group Size = 2



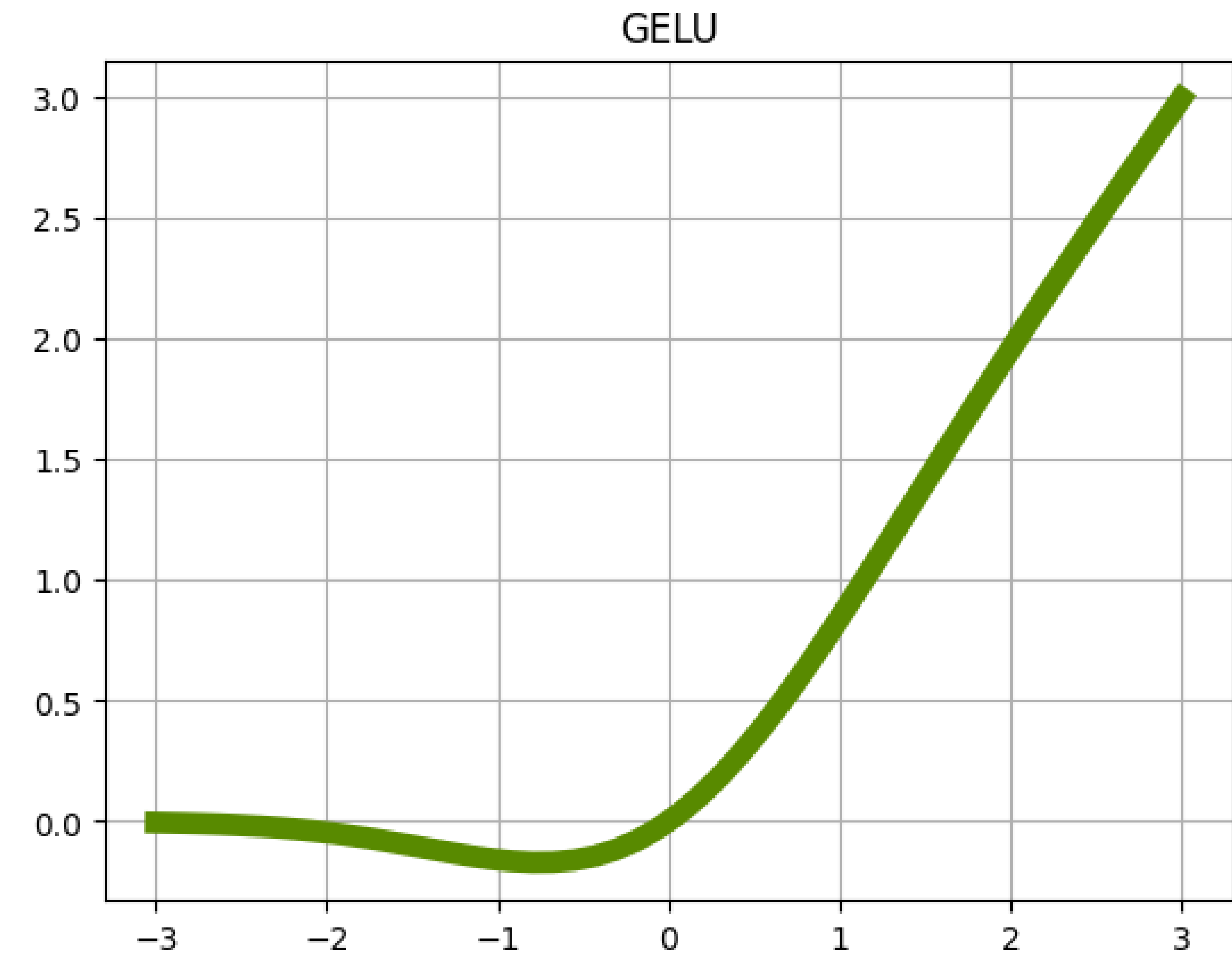
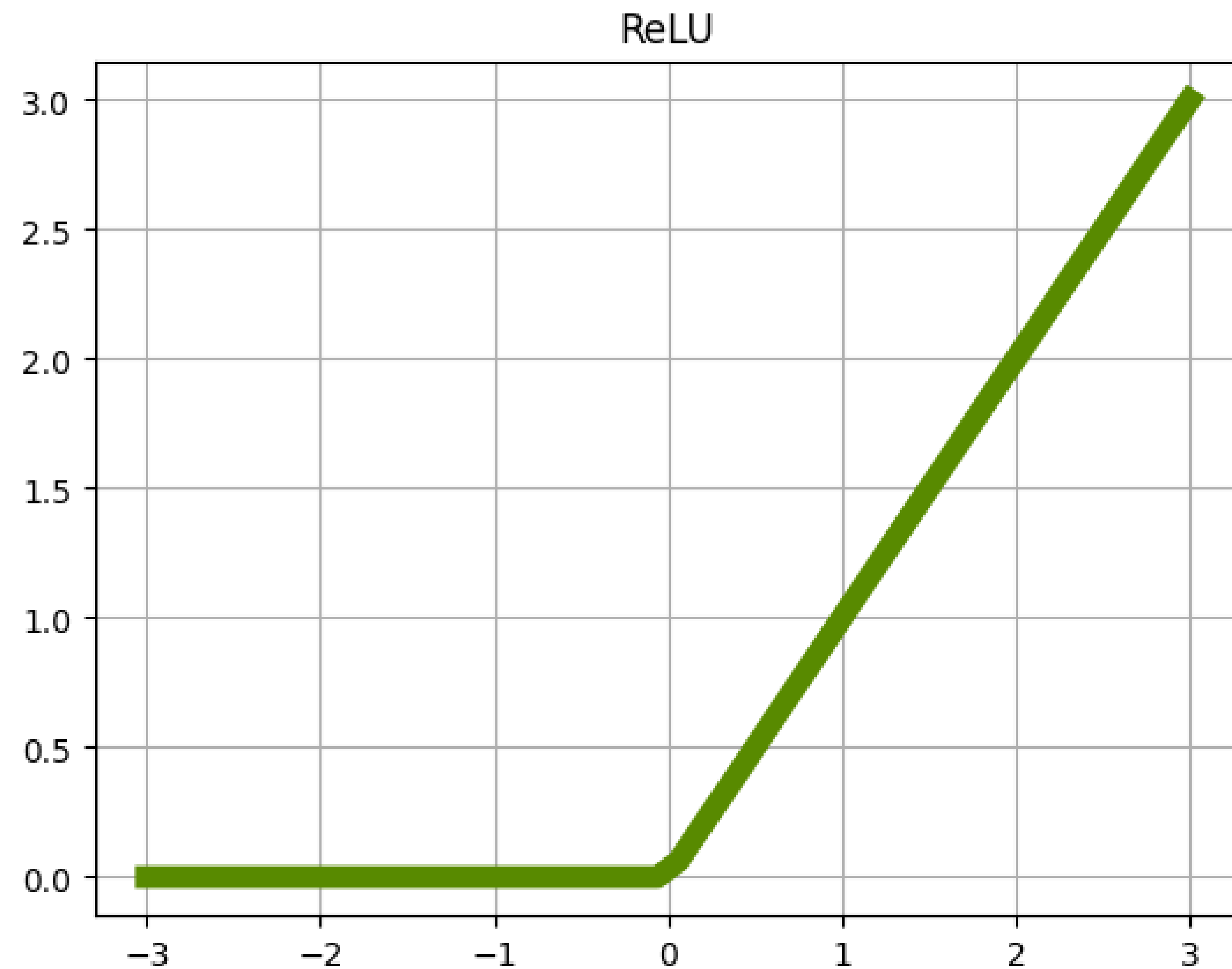
Group Normalization

Group Size = 3

Gaussian Error Linear Unit (GELU)

GELU

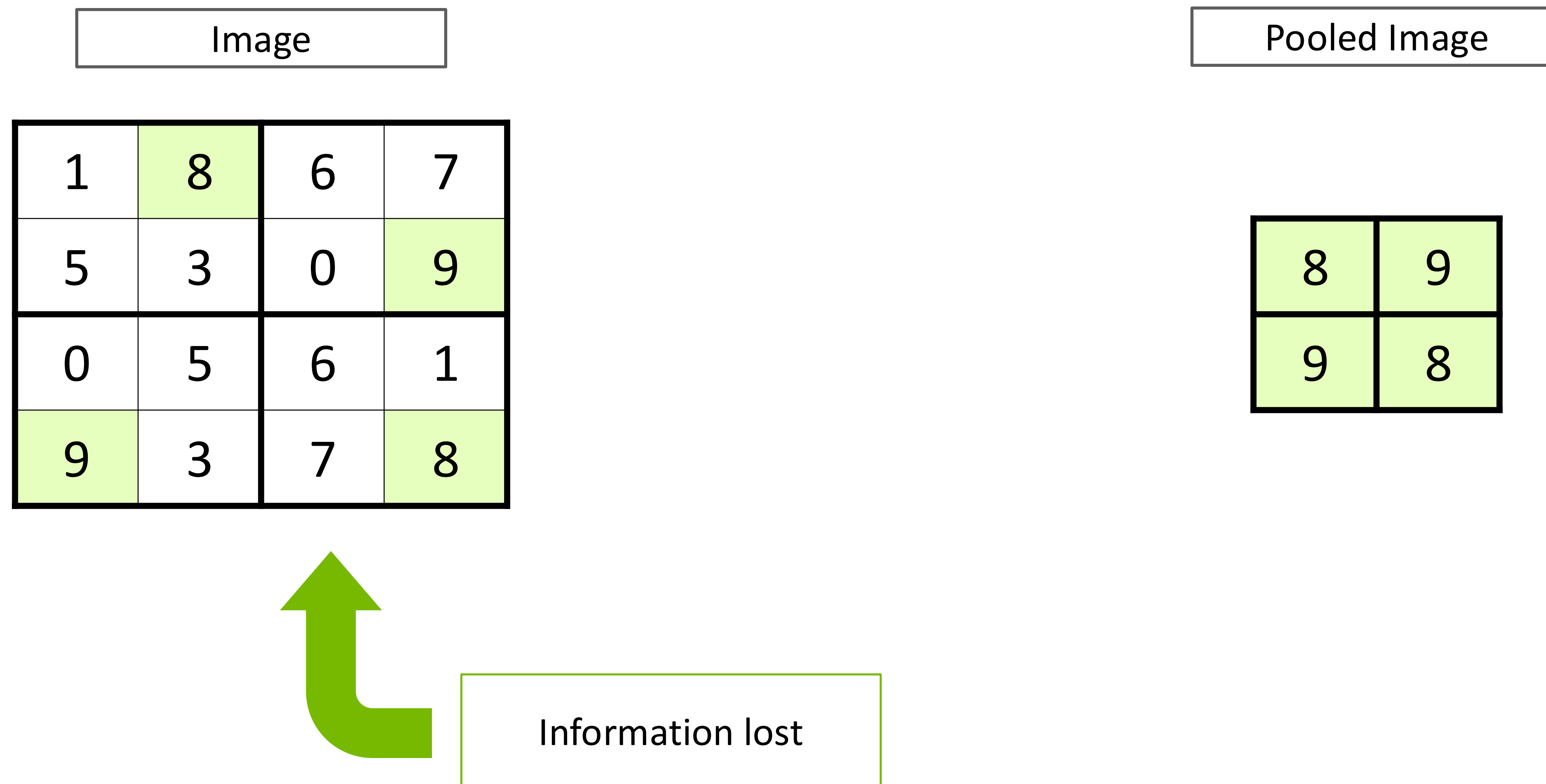
Gaussian Error Linear Unit



Rearrange Pooling

Pooling

Max Pooling Review



Einops

Better Dimension Manipulation

```
Rearrange(  
    "c (h p1) (w p2) -> (c p1 p2) h w",  
    p1=2, p2=2  
)
```

Cut image into strips
and stack

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16

Einops: Einstein Operations

Flexible and powerful tensor operations for readable and reliable code.

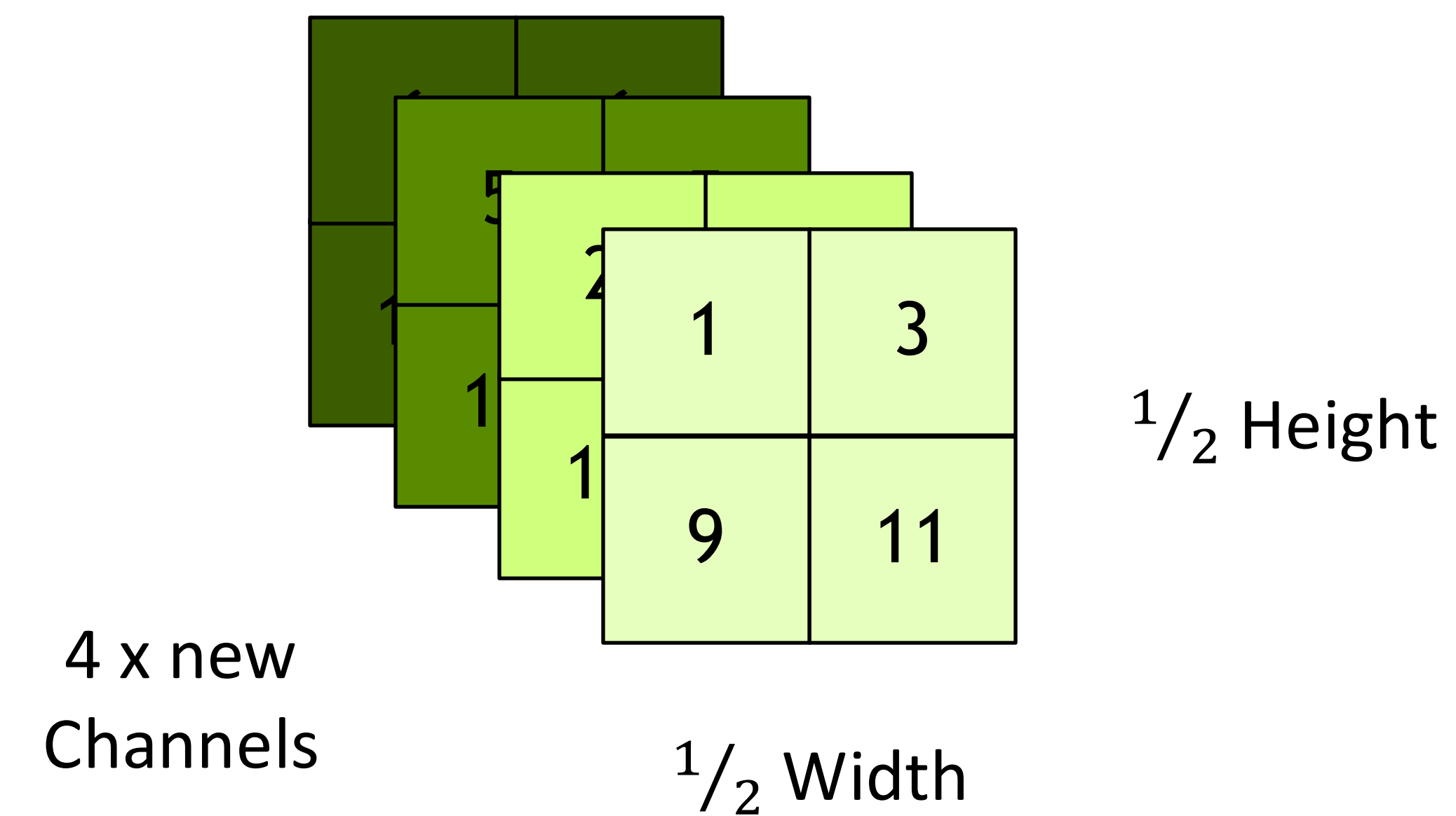
Supports numpy, pytorch, tensorflow, jax, and others.

Einops

Better Dimension Manipulation

```
Rearrange(  
    "c (h p1) (w p2) -> (c p1 p2) h w",  
    p1=2, p2=2  
)
```

Cut image into strips
and stack



Einops

Order Matters

```
Rearrange(  
  "c (h p1) w -> (c p1) h w", p1=2  
)
```


```
Rearrange(  
  "c (p1 h) w -> (c p1) h w", p1=2  
)
```


Sinusoidal Position Embeddings

Time as a Sequence

How to Represent Time as Discrete Steps?

t

As a one-hot encoding?

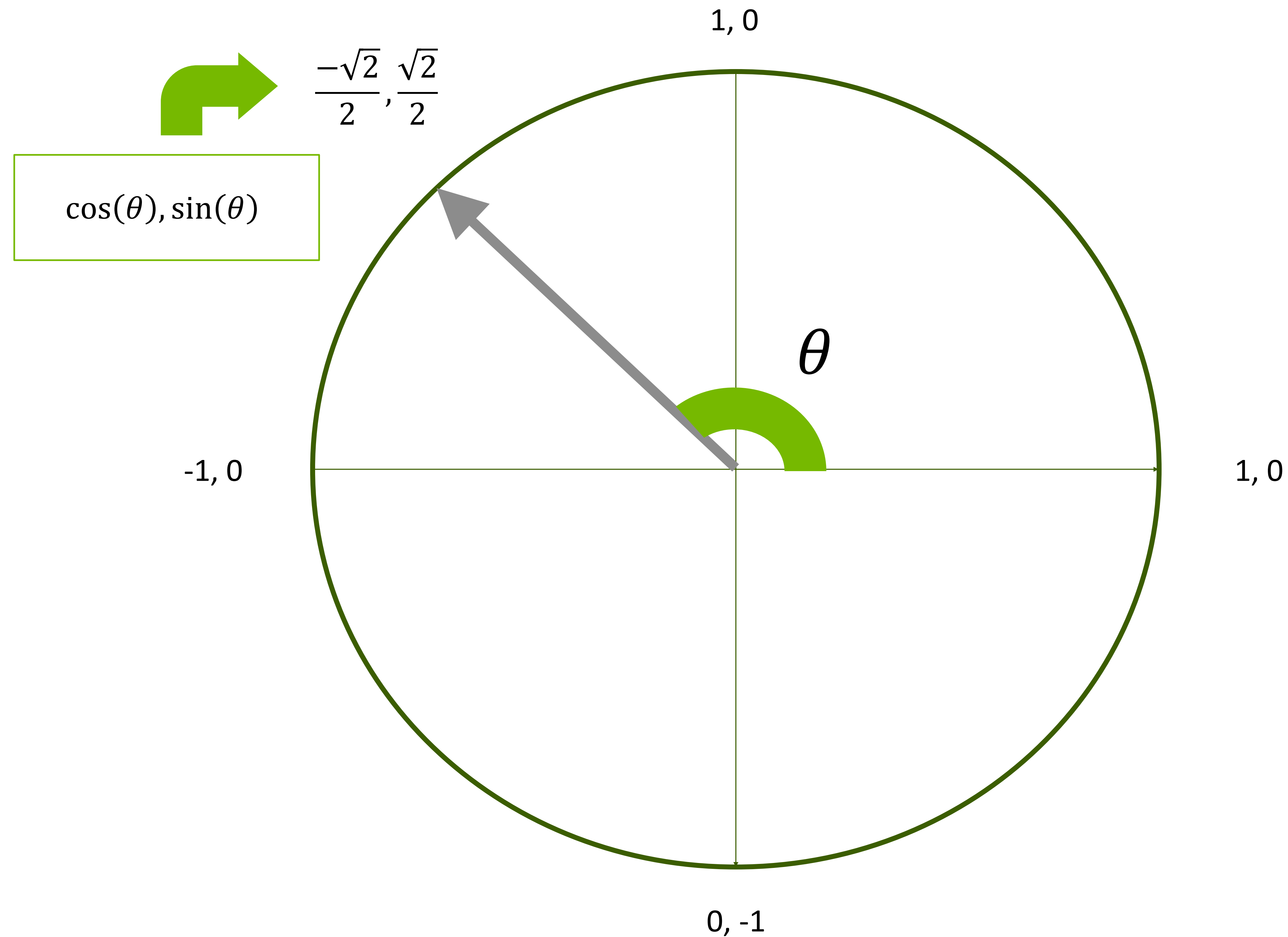
t = 7

As binary?

0	1	2	3	4	5	6	7	8	9	...
0	0	0	0	0	0	0	1	0	0	...
0	0	0	0	0	0	0	0	1	1	...
0	0	0	0	1	1	1	1	0	0	...
0	0	1	1	0	0	1	1	0	0	...
0	1	0	1	0	1	0	1	0	1	...

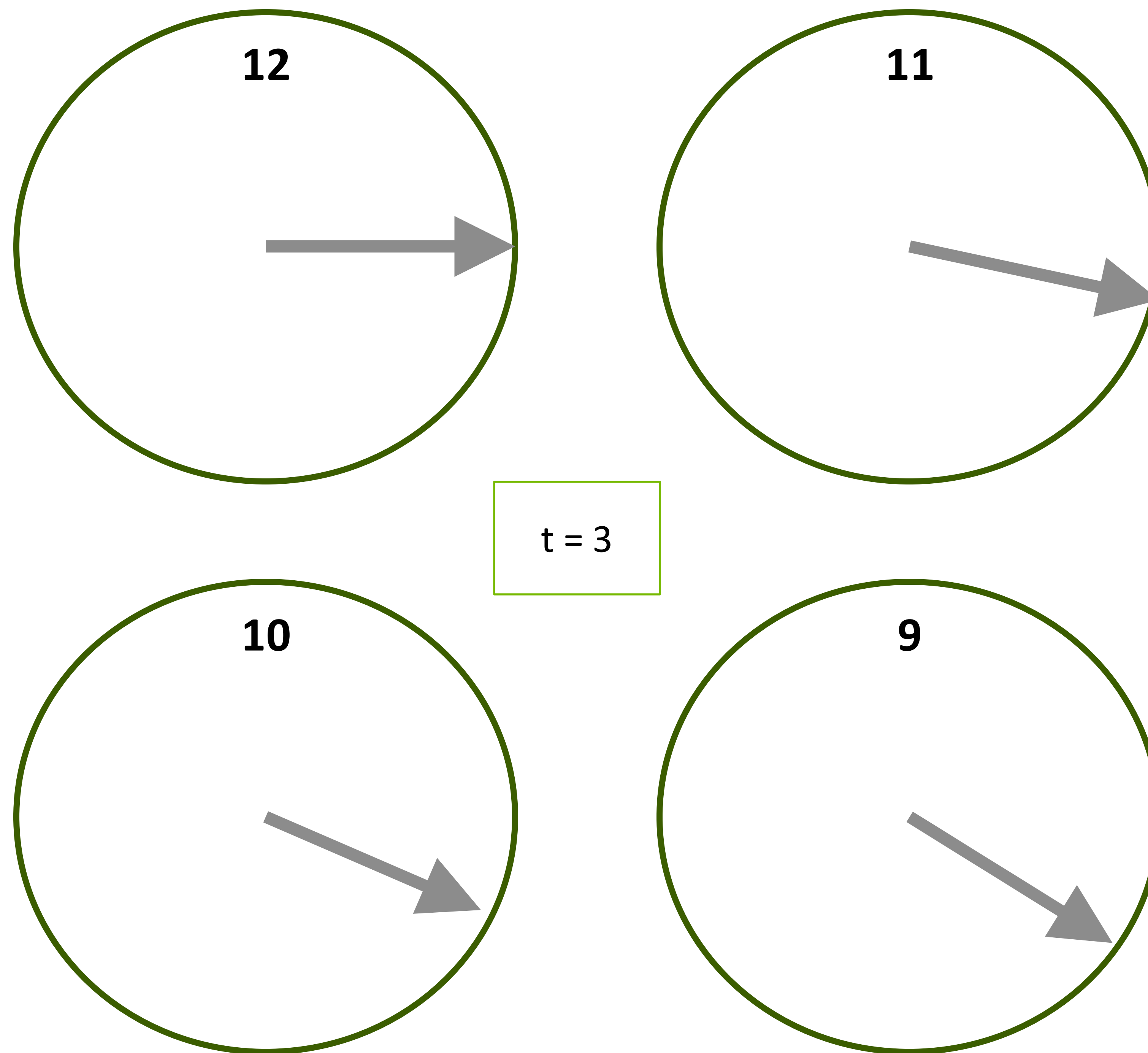
Time as a Sequence

As a Unit Circle?



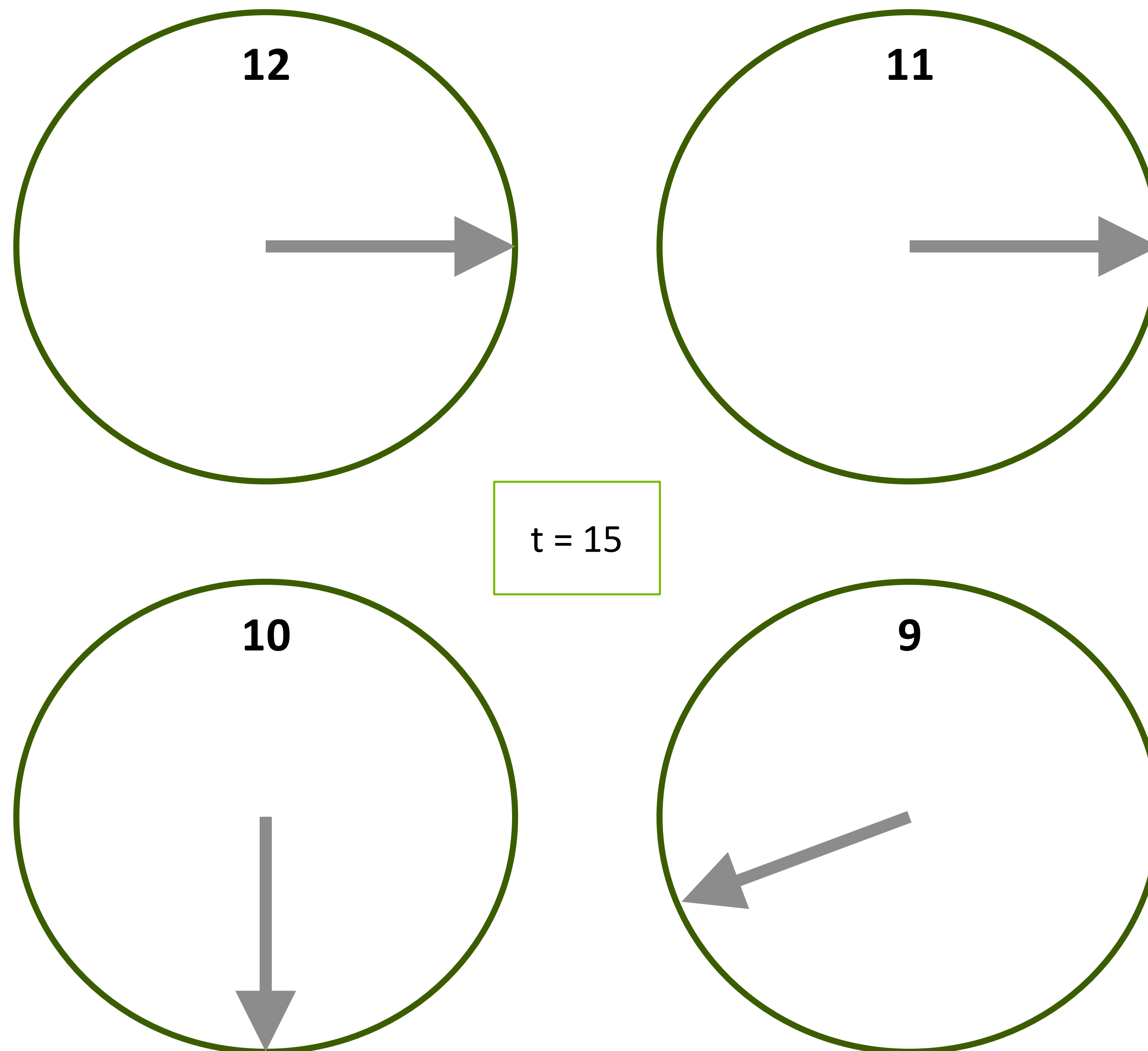
Time as a Sequence

As Unit Circles?



Time as a Sequence

As Unit Circles?

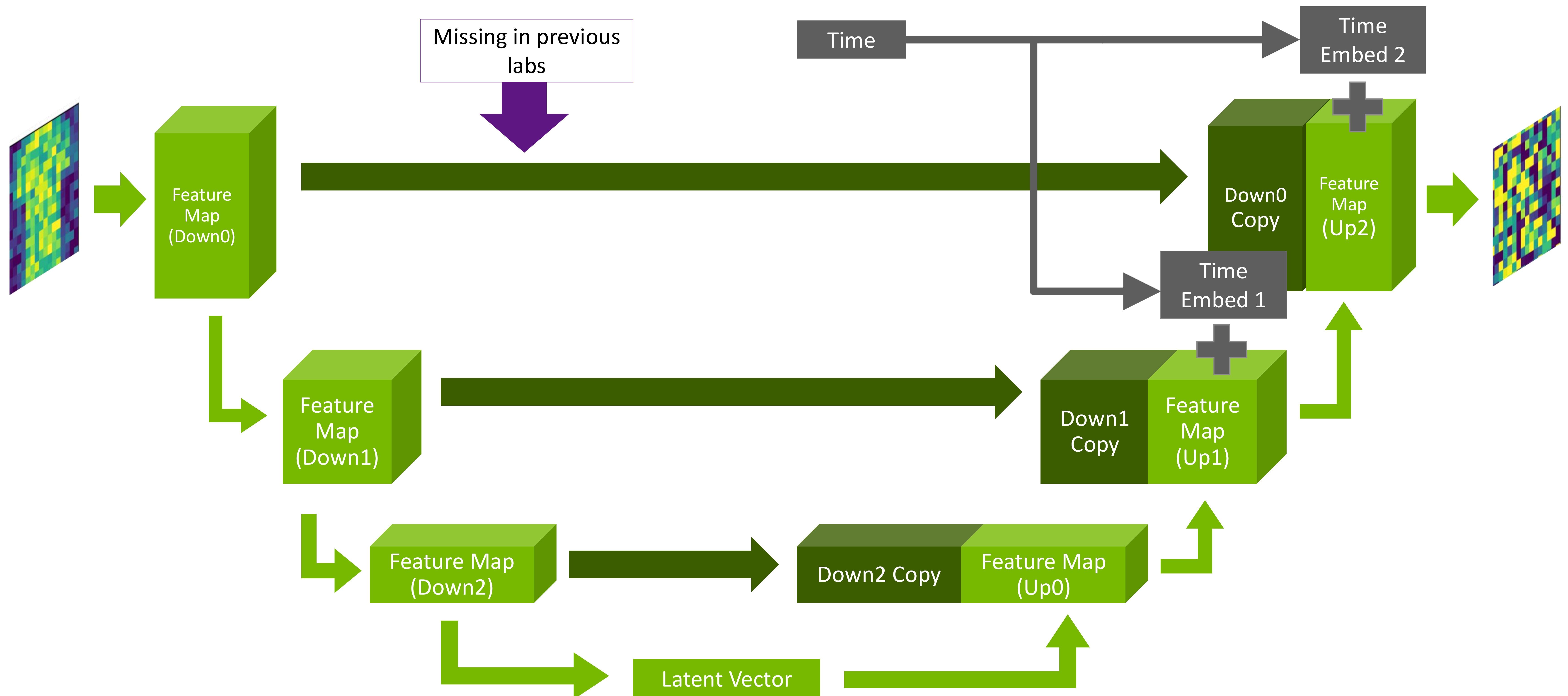


A bunch of abstract clocks with different numbers on them

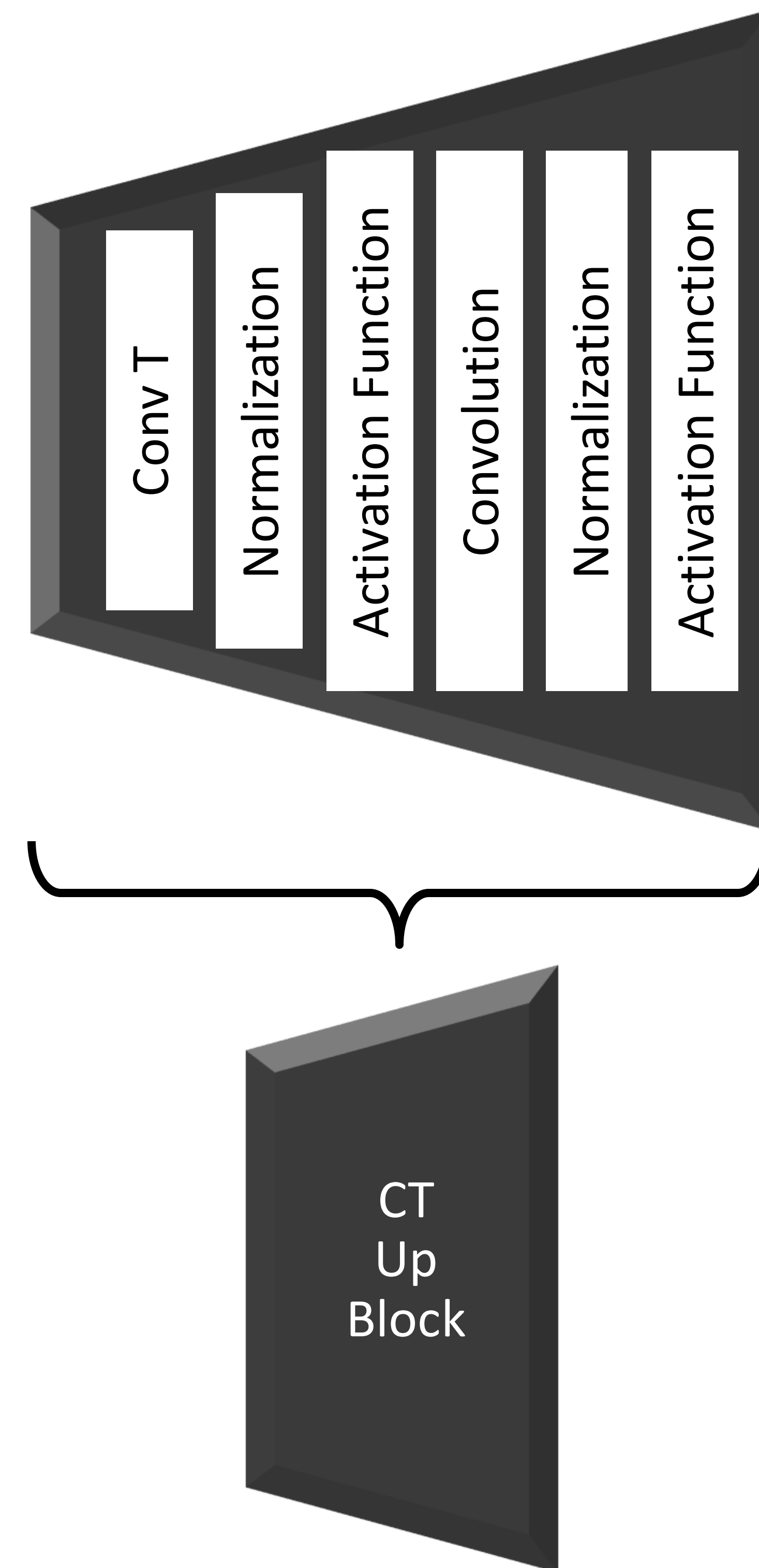
Deeper Networks

The U-Net Architecture

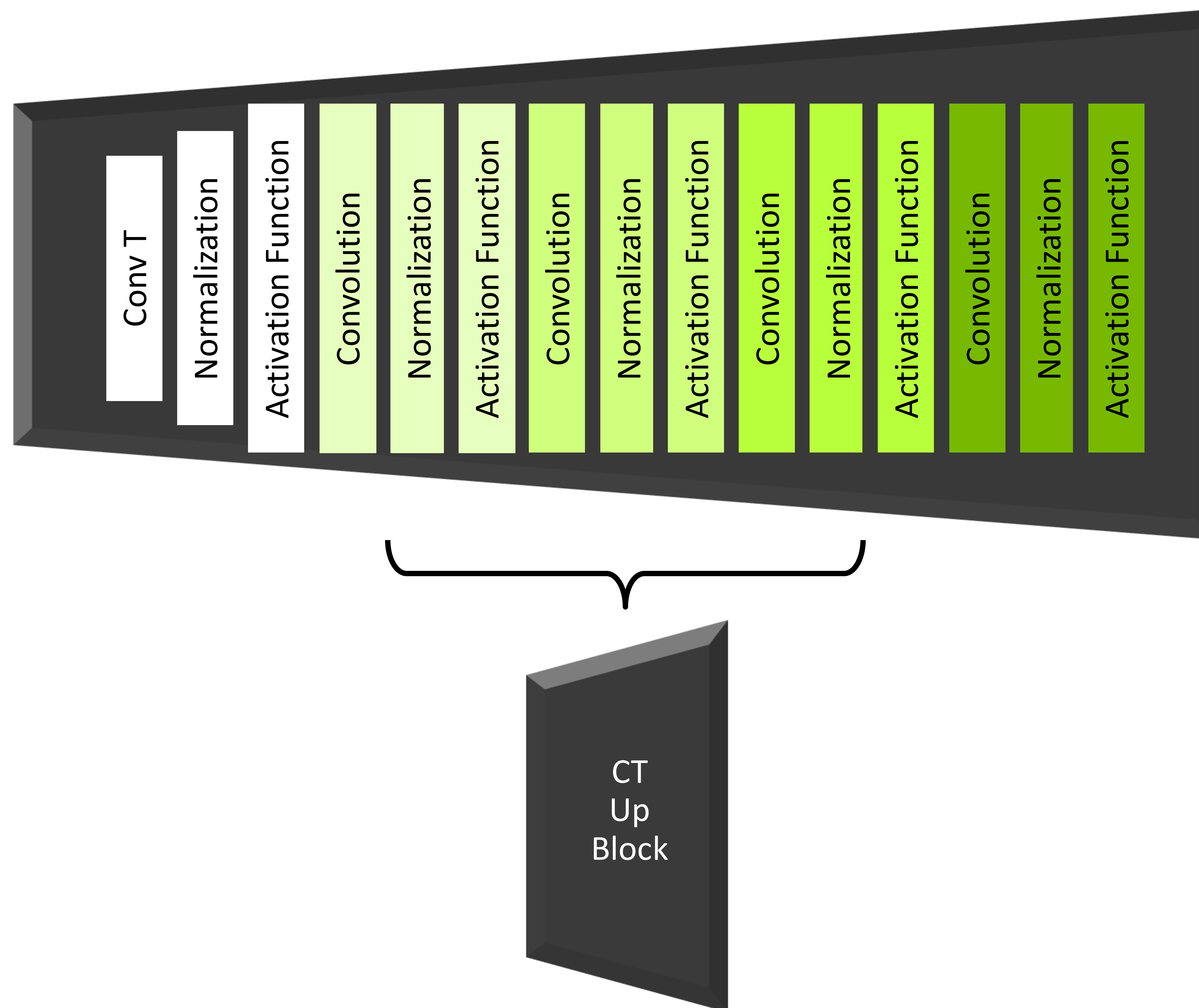
Adding Time



Adding Depth



Adding Depth



Let's get started!

3: Optimizations

Generative AI with Diffusion Models

 Products Solutions Industries For You

Shop Drivers Support    Min-Yuh Day

Deep Learning Institute Find Training Self Paced Courses Instructor-Led Workshops Educator Programs Enterprise Solutions Certification Resources

Course Progress Bookmarks Updates

Generative AI with Diffusion Models Start Here 3: Optimizations

Generative AI with Diffusion Models

Start Here

0: Server Access

1: From U-Nets to Diffusion

2: Denoising Diffusion Probabilistic Models

3: Optimizations

4: Classifier-Free Diffusion Guidance

5: CLIP

6: Wrap-up and Assessment

Next Steps

Feedback

Previous

Next

8:25 / 8:28

Speed 1.0x





 DEEP LEARNING INSTITUTE

**Generative AI with Diffusion Models**
Part 3: Optimizations

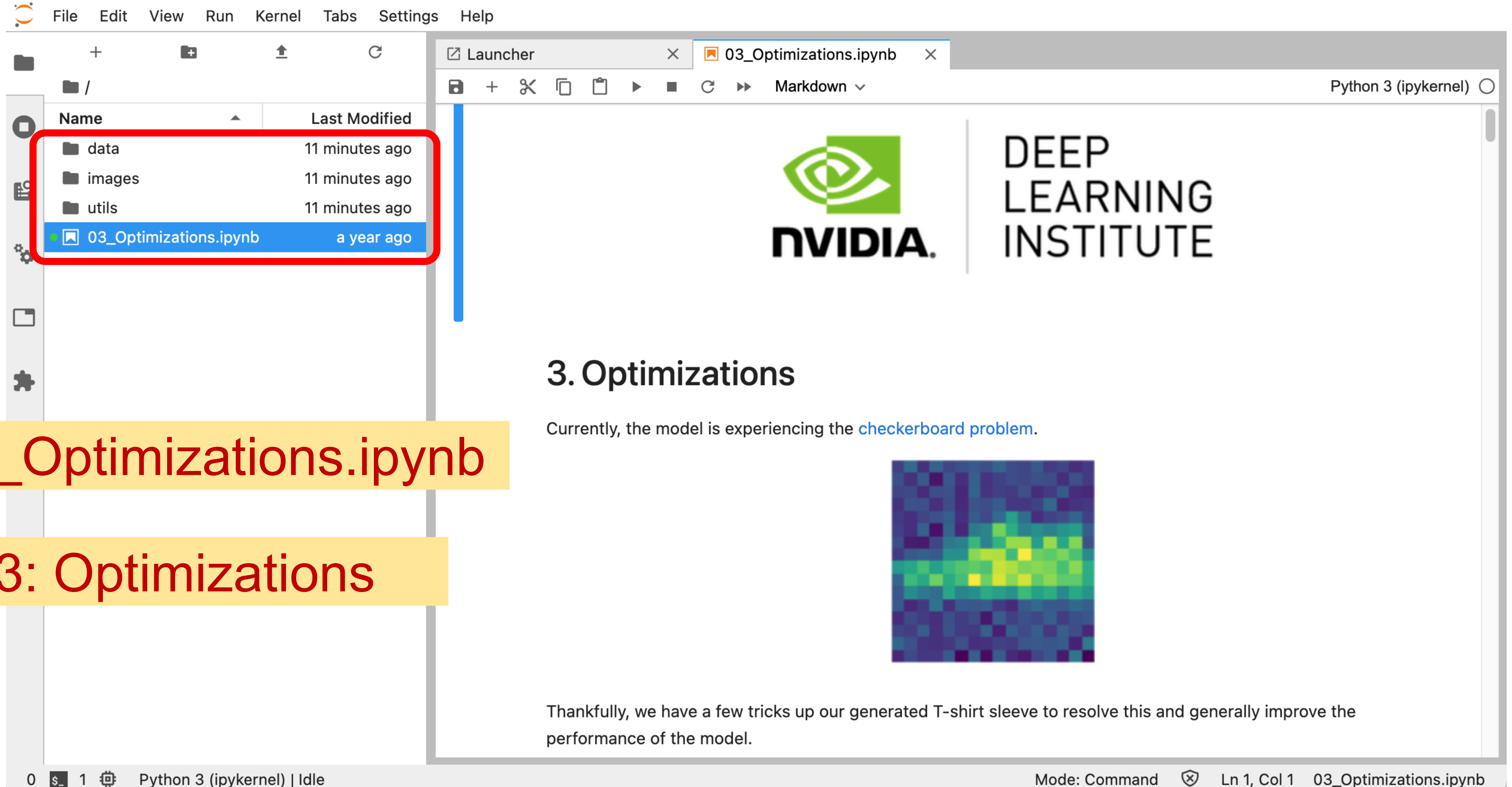
LOADING 

 STOP TASK

START

3: Optimizations

Generative AI with Diffusion Models



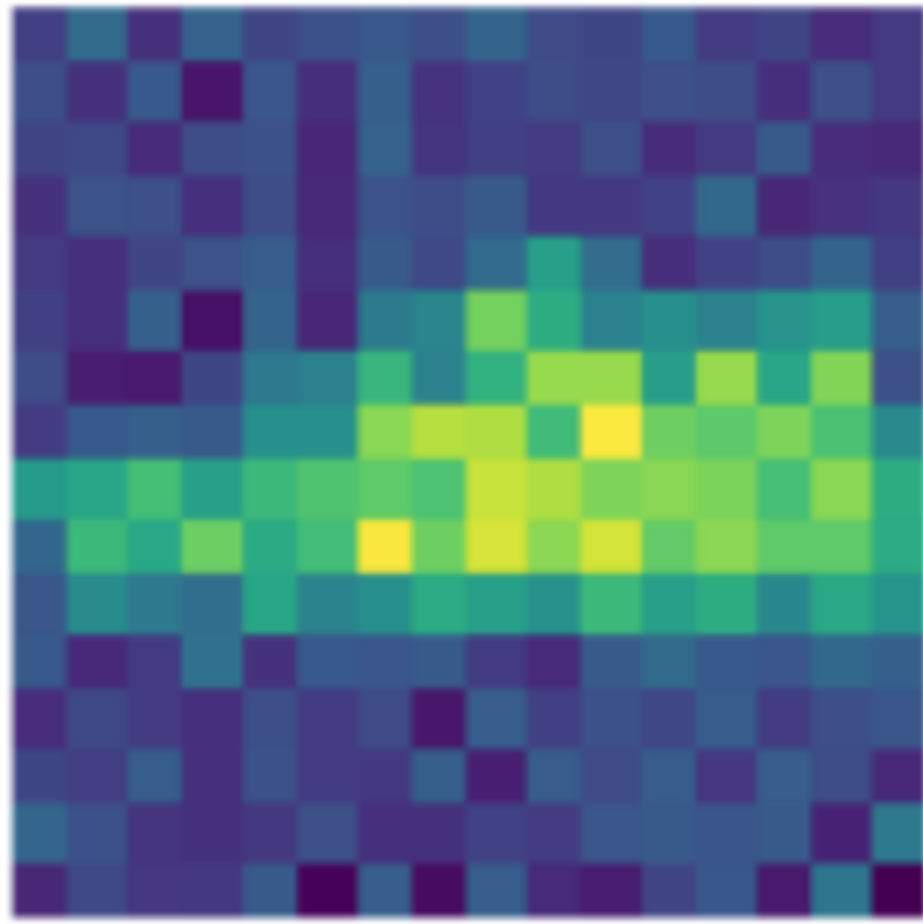
File Edit View Run Kernel Tabs Settings Help

Launcher 03_Optimizations.ipynb Python 3 (ipykernel)

data 11 minutes ago
images 11 minutes ago
utils 11 minutes ago
03_Optimizations.ipynb a year ago

3. Optimizations

Currently, the model is experiencing the [checkerboard problem](#).



Thankfully, we have a few tricks up our generated T-shirt sleeve to resolve this and generally improve the performance of the model.

0 \$ 1 Python 3 (ipykernel) | Idle Mode: Command Ln 1, Col 1 03_Optimizations.ipynb

03_Optimizations.ipynb

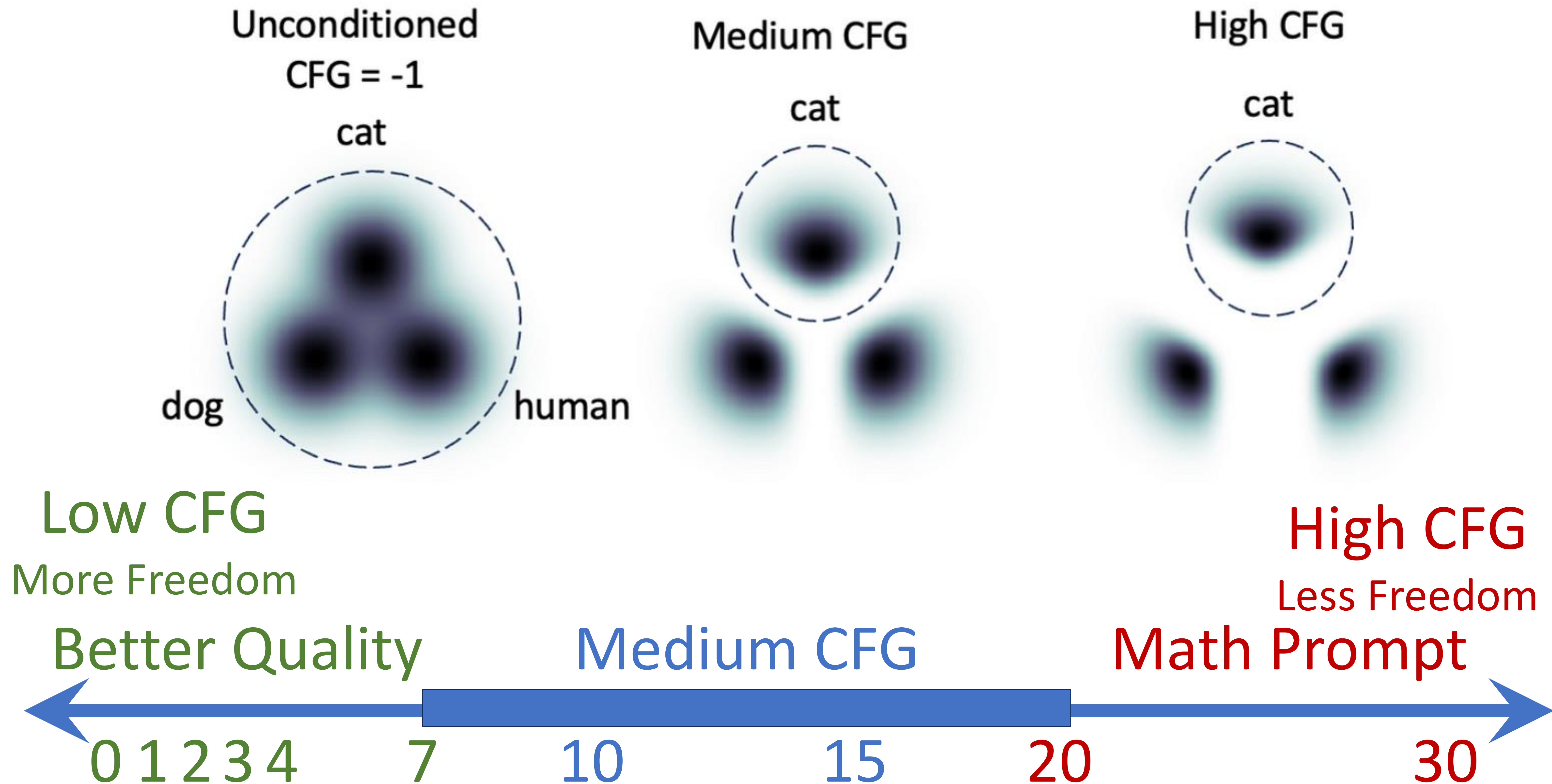
3: Optimizations

Generative AI with Diffusion Models

- Part 1: From U-Nets to Diffusion
- Part 2: Denoising Diffusion Probabilistic Models
- **Part 3: Optimizations**
- **Part 4: Classifier-Free Diffusion Guidance (CFG)**
- Part 5: CLIP (Contrastive Language-Image Pre-Training)
- Part 6: Wrap-up & Assessment

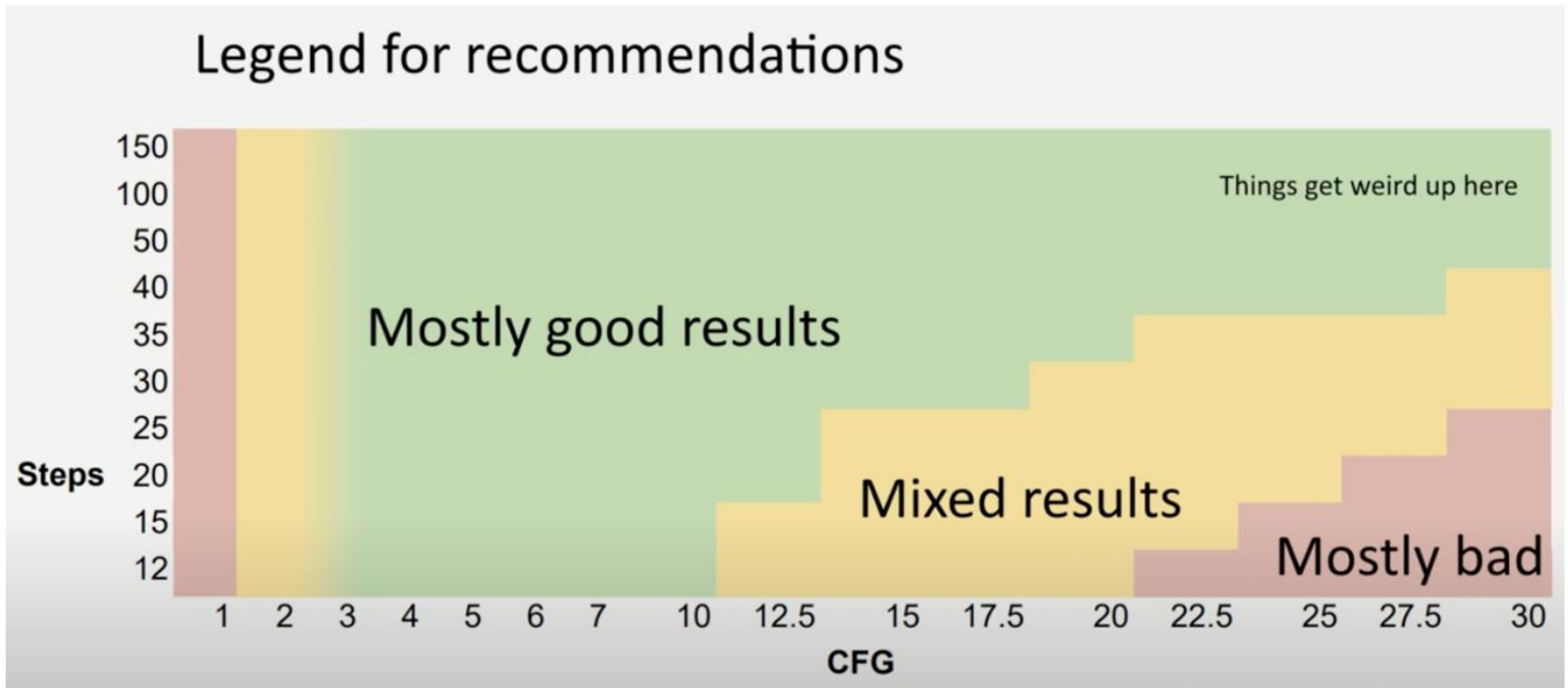
Classifier-Free Diffusion Guidance (CFG)

CFG Scale (Condition)



Classifier-Free Diffusion Guidance (CFG)

CFG and Steps



Source: SiliconThaumaturgy (2023), Stable Diffusion Deep Dive - CFG - Don't Accidentally Fry Your Images,
<https://www.youtube.com/watch?v=kuhO9zAzetk>



Generative AI with Diffusion Models

Part 4: Classifier-Free Diffusion Guidance

Agenda

- Part 1: From U-Nets to Diffusion

- Part 2: Denoising Diffusion Probabilistic Models

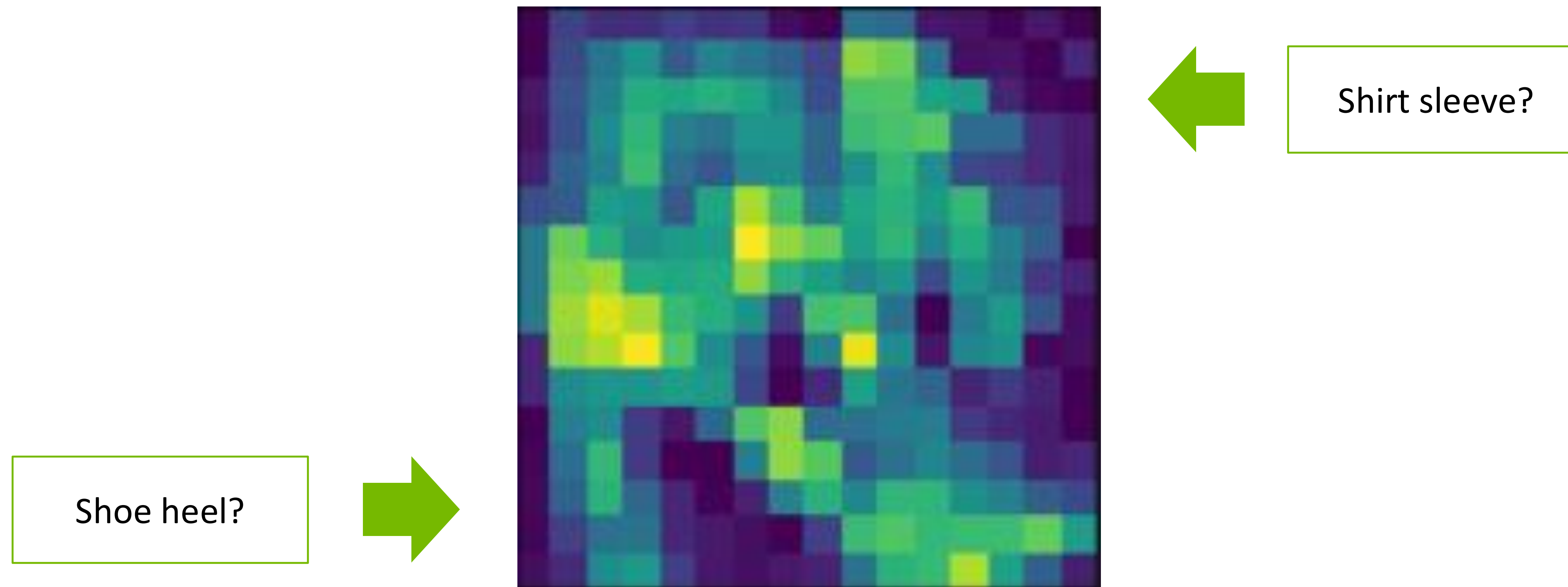
- Part 3: Optimizations

- Part 4: Classifier-Free Diffusion Guidance
- Part 5: CLIP

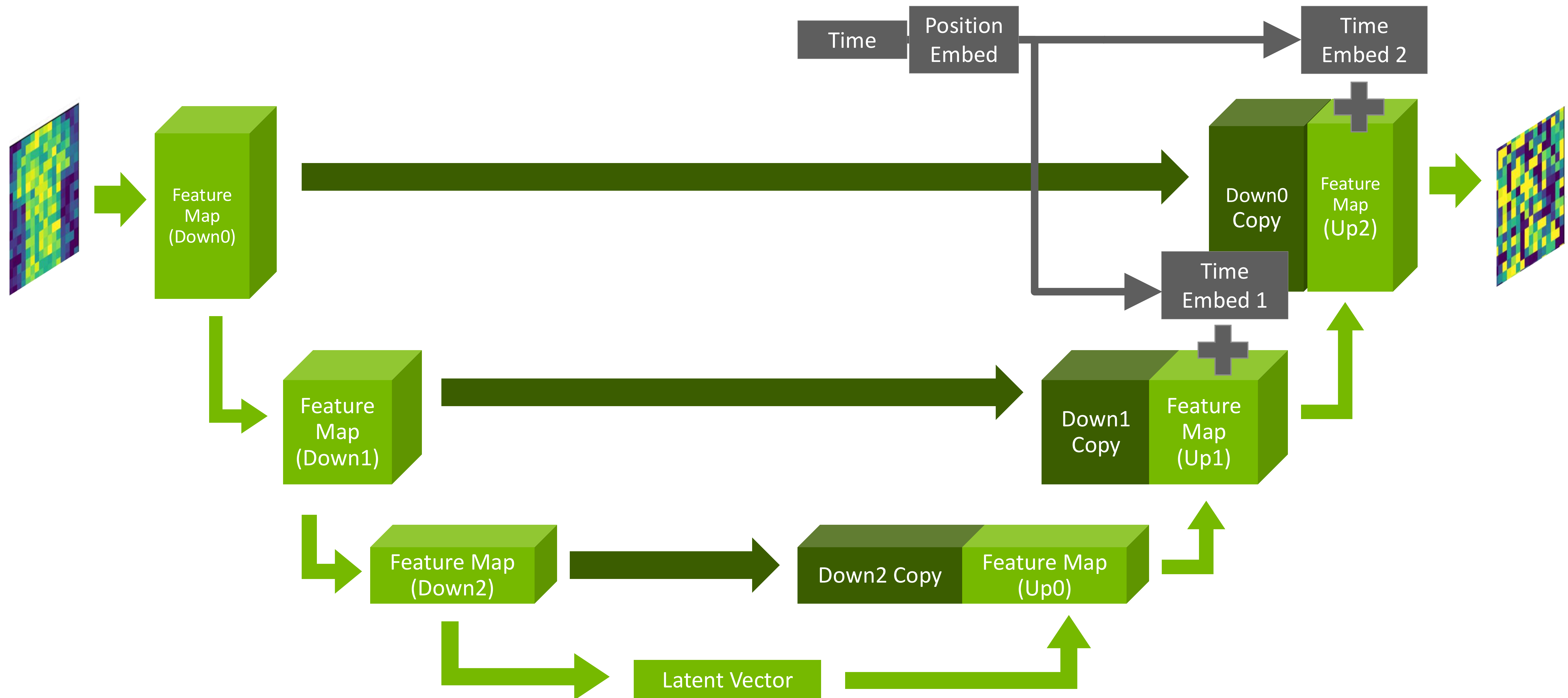
- Part 6: Wrap-up & Assessment

Adding Context

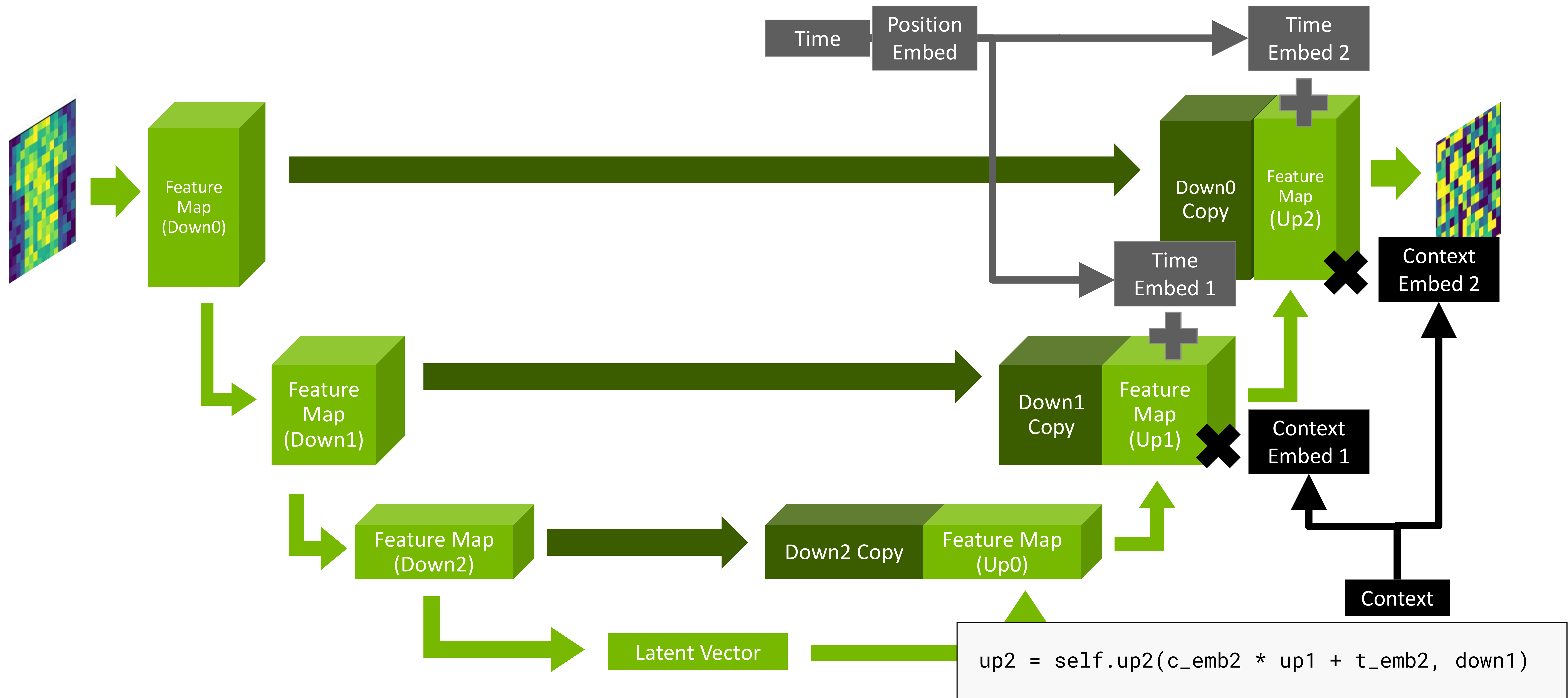
Adding Context



Adding Context

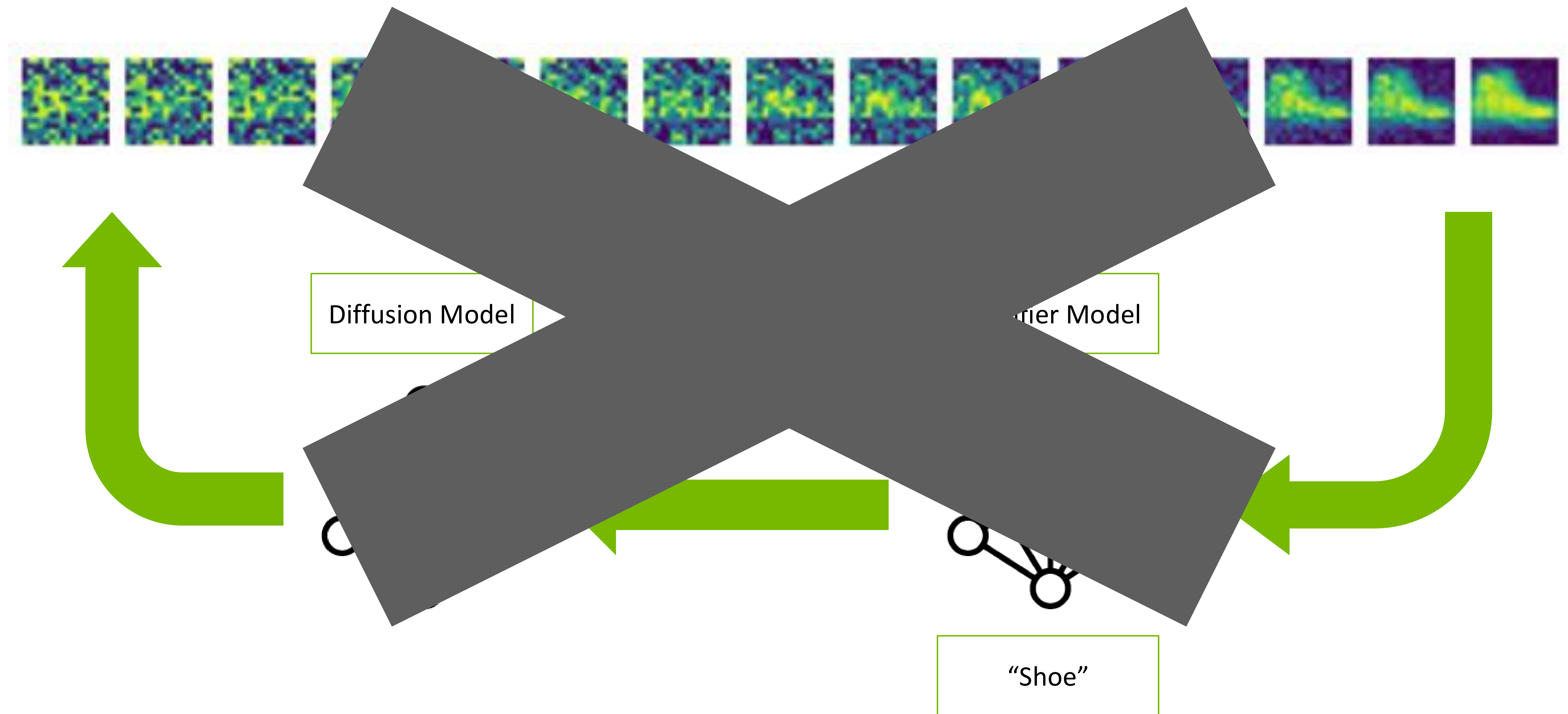


Adding Context



Classifier-Free Diffusion Guidance

Make a Classifier Model?



Classifier-Free Diffusion Guidance

CLASSIFIER-FREE DIFFUSION GUIDANCE

Jonathan Ho & Tim Salimans

Google Research, Brain team

`{jonathanho,salimans}@google.com`

ABSTRACT

Classifier guidance is a recently introduced method to trade off mode coverage and sample fidelity in conditional diffusion models post training, in the same spirit as low temperature sampling or truncation in other types of generative models. Classifier guidance combines the score estimate of a diffusion model with the gradient of an image classifier and thereby requires training an image classifier separate from the diffusion model. It also raises the question of whether guidance

Classifier Free Diffusion Guidance

Bernoulli Masks

“sneaker”

one-hot encoding

0	0	0	0	0	0	0	1	0	0
---	---	---	---	---	---	---	---	---	---



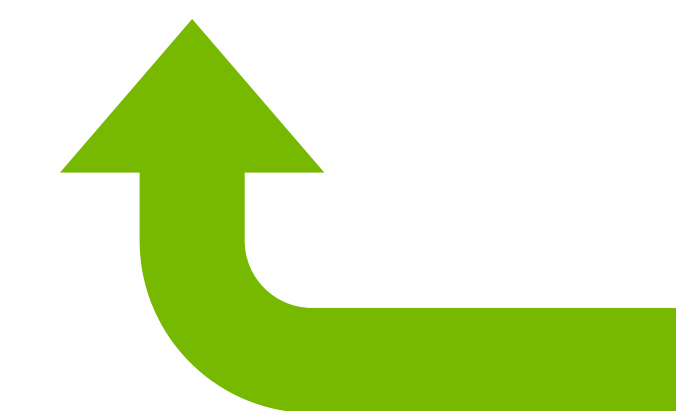
Bernoulli mask

$p = .9$

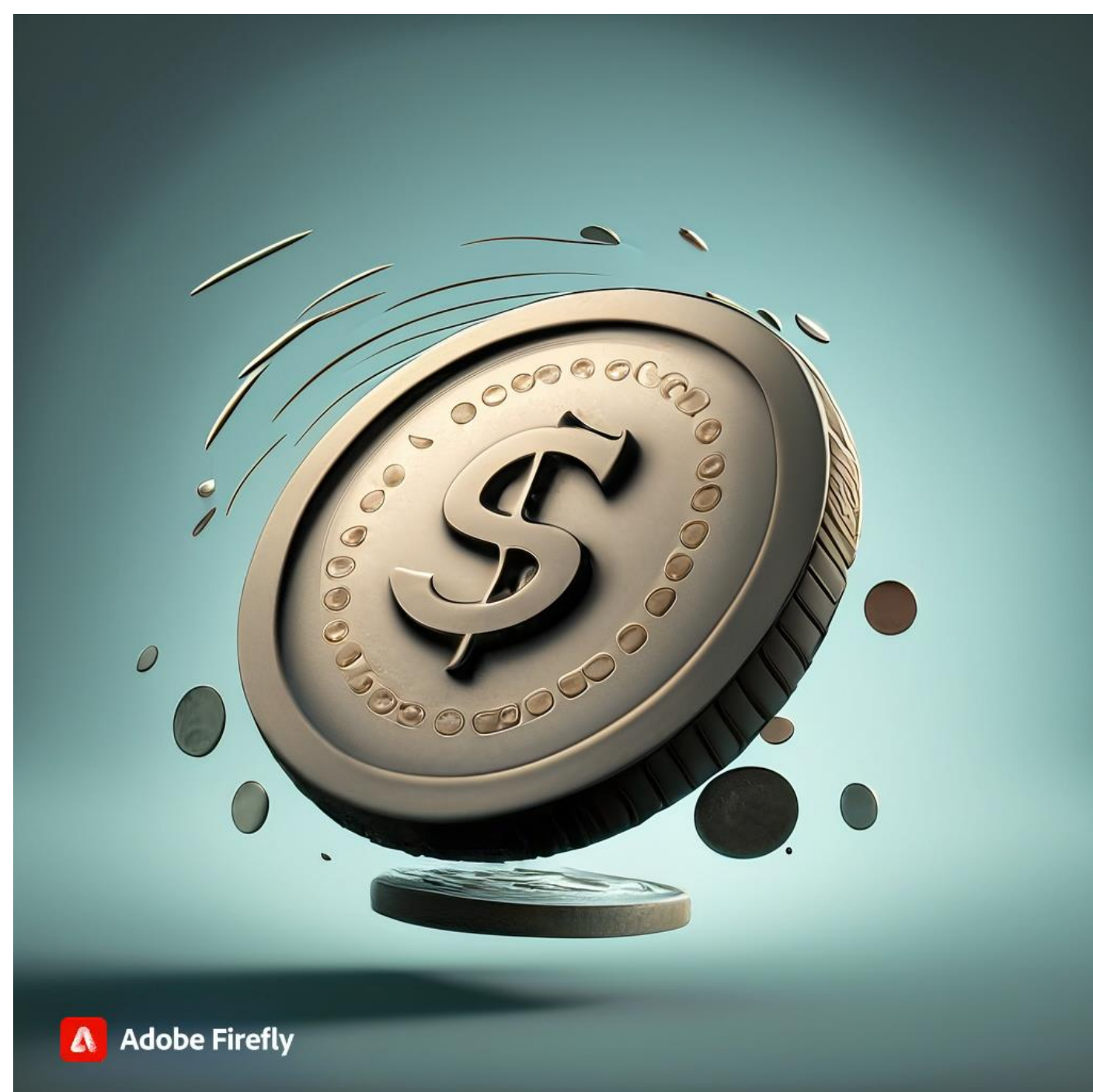
1	1	0	1	0	1	1	1	1	1
---	---	---	---	---	---	---	---	---	---



0	0	0	0	0	0	0	1	0	0
---	---	---	---	---	---	---	---	---	---



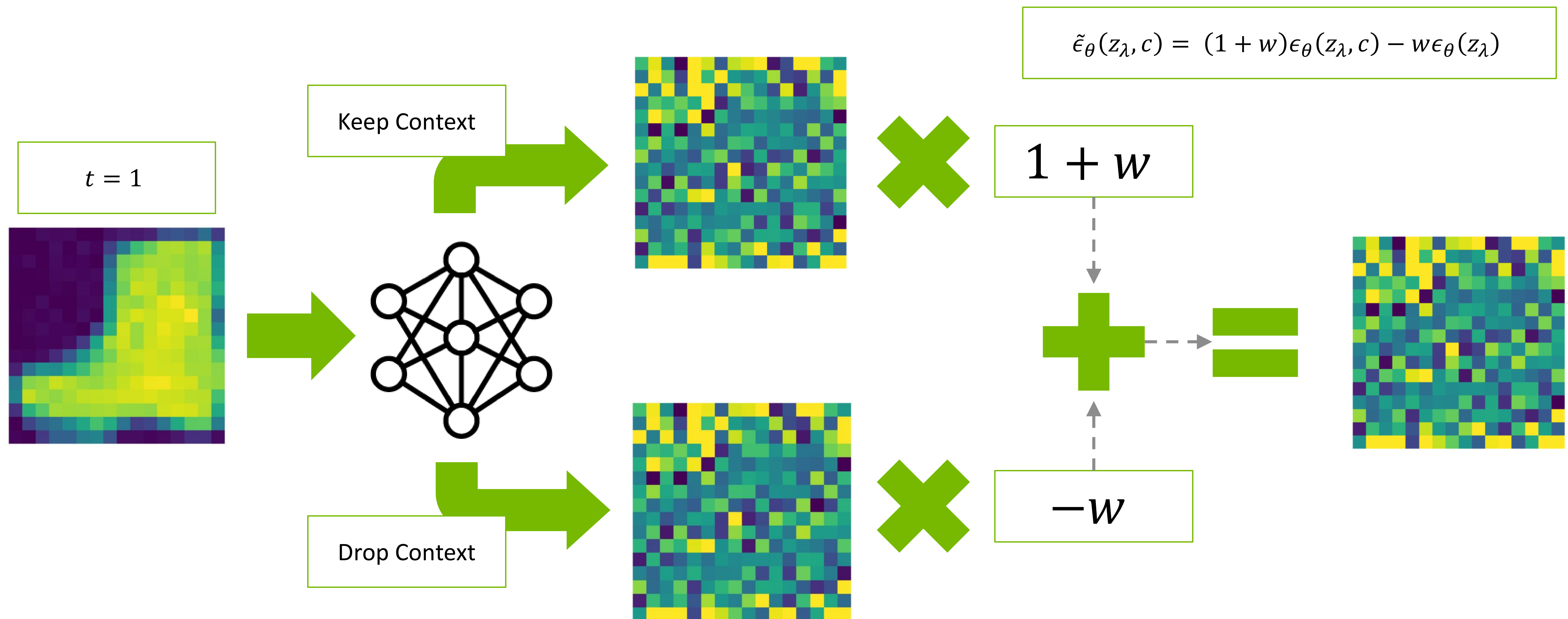
$1 - p$ chance class will be dropped



A weighted coin flipping through the air like a cartoon

Classifier Free Diffusion Guidance

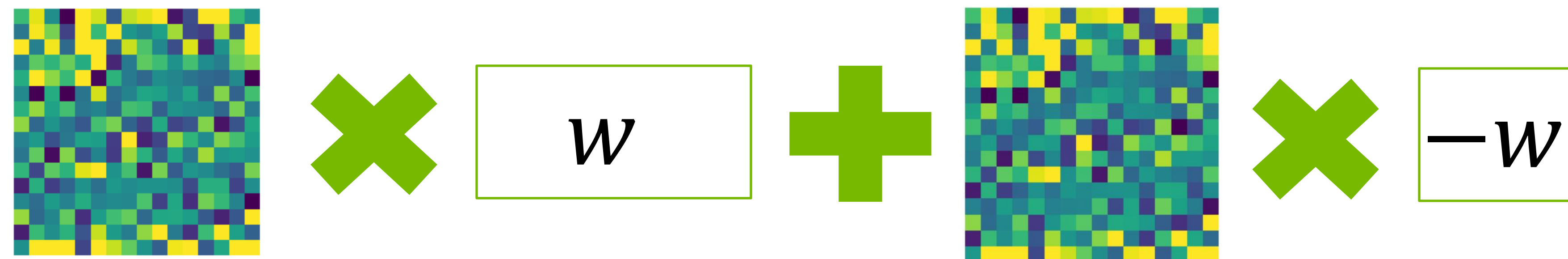
Weighted Reverse Diffusion



Classifier Free Diffusion Guidance

Weighted Reverse Diffusion

$$(\text{Noise Map})w)\epsilon_{\theta}(z_{\lambda}, c) - w\epsilon_{\theta}(z_{\lambda})$$

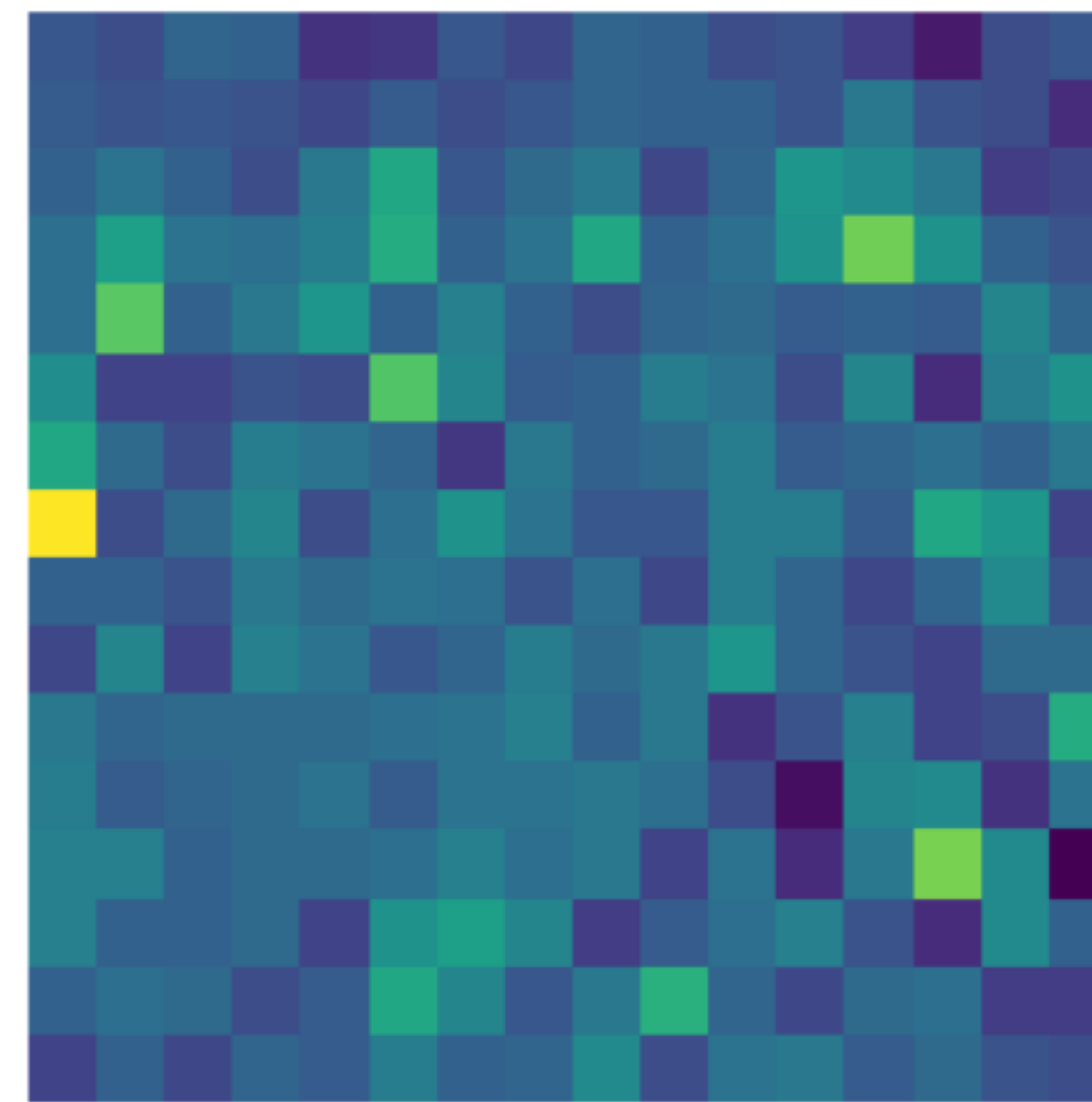


$w = 0$



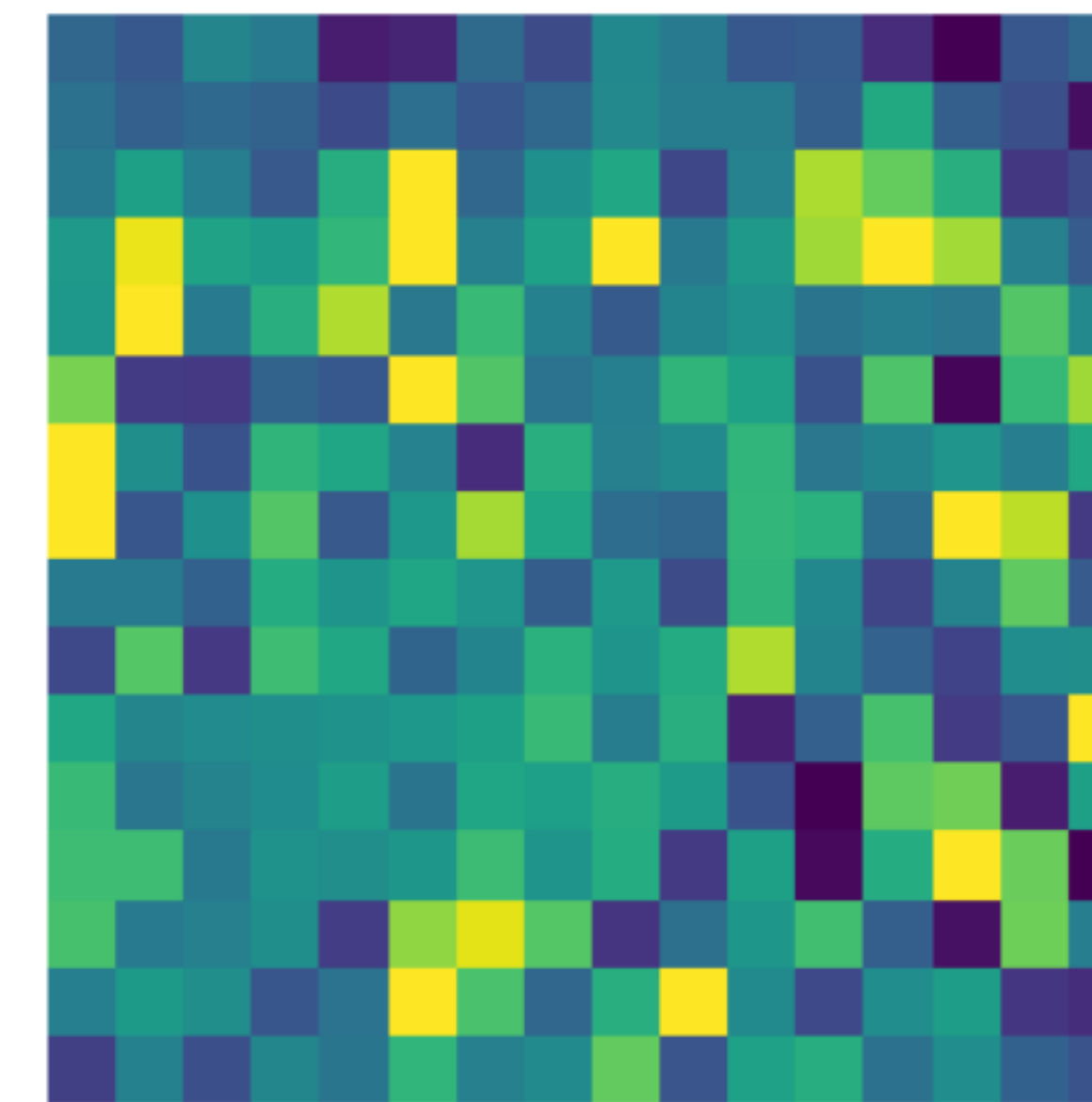
No context feature
exaggeration

$w = 1$



Some context feature
exaggeration

$w = 10$



Huge context feature
exaggeration

Classifier Free Diffusion Guidance

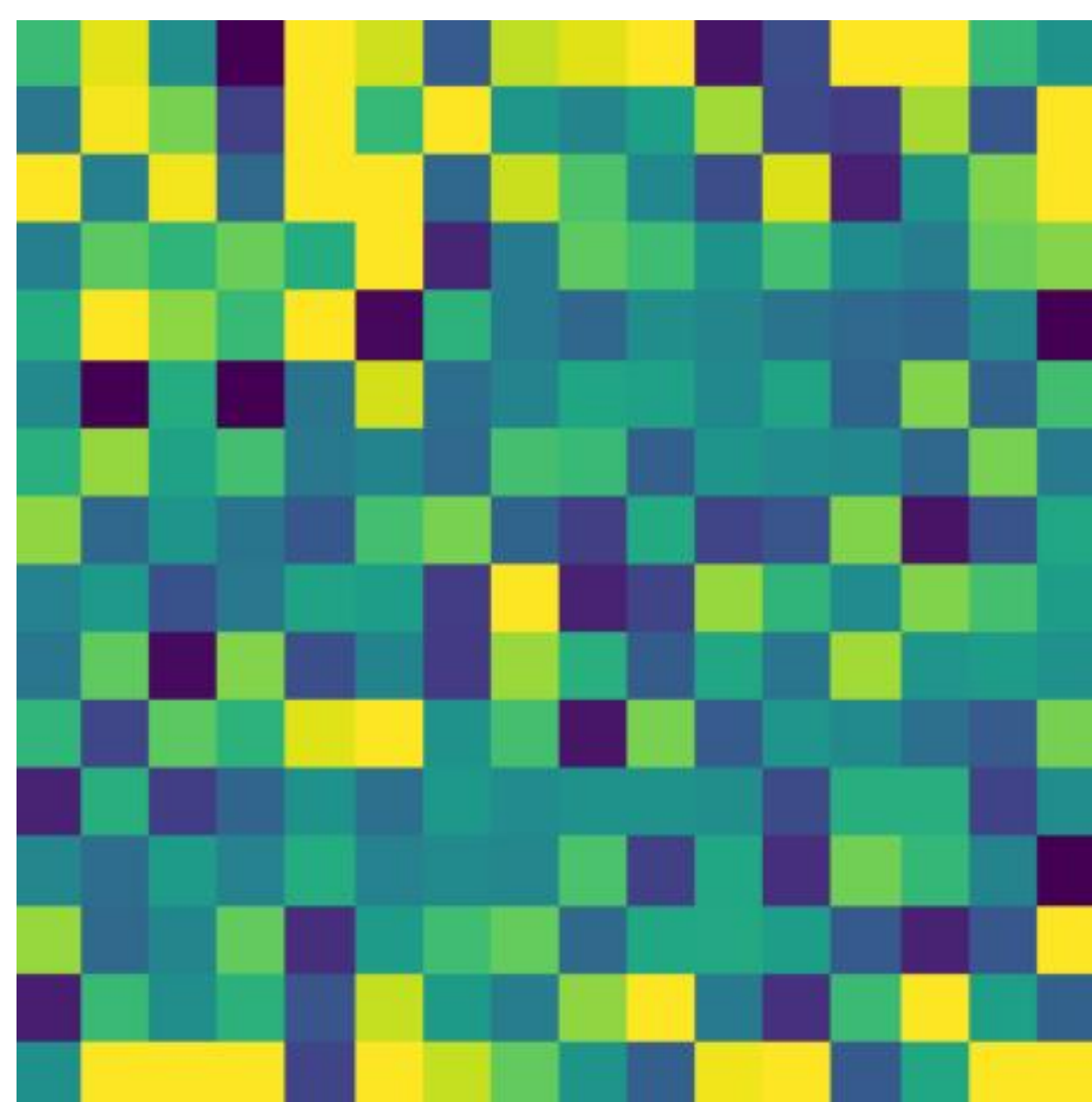
Weighted Reverse Diffusion

$$(1 + w)\epsilon_{\theta}(z_{\lambda}, c) - w\epsilon_{\theta}(z_{\lambda})$$



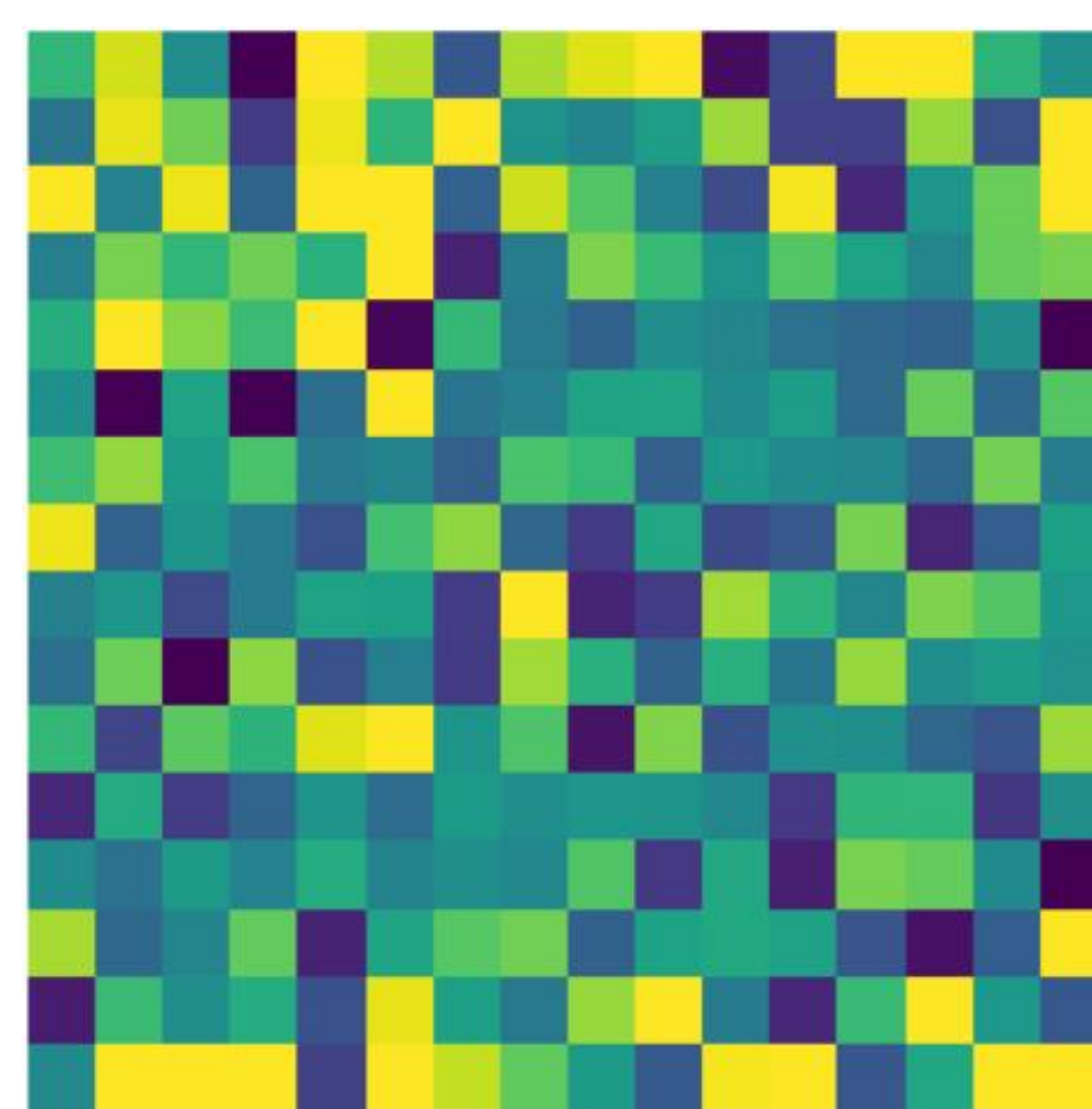
The diagram illustrates the formula for weighted reverse diffusion. It shows a noise map (a 16x16 grid of colored squares) multiplied by $1 + w$, followed by a plus sign, and then another noise map multiplied by $-w$. The multiplication is represented by a large green 'X' symbol, and the addition is represented by a large green '+' symbol.

$w = 0$



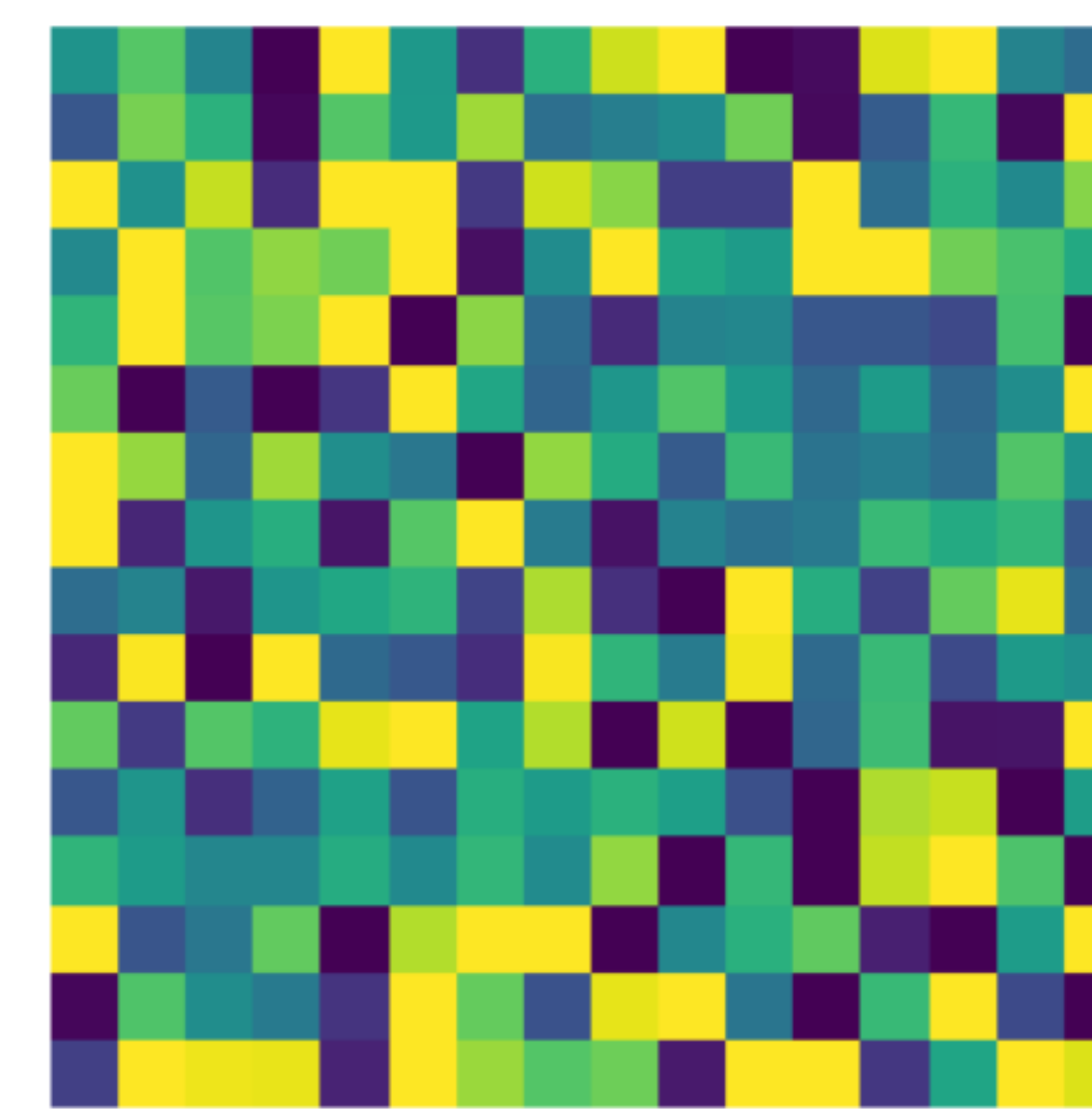
No context feature
exaggeration

$w = 1$



Some context feature
exaggeration

$w = 10$



Huge context feature
exaggeration

Classifier Free Diffusion Guidance

Weighted Reverse Diffusion



Modified TF Flowers

Modified TF Flowers

Modified for Image Generation



Photo by _e.t

Let's get started!

4: Classifier-Free Diffusion Guidance (CFG)

Generative AI with Diffusion Models


[Products](#)
[Solutions](#)
[Industries](#)
[For You](#)

[Shop](#)
[Drivers](#)
[Support](#)
[Q](#)
[🛒](#)
[Min-Yuh Day](#)

Deep Learning Institute
[Find Training](#)
[Self Paced Courses](#)
[Instructor-Led Workshops](#)
[Educator Programs](#)
[Enterprise Solutions](#)
[Certification](#)
[Resources](#)

[Course](#)
[Progress](#)
[Bookmarks](#)
[Updates](#)

[Generative AI with Diffusion Models](#)
[Start Here](#)
[4: Classifier-Free Diffusion Guidance](#)

Generative AI with Diffusion Models

Start Here

0: Server Access

1: From U-Nets to Diffusion

2: Denoising Diffusion Probabilistic Models

3: Optimizations

4: Classifier-Free Diffusion Guidance

5: CLIP

6: Wrap-up and Assessment

Next Steps

Feedback

Previous

Next

6:30 / 6:32

Video position. Press space to toggle playback


DEEP LEARNING INSTITUTE


Generative AI with Diffusion Models
Part 4: Classifier-Free Diffusion Guidance

LOADING

STOP TASK

4: Classifier-Free Diffusion Guidance (CFG)

START

File Edit View Run Kernel Tabs Settings Help

+

+

↑

↺

/

Name	Last Modified
04_images	3 minutes ago
data	3 minutes ago
images	3 minutes ago
utils	3 minutes ago
04_Classifier_Free_Diffusion.ipynb	a year ago

04_Classifier_Free_Diffusion.ipynb

4: Classifier-Free Diffusion Guidance (CFG)

Python 3 (ipykernel) | Idle

Launcher

04_Classifier_Free_Diffusion.ipynb

Python 3 (ipykernel)

Markdown



DEEP LEARNING INSTITUTE

4. Classifier-Free Diffusion Guidance

So far, we have been able to train a model to generate images of clothing using the [FashionMNIST](#) dataset. However, there is no way for the user to specify what kind of images should be generated. Let's fix that by using [Classifier-Free Diffusion Guidance](#), a relatively simple way to create a [Conditional Diffusion Model](#).

Learning Objectives

The goals of this notebook are to:

- Add categorical embeddings to a U-Net
- Train a model with a Bernoulli mask
- Add a weight the reverse diffusion process
- Practice learnings on a more challenging dataset

Mode: Command Ln 1, Col 1 04_Classifier_Free_Diffusion.ipynb

Generative AI with Diffusion Models

 Products Solutions Industries For You

Shop Drivers Support    Min-Yuh Day

Deep Learning Institute Find Training Self Paced Courses Instructor-Led Workshops Educator Programs Enterprise Solutions Certification Resources

Generative AI with Diffusion Models

Course Progress Bookmarks Updates

Generative AI with Diffusion Models Start Here 6: Wrap-up and Assessment

Generative AI with Diffusion Models

Start Here

0: Server Access

1: From U-Nets to Diffusion

2: Denoising Diffusion Probabilistic Models

3: Optimizations

4: Classifier-Free Diffusion Guidance

5: CLIP

6: Wrap-up and Assessment

Next Steps

Feedback

Previous

Next



0:00 / 8:11

Speed 1.0x

 DEEP LEARNING INSTITUTE

This Lab 0 : 14 : 10 / 2 : 00 : 00

Course 8 : 03 : 08 / 24 : 00 : 00

 LAUNCH

 STOP TASK

 ASSESS TASK

**Generative AI with Diffusion Models**
Part 6: Wrap-up & Assessment

6: Wrap-up and Assessment

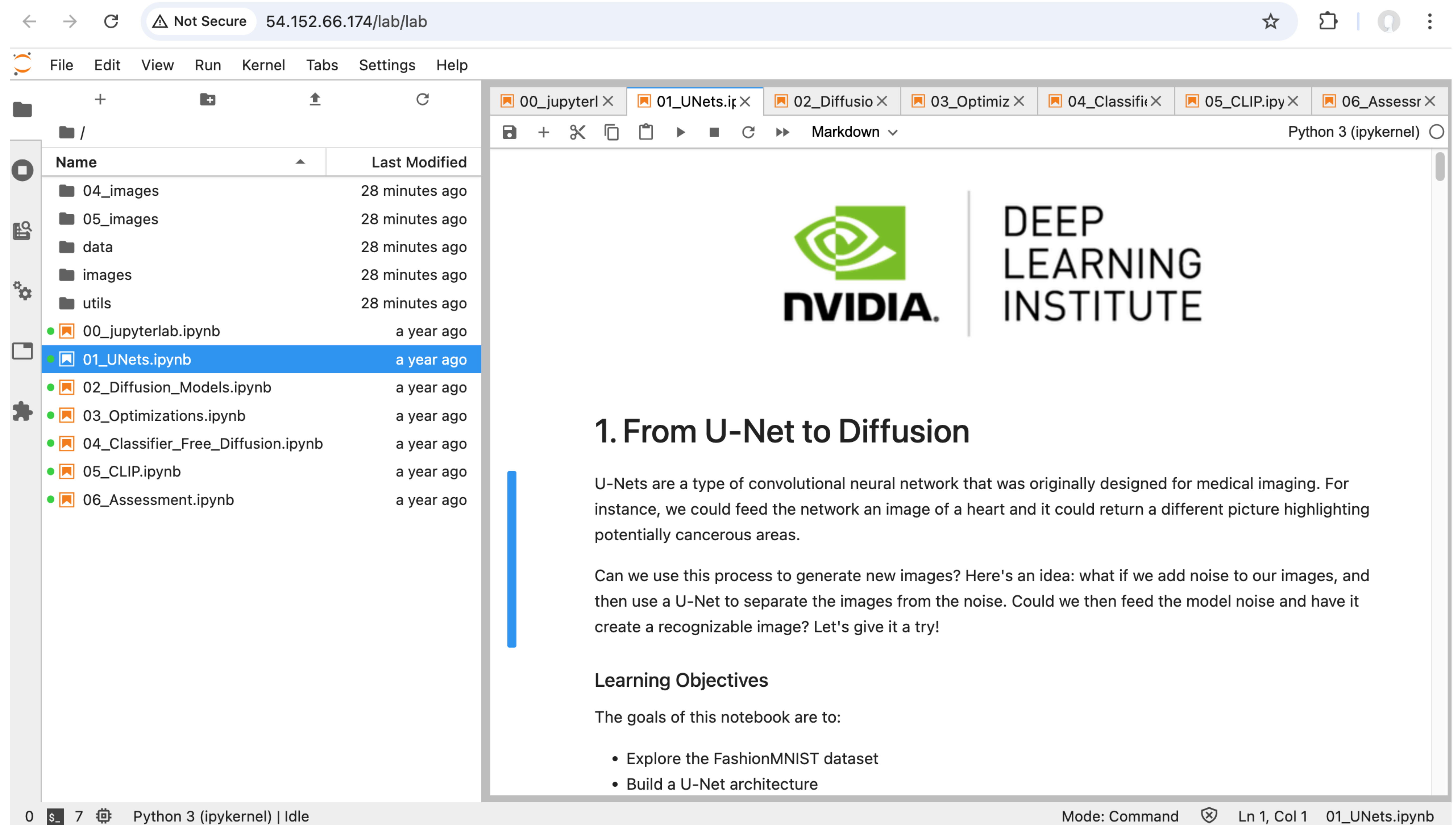
NVIDIA Generative AI with Diffusion Models

The screenshot displays the JupyterLab web interface. On the left is a file browser sidebar showing a directory structure with folders like '04_images', '05_images', 'data', 'images', and 'utils', and notebooks like '00_jupyterlab.ipynb' through '06_Assessment.ipynb'. The top features a menu bar with options: File, Edit, View, Run, Kernel, Tabs, Settings, and Help. The main work area shows a 'Launcher' page with the NVIDIA Deep Learning Institute logo and the title 'JupyterLab'. Below the logo, a paragraph explains that JupyterLab is used to manage the environment and provides access to interactive notebooks, a terminal window, and a dashboard. A green arrow points to the menu bar in the inset image.

Menu bar

https://iupvterlab.readthedocs.io/en/stable/user/interface.html

NVIDIA Generative AI with Diffusion Models



The screenshot shows a JupyterLab interface with a file explorer on the left and a notebook editor on the right. The file explorer lists several folders and files, with '01_UNets.ipynb' selected. The notebook editor shows the content of '01_UNets.ipynb', which includes the NVIDIA logo and the text 'DEEP LEARNING INSTITUTE'.

1. From U-Net to Diffusion

U-Nets are a type of convolutional neural network that was originally designed for medical imaging. For instance, we could feed the network an image of a heart and it could return a different picture highlighting potentially cancerous areas.

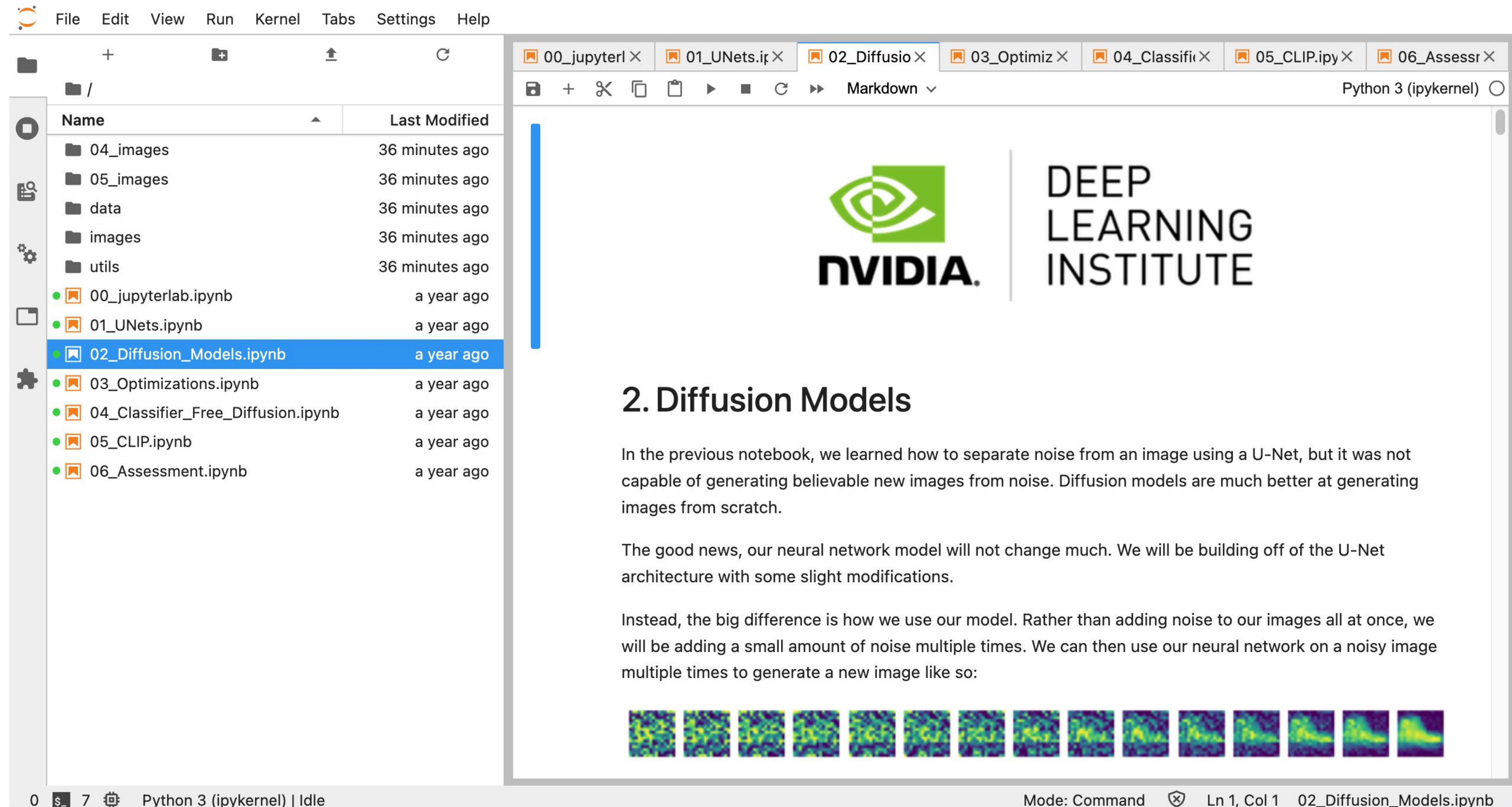
Can we use this process to generate new images? Here's an idea: what if we add noise to our images, and then use a U-Net to separate the images from the noise. Could we then feed the model noise and have it create a recognizable image? Let's give it a try!

Learning Objectives

The goals of this notebook are to:

- Explore the FashionMNIST dataset
- Build a U-Net architecture

NVIDIA Generative AI with Diffusion Models



File Edit View Run Kernel Tabs Settings Help

04_images 36 minutes ago
05_images 36 minutes ago
data 36 minutes ago
images 36 minutes ago
utils 36 minutes ago
00_jupyterlab.ipynb a year ago
01_UNets.ipynb a year ago
02_Diffusion_Models.ipynb a year ago
03_Optimizations.ipynb a year ago
04_Classifier_Free_Diffusion.ipynb a year ago
05_CLIP.ipynb a year ago
06_Assessment.ipynb a year ago

00_jupyterl × 01_UNets.ip × 02_Diffusio × 03_Optimiz × 04_Classifi × 05_CLIP.ipy × 06_Assessr ×

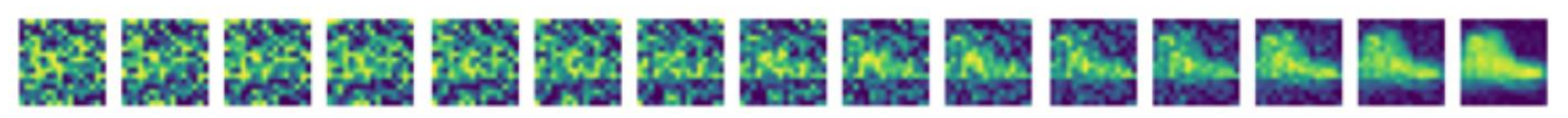
Python 3 (ipykernel)

2. Diffusion Models

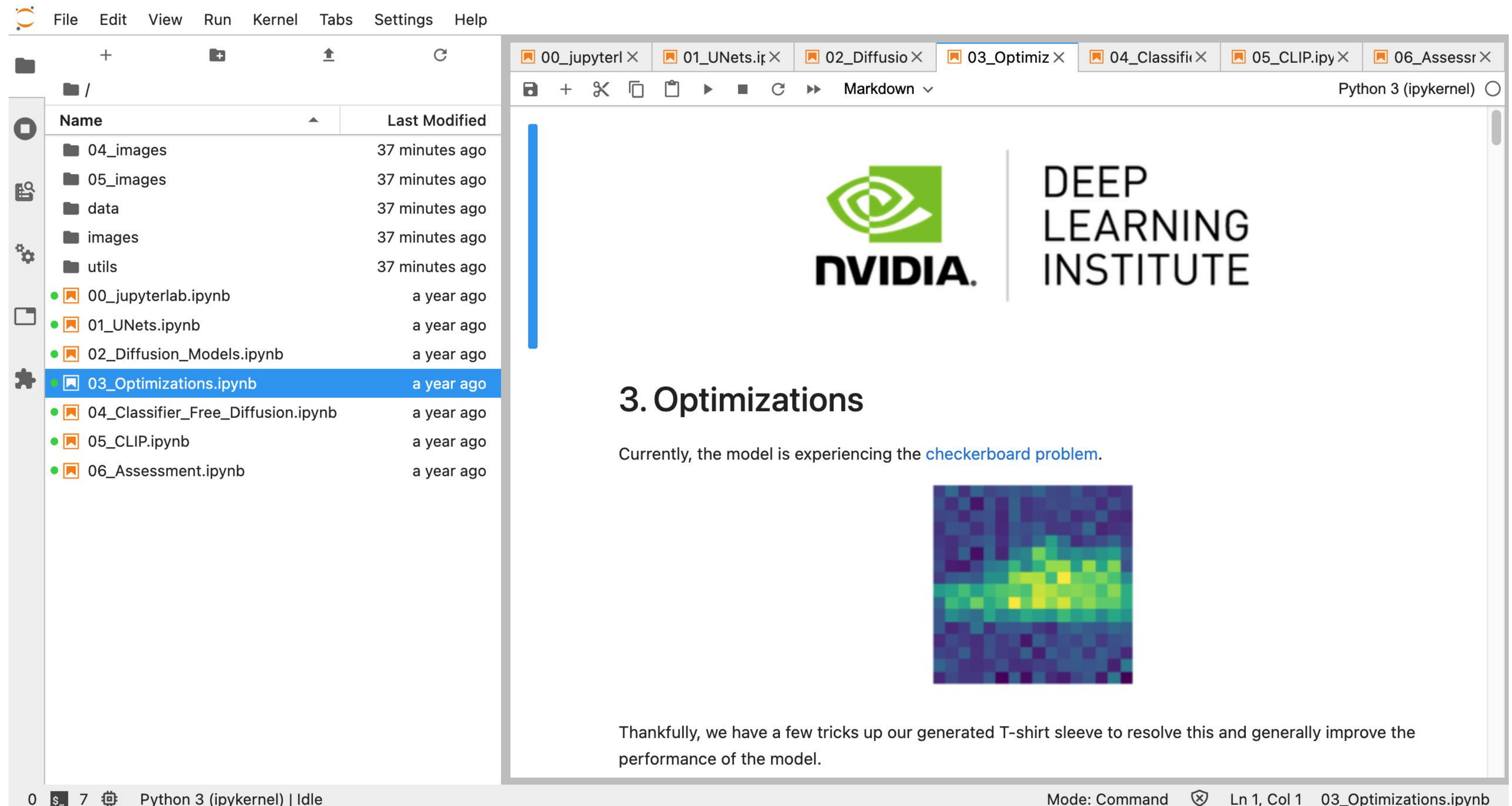
In the previous notebook, we learned how to separate noise from an image using a U-Net, but it was not capable of generating believable new images from noise. Diffusion models are much better at generating images from scratch.

The good news, our neural network model will not change much. We will be building off of the U-Net architecture with some slight modifications.

Instead, the big difference is how we use our model. Rather than adding noise to our images all at once, we will be adding a small amount of noise multiple times. We can then use our neural network on a noisy image multiple times to generate a new image like so:



NVIDIA Generative AI with Diffusion Models



The screenshot displays a JupyterLab environment. On the left, a file explorer shows a directory structure with folders like '04_images', '05_images', 'data', 'images', and 'utils', and notebooks '00_jupyterlab.ipynb' through '06_Assessment.ipynb'. The notebook '03_Optimizations.ipynb' is selected and open in the main area. The notebook content features the NVIDIA and Deep Learning Institute logos at the top. Below them is the section header '3. Optimizations'. The text states: 'Currently, the model is experiencing the [checkerboard problem](#).' This is followed by a heatmap visualization showing a pattern of high and low values. The text concludes: 'Thankfully, we have a few tricks up our generated T-shirt sleeve to resolve this and generally improve the performance of the model.'

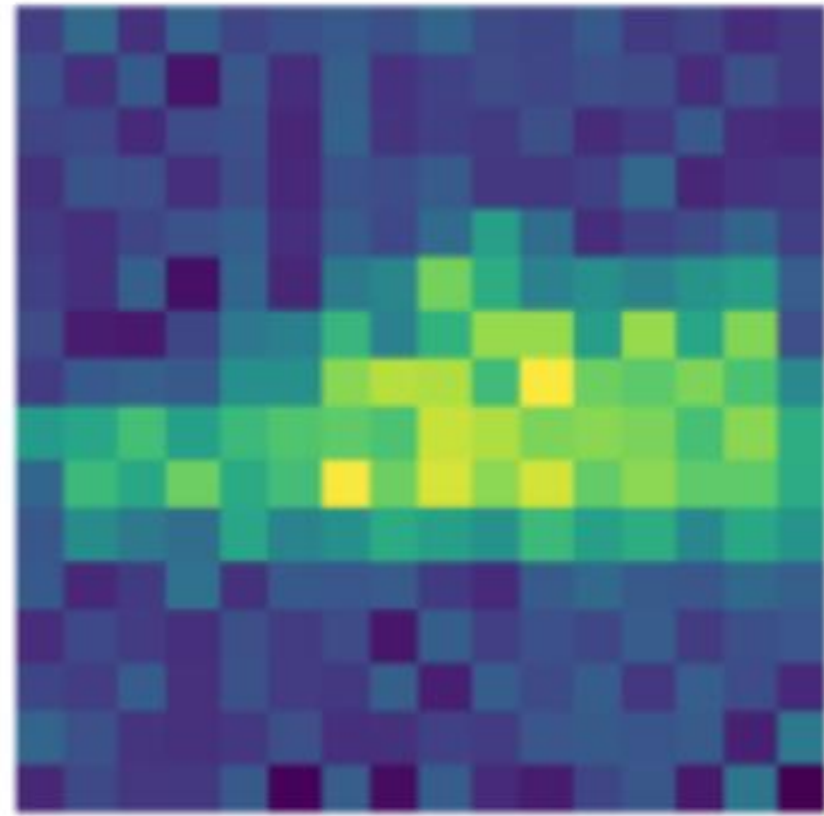
Name	Last Modified
04_images	37 minutes ago
05_images	37 minutes ago
data	37 minutes ago
images	37 minutes ago
utils	37 minutes ago
00_jupyterlab.ipynb	a year ago
01_UNets.ipynb	a year ago
02_Diffusion_Models.ipynb	a year ago
03_Optimizations.ipynb	a year ago
04_Classifier_Free_Diffusion.ipynb	a year ago
05_CLIP.ipynb	a year ago
06_Assessment.ipynb	a year ago

03_Optimizations.ipynb

DEEP LEARNING INSTITUTE

3. Optimizations

Currently, the model is experiencing the [checkerboard problem](#).



Thankfully, we have a few tricks up our generated T-shirt sleeve to resolve this and generally improve the performance of the model.

NVIDIA Generative AI with Diffusion Models

The screenshot displays a JupyterLab environment. On the left, a file browser shows a directory structure with folders '04_images', '05_images', 'data', 'images', and 'utils', all last modified '37 minutes ago'. Below these are notebooks '00_jupyterlab.ipynb' through '06_Assessment.ipynb', all last modified 'a year ago'. The notebook '04_Classifier_Free_Diffusion.ipynb' is selected. The main editor area shows the notebook content, which includes the NVIDIA and Deep Learning Institute logos, the section title '4. Classifier-Free Diffusion Guidance', and introductory text about training models on the FashionMNIST dataset. The text mentions 'Classifier-Free Diffusion Guidance' as a way to create a 'Conditional Diffusion Model'. Below this, 'Learning Objectives' are listed: adding categorical embeddings to a U-Net, training with a Bernoulli mask, adding a weight to the reverse diffusion process, and practicing on a more challenging dataset. The notebook concludes with the instruction to load necessary libraries and dataset information.

Name	Last Modified
04_images	37 minutes ago
05_images	37 minutes ago
data	37 minutes ago
images	37 minutes ago
utils	37 minutes ago
00_jupyterlab.ipynb	a year ago
01_UNets.ipynb	a year ago
02_Diffusion_Models.ipynb	a year ago
03_Optimizations.ipynb	a year ago
04_Classifier_Free_Diffusion.ipynb	a year ago
05_CLIP.ipynb	a year ago
06_Assessment.ipynb	a year ago

4. Classifier-Free Diffusion Guidance

So far, we have been able to train a model to generate images of clothing using the [FashionMNIST](#) dataset. However, there is no way for the user to specify what kind of images should be generated. Let's fix that by using [Classifier-Free Diffusion Guidance](#), a relatively simple way to create a [Conditional Diffusion Model](#).

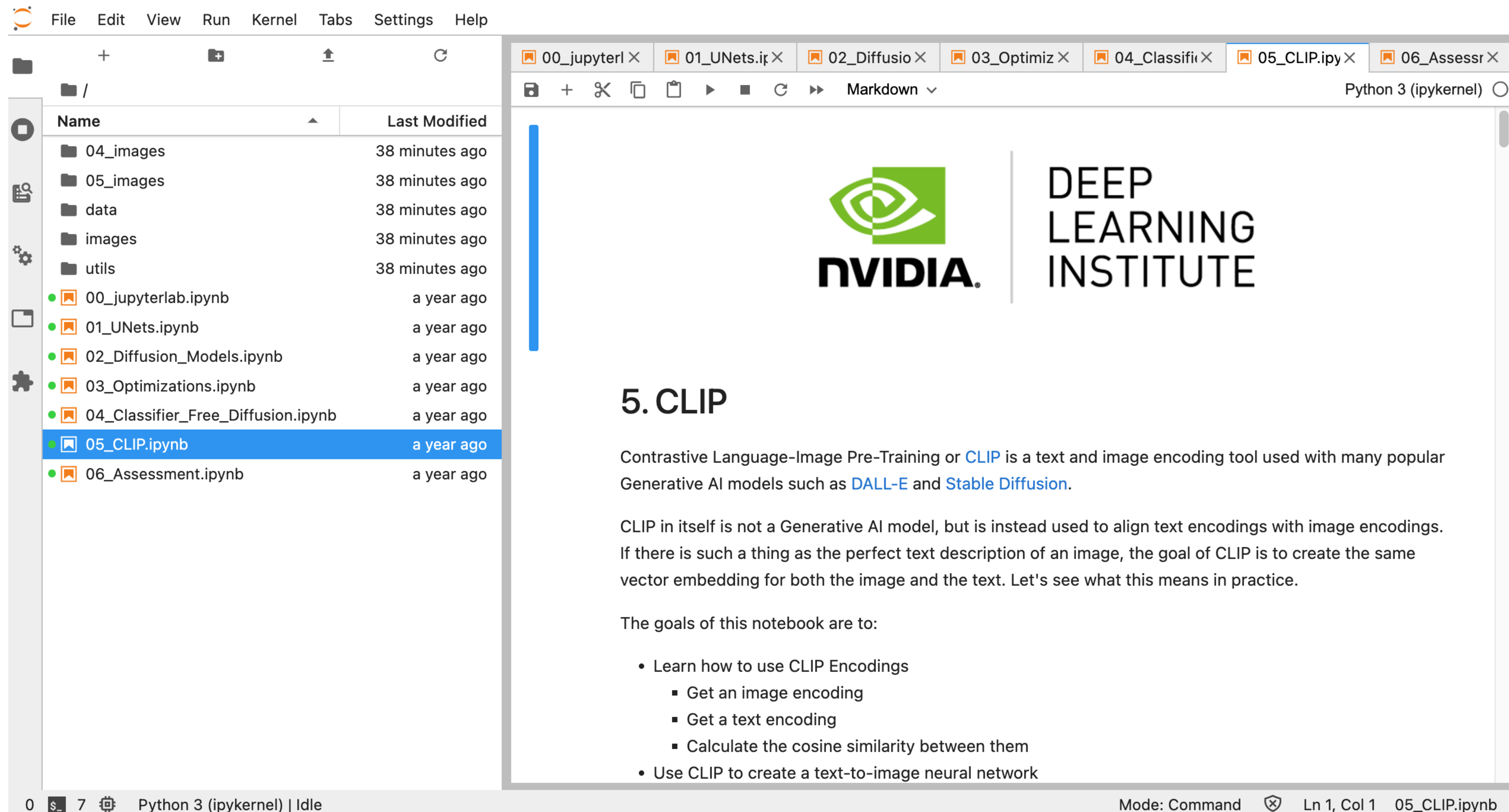
Learning Objectives

The goals of this notebook are to:

- Add categorical embeddings to a U-Net
- Train a model with a Bernoulli mask
- Add a weight to the reverse diffusion process
- Practice learnings on a more challenging dataset

Before we get started, let's load the necessary libraries and dataset information.

NVIDIA Generative AI with Diffusion Models



File Edit View Run Kernel Tabs Settings Help

04_images 38 minutes ago
05_images 38 minutes ago
data 38 minutes ago
images 38 minutes ago
utils 38 minutes ago
00_jupyterlab.ipynb a year ago
01_UNets.ipynb a year ago
02_Diffusion_Models.ipynb a year ago
03_Optimizations.ipynb a year ago
04_Classifier_Free_Diffusion.ipynb a year ago
05_CLIP.ipynb a year ago
06_Assessment.ipynb a year ago

00_jupyterl × 01_UNets.ip × 02_Diffusio × 03_Optimiz × 04_Classific × 05_CLIP.ipyn × 06_Assessr ×

Python 3 (ipykernel)

5. CLIP

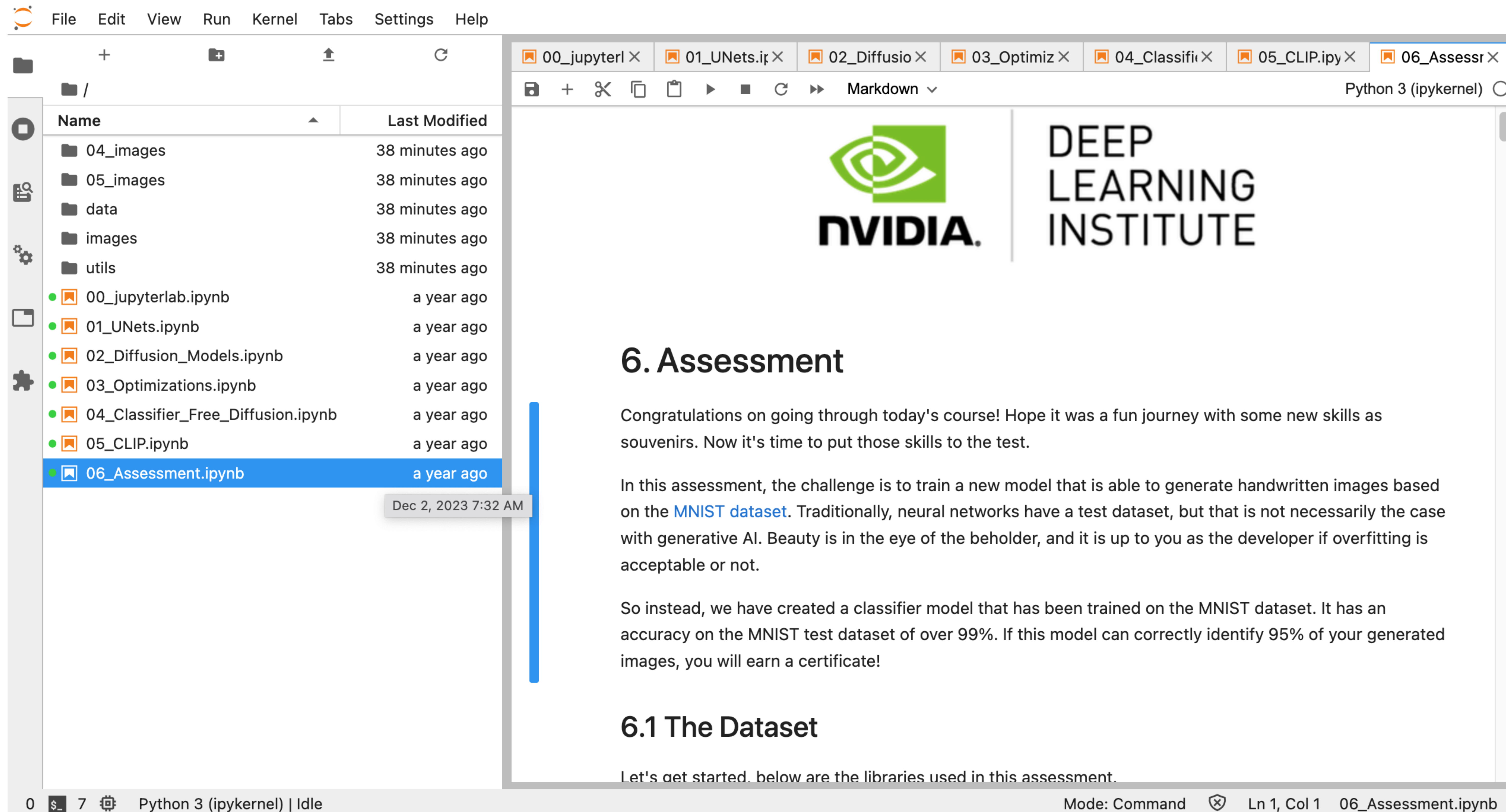
Contrastive Language-Image Pre-Training or [CLIP](#) is a text and image encoding tool used with many popular Generative AI models such as [DALL-E](#) and [Stable Diffusion](#).

CLIP in itself is not a Generative AI model, but is instead used to align text encodings with image encodings. If there is such a thing as the perfect text description of an image, the goal of CLIP is to create the same vector embedding for both the image and the text. Let's see what this means in practice.

The goals of this notebook are to:

- Learn how to use CLIP Encodings
 - Get an image encoding
 - Get a text encoding
 - Calculate the cosine similarity between them
- Use CLIP to create a text-to-image neural network

NVIDIA Generative AI with Diffusion Models



File Edit View Run Kernel Tabs Settings Help

04_images 38 minutes ago
05_images 38 minutes ago
data 38 minutes ago
images 38 minutes ago
utils 38 minutes ago
00_jupyterlab.ipynb a year ago
01_UNets.ipynb a year ago
02_Diffusion_Models.ipynb a year ago
03_Optimizations.ipynb a year ago
04_Classifier_Free_Diffusion.ipynb a year ago
05_CLIP.ipynb a year ago
06_Assessment.ipynb a year ago

Dec 2, 2023 7:32 AM

00_jupyterl X 01_UNets.ip X 02_Diffusio X 03_Optimiz X 04_Classifi X 05_CLIP.ip X 06_Assessr X

Python 3 (ipykernel)

6. Assessment

Congratulations on going through today's course! Hope it was a fun journey with some new skills as souvenirs. Now it's time to put those skills to the test.

In this assessment, the challenge is to train a new model that is able to generate handwritten images based on the [MNIST dataset](#). Traditionally, neural networks have a test dataset, but that is not necessarily the case with generative AI. Beauty is in the eye of the beholder, and it is up to you as the developer if overfitting is acceptable or not.

So instead, we have created a classifier model that has been trained on the MNIST dataset. It has an accuracy on the MNIST test dataset of over 99%. If this model can correctly identify 95% of your generated images, you will earn a certificate!

6.1 The Dataset

Let's get started. below are the libraries used in this assessment.

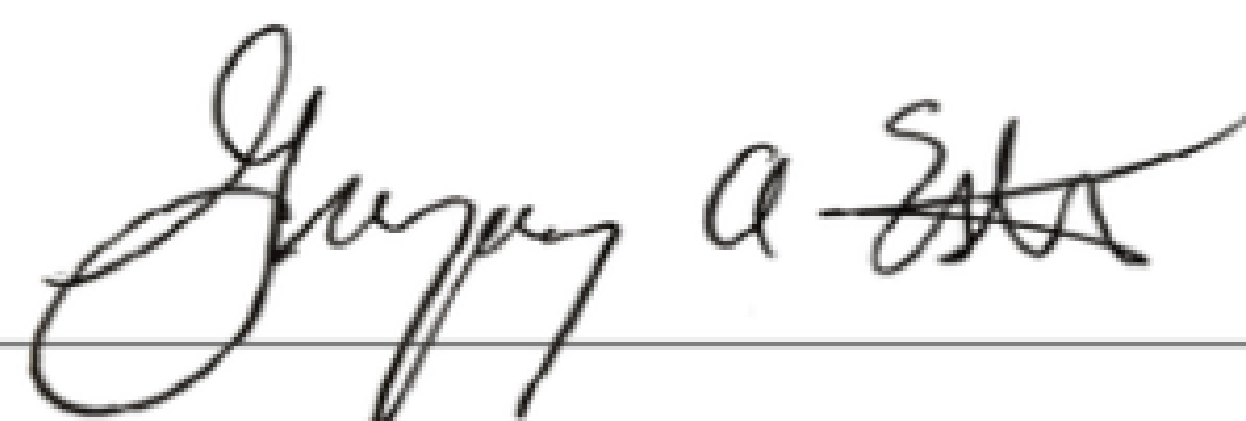
Certificate of Competency

This certificate is awarded to

Min-Yuh Day

for demonstrating competence in the completion of

Generative AI with Diffusion Models



Greg Estes

Vice President, NVIDIA

Issue Date: : February 28, 2025

Certification ID: q3oo-oBhTQKtyCCote2E-Q

References

- NVIDIA DLI (2024), Building RAG Agents with LLMs,
https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-15+V1
- NVIDIA DLI (2024), Generative AI with Diffusion Models,
https://learn.nvidia.com/courses/course-detail?course_id=course-v1:DLI+S-FX-14+V1