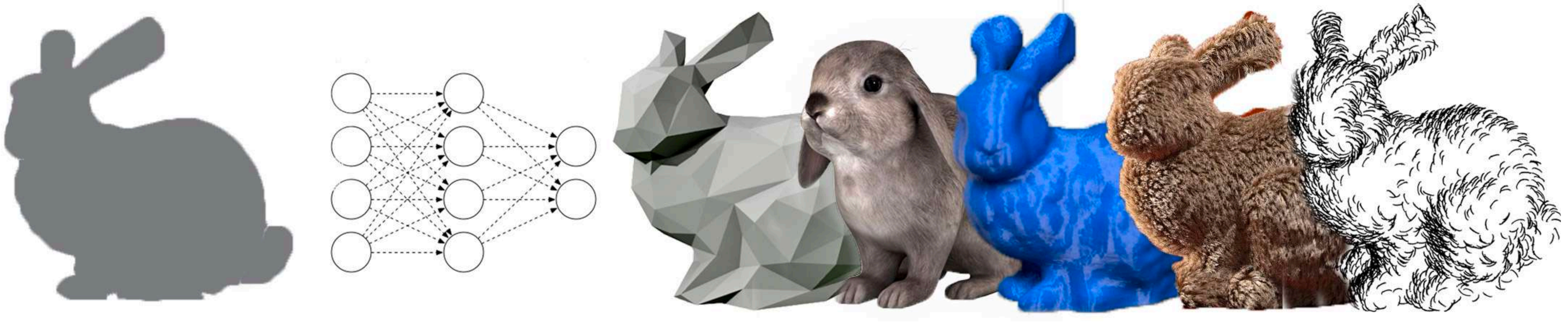
COMP0169: Machine Learning for Visual Computing

Introduction

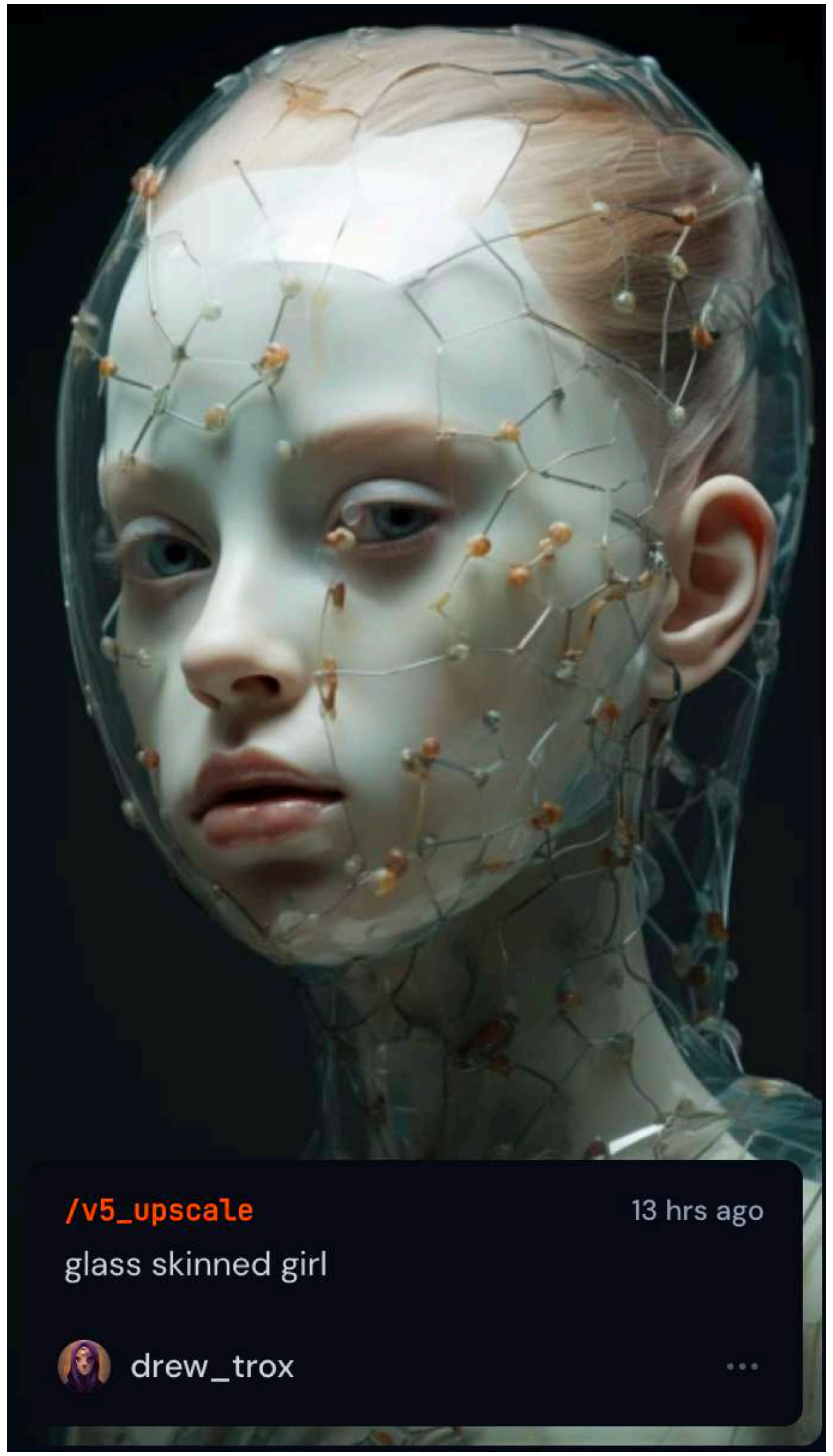



UCL

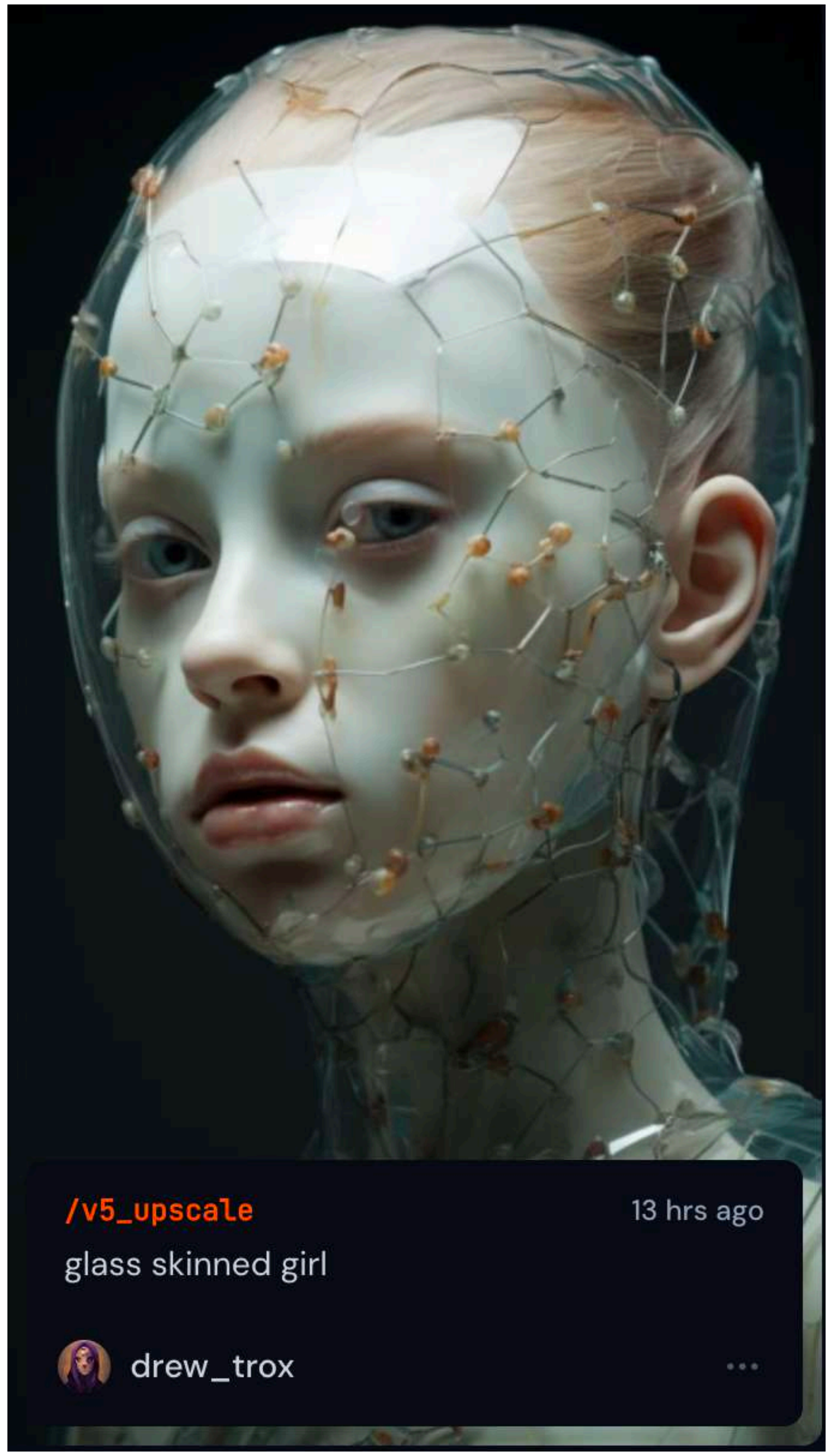
Lectures will be Recorded

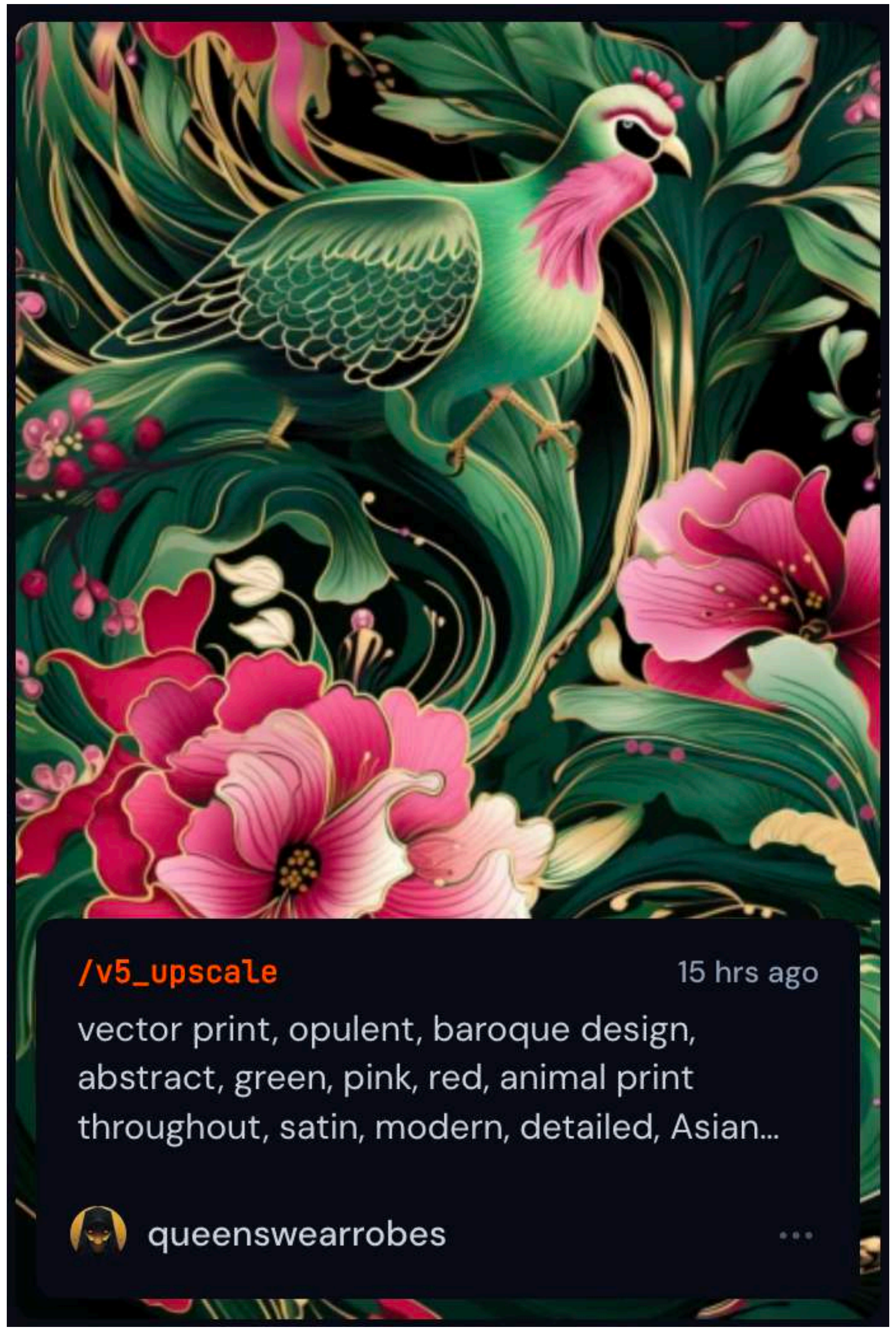
MidJourney Gallery

MidJourney Gallery

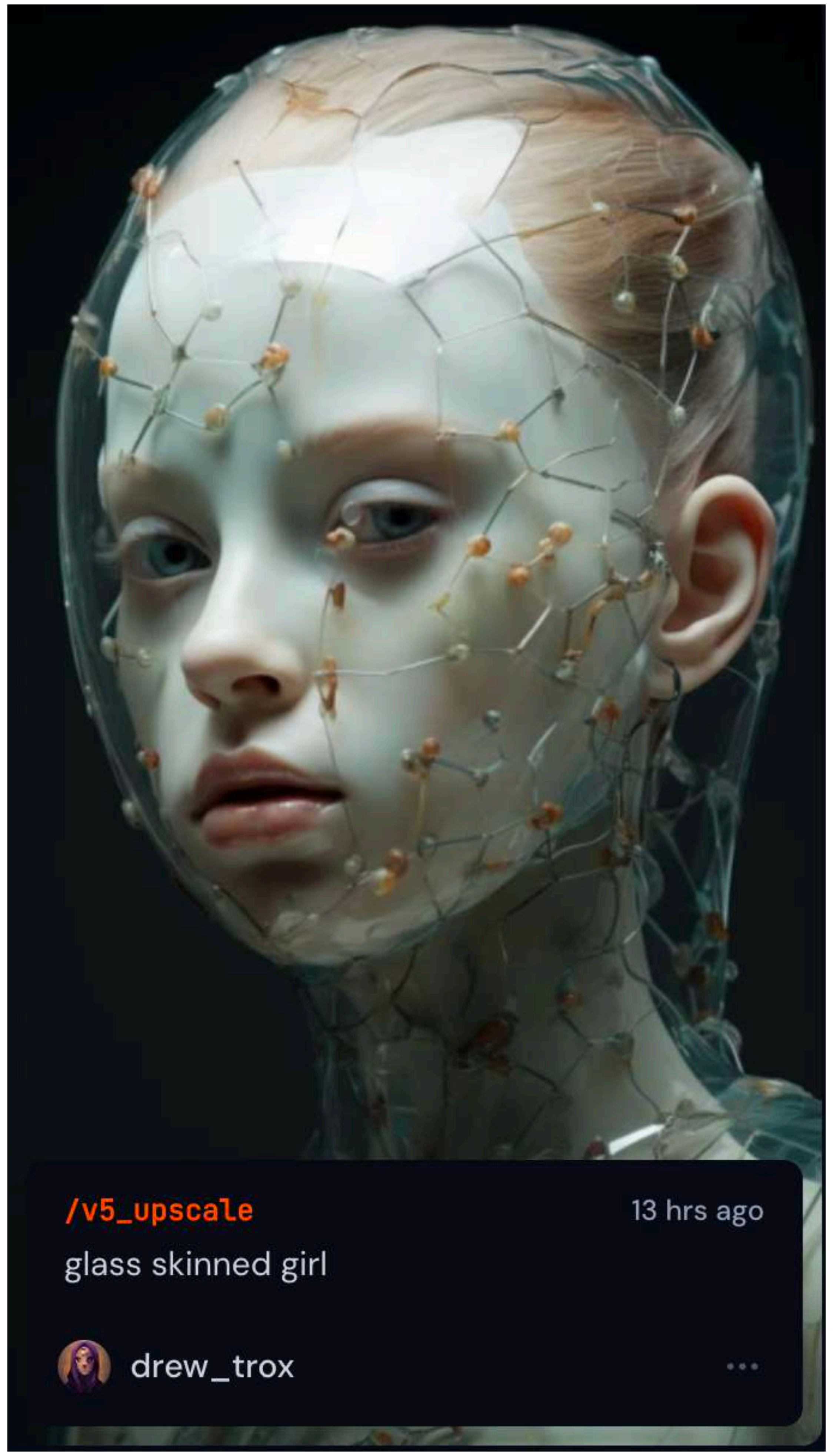


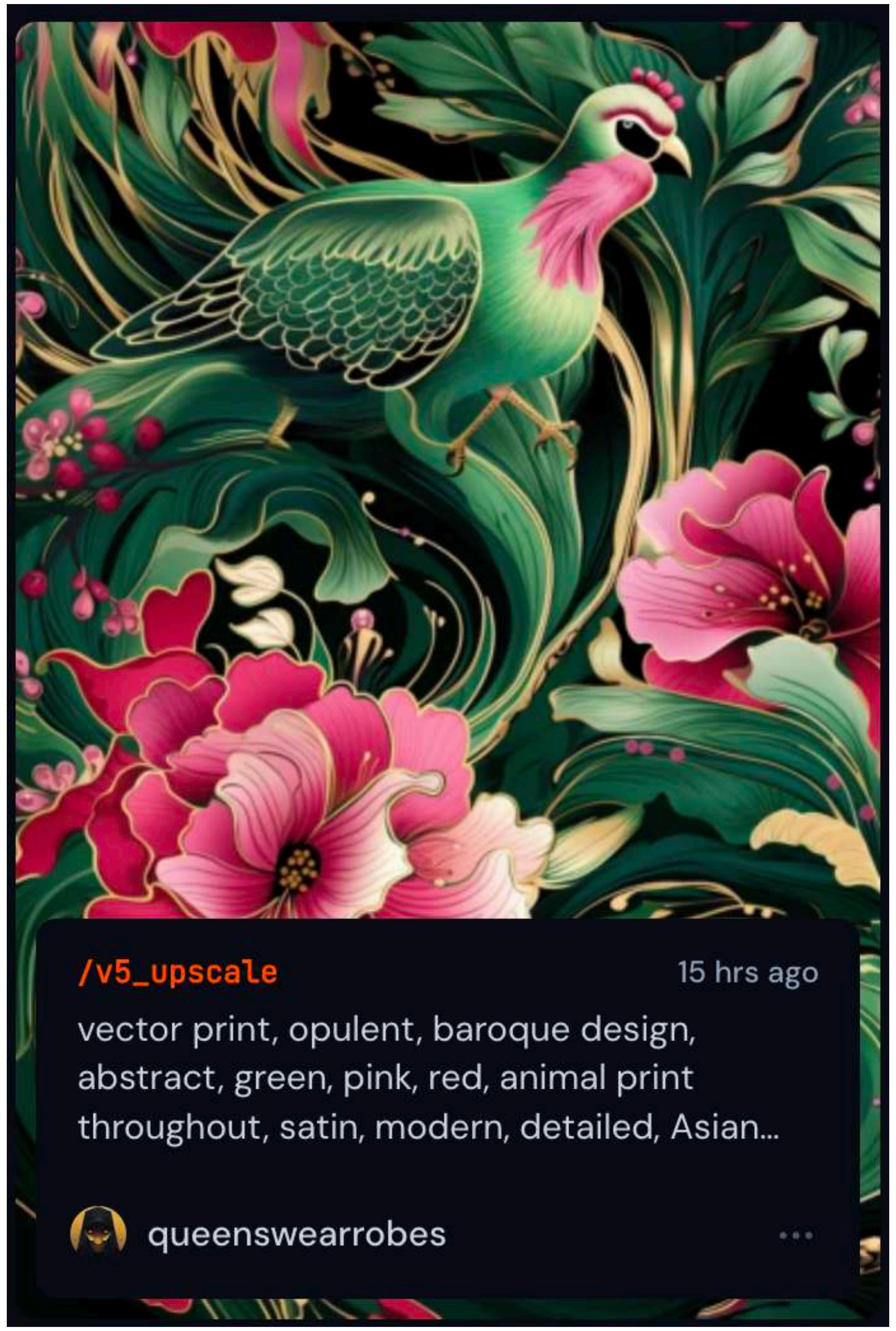
MidJourney Gallery

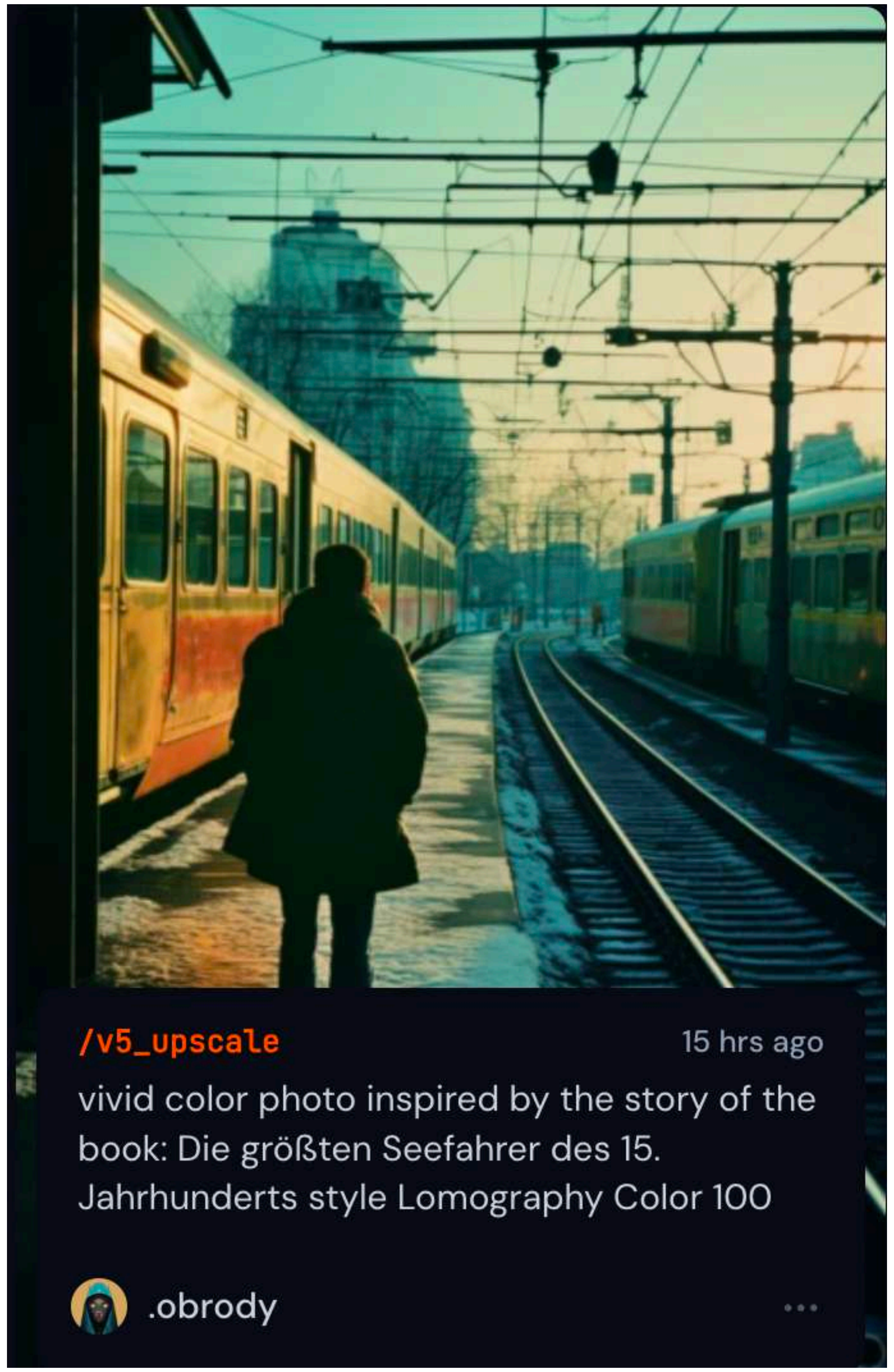




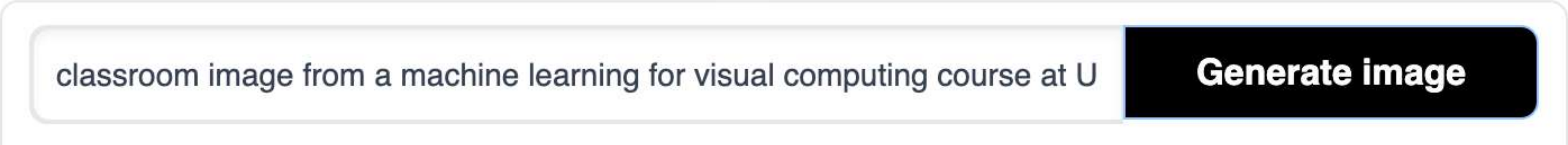
MidJourney Gallery



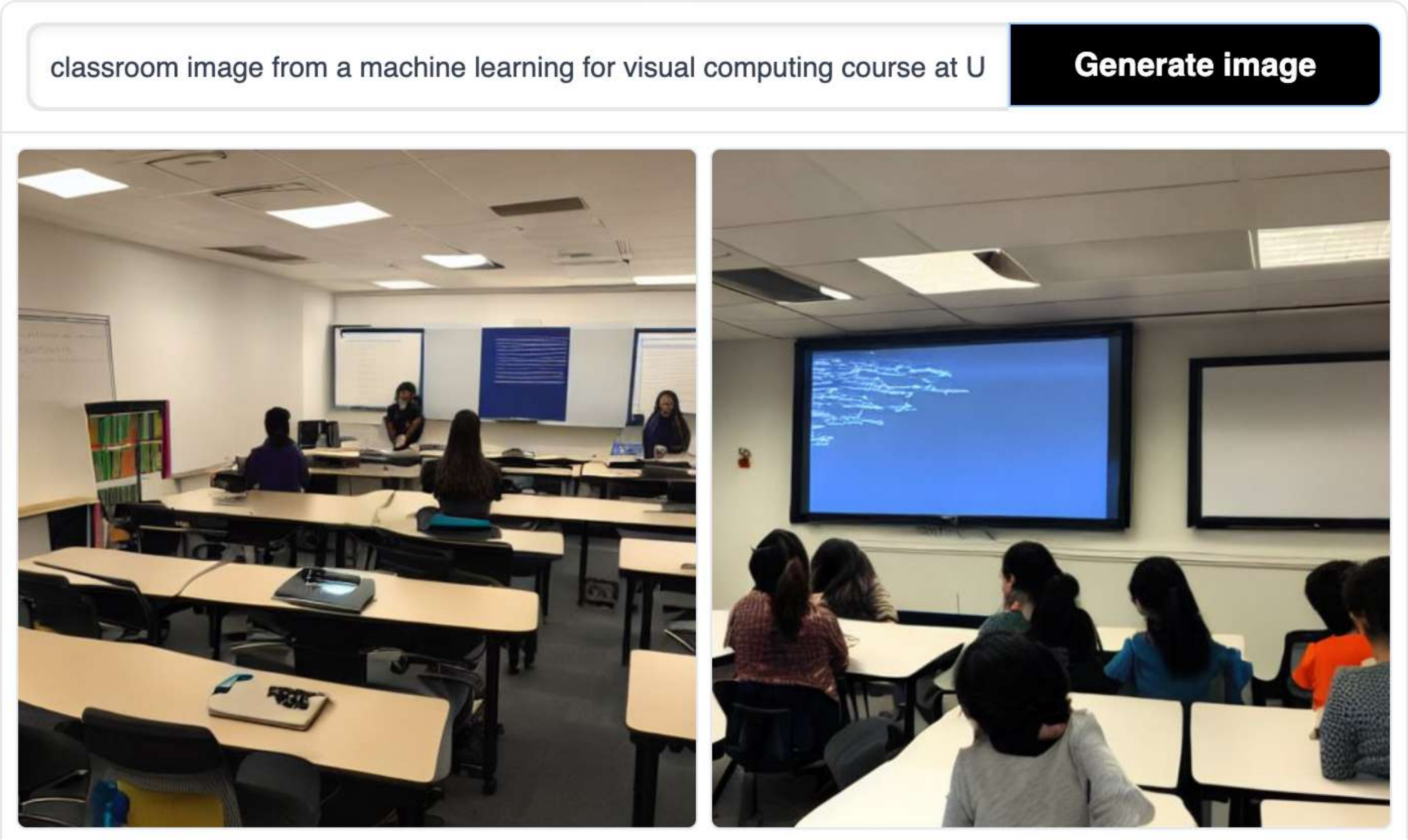




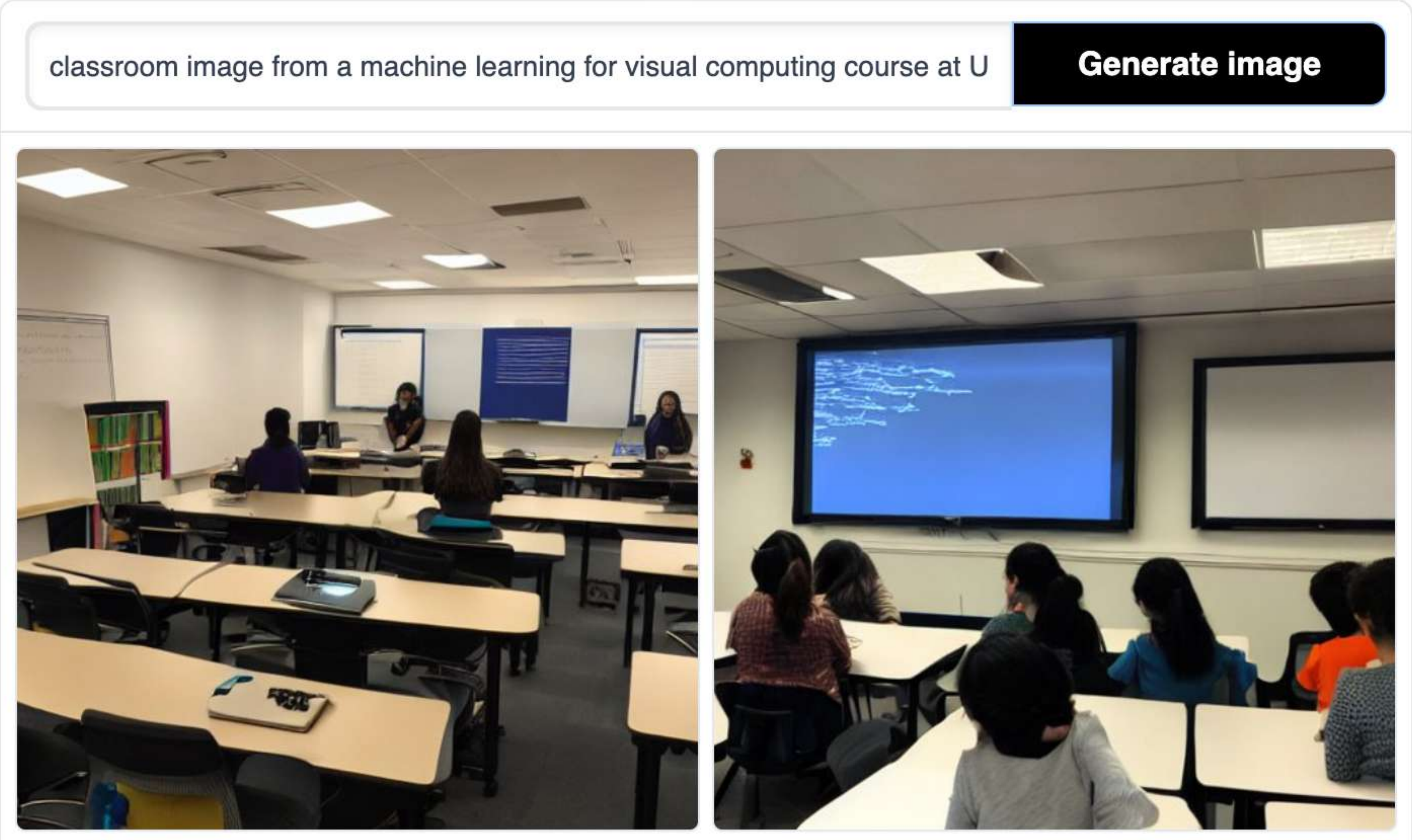
Multimodal Generation (Stable Diffusion)

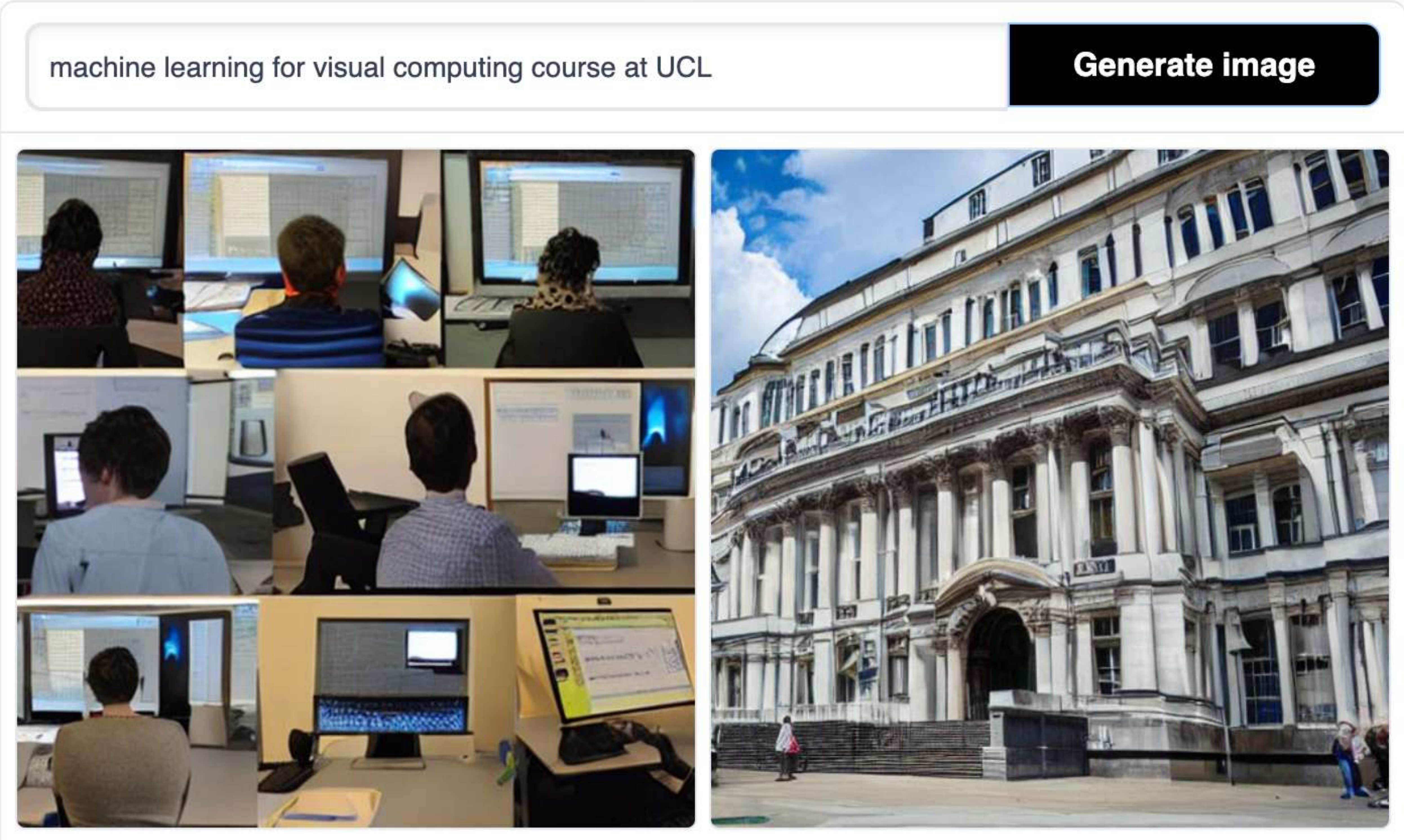


Multimodal Generation (Stable Diffusion)

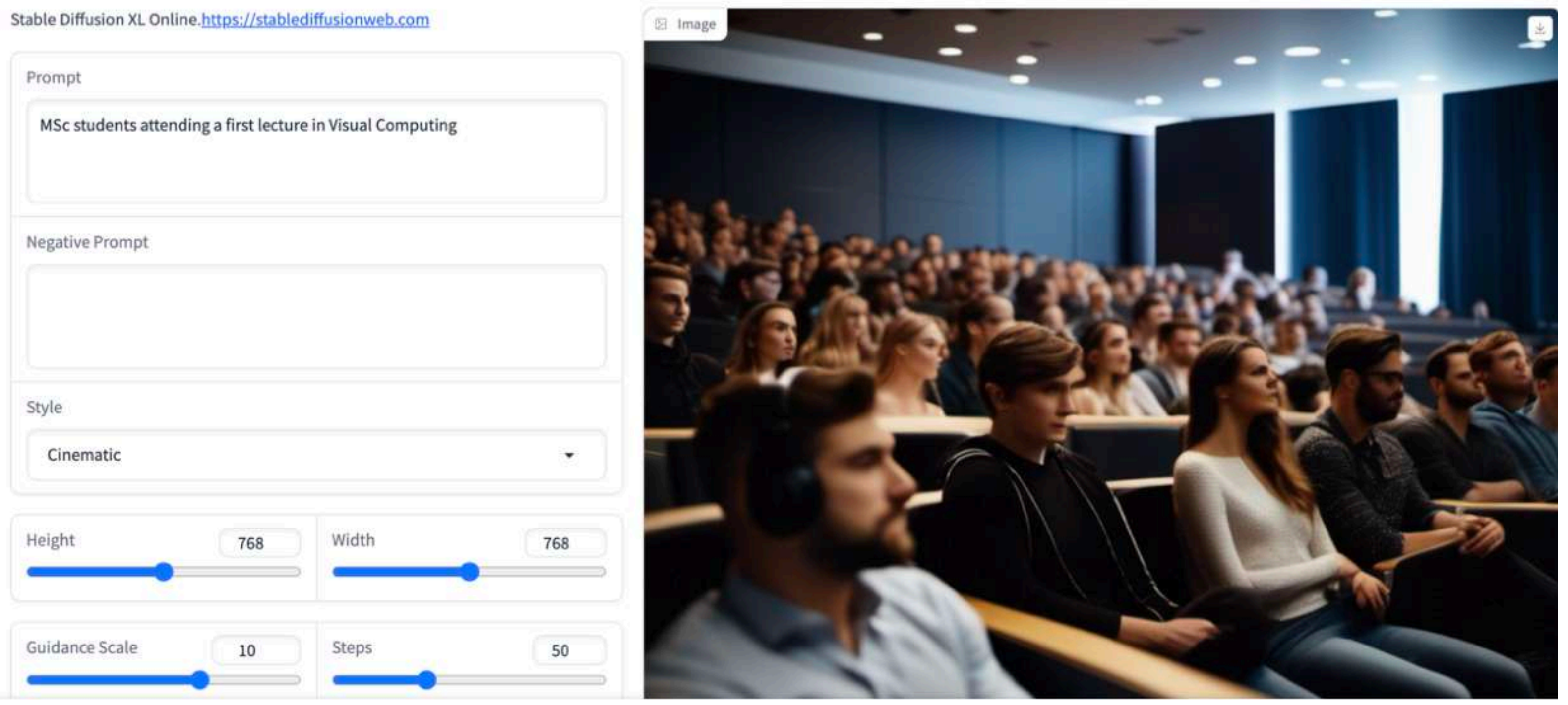


Multimodal Generation (Stable Diffusion)





Stable Diffusion XL



Stable Diffusion XL

Stable Diffusion XL Online.https://stablediffusionweb.com

Prompt

MSc students attending a first lecture in Visual Computing

Negative Prompt

Style

Cinematic

Height 768

Width 768

Guidance Scale 10

Steps 50

People

Niloy Mitra

People

Niloy Mitra **Tobias Ritschel**

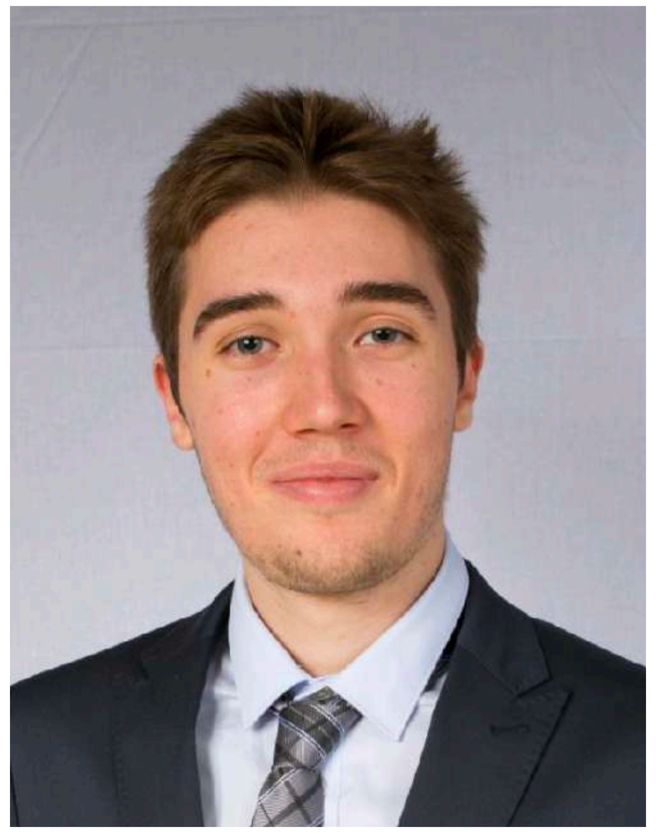
People

Niloy Mitra

Tobias Ritschel

Chen

Maria



Remy

Sanjeev

Course Overview

Course Overview

- Introduction (machine learning motivation and terminology) + Regression

Course Overview

- Introduction (machine learning motivation and terminology) + Regression
- Visual media types and representations

Course Overview

- Introduction (machine learning motivation and terminology) + Regression
- Visual media types and representations
- SGD + optimisation

Course Overview

- Introduction (machine learning motivation and terminology) + Regression
- Visual media types and representations
- SGD + optimisation
- MLPs for overfitting

Course Overview

- Introduction (machine learning motivation and terminology) + Regression
- Visual media types and representations
- SGD + optimisation
- MLPs for overfitting
- CNNs

Course Overview

- Introduction (machine learning motivation and terminology) + Regression
- Visual media types and representations
- SGD + optimisation
- MLPs for overfitting
- CNNs
- Classification networks and feature visualisation

Course Overview

- Introduction (machine learning motivation and terminology) + Regression
- Visual media types and representations
- SGD + optimisation
- MLPs for overfitting
- CNNs
- Classification networks and feature visualisation
- Style and transformers

Course Overview

- Introduction (machine learning motivation and terminology) + Regression
- Visual media types and representations
- SGD + optimisation
- MLPs for overfitting
- CNNs
- Classification networks and feature visualisation
- Style and transformers
- Temporal sequences

Course Overview

- Introduction (machine learning motivation and terminology) + Regression
- Visual media types and representations
- SGD + optimisation
- MLPs for overfitting
- CNNs
- Classification networks and feature visualisation
- Style and transformers
- Temporal sequences
- Generative models I (AE, VAE, GAN)

Course Overview

- Introduction (machine learning motivation and terminology) + Regression
- Visual media types and representations
- SGD + optimisation
- MLPs for overfitting
- CNNs
- Classification networks and feature visualisation
- Style and transformers
- Temporal sequences
- Generative models I (AE, VAE, GAN)
- Generative models II (diffusion models, language models)

Two-way Communication

Two-way Communication

- ***Course is being restructured this term!***

Two-way Communication

- ***Course is being restructured this term!***

- Our aim is to introduce you to the world of DL+CG

Two-way Communication

- ***Course is being restructured this term!***

- Our aim is to introduce you to the world of DL+CG

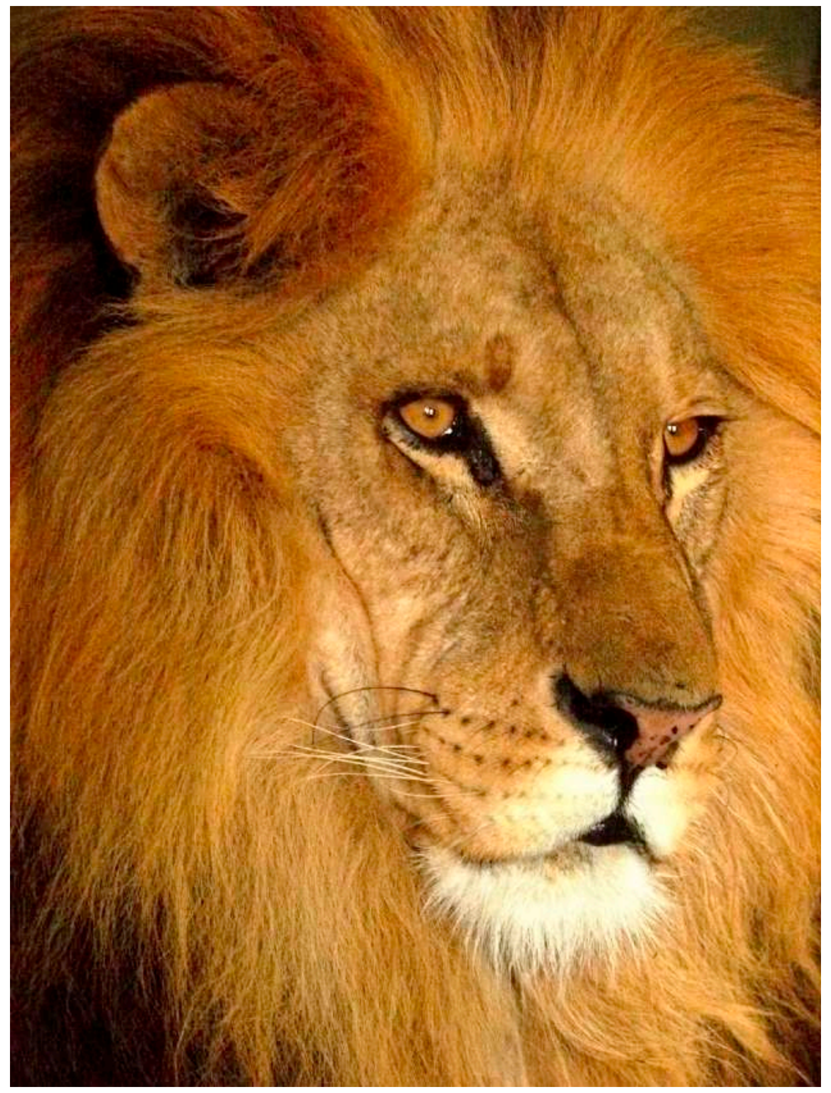
- You are invited/encouraged to give feedback
  - On-line forum (via Moodle)
  - Speak up. Please send us your criticism/comments/suggestions
  - Ask questions, please!

- **Thanks to many people who helped so far with slides/comments**

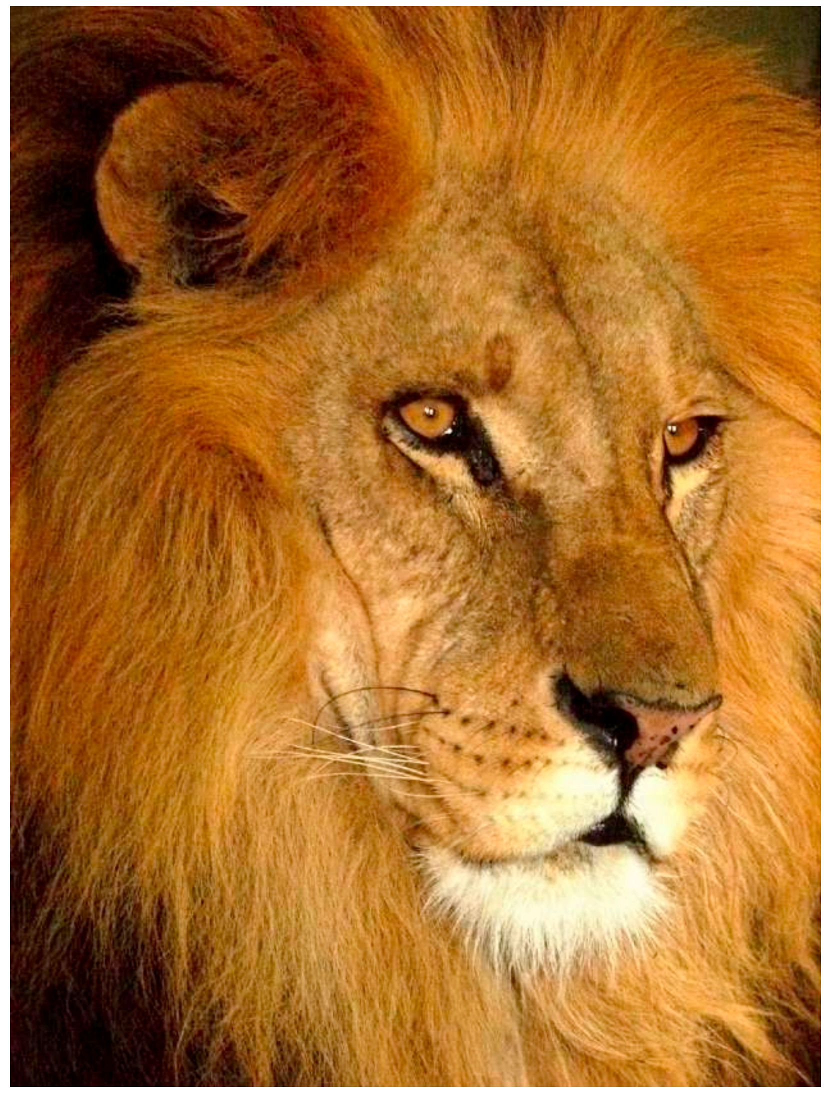
Media Types

Media Types

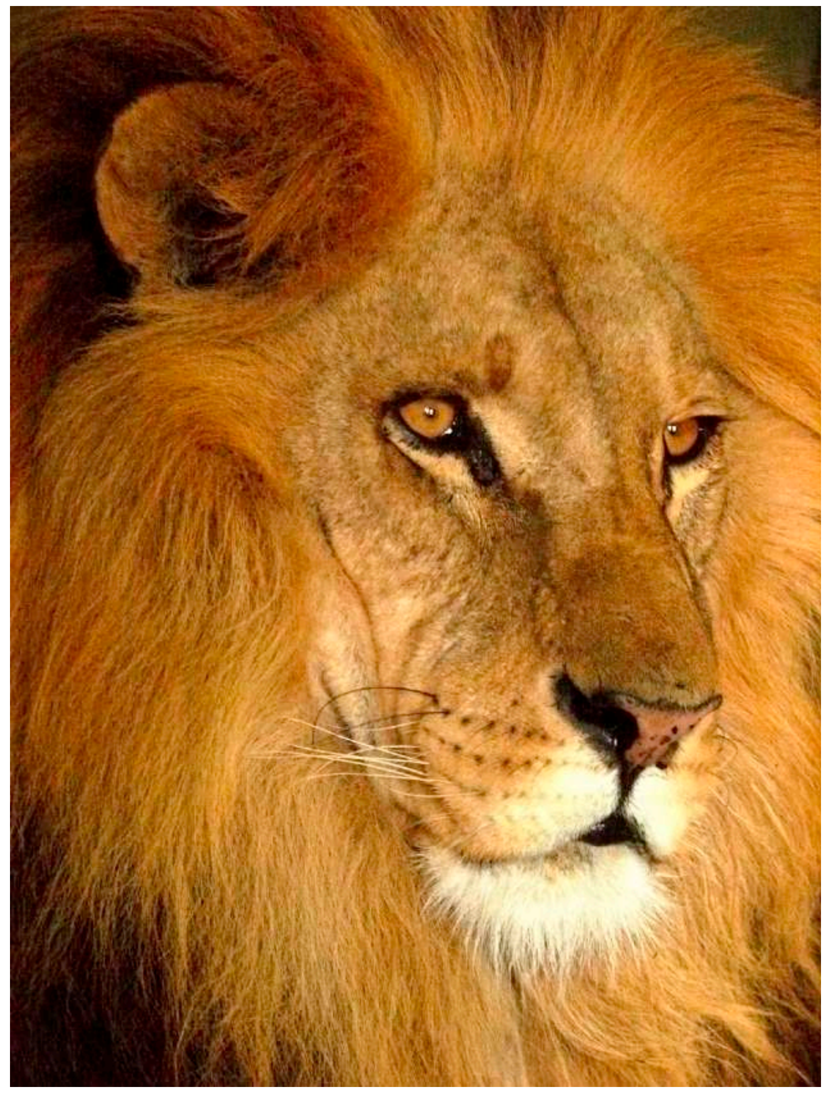
Media Types

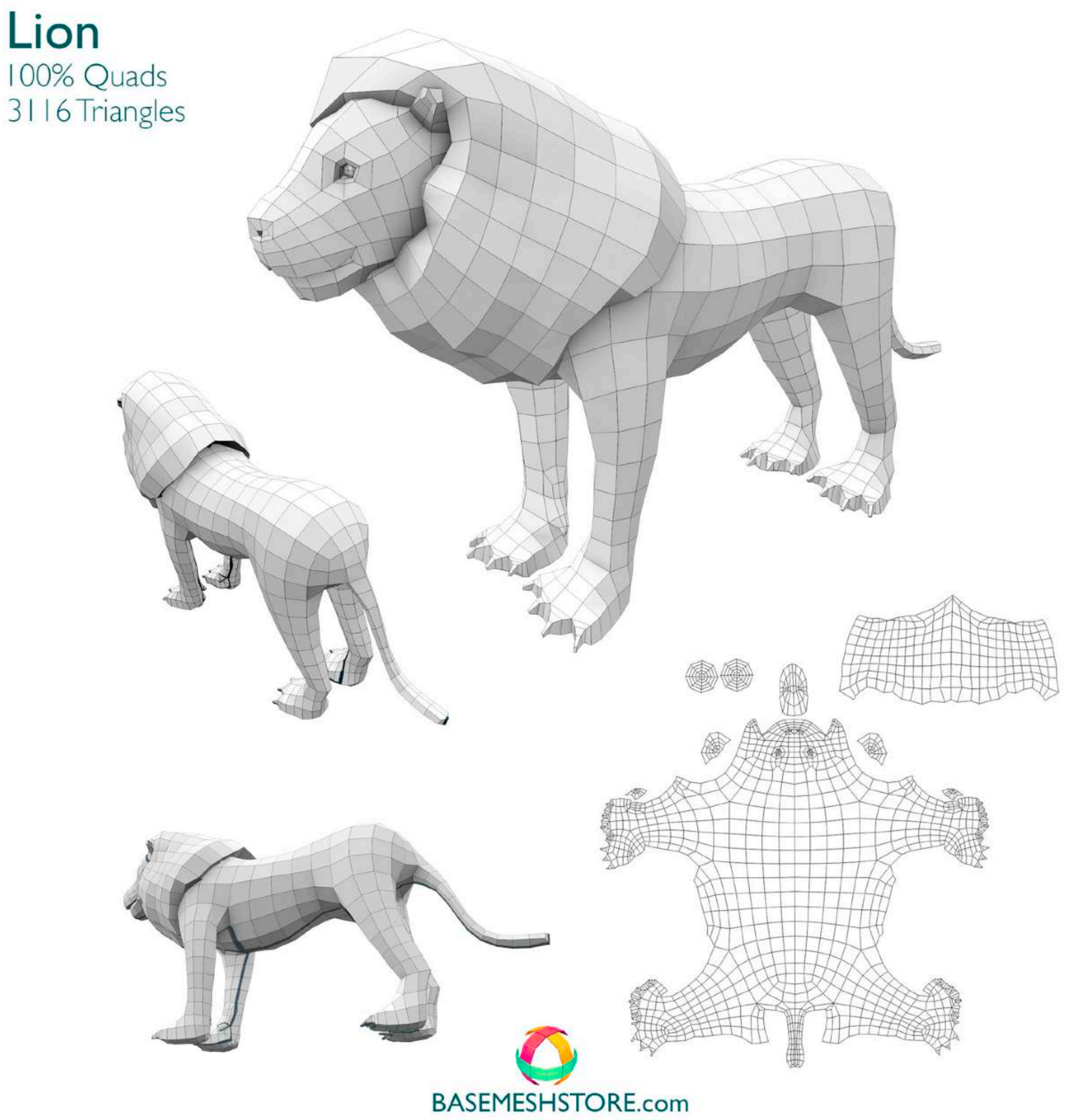


Media Types


Media Types





Representations in CG

Representations in CG

- **Images, videos** (e.g., pixel grid, pixel grid over time)
- **Volume** (e.g., voxel grid)

Representations in CG

- **Images, videos** (e.g., pixel grid, pixel grid over time)
- **Volume** (e.g., voxel grid)
- **Meshes** (e.g., vertices/edges/faces)

Representations in CG

- **Images, videos** (e.g., pixel grid, pixel grid over time)
- **Volume** (e.g., voxel grid)
- **Meshes** (e.g., vertices/edges/faces)
- **Animation** (e.g., skeletal positions over time; cloth dynamics over time)

Representations in CG

- **Images, videos** (e.g., pixel grid, pixel grid over time)
- **Volume** (e.g., voxel grid)
- **Meshes** (e.g., vertices/edges/faces)
- **Animation** (e.g., skeletal positions over time; cloth dynamics over time)
- **Pointclouds** (e.g., point arrays)

Representations in CG

- **Images, videos** (e.g., pixel grid, pixel grid over time)
- **Volume** (e.g., voxel grid)
- **Meshes** (e.g., vertices/edges/faces)
- **Animation** (e.g., skeletal positions over time; cloth dynamics over time)
- **Pointclouds** (e.g., point arrays)
- **Physics simulations** (e.g., fluid flow over space/time)

Common Analysis Tasks

Common Analysis Tasks

Urban scene

Common Analysis Tasks

Urban scene

People

Common Analysis Tasks

Urban scene

People

"University building with students walking around"

Common Analysis Tasks

Urban scene

People

"University building with students walking around"

Generative editing and Synthesis

Example Synthesis Tasks

Example Synthesis Tasks

Example Synthesis Tasks

Mappings in Computer Graphics

- Feature detection (image features, point features) $\mathbb{R}^{m\times m} \rightarrow \mathbb{R}^{k}$
- Denoising, Smoothing, etc. $\mathbb{R}^{m\times m} \rightarrow \mathbb{R}^{m\times m}$
- Embedding, Distance computation $\mathbb{R}^{m\times m, m\times m} \rightarrow \mathbb{R}^{d}$
- Rendering $\mathbb{R}^{m\times m} \rightarrow \mathbb{R}^{m\times m}$
- Animation $\mathbb{R}^{3m\times t} \rightarrow \mathbb{R}^{3m}$
- Physical simulation $\mathbb{R}^{3m\times t} \rightarrow \mathbb{R}^{3m}$
- Generative models $\mathbb{R}^{d} \rightarrow \mathbb{R}^{m\times m}$

Image Classification

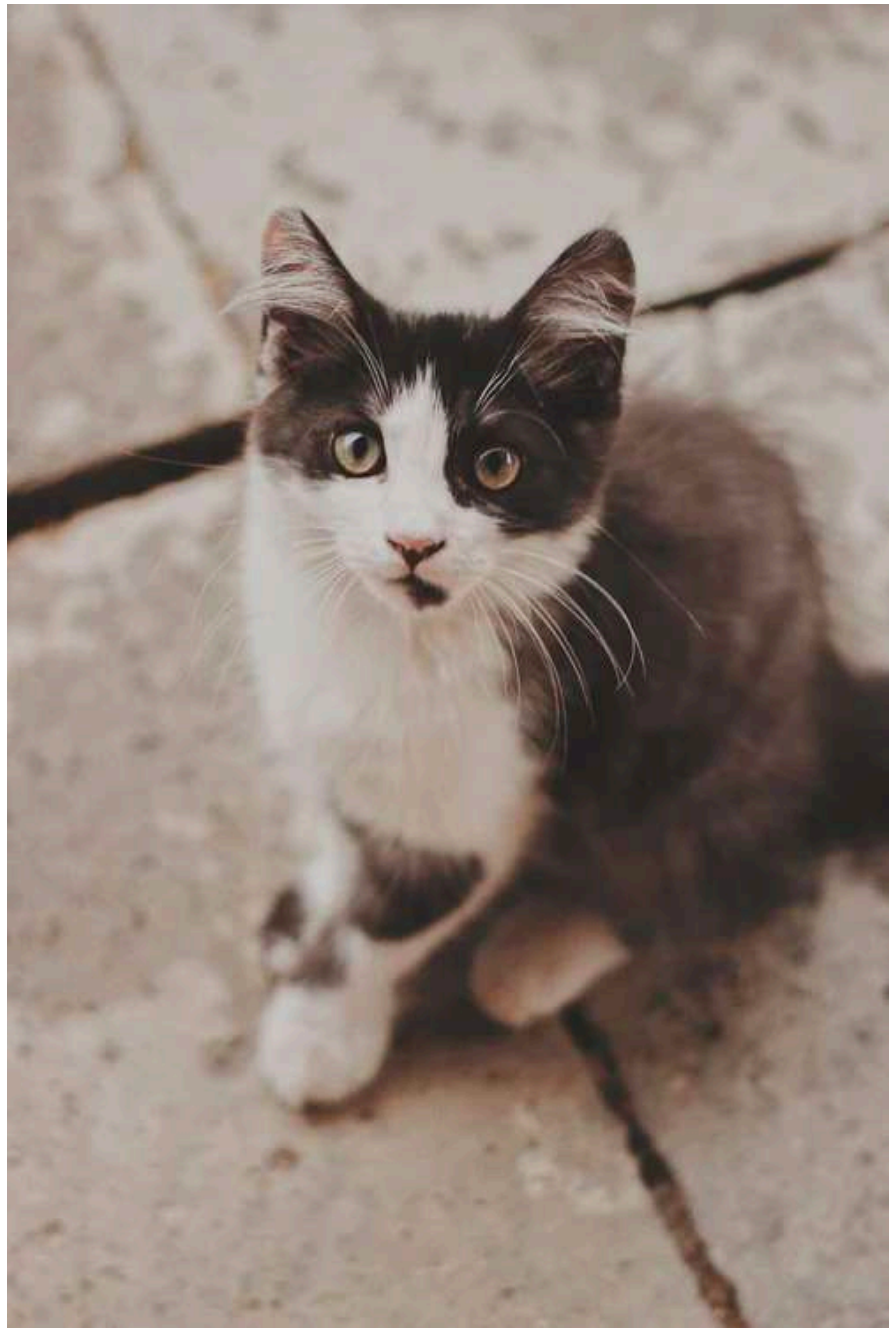
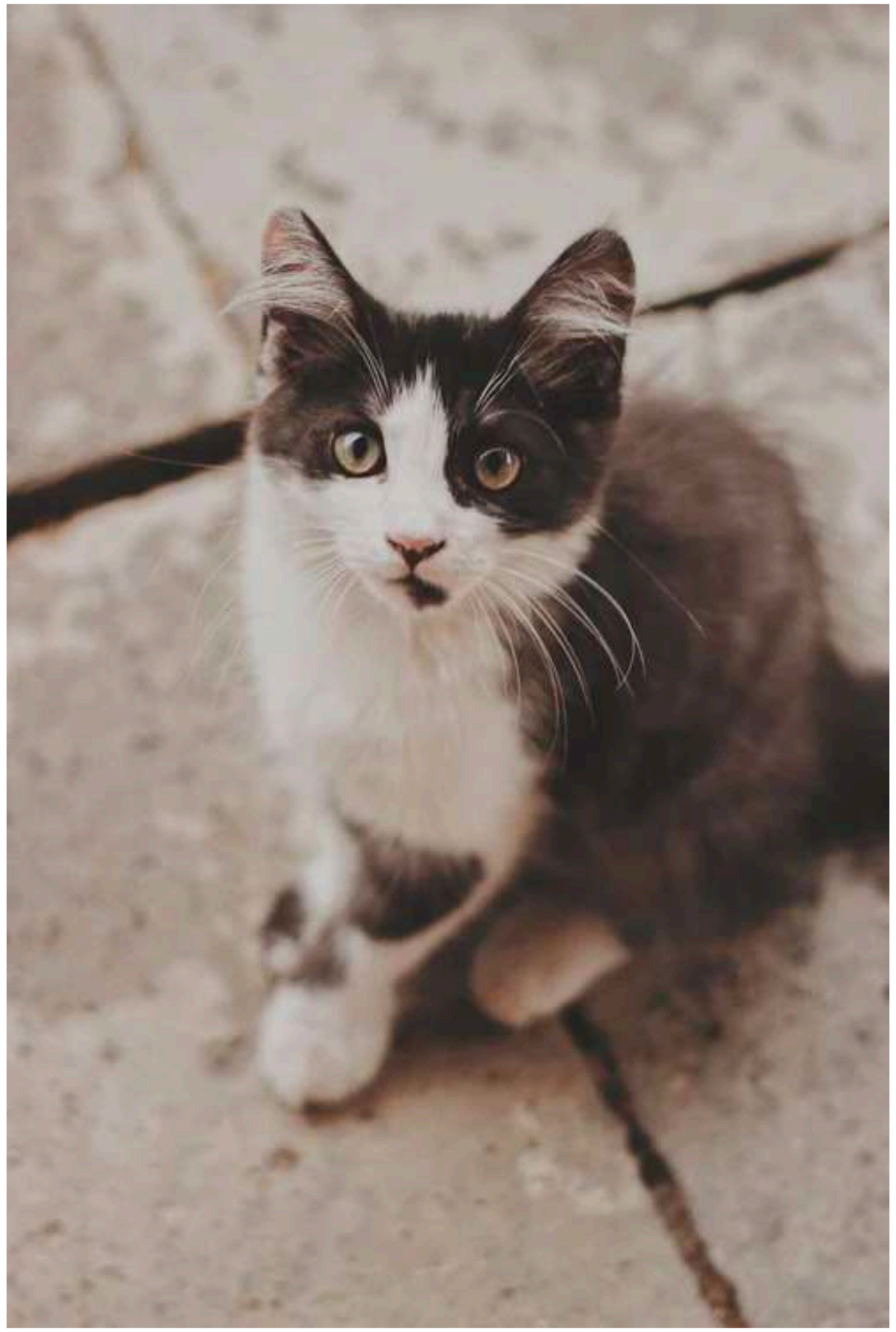
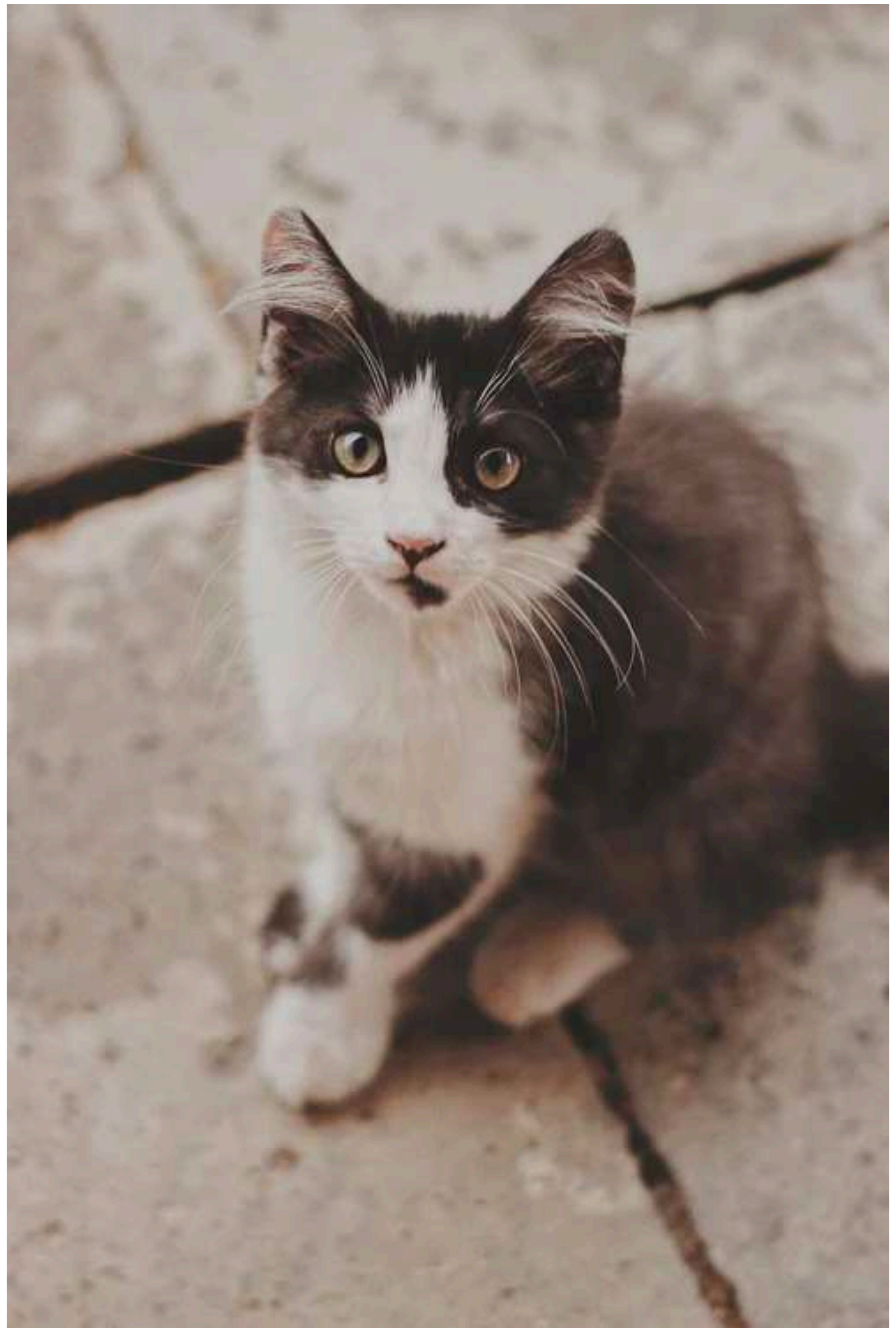


Image Classification



Human
Cat
Dog
Tiger
Chair

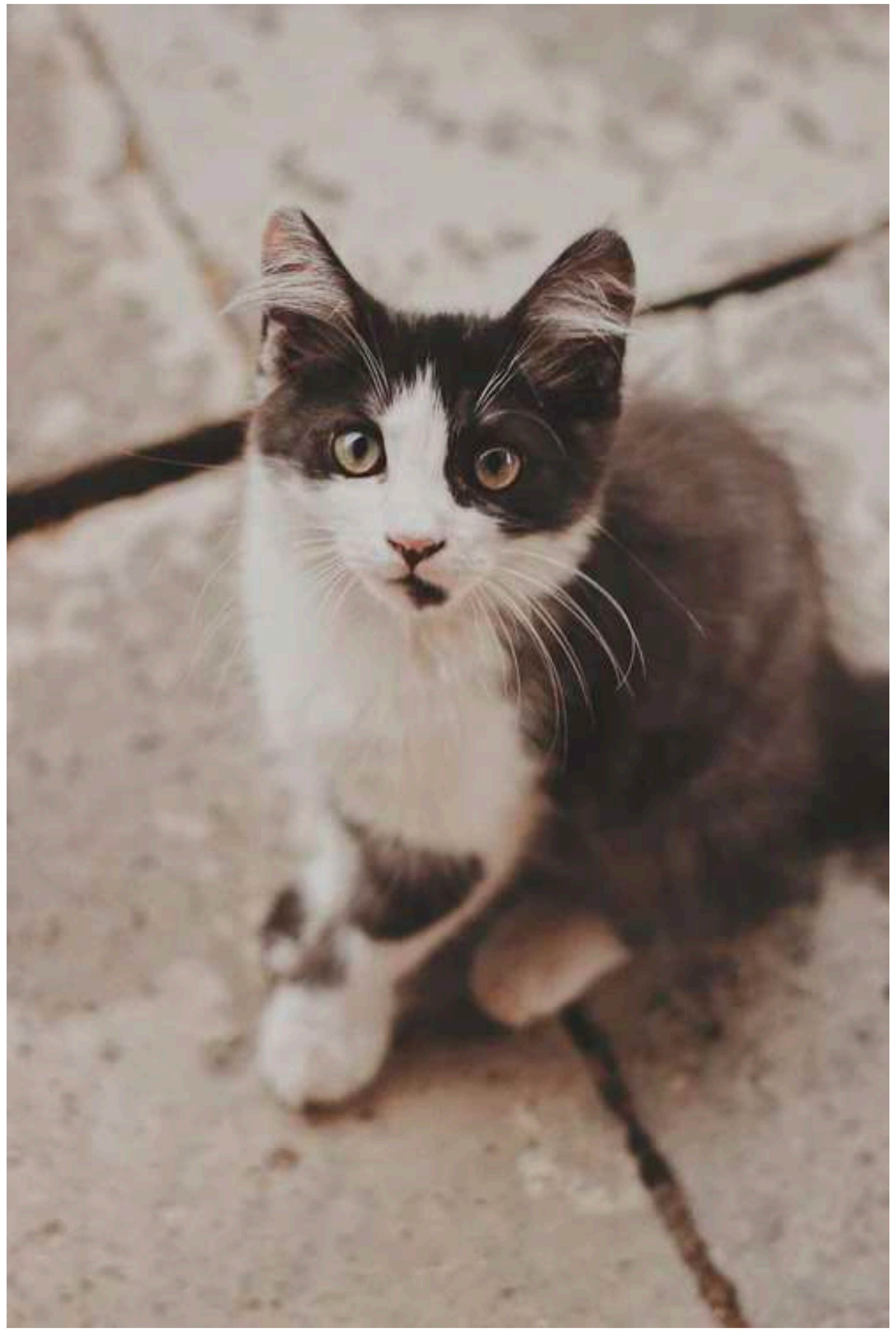
Image Classification



$$\mathbb{W} \times \mathbb{H} \times 3 \rightarrow \mathbb{Z}$$

Human
Cat
Dog
Tiger
Chair

Image Classification



$$\mathbb{W} \times \mathbb{H} \times 3 \rightarrow \mathbb{Z}$$

Human
Cat
Dog
Tiger
Chair

How will random guess or 'chance' fair?

Image as a Vector

35	10	16	100
30	152	62	30
16	1	200	106
12	98	58	167

Image as a Vector

35	10	16	100
30	152	62	30
16	1	200	106
12	98	58	167

$\rightarrow$

35
10
16
100
30
152
62
30
16
1
200
106
12
98
58
167

Notations: Vectors and Matrices

Notations: Vectors and Matrices

vector $\mathbf{x}$

Notations: Vectors and Matrices

vector $\mathbf{x}$

matrix $\mathbf{A}_{m \times n} = [\mathbf{a}_1 \ldots \mathbf{a}_n]$

Notations: Vectors and Matrices

vector	$\mathbf{x}$
matrix	$\mathbf{A}_{m \times n} = [\mathbf{a}_1 \ldots \mathbf{a}_n]$
linear equation	$\mathbf{Ax} = \mathbf{b}$

Notations: Vectors and Matrices

vector	$\mathbf{x}$
matrix	$\mathbf{A}_{m \times n} = [\mathbf{a}_1 \ldots \mathbf{a}_n]$
linear equation	$\mathbf{A}\mathbf{x} = \mathbf{b}$
inner product	$< \mathbf{x}, \mathbf{y} >= \mathbf{x}^T \mathbf{y}$

Notations: Vectors and Matrices

vector $\mathbf{x}$

matrix $\mathbf{A}_{m\times n} = [\mathbf{a}_1 \ldots \mathbf{a}_n]$

linear equation $\mathbf{A}\mathbf{x} = \mathbf{b}$

inner product $< \mathbf{x}, \mathbf{y} >= \mathbf{x}^T\mathbf{y}$

$$\|\mathbf{x}\| = \sqrt{\mathbf{x}^T\mathbf{x}}$$

$$\mathbf{x}^T\mathbf{y} = \|\mathbf{x}\|\|\mathbf{y}\|\cos(\theta)$$

Notations: Vectors and Matrices

- linear ***independence***; ***rank*** of a matrix
- ***span*** of a matrix

vector	$\mathbf{x}$
matrix	$\mathbf{A}_{m \times n} = [\mathbf{a}_1 \ldots \mathbf{a}_n]$
linear equation	$\mathbf{A}\mathbf{x} = \mathbf{b}$
inner product	$< \mathbf{x}, \mathbf{y} >= \mathbf{x}^T \mathbf{y}$

$$\|\mathbf{x}\| = \sqrt{\mathbf{x}^T \mathbf{x}}$$

$$\mathbf{x}^T \mathbf{y} = \|\mathbf{x}\| \|\mathbf{y}\| \cos(\theta)$$

Notations: Vectors and Matrices (cont.)

Notations: Vectors and Matrices (cont.)

$$\|\mathbf{x}\|_p = \left(|x_1|^p + |x_2|^p + \dots\right)^{1/p}$$

Notations: Vectors and Matrices (cont.)

$$\|\mathbf{x}\|_p = \left(|x_1|^p + |x_2|^p + \dots\right)^{1/p}$$

$$\|\mathbf{x}\|_p = \max\{|x_1|, |x_2|, \dots\} \qquad p = \infty$$

Notations: Vectors and Matrices (cont.)

$$\|\mathbf{x}\|_p = \left(|x_1|^p + |x_2|^p + \dots\right)^{1/p}$$

$$\|\mathbf{x}\|_p = \max\{|x_1|, |x_2|, \dots\} \qquad p = \infty$$

$$L_1, L_2, L_p, L_\infty$$

Notations: Vectors and Matrices (cont.)

$$\|\mathbf{x}\|_p = (|x_1|^p + |x_2|^p + \dots)^{1/p}$$

$$\|\mathbf{x}\|_p = \max\{|x_1|, |x_2|, \dots\} \qquad p = \infty$$

$$L_1, L_2, L_p, L_\infty$$

range $\mathcal{R}(\mathbf{A}) = \{\mathbf{A}\mathbf{x} : \mathbf{x} \in \mathbb{R}^n\}$

Notations: Vectors and Matrices (cont.)

$$\|\mathbf{x}\|_p = (|x_1|^p + |x_2|^p + \dots)^{1/p}$$

$$\|\mathbf{x}\|_p = \max\{|x_1|, |x_2|, \dots\} \qquad p = \infty$$

$$L_1, L_2, L_p, L_\infty$$

range $\mathcal{R}(\mathbf{A}) = \{\mathbf{Ax} : \mathbf{x} \in \mathbb{R}^n\}$

null space $\mathcal{N}(\mathbf{A}) = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{Ax} = 0\}$

Eigenvectors and Eigenvalues

$$\mathbf{y} = \mathbf{A}\mathbf{x}$$

Eigenvectors and Eigenvalues

$$\mathbf{y} = \mathbf{A}\mathbf{x}$$

$$\mathbf{A}\mathbf{e}_i = \lambda_i \mathbf{e}_i$$

Eigenvectors and Eigenvalues

$$\mathbf{y} = \mathbf{A}\mathbf{x}$$

$$\mathbf{A}\mathbf{e}_i = \lambda_i \mathbf{e}_i$$

$$\mathbf{T} = [\mathbf{v}_1\ \mathbf{v}_2 \ldots]$$

$$\mathbf{T}^{-1} A \mathbf{T} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots)$$

Eigenvectors and Eigenvalues

- All eigenvalues of symmetric (real) matrices are real.
- Any real symmetric nxn matrix has a set of **n** mutually orthogonal eigenvectors.

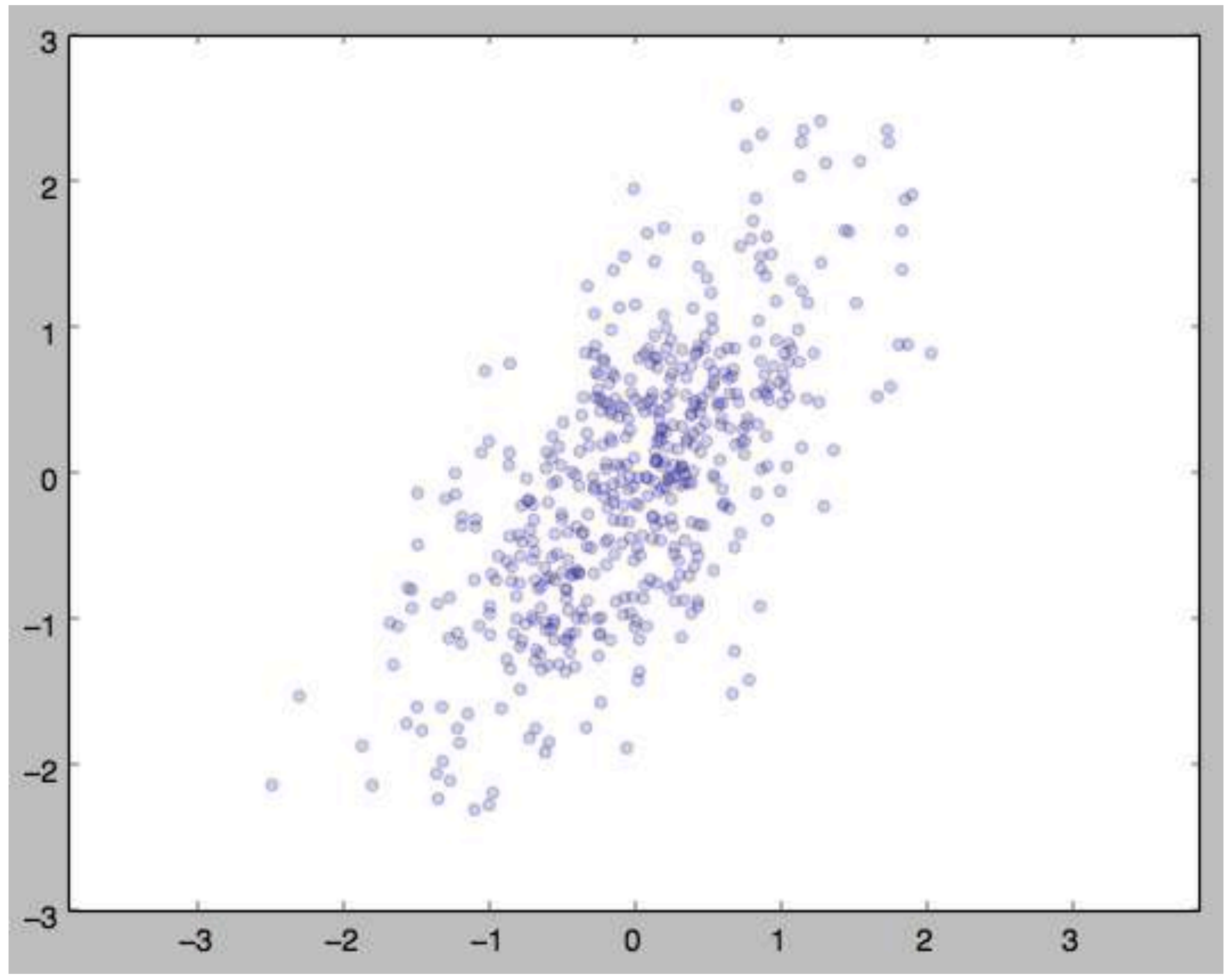
$$\mathbf{y} = \mathbf{A}\mathbf{x}$$

$$\mathbf{A}\mathbf{e}_i = \lambda_i \mathbf{e}_i$$

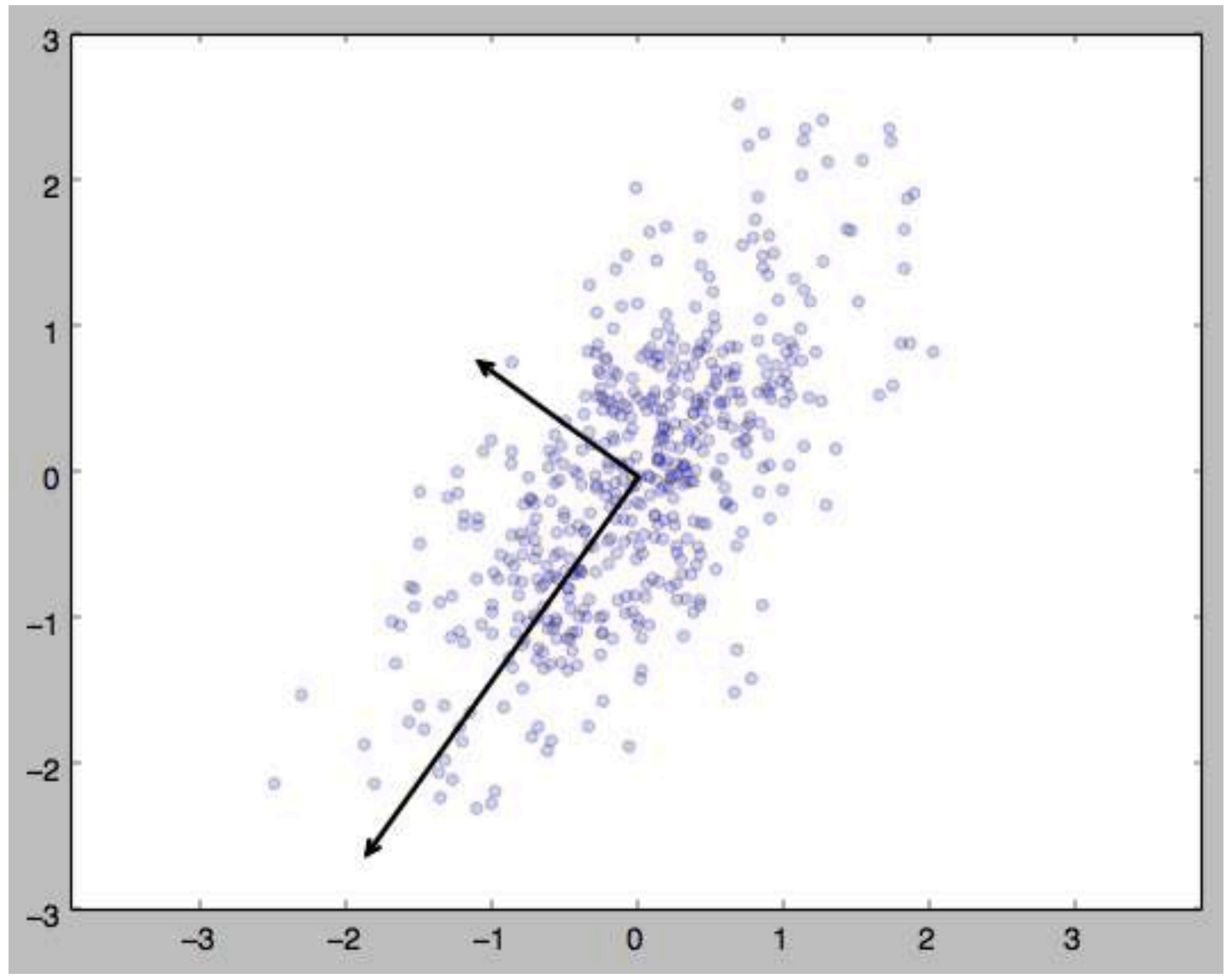
$$\mathbf{T} = [\mathbf{v}_1 \ \mathbf{v}_2 \ldots]$$

$$\mathbf{T}^{-1} A \mathbf{T} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots)$$

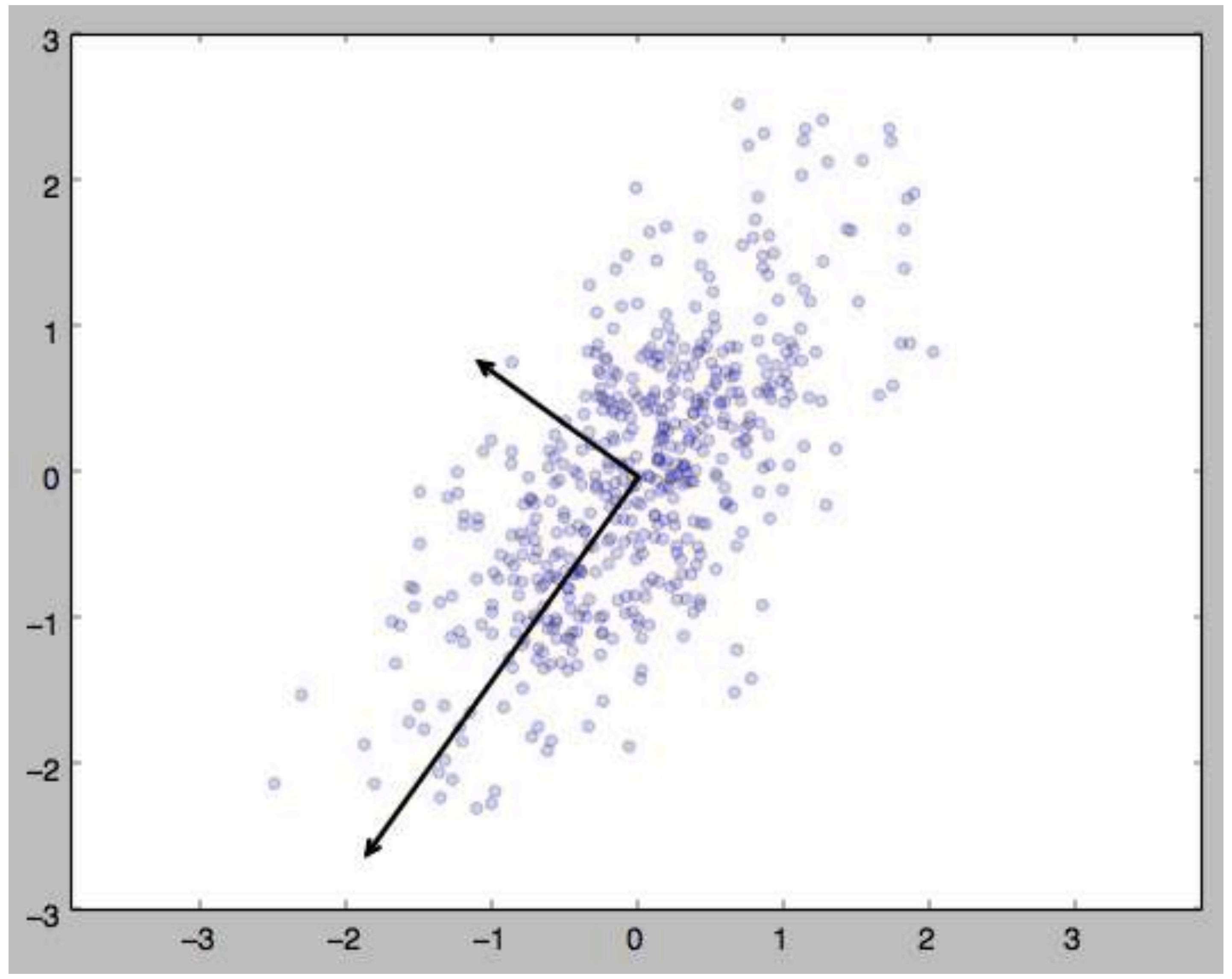
Code Example



Code Example



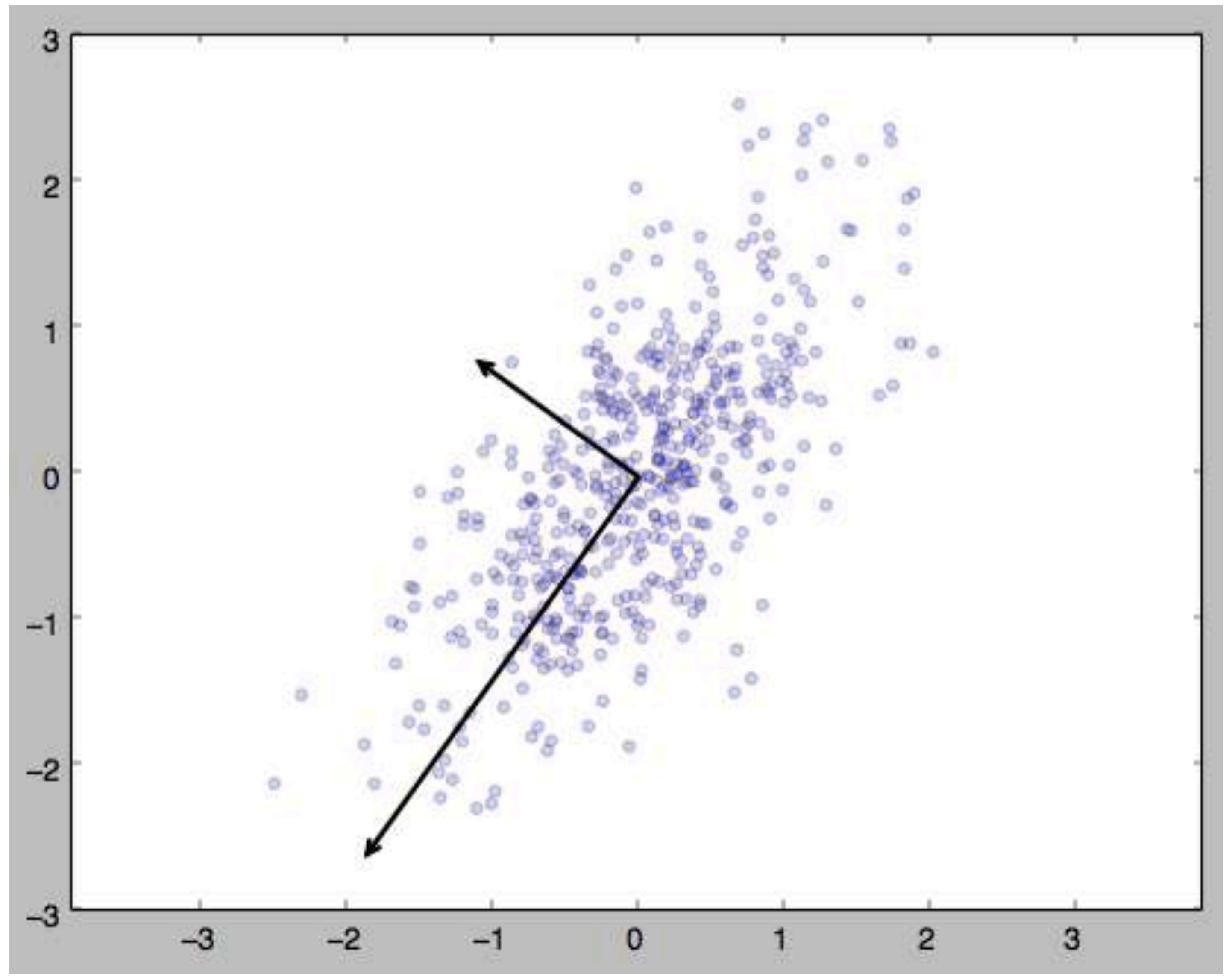
Code Example




```
rng = np.random.RandomState(10)
X = np.dot(rng.rand(2, 2), rng.randn(2, 500)).T

mean_vec = np.mean(X, axis=0)
cov_mat = (X - mean_vec).T.dot((X - mean_vec)) / (X.shape[0]-1)
eig_vals, eig_vecs = np.linalg.eig(cov_mat)
```

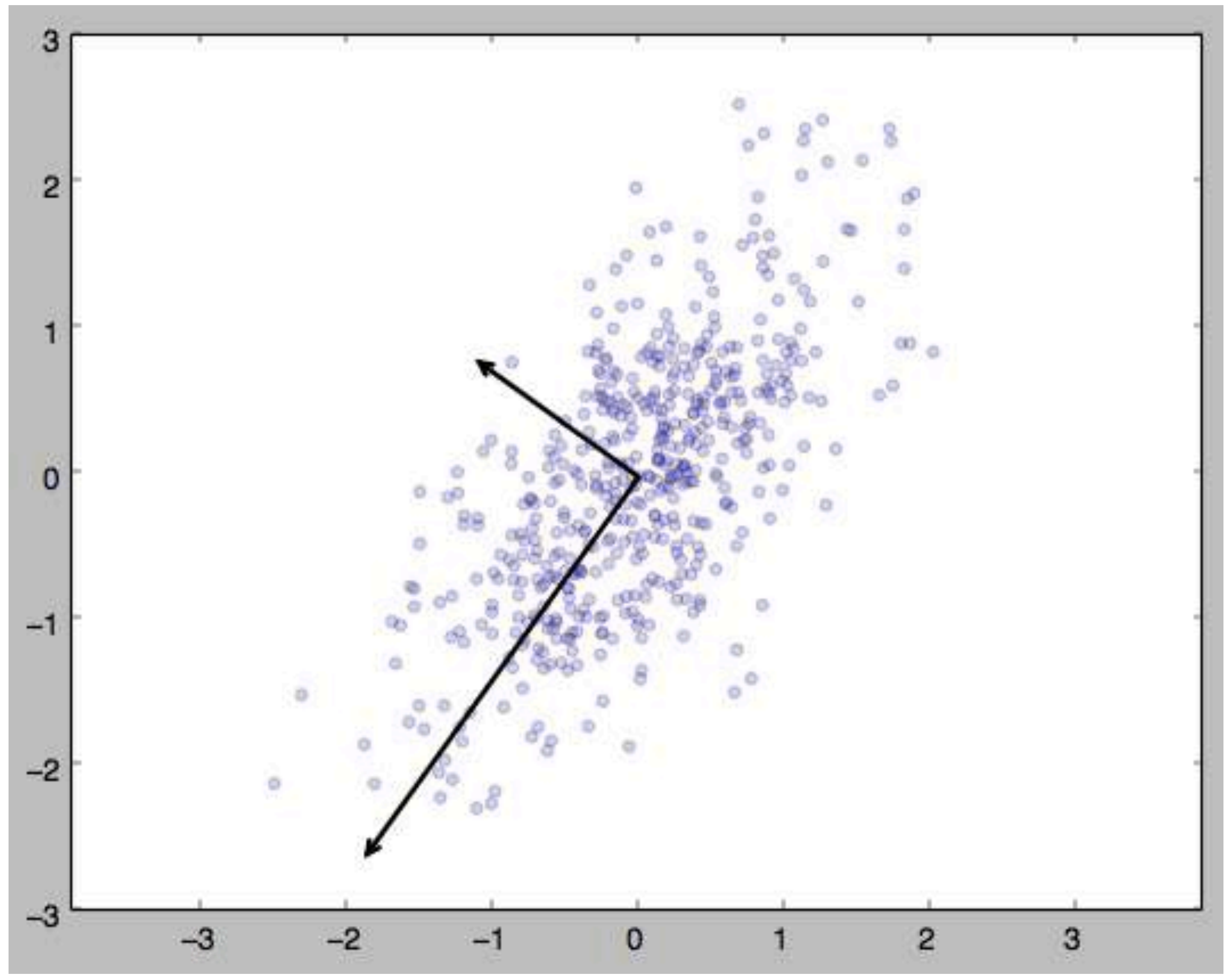
Code Example




```
rng = np.random.RandomState(10)
X = np.dot(rng.rand(2, 2), rng.randn(2, 500)).T

mean_vec = np.mean(X, axis=0)
cov_mat = (X - mean_vec).T.dot((X - mean_vec)) / (X.shape[0]-1)
eig_vals, eig_vecs = np.linalg.eig(cov_mat)
```

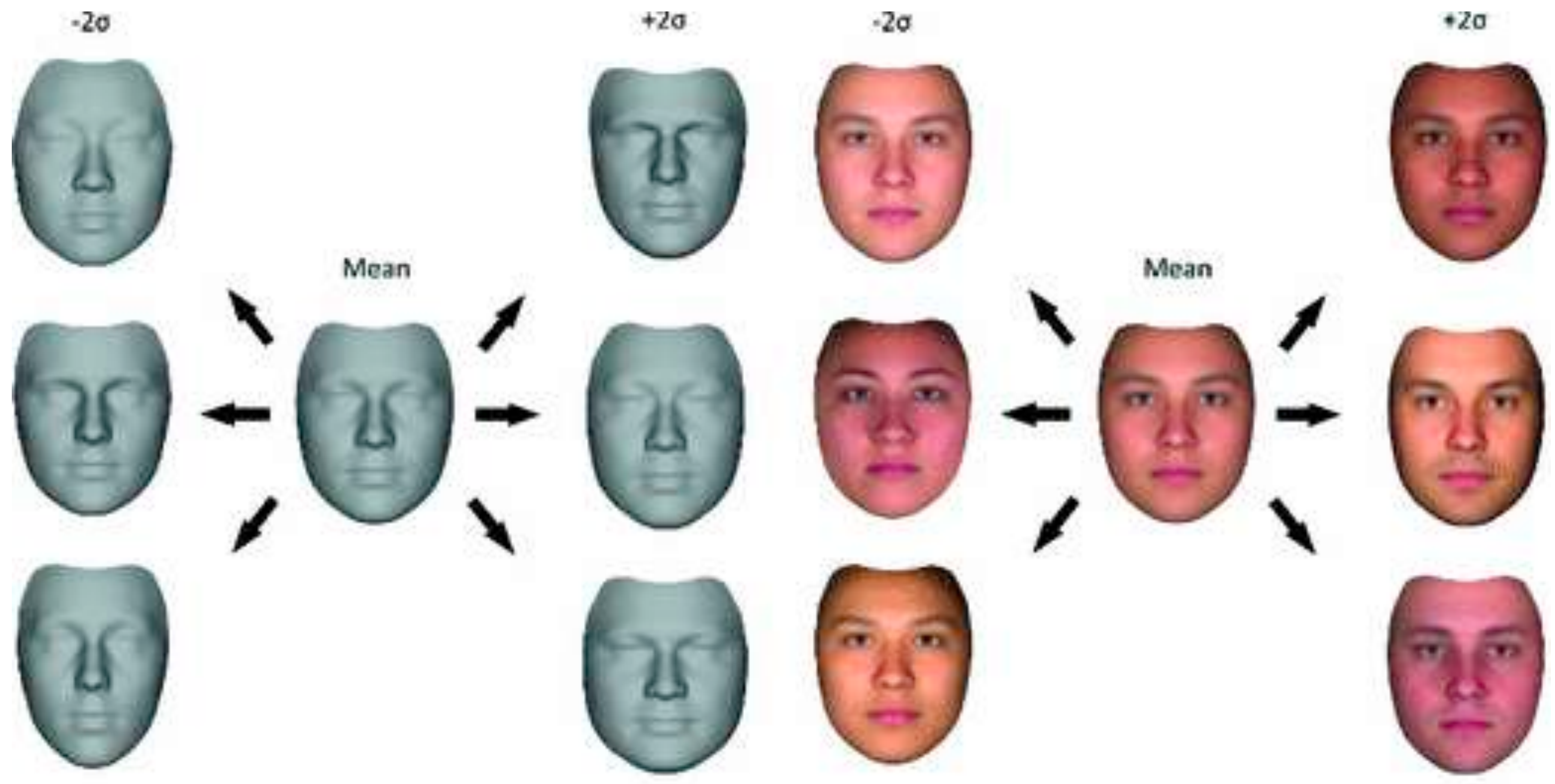
Code Example




```
rng = np.random.RandomState(10)
X = np.dot(rng.rand(2, 2), rng.randn(2, 500)).T

mean_vec = np.mean(X, axis=0)
cov_mat = (X - mean_vec).T.dot((X - mean_vec)) / (X.shape[0]-1)
eig_vals, eig_vecs = np.linalg.eig(cov_mat)
```

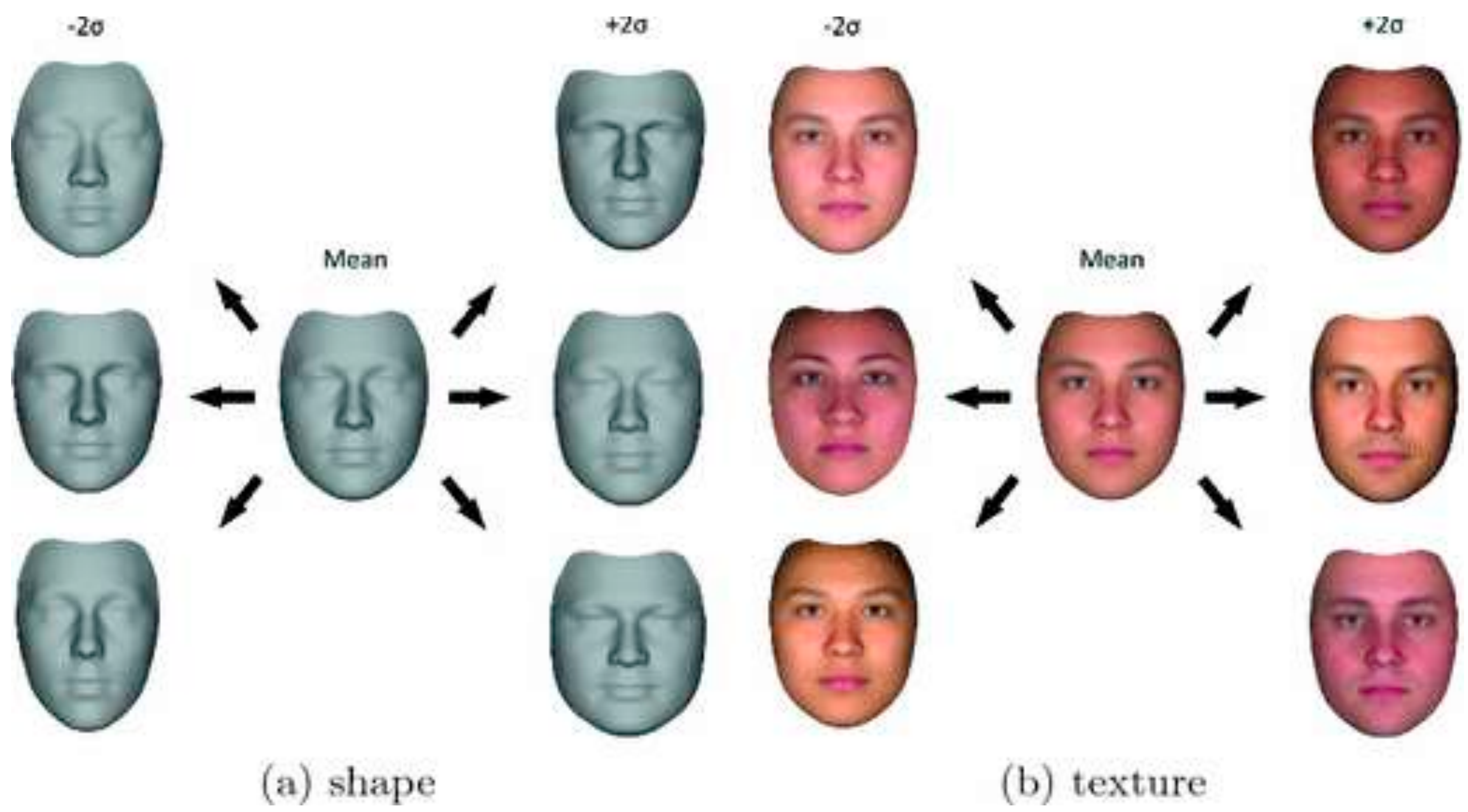
Morphable Faces



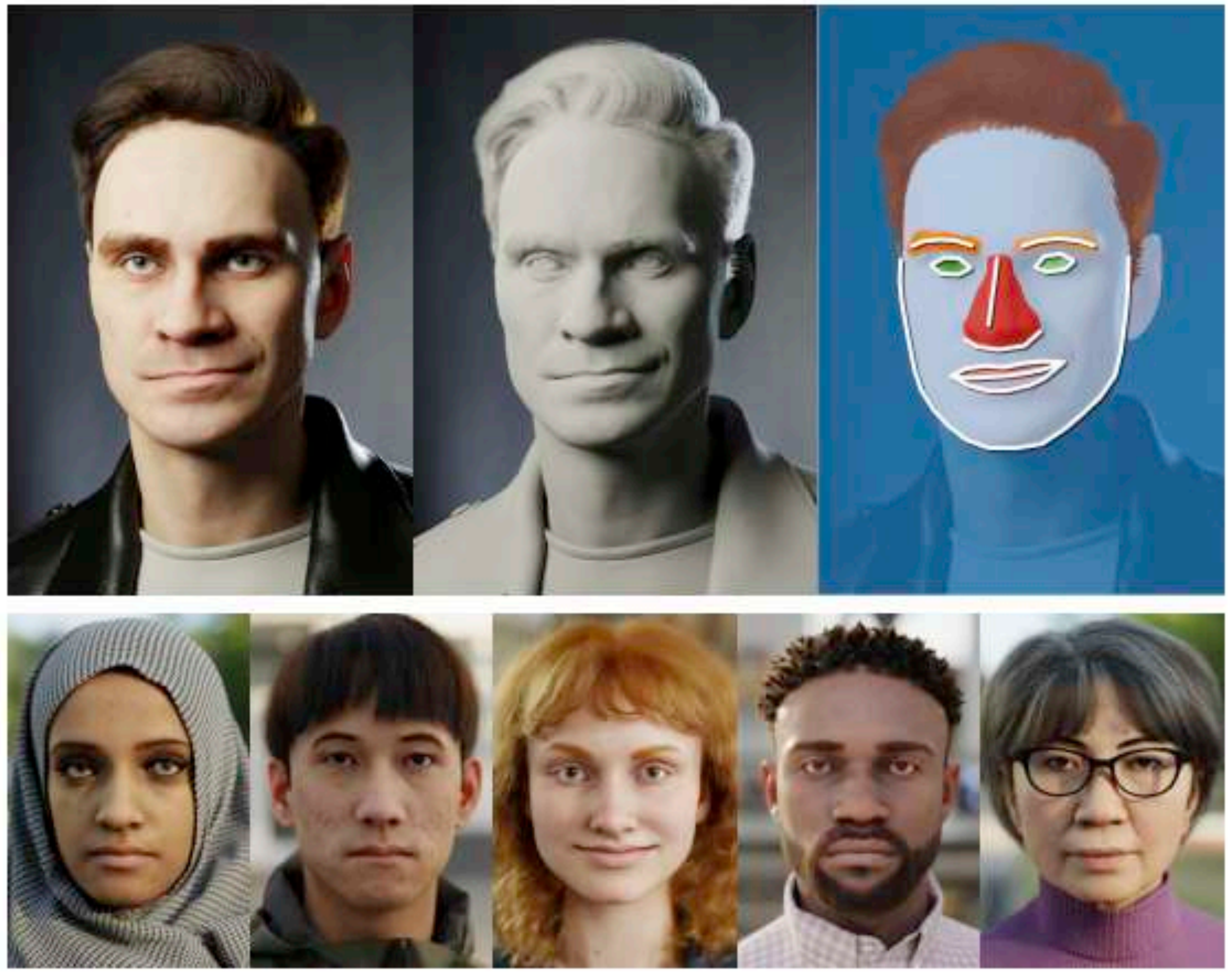
(a) shape (b) texture

Video at http://gravis.dmi.unibas.ch/Sigg99.html

Morphable Faces



Video at http://gravis.dmi.unibas.ch/Sigg99.html



Wood et al., 2021 Arxiv,
Fake it till you make it: face analysis in the wild using synthetic data alone

Singular Value Decomposition (SVD)

- Very useful for matrix manipulation
- Used for robust numerical computation

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$$

特征值分解

$$\mathbf{A} = \mathbf{A}^T = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^T$$

Singular Value Decomposition (SVD)

- Very useful for matrix manipulation
- Used for robust numerical computation

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$$

rotation

$$\mathbf{A} = \mathbf{A}^T = \mathbf{U}\boldsymbol{\Sigma}\mathbf{U}^T$$

Singular Value Decomposition (SVD)

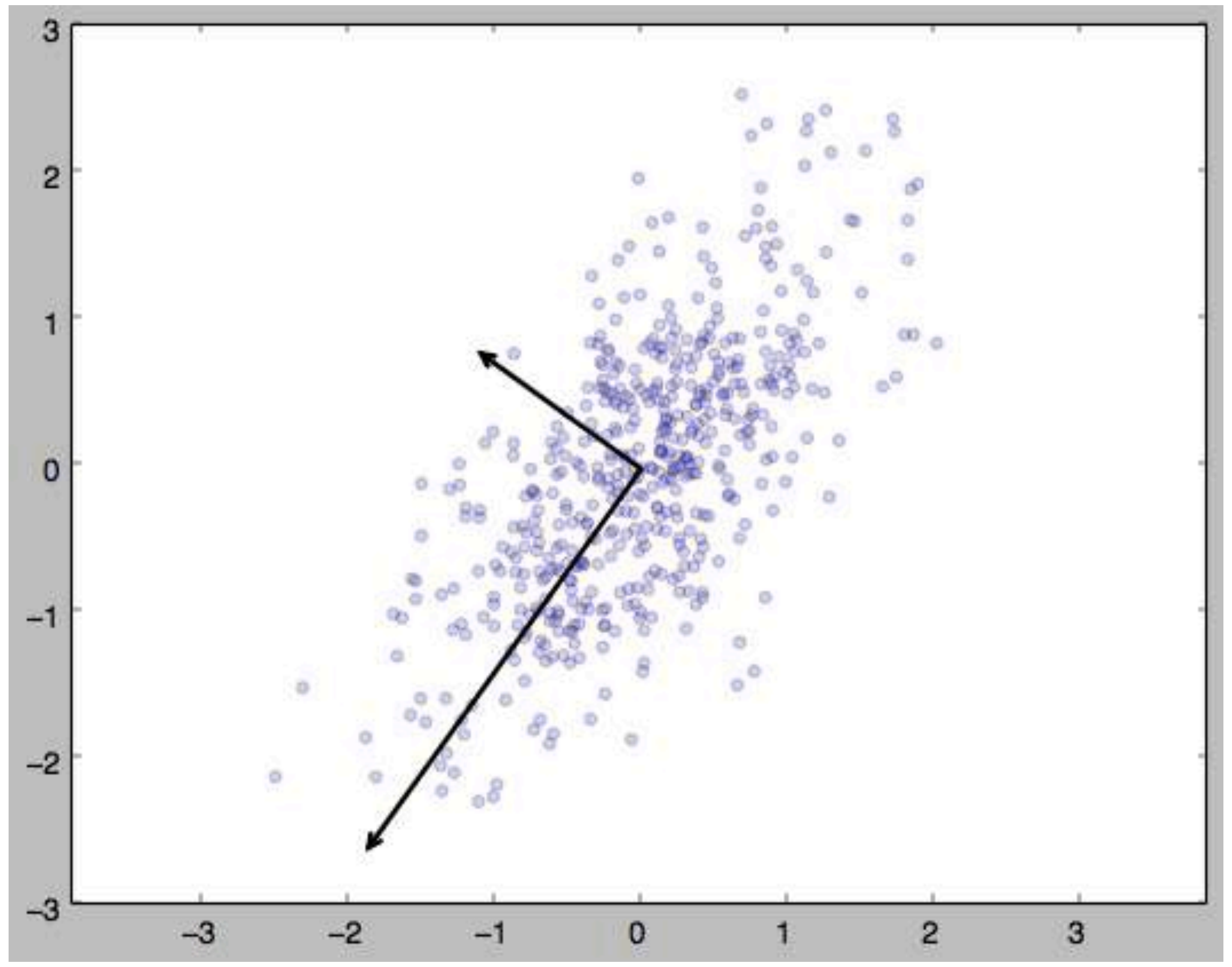
- Very useful for matrix manipulation
- Used for robust numerical computation

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$$

scaling rotation

$$\mathbf{A} = \mathbf{A}^T = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^T$$

Code Example



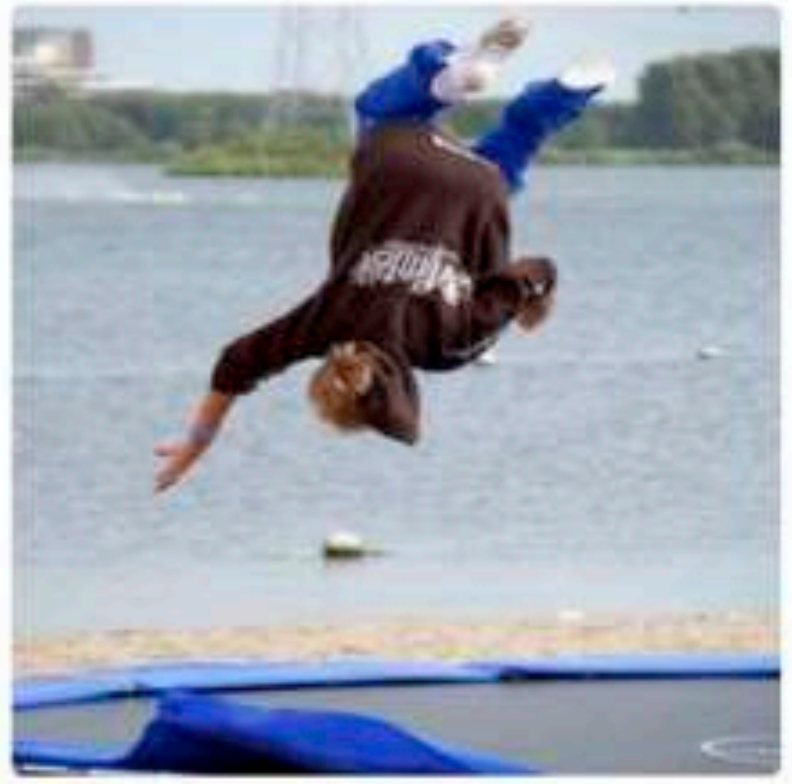

```
mean_vec = np.mean(X, axis=0)
cov_mat = (X - mean_vec).T.dot((X - mean_vec)) / (X.shape[0]-1)
matU, sigma, matV = np.linalg.svd(cov_mat)
```

Image Captioning

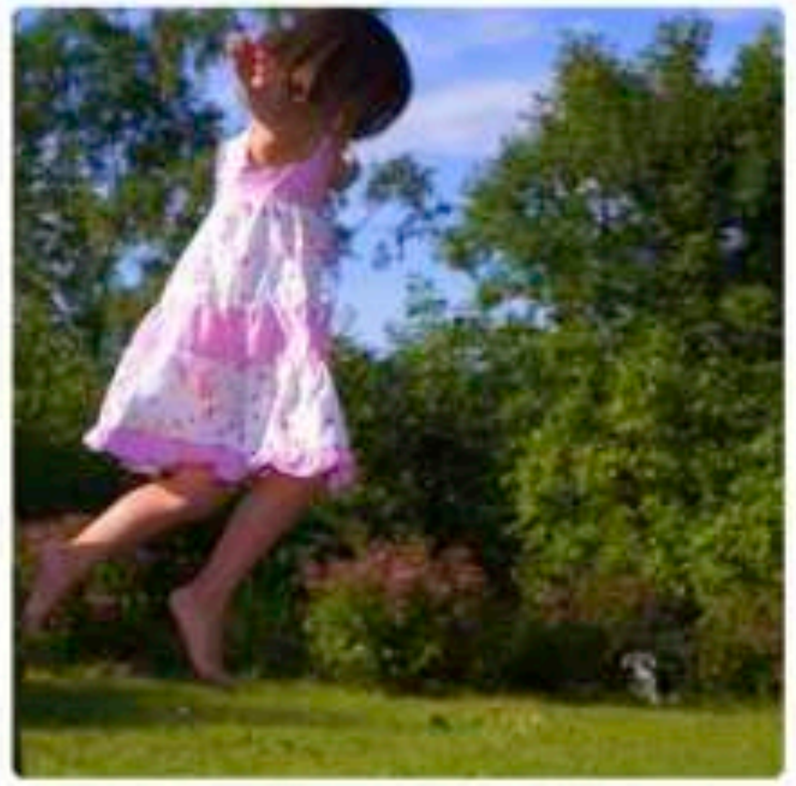
"man in black shirt is playing guitar."

"construction worker in orange safety vest is working on road."

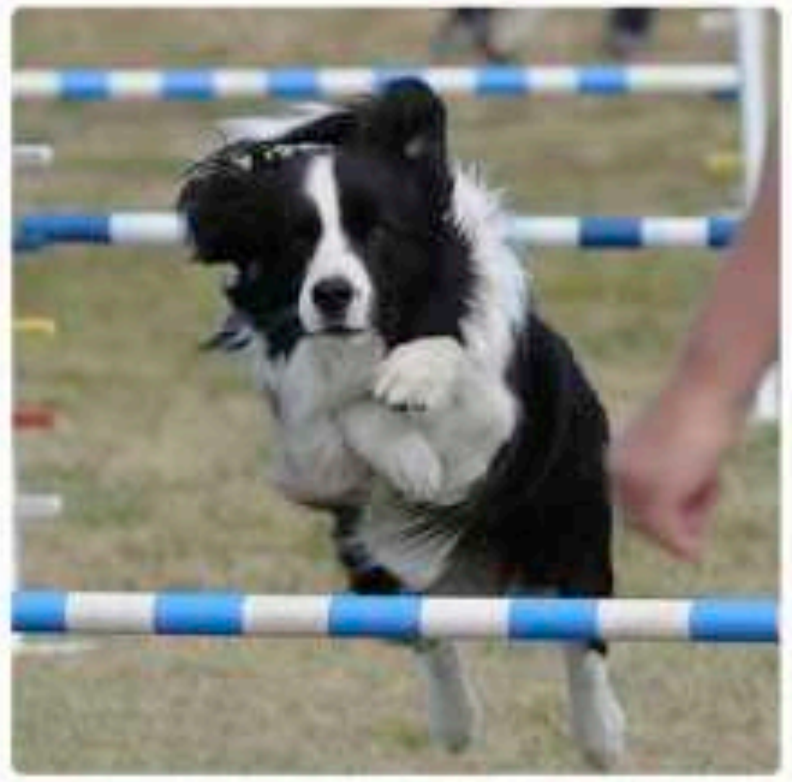
"two young girls are playing with lego toy."



"boy is doing backflip on wakeboard."



"girl in pink dress is jumping in air."



"black and white dog jumps over bar."

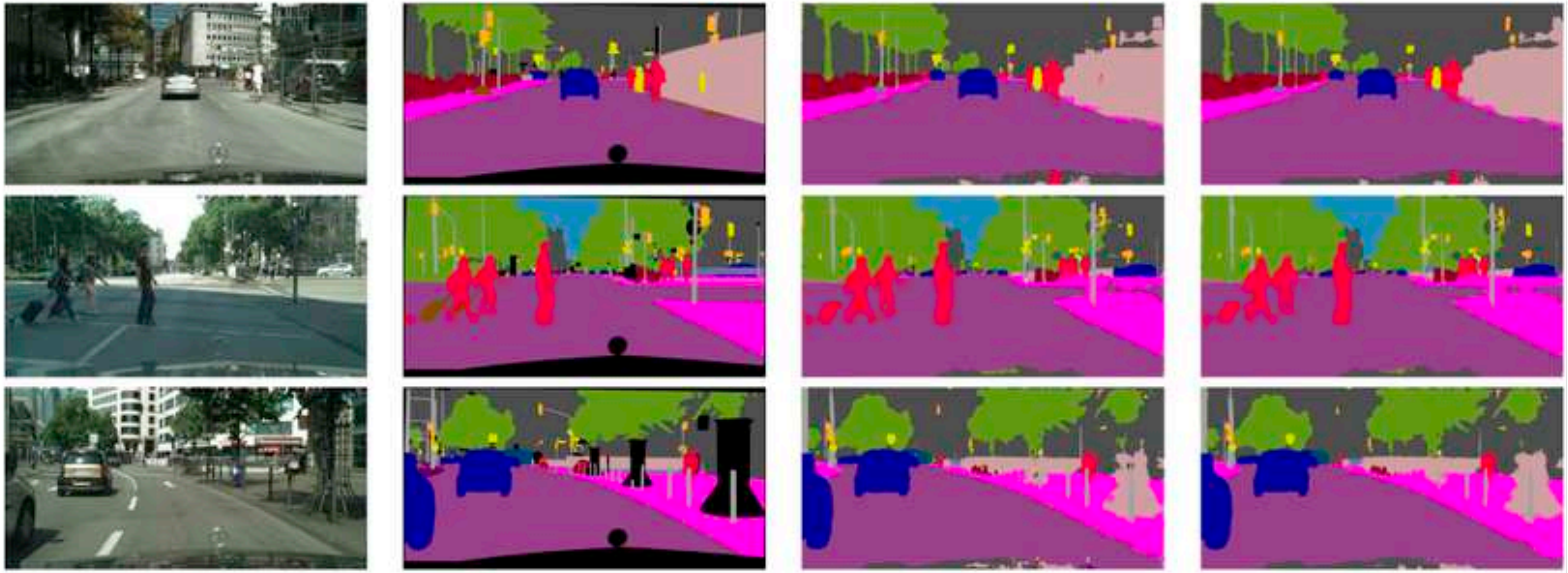
"young girl in pink shirt is swinging on swing."

"man in blue wetsuit is surfing on wave."

Andrej Karpathy, Fei-Fei Li

Semantic Segmentation

⇒ diff colour ⇒
diff stuff



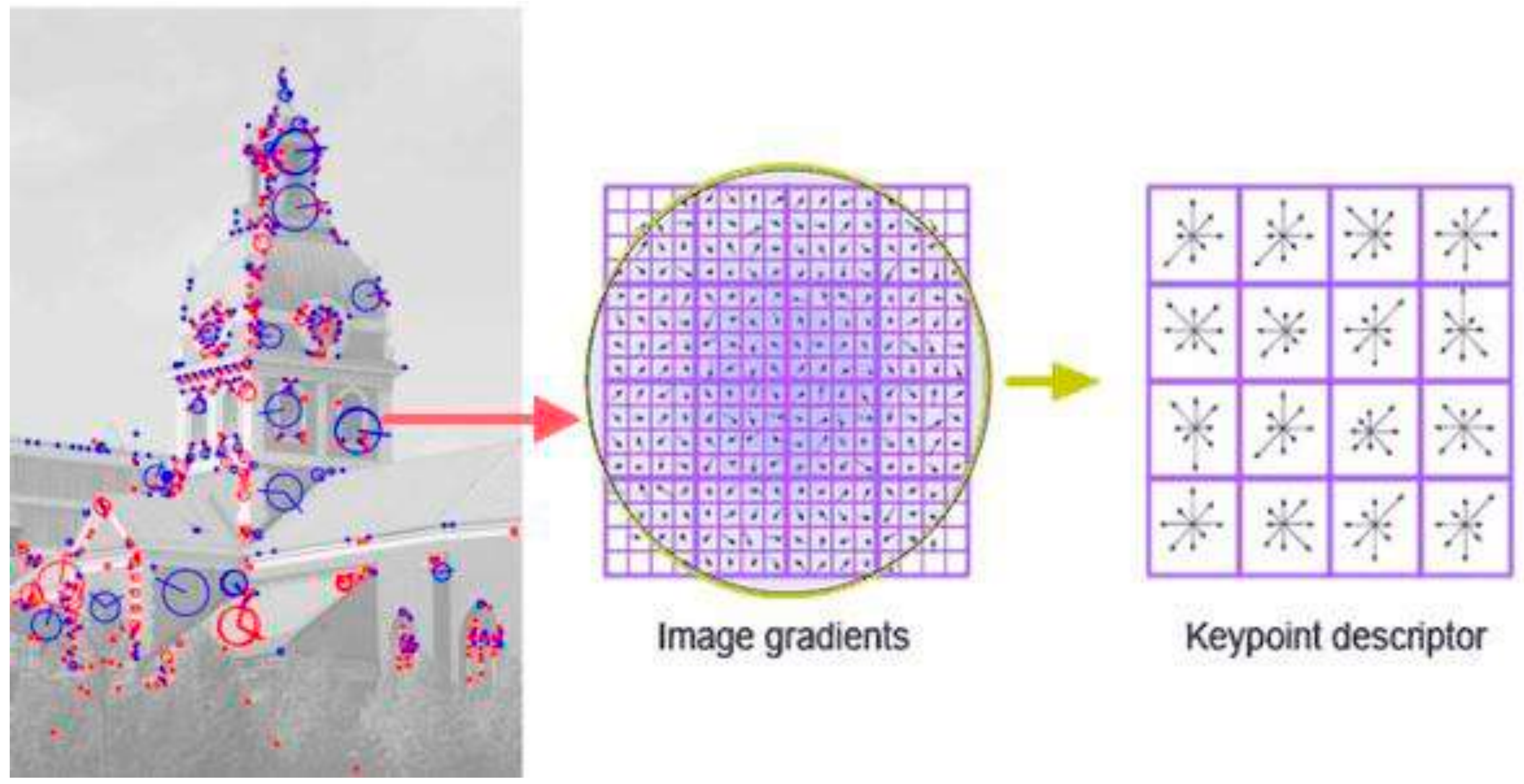
B. Sahu

3D Content Creation

Kelly, Guerrero, Steed, Wonka, Mitra

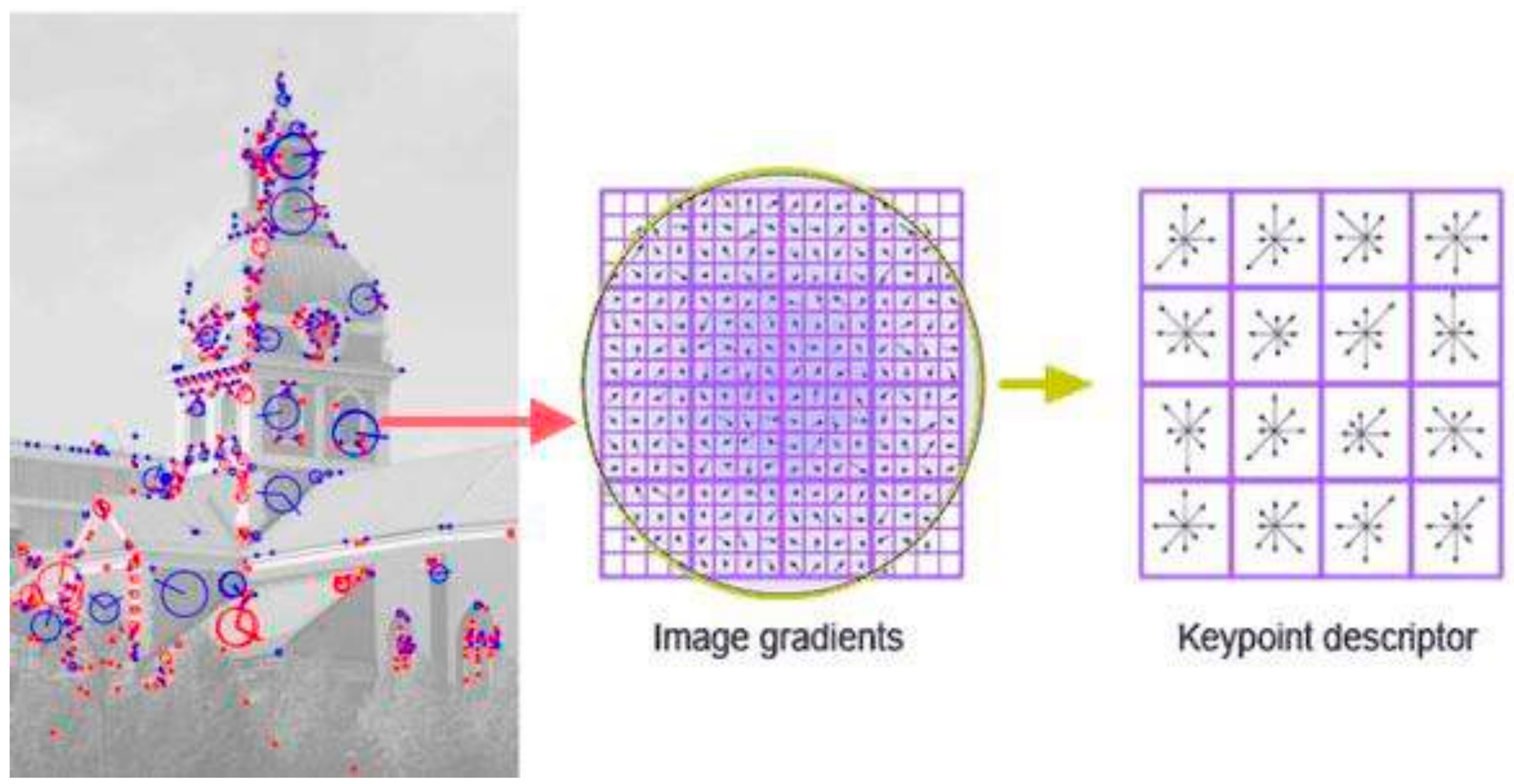
Features: Hand-crafted versus Data-driven

Features: Hand-crafted versus Data-driven

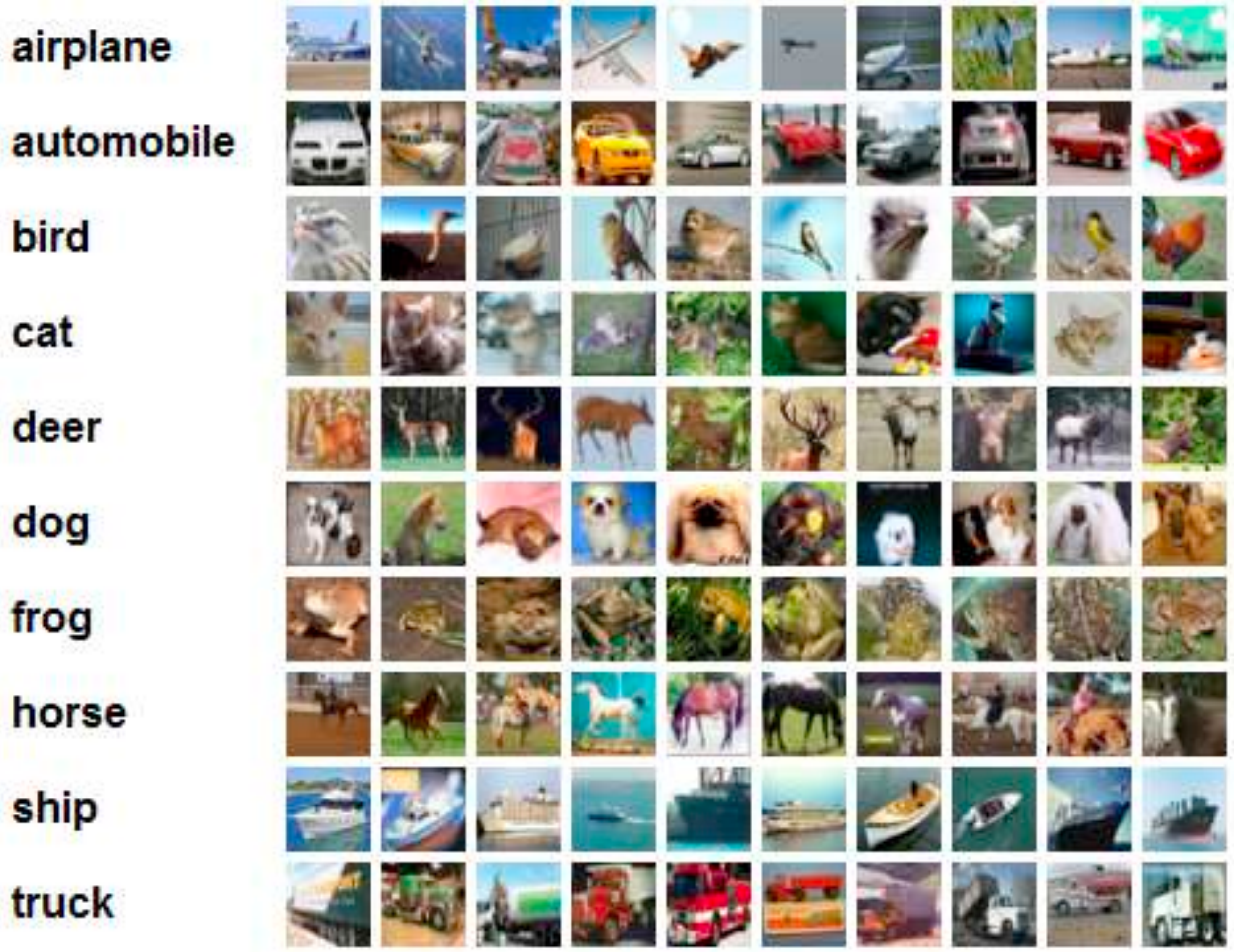


HOG or SIFT Features

Features: Hand-crafted versus Data-driven



HOG or SIFT Features



CIFAR10 dataset

Data-driven Algorithms (Supervised)

Labelled data
(supervision data)

Data-driven Algorithms (Supervised)

Labelled data
(supervision data)

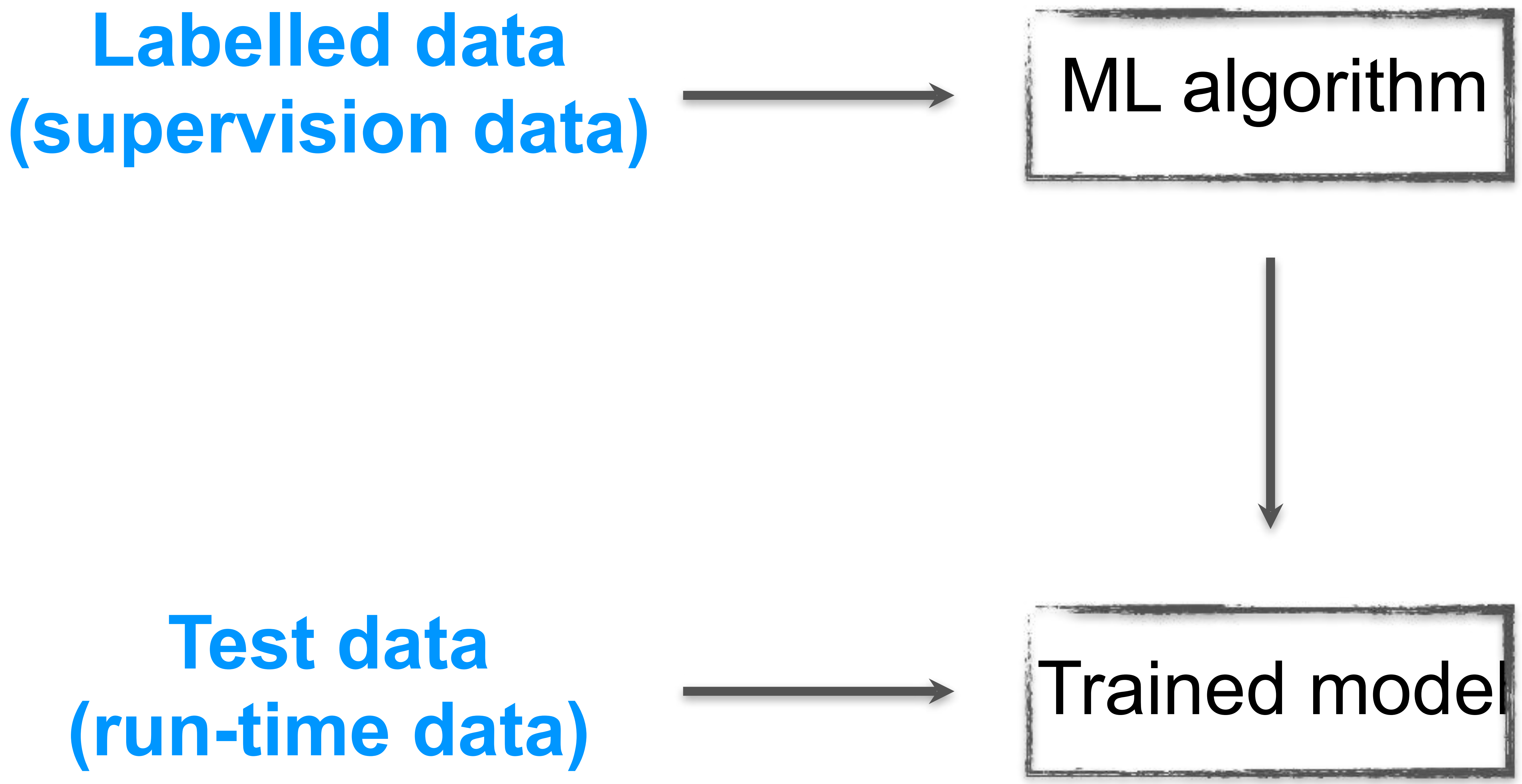
→ ML algorithm

↓

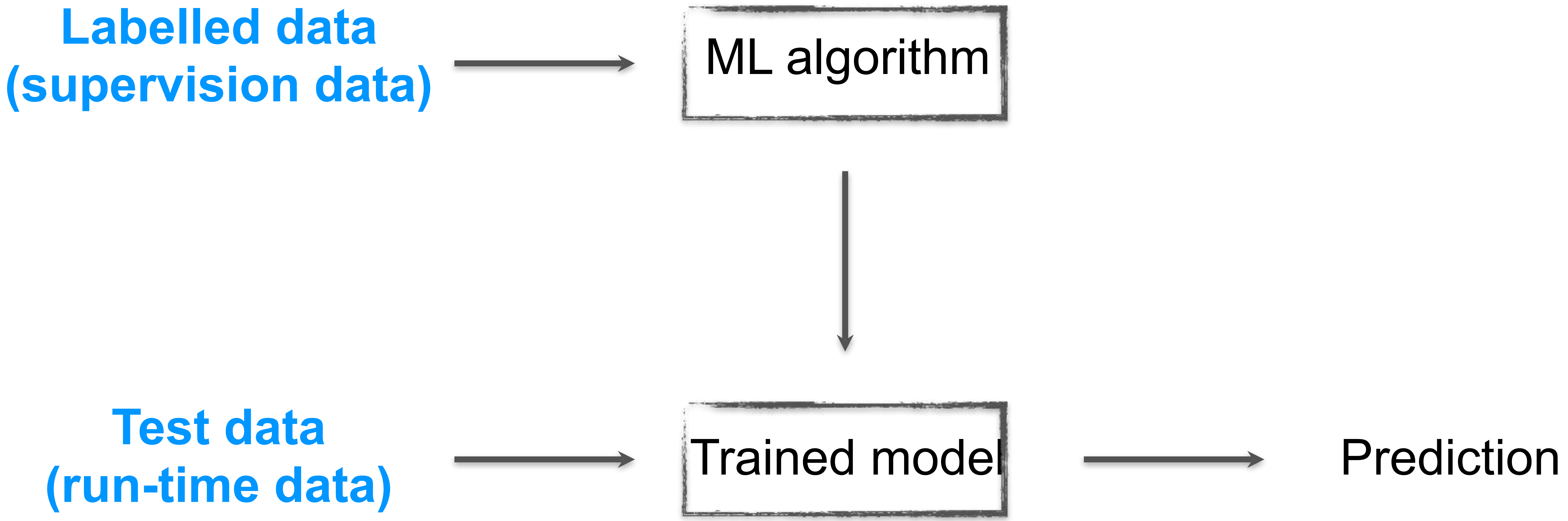
Trained model

UCL

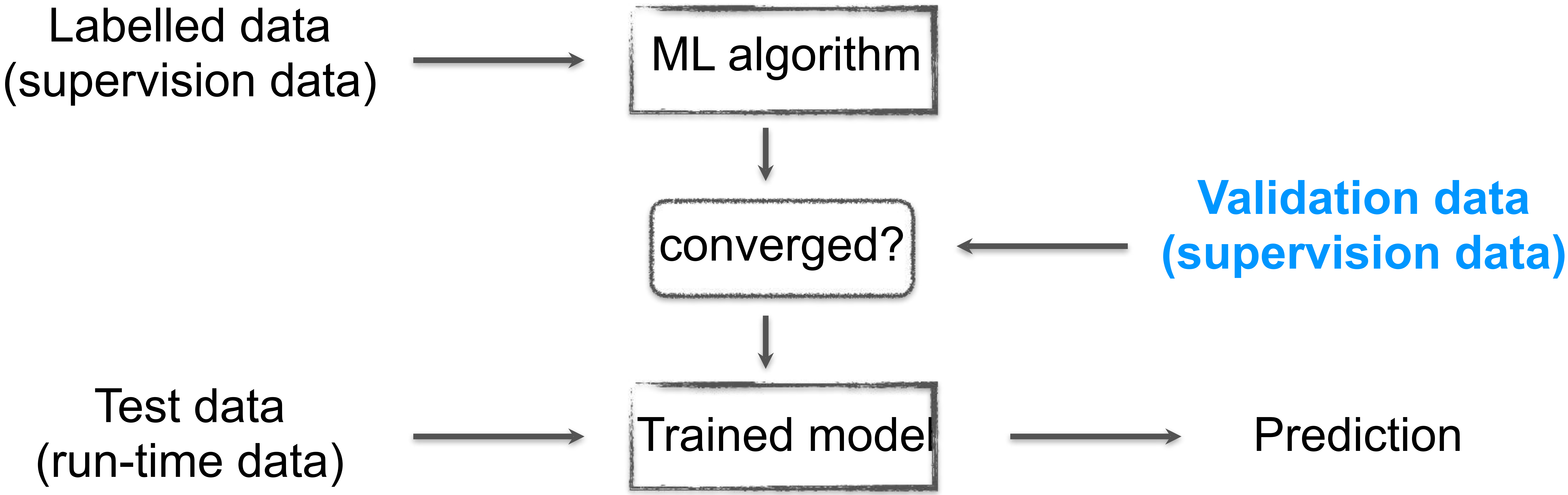
Data-driven Algorithms (Supervised)



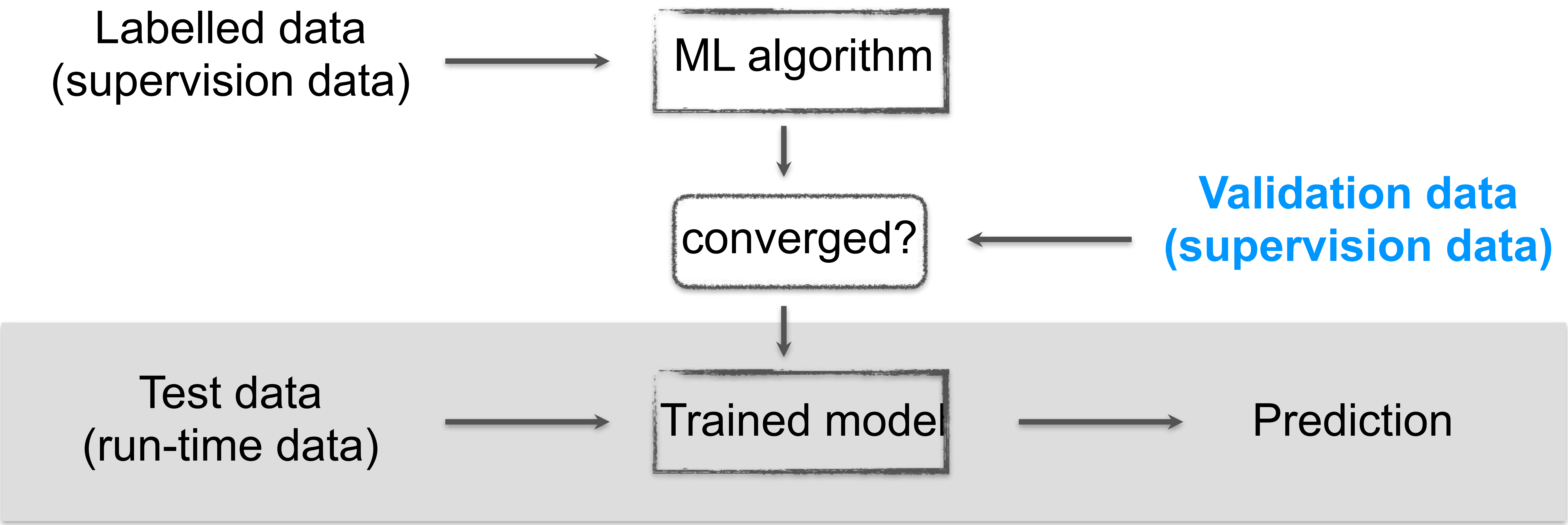
Data-driven Algorithms (Supervised)



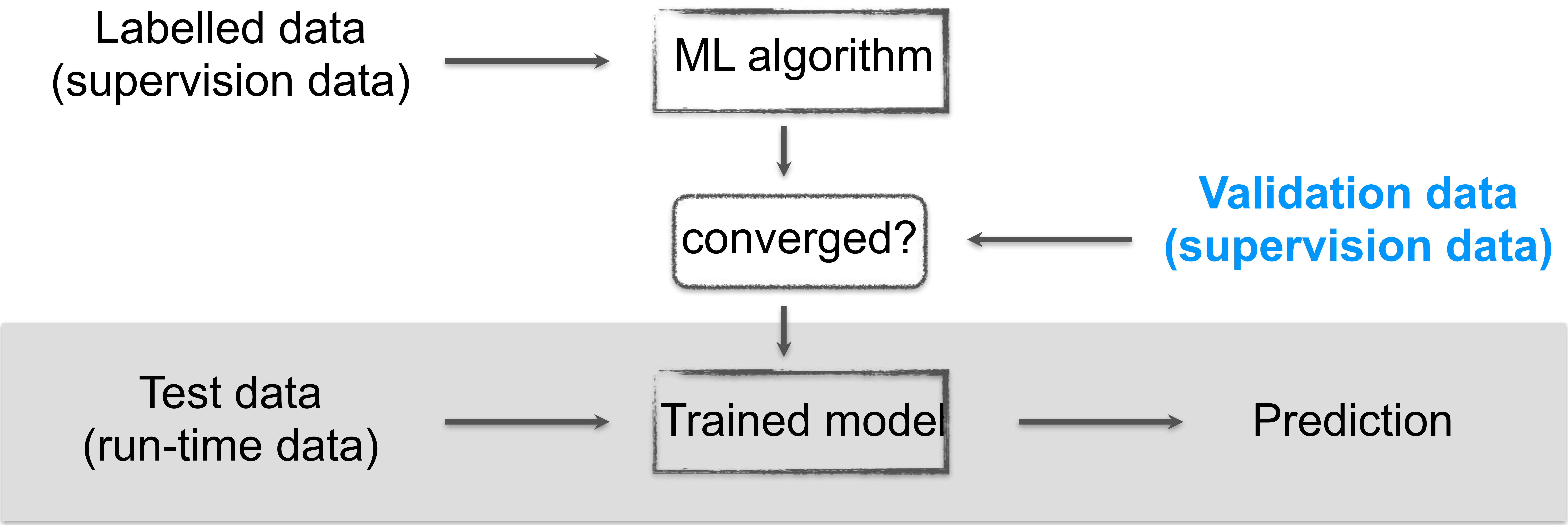
Data-driven Algorithms (Supervised)



Data-driven Algorithms (Supervised)

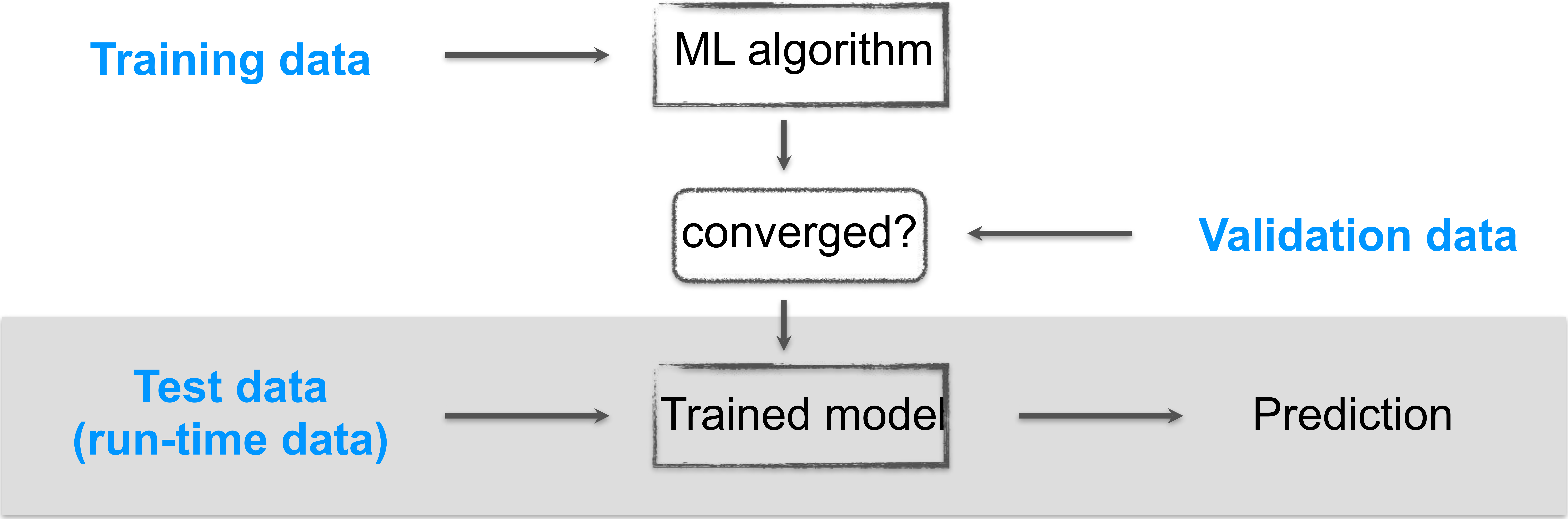


Data-driven Algorithms (Supervised)



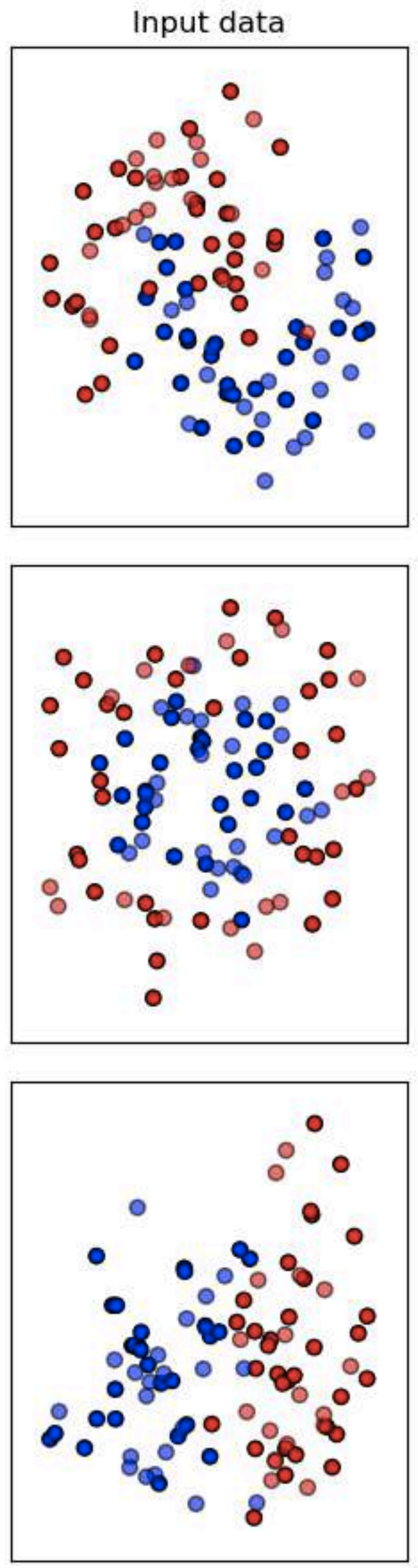
Implementation details: Training: 70%; Validation: 15%; Test 15%

Data-driven Algorithms (Unsupervised)



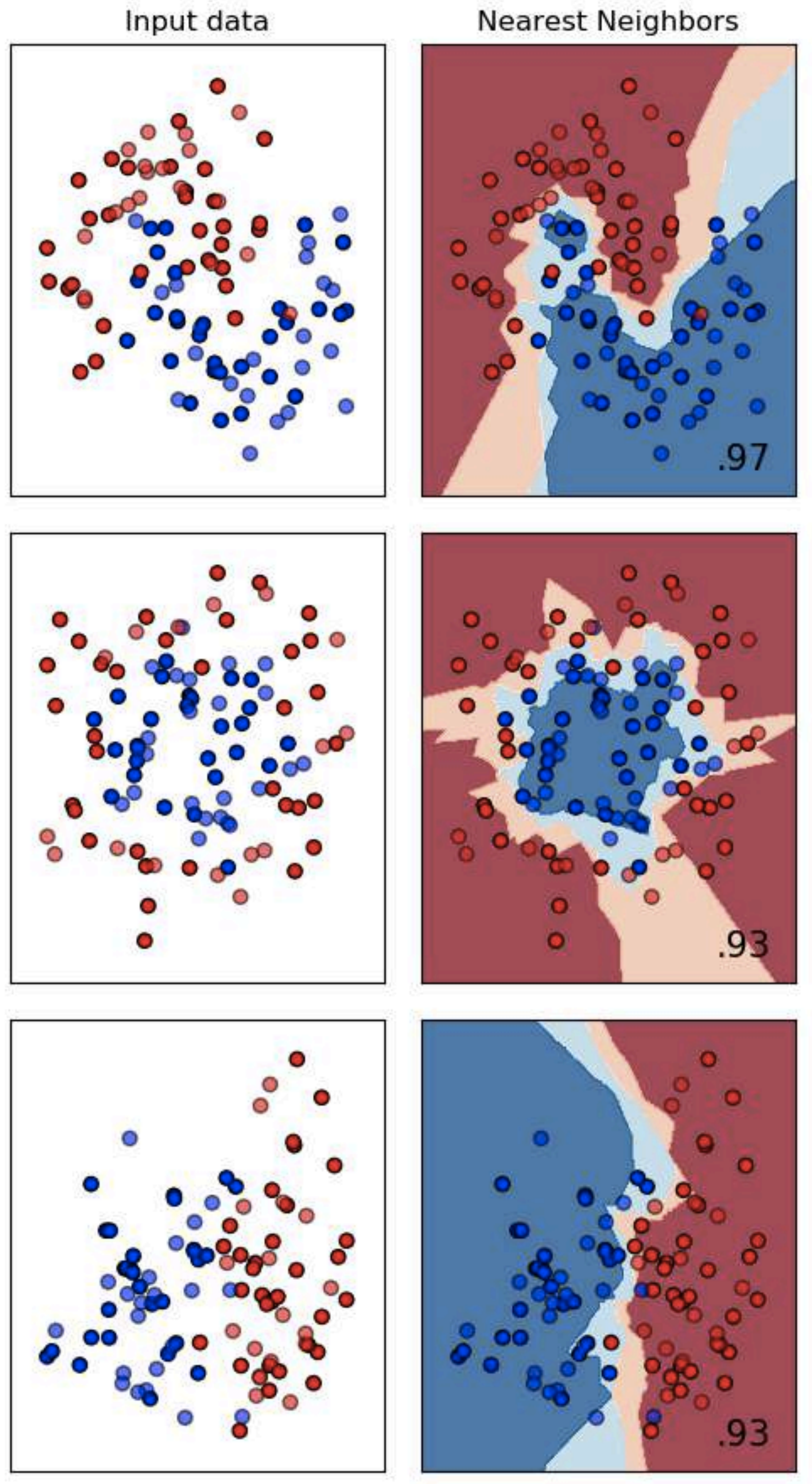
Implementation details: Training: 70%; Validation: 15%; Test 15%

Various ML Approaches (Supervised approaches)



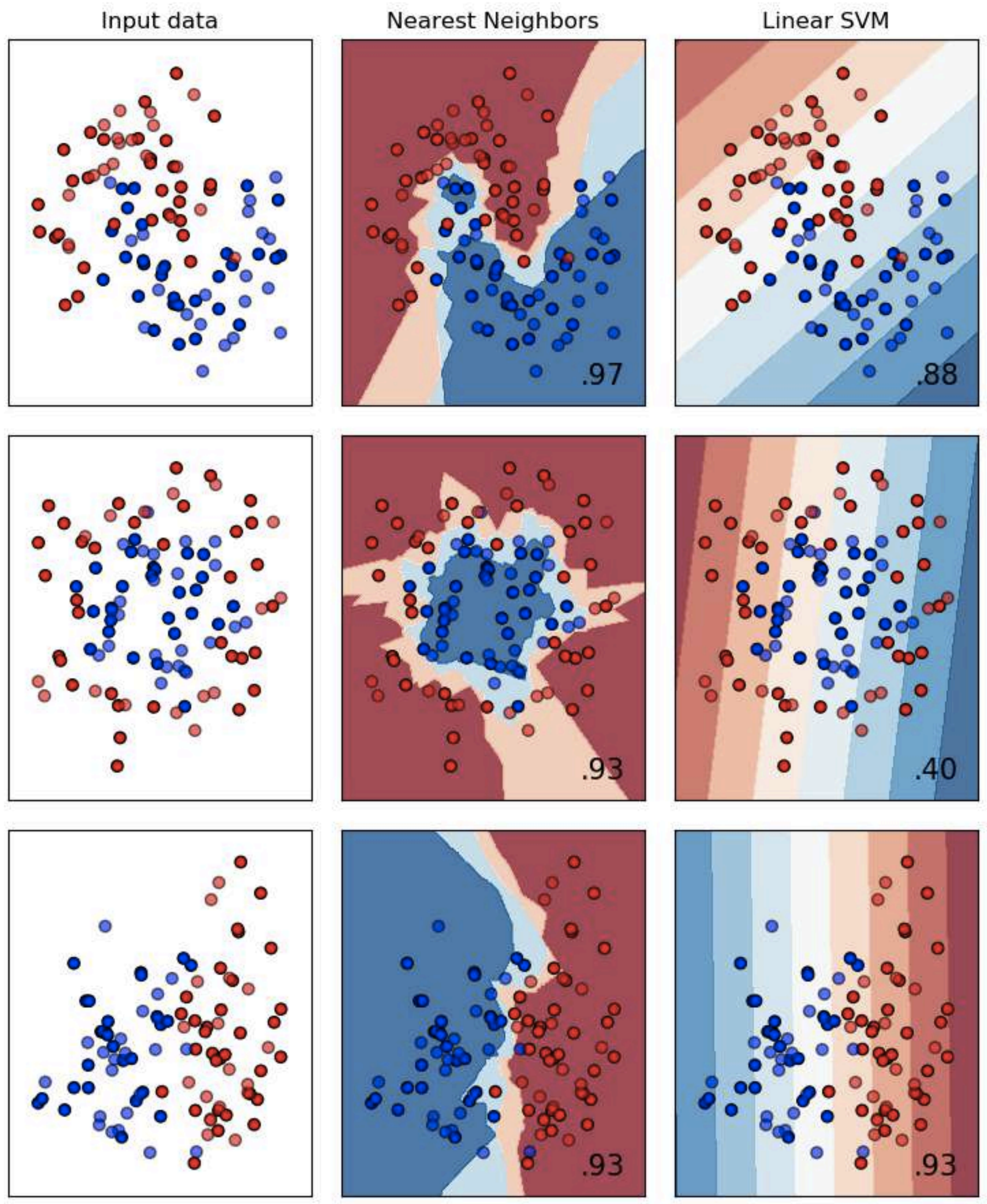
http://scikit-learn.org/stable/auto_examples/classification/plot_classifier_comparison.html

Various ML Approaches (Supervised approaches)



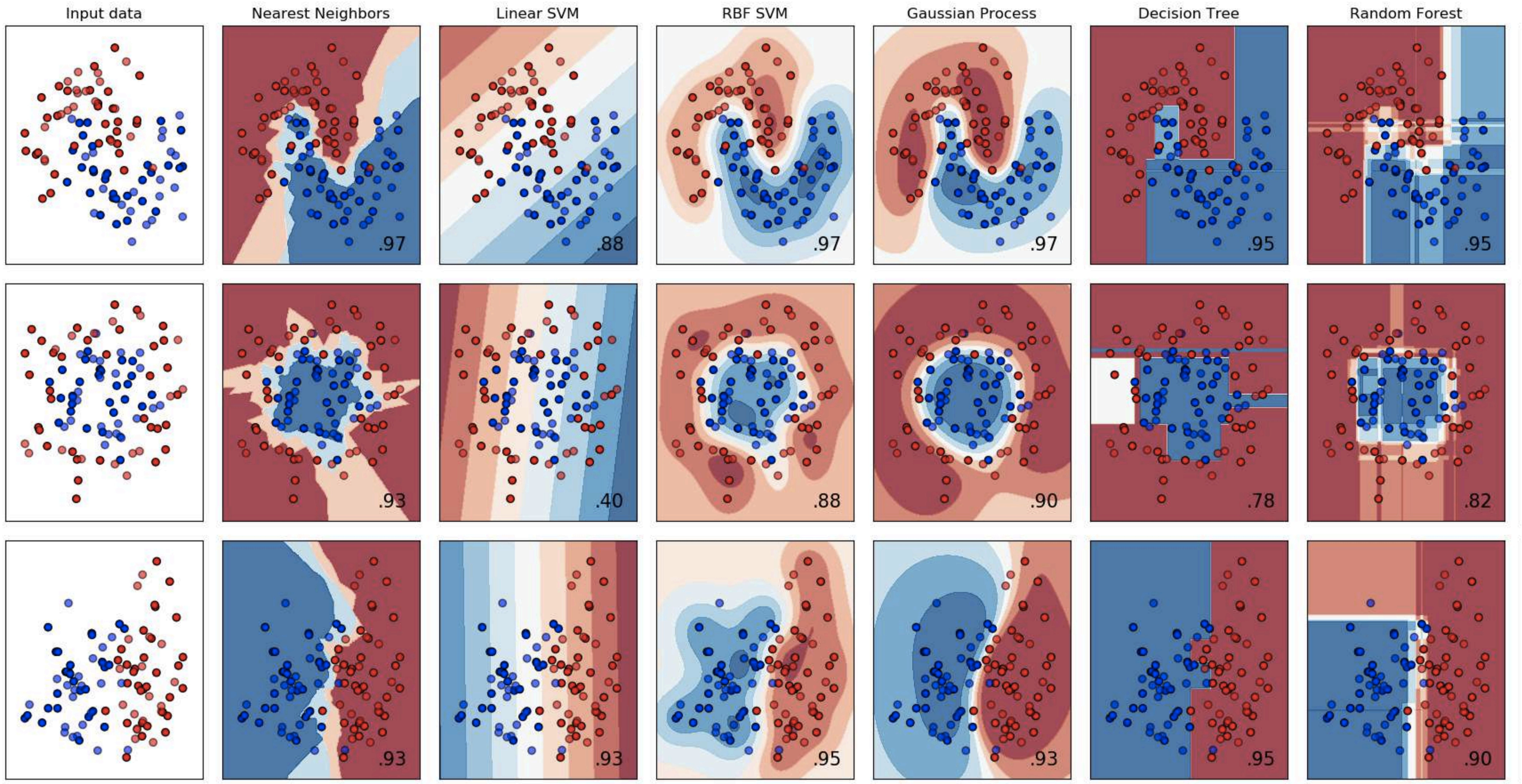
http://scikit-learn.org/stable/auto_examples/classification/plot_classifier_comparison.html

Various ML Approaches (Supervised approaches)



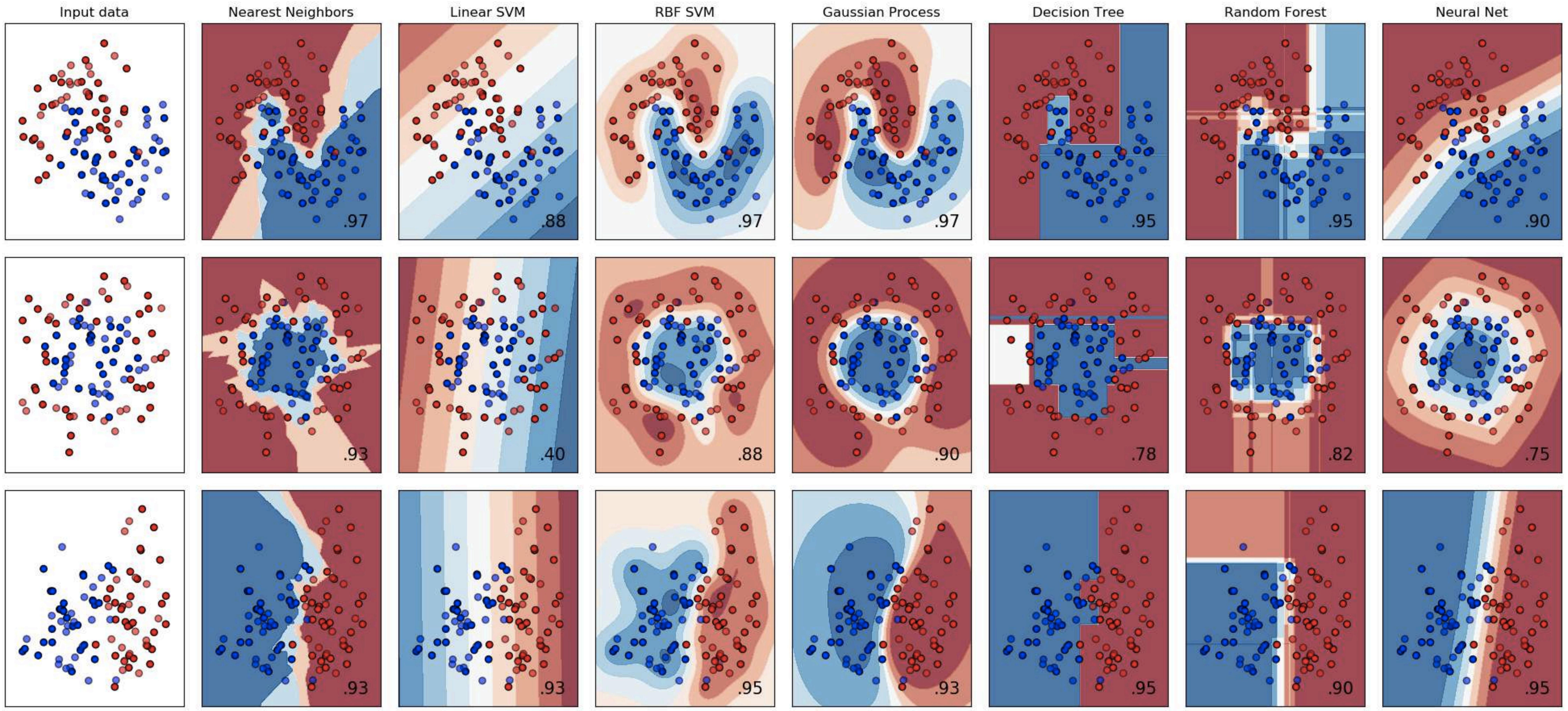
http://scikit-learn.org/stable/auto_examples/classification/plot_classifier_comparison.html

Various ML Approaches (Supervised approaches)



http://scikit-learn.org/stable/auto_examples/classification/plot_classifier_comparison.html

Various ML Approaches (Supervised approaches)



http://scikit-learn.org/stable/auto_examples/classification/plot_classifier_comparison.html

Rise of Learning

- 1958: Perceptron
- 1974: Backpropagation
- 1981: Hubel & Wiesel wins Nobel prize for ‘visual system’
- 1990s: SVM era
- 1998: CNN used for handwriting analysis
- **2012: AlexNet wins ImageNet**

What is Special about Graphics?

What is Special about Graphics?

- **Image Processing** (image translation tasks)

- Many sources of input data — **model building** (e.g., images, scanners, motion capture)

What is Special about Graphics?

- **Image Processing** (image translation tasks)
- Many sources of input data — **model building** (e.g., images, scanners, motion capture)
- Many sources of **synthetic data** — can serve as supervision data (e.g., rendering, animation)

What is Special about Graphics?

- **Image Processing** (image translation tasks)
- Many sources of input data — **model building** (e.g., images, scanners, motion capture)
- Many sources of **synthetic data** — can serve as supervision data (e.g., rendering, animation)

What is Special about Graphics?

- **Image Processing** (image translation tasks)
- Many sources of input data — **model building** (e.g., images, scanners, motion capture)
- Many sources of **synthetic data** — can serve as supervision data (e.g., rendering, animation)
- Light transport models; geometric invariants

What is Special about Graphics?

- **Image Processing** (image translation tasks)
- Many sources of input data — **model building** (e.g., images, scanners, motion capture)
- Many sources of **synthetic data** — can serve as supervision data (e.g., rendering, animation)
- Light transport models; geometric invariants
- Many problems in **generative models**

Examples in Graphics

Geometry

Image manipulation

Animation

Rendering

Examples in Graphics

Geometry

Procedural modelling

Mesh segmentation

Learning deformations

Colorization

Sketch simplification

Image manipulation

Animation

Boxification

BRDF estimation

Fluid

Real-time rendering

Animation

Rendering

Denoising

Facial animation

PCD processing

Examples in Graphi

Sketch simplification

Colorization

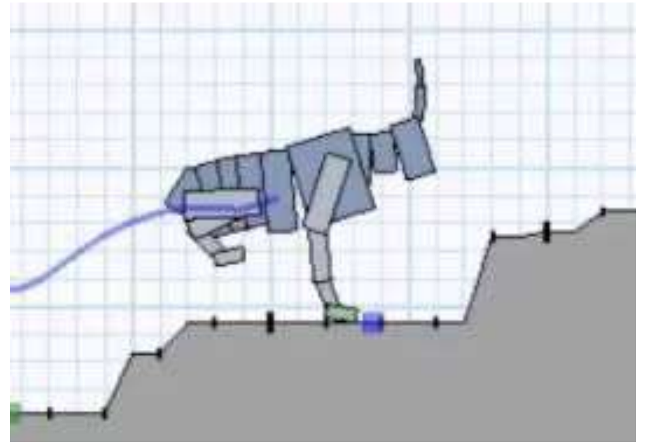
Procedural modelling

Mesh segmentation

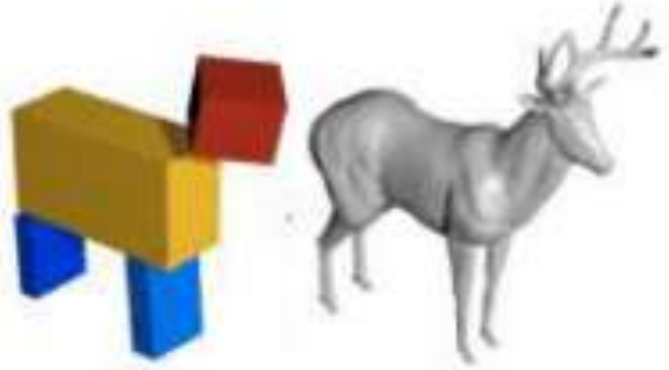
Learning deformations

BRDF estimation

Fluid



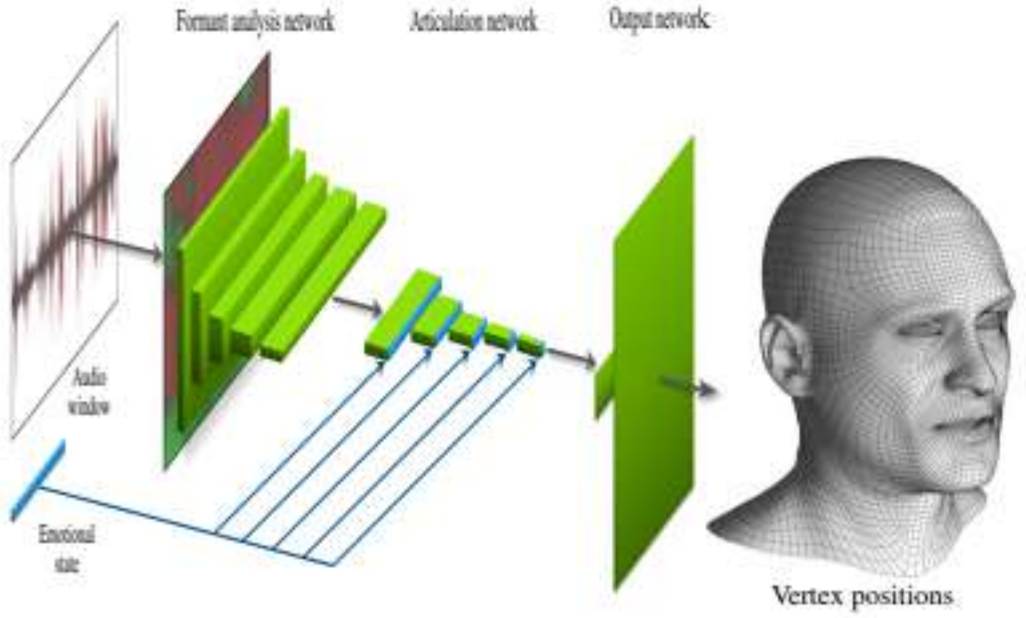
Animation



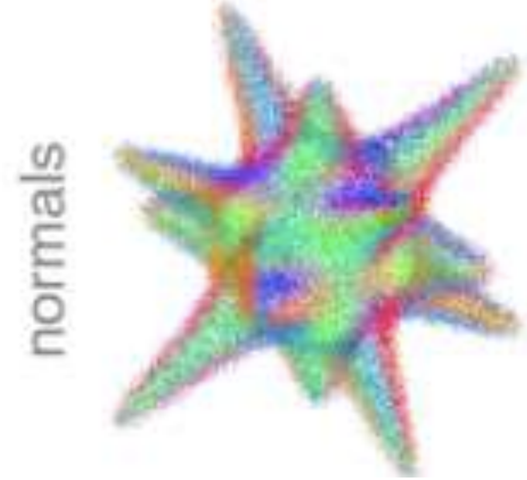
Boxification

Real-time rendering

Denoising



Facial animation



PCD processing

Course Logistics

- Lectures + QA sessions + lab sessions
- Grading: 1 coursework (50%) (**to be announced**)
- Final evaluation
- Expect to ***spend a lot of time*** on this course.
- Lecture slides and video — available on Moodle
- Reachable via email or via appointment
- Coding platform: Python and using Google Colab
- Further reading: recent research papers (check our lab pages!); CS223n (Stanford); EECS 498 (UMich), ML for CG courses (Eurographics and Siggraph)