LECTURER: Nghia Duong-Trung

ARTIFICIAL INTELLIGENCE

TOPIC OUTLINE

History of AI

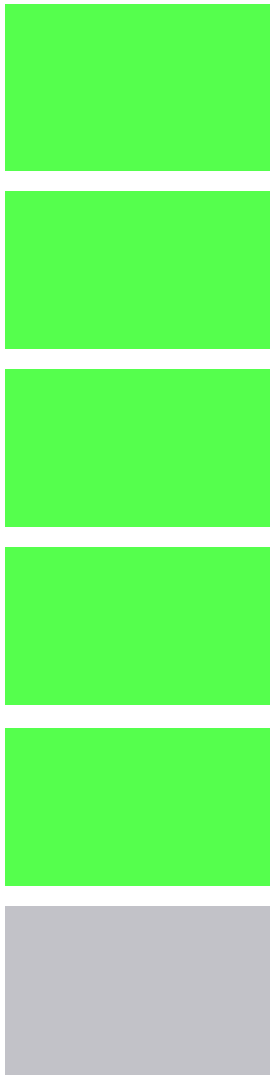
Modern AI Systems

Reinforcement Learning

Natural Language Processing – Part 1

Natural Language Processing – Part 2

Computer Vision



- Course book: Artificial Intelligence_DLBDSEAIS01, provided by IU, myStudies
- Reading list provided by IU, myStudies
- The amount of slides content is based on the course book.
- Additional teaching materials:

https://github.com/duongtrung/IU-ArtificialIntelligenceCourse

DISCLAIMER

- This is the modified version of the IU slides.
- I used it for my lectures at IU only.

UNIT 6

COMPUTER VISION

STUDY GOALS

On completion of this unit, you will be able to …

… define computer vision.

… explain how to represent images as pixels.

… distinguish between detection, description, and matching of features.

… correct distortion with calibration methods.

EXPLAIN SIMPLY

1. What are the typical computer vision tasks?

2. How are filters used in computer vision?

3. How does camera calibration work?

HISTORICAL DEVELOPMENTS

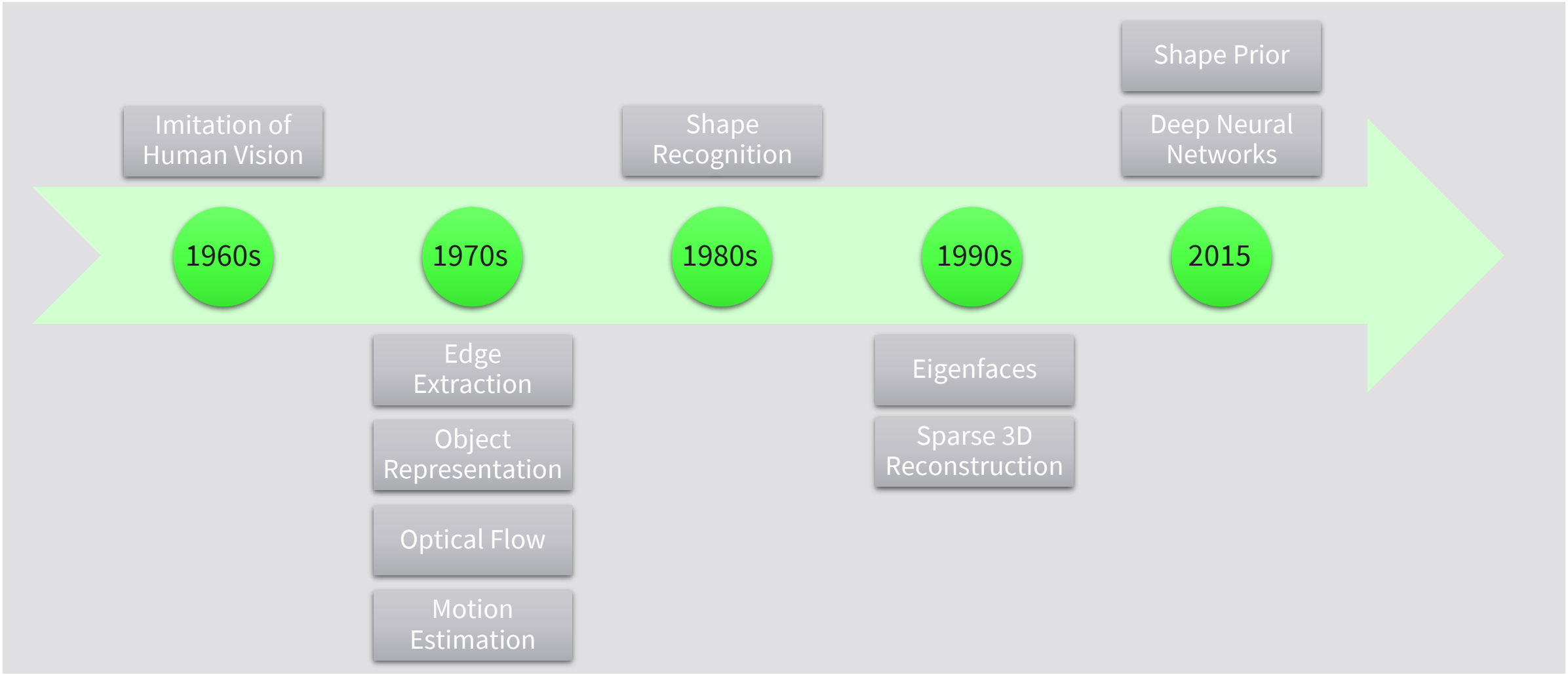


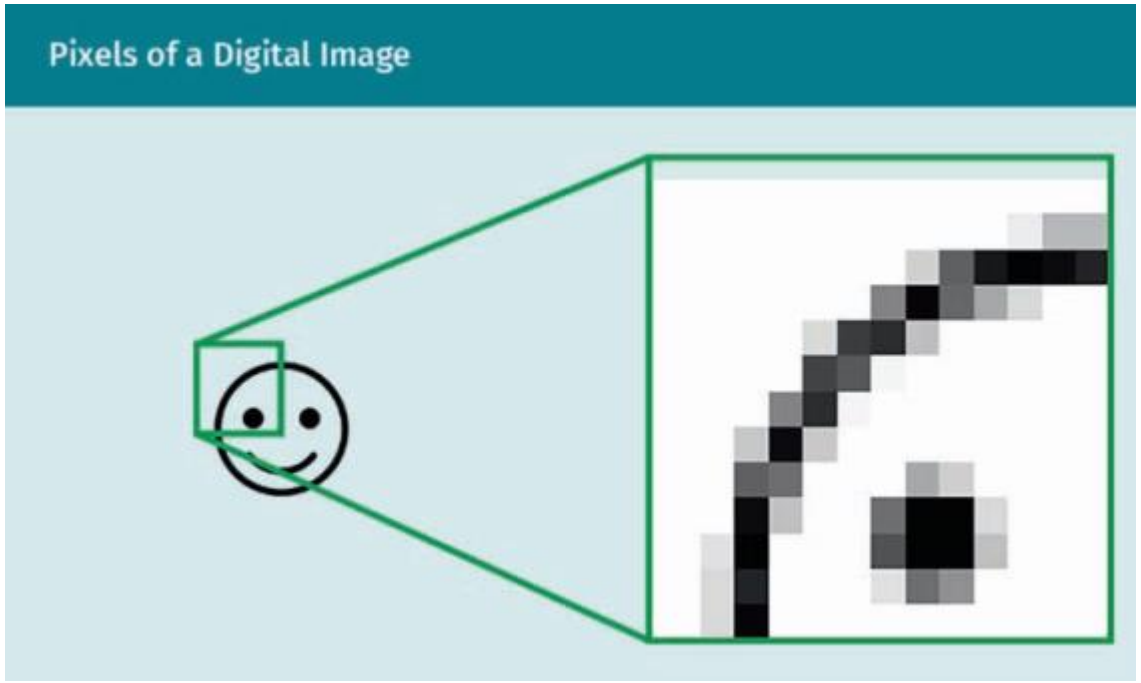
Image Source: Custom Depiction

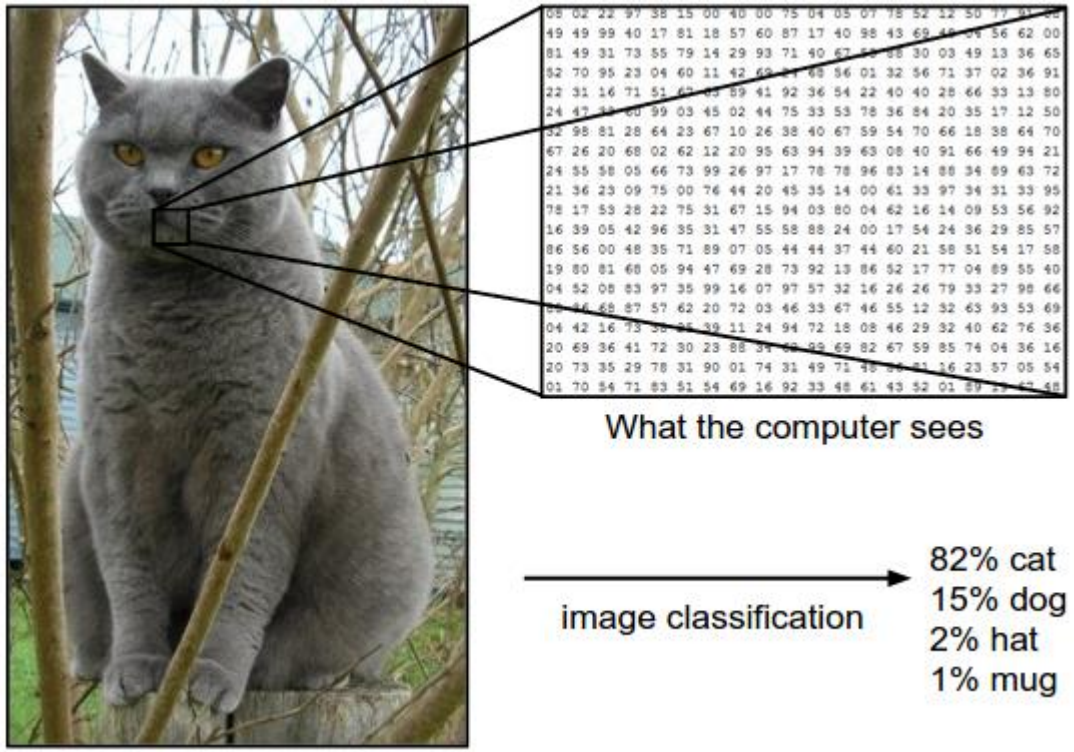
Term: pixel

color representations

monochrome

real, high, true, deep color





TYPICAL COMPUTER VISION TASKS

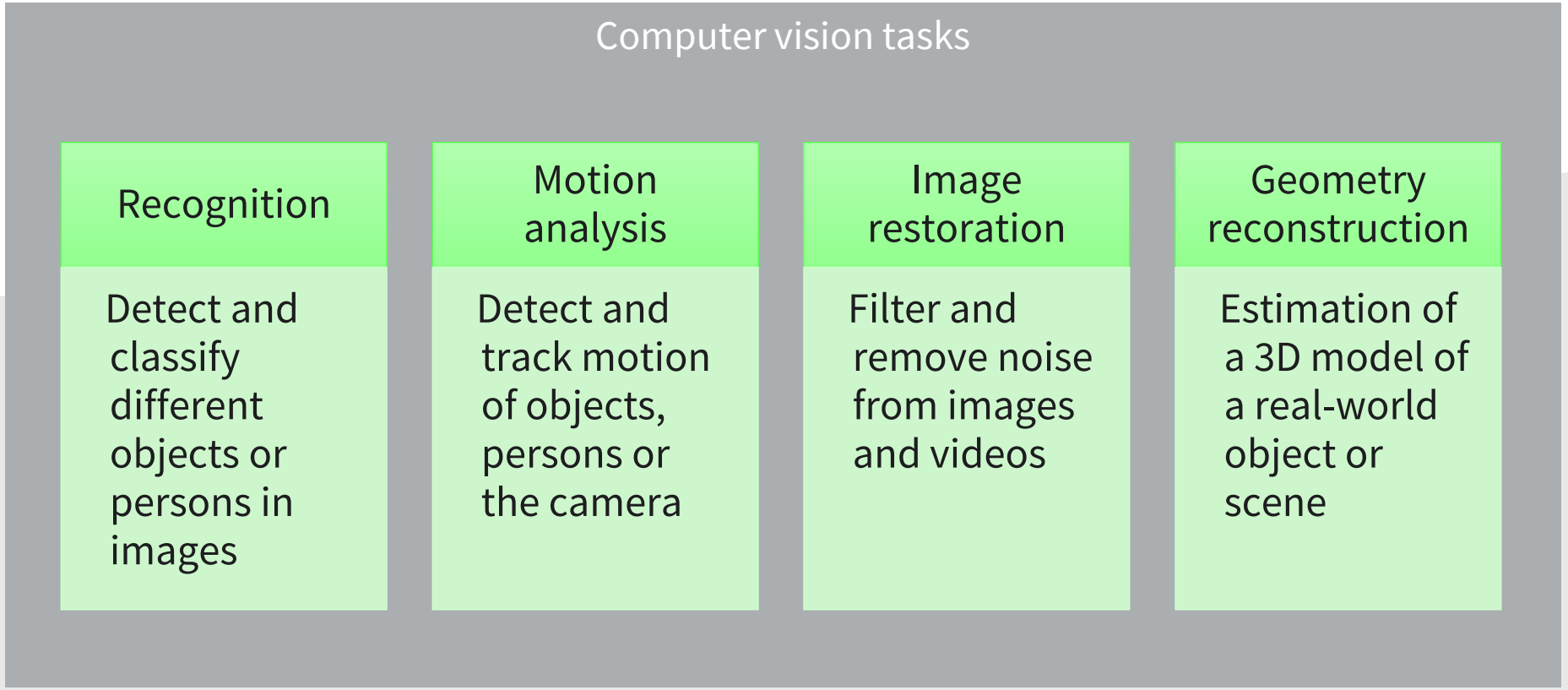


Image Source:

10 INTERESTING COMPUTER VISION TASKS: OBJECT DETECTION

- identify and locate an object of particular interest from still image or video data.

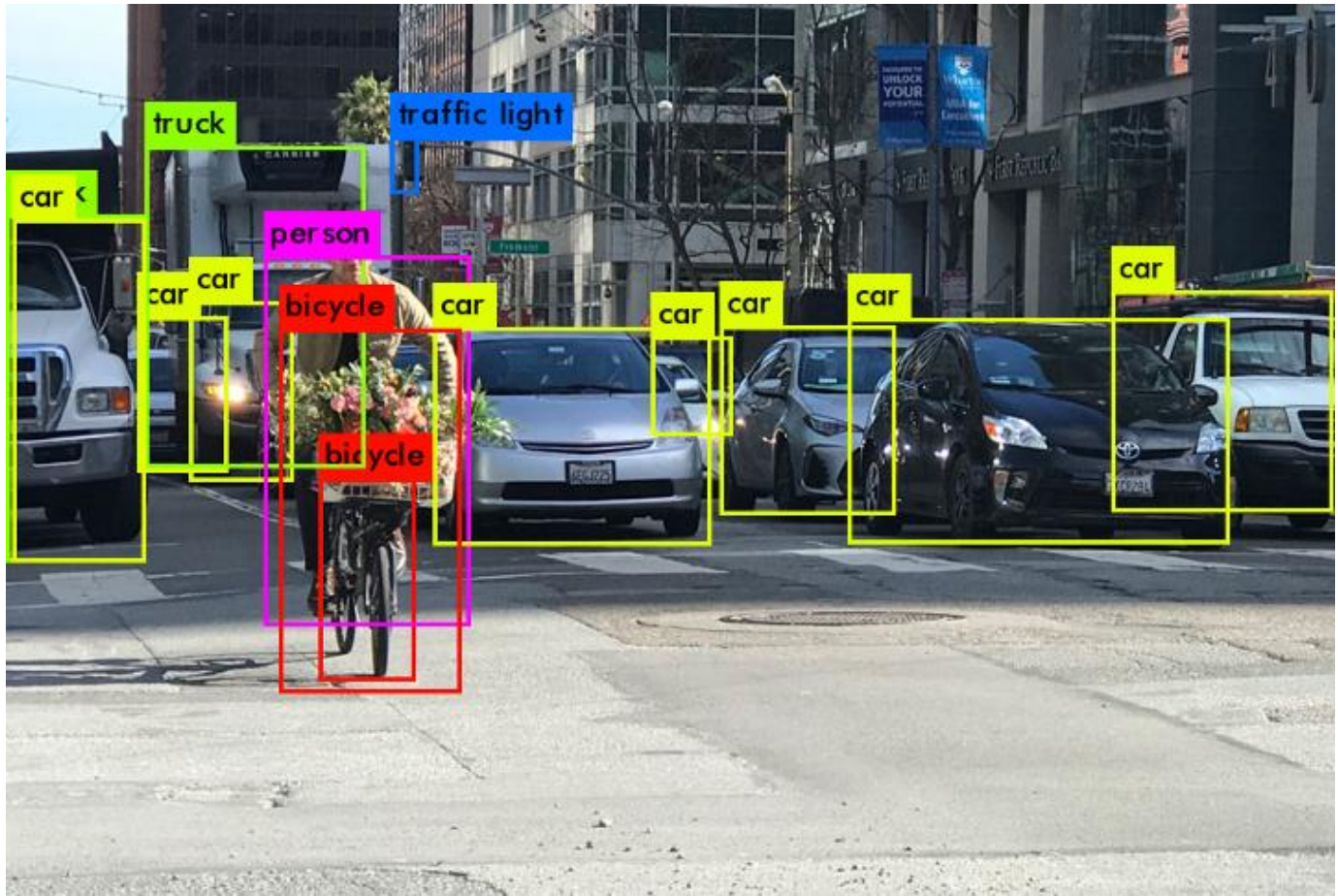


Image source: towardsdev.com

10 INTERESTING COMPUTER VISION TASKS: IMAGE SEGMENTATION

- The process of partitioning an image into multiple segments. the aim is to transform the image in a representation that is easier to analyze, there are many flavors of image segmentation

Image source: towardsdev.com

10 INTERESTING COMPUTER VISION TASKS: POSE ESTIMATION

Estimating the position and the orientation of the joints of a human body. this has a wide range of applications in AR/VR (augmented and virtual reality)

Image source: towardsdev.com

10 INTERESTING COMPUTER VISION TASKS: IMAGE GENERATION

Generating new images from an existing data set, this task has been made popular by the use Generative adversarial networks (GANs) to generate seemingly real images or videos which were actually fake.

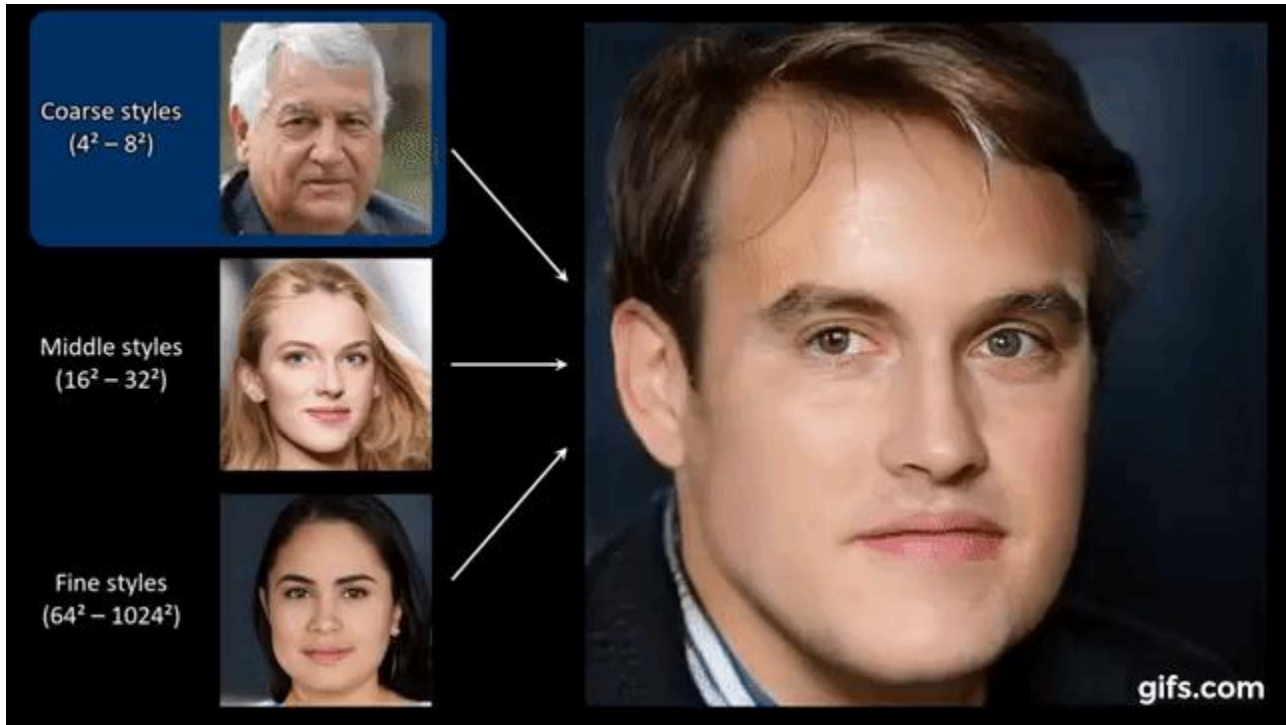


Image source: towardsdev.com

Identifying classes of objects within the image

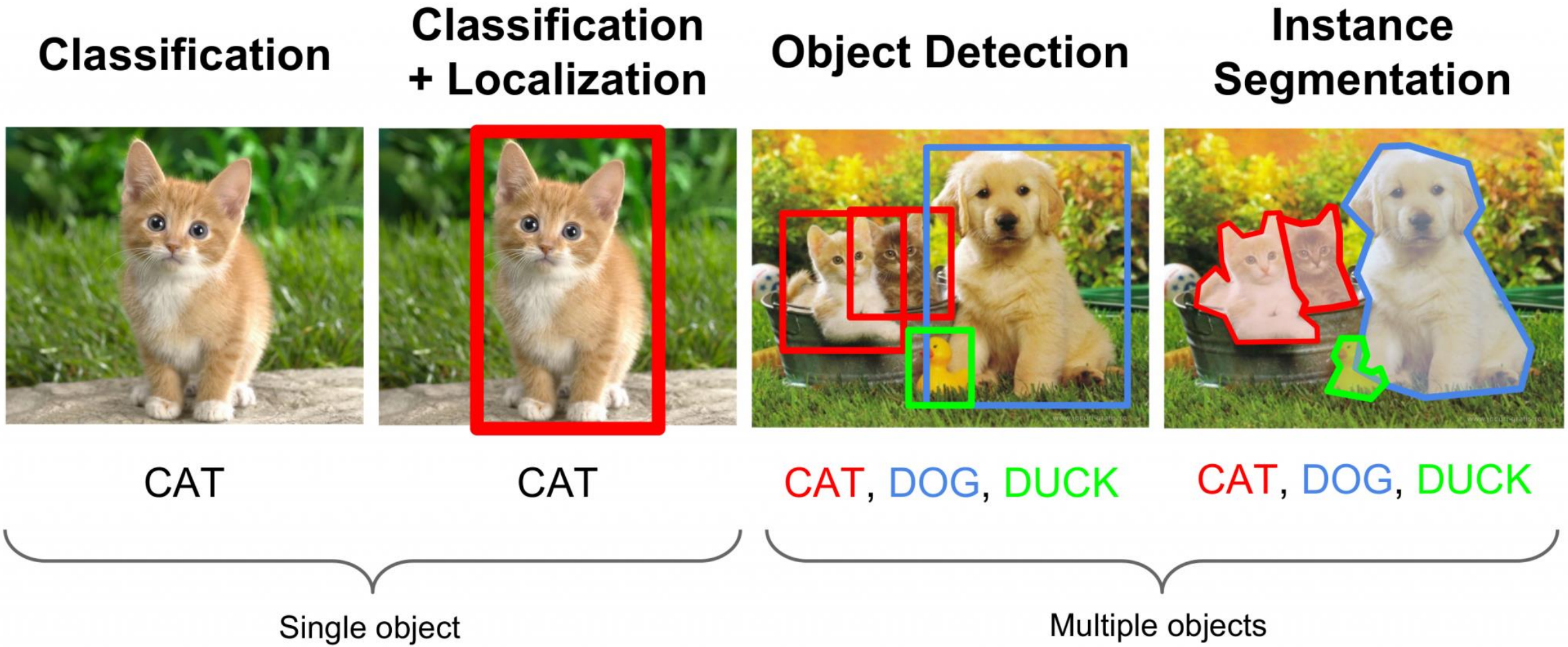


Image source: analyticsindiamag.com

maintaining a predictive model performance even if the data distribution the model was trained on has changed. domain adaptation uses labeled data in one or more data distribution and tries to solve a new task with data belonging to a different distribution

Image source: towardsdev.com

10 INTERESTING COMPUTER VISION TASKS: DE-NOISING

predicting the original image by removing noise from a noise-contaminated image

Image source: towardsdev.com

10 INTERESTING COMPUTER VISION TASKS: SUPER-RESOLUTION

taking an input from a low resolution image and up-scaling it to a higher resolution

Image source: towardsdatascience.com

10 INTERESTING COMPUTER VISION TASKS: 3D FACE ANIMATION

reconstructing an animated 3D human face from a still image

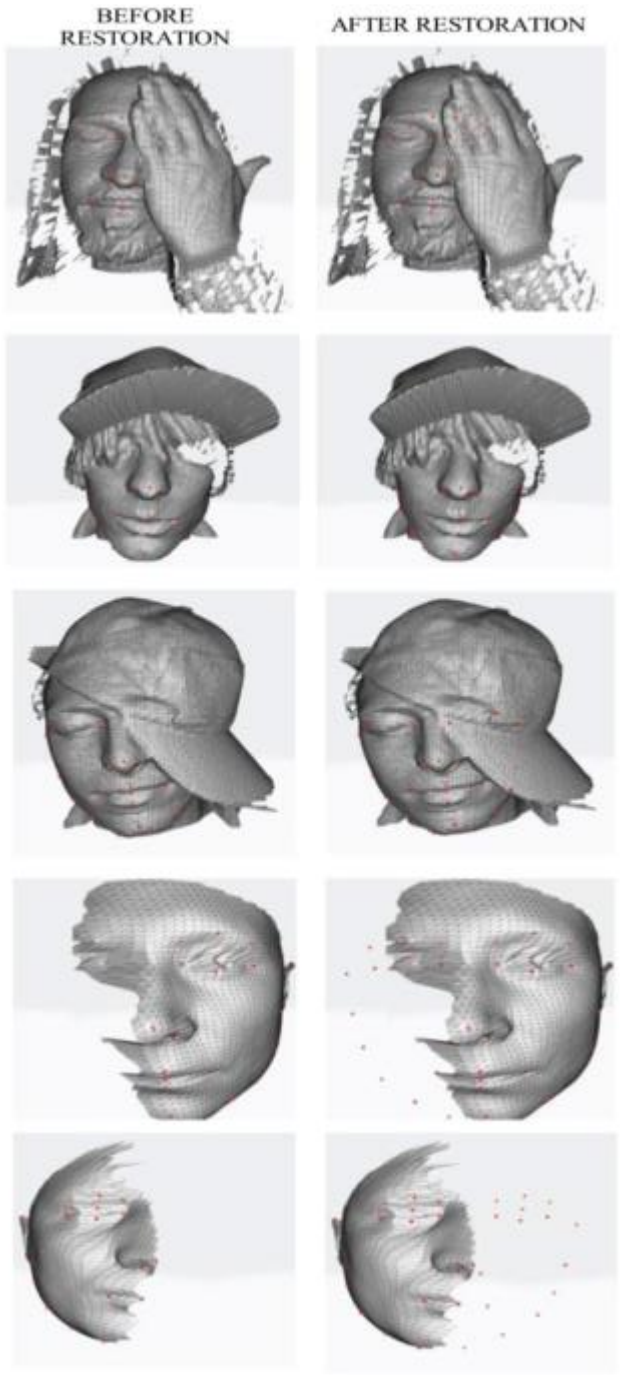


Image source: towardsdev.com, https://link.springer.com/article/10.1007/s11831-021-09705-4

predicting the movement of an object from observations made in a video stream

Image source: towardsdev.com

IMAGES AS FUNCTIONS

A function that can map a two-dimensional coordinate (x,y) to a specific color value.

On the x-axis we begin on the left with a value of 0 and continue to the right until the maximum width of an image is reached.

On the y-axis, we begin with 0 at the top and reach the height of an image at the bottom.

f(42,100)

Filter:

Function which receives an image as an input, applies modifications, and returns the filtered image as an output.

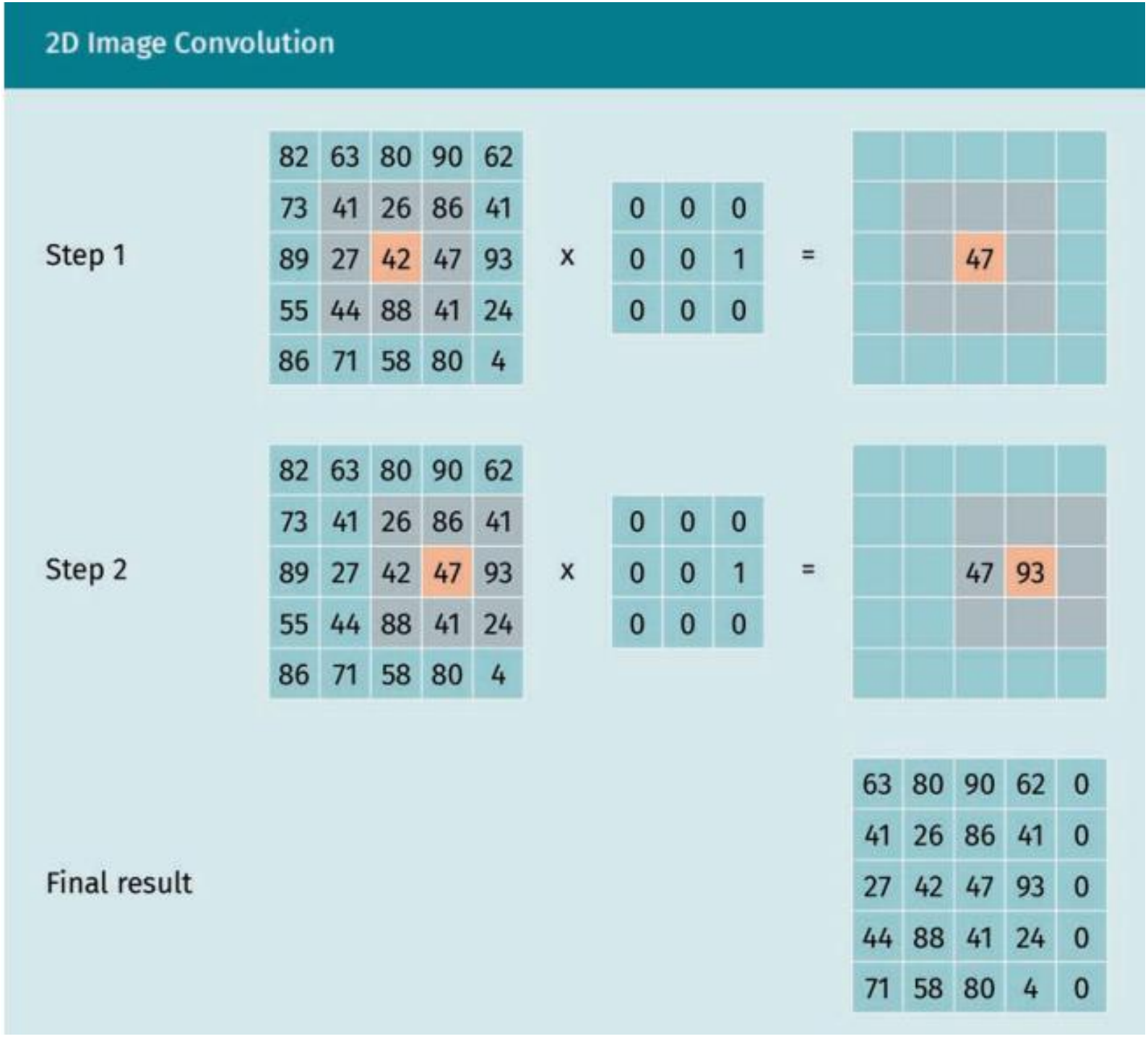
Example: 2D image convolution

Convolution of an image I with a kernel k with a size of n and a center coordinate a:

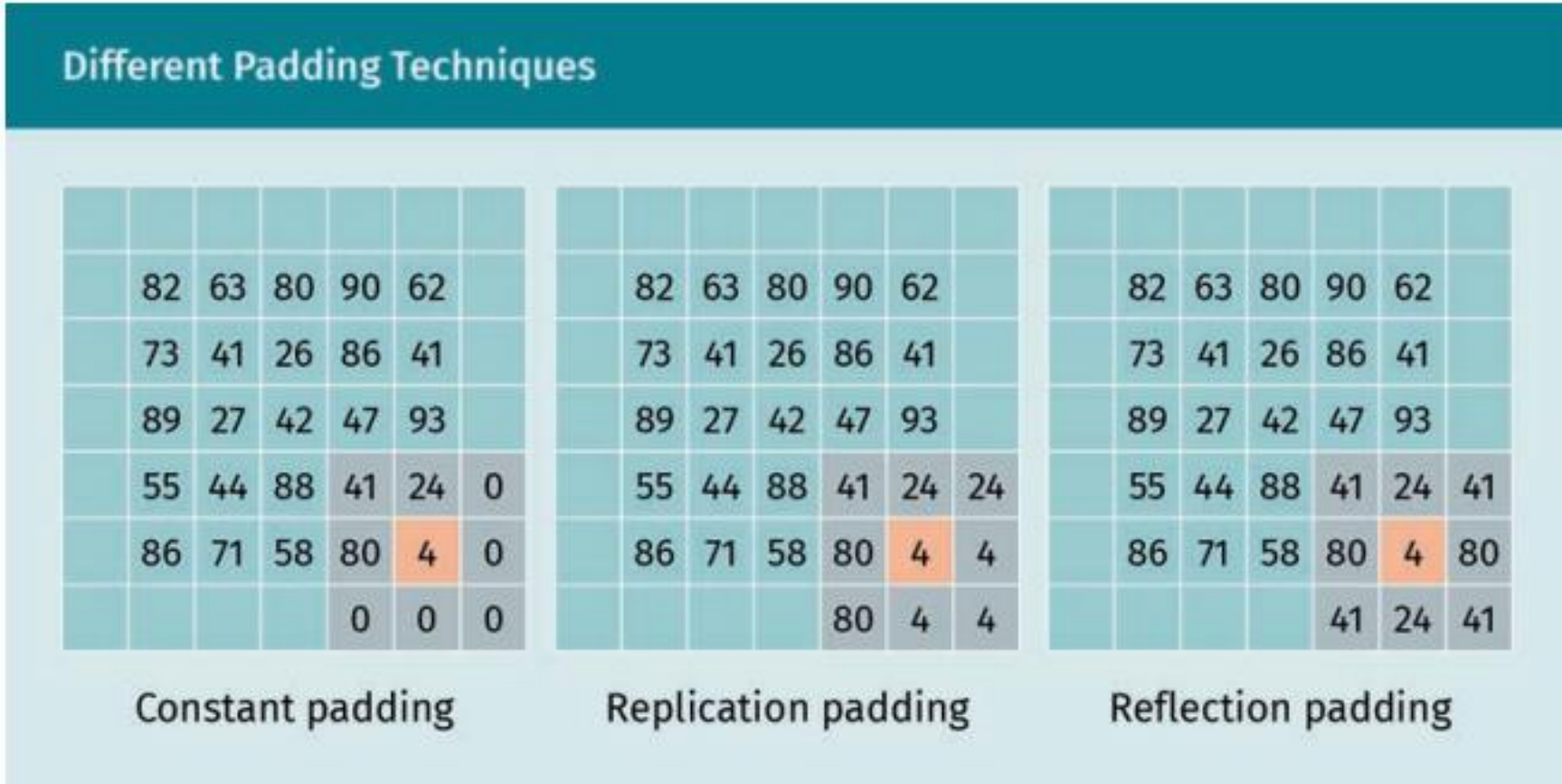
$$\mathrm{I} \cdot (x, y) = \sum_{i=1}^{n} \sum_{i=j}^{n} \mathrm{I}(x - 1 + a, y - j + a) k(i, j)$$

$\mathrm{I} \cdot (x, y)$: value of the resulting image $\mathrm{I} \cdot$ at position (x, y), I is the original image.

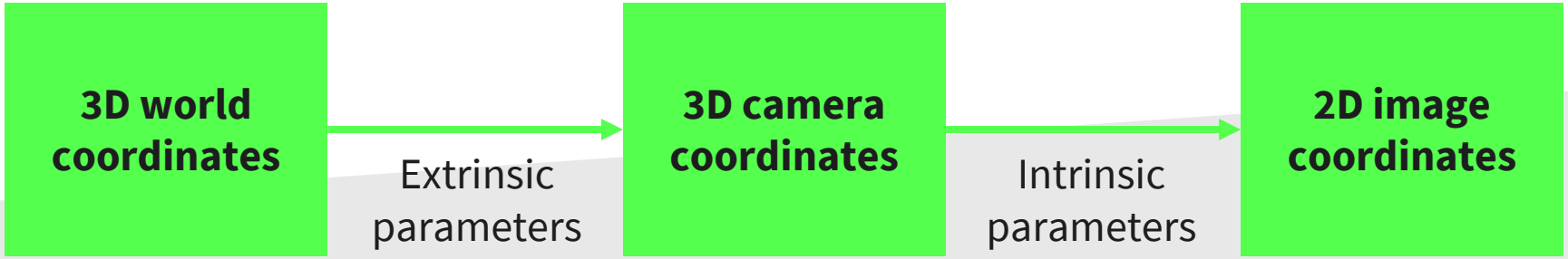
FILTERS


2D Image Convolution
Step 1
82 63 80 90 62
73 41 26 86 41
89 27 42 47 93
55 44 88 41 24
86 71 58 80 4
x
0 0 0
0 0 1
0 0 0
=
47
Step 2
82 63 80 90 62
73 41 26 86 41
89 27 42 47 93
55 44 88 41 24
86 71 58 80 4
x
0 0 0
0 0 1
0 0 0
=
47 93
Final result
63 80 90 62 0
41 26 86 41 0
27 42 47 93 0
44 88 41 24 0
71 58 80 4 0

PADDING TECHNIQUES



The calibration makes it possible to extract distortion from the images



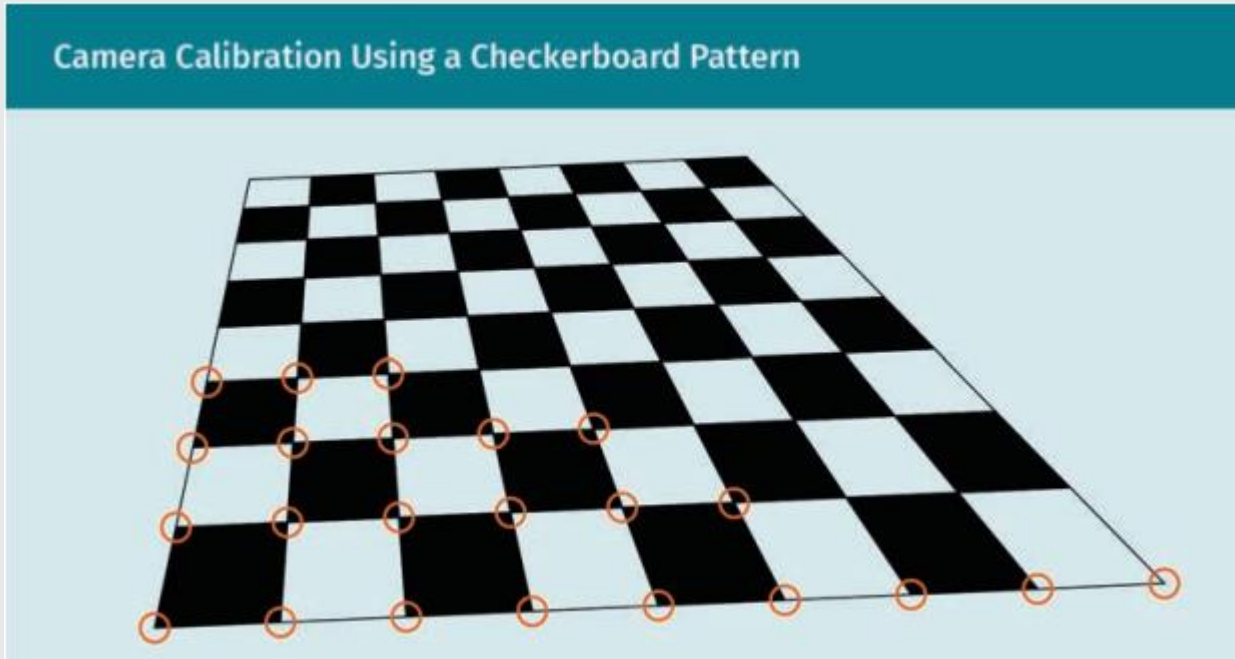


Image Source: Custom Depiction

Feature Detection

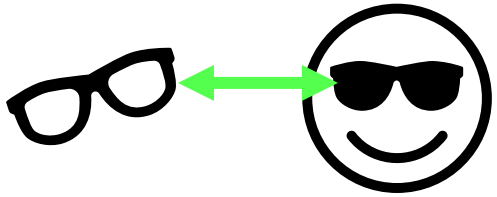
- Blobs (blue)
- Edges (red)
- Corners (yellow)

Feature Description

Semantic information

- Common algorithms:
  - BRIEF
  - ORB
  - SIFT
  - SURF

Feature Matching



- Identification of similar features in different images

https://medium.com/data-breach/introduction-to-surf-speeded-up-robust-features-c7396d6e7c4e
https://docs.opencv.org/3.4/df/dd2/tutorial_py_surf_intro.html

Image Source: Custom Depiction

IMPORTANT CHARACTERISTICS OF FEATURE DETECTION ALGORITHMS

- Robustness
- Accuracy
- Repeatability
- Generality
- Efficiency
- Quantity

SEMANTIC SEGMENTATION

Original image

Segmentation map

Segmentation overlay

Background Chair Coffee table

REVIEW STUDY GOALS

On completion of this unit, you will be able to …

… define computer vision.

… explain how to represent images as pixels.

… distinguish between detection, description, and matching of features.

… correct distortion with calibration methods.

SESSION 6

TRANSFER TASK

1.

a) Given the following image matrix *I* and the kernel matrix *k*, compute the values of the resulting image after the convolution. What does the kernel matrix do?

I =

82	63	80	90	62
73	41	26	86	41
89	27	42	47	93
55	44	88	41	24
86	71	58	80	4

k =

0	1	0
0	0	0
0	0	0

1.

b) What type of filter would the kernel matrix k2 apply to an image?

c) How would the filtered image change if you applied kernel matrix k3 instead?

$$k2 = \frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$k3 = \frac{1}{25}\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

TRANSFER TASK

2. For camera calibration, it is important to identify the external and internal parameters.
Which extrinsic and intrinsic parameters do you know?

3. Think about possible use cases for semantic image segmentation. Where could it be used?

TRANSFER TASK
PRESENTATION OF THE RESULTS

Please present your results.

The results will be discussed in plenary.

1. a) Image convolution

82	63	80	90	62
73	41	26	86	41
89	27	42	47	93
55	44	88	41	24
86	71	58	80	4

X

0	1	0
0	0	0
0	0	0

=

63	80	90	62	0
41	26	86	41	0
27	42	47	93	0
44	88	41	24	0
71	58	80	4	0

The kernel matrix shifts the image matrix to the left.

1.

b) Matrix k2 will apply a smoothing to an image.

c) When using the kernel matrix k3, the image will be smoothed even more.

https://www.codingame.com/playgrounds/2524/basic-image-manipulation/filtering

https://www.ni.com/de-de/innovations/white-papers/06/image-filtering-overview.html

https://ai.stanford.edu/~syyeung/cvweb/tutorial1.html

https://en.wikipedia.org/wiki/Kernel_(image_processing)

2. Camera calibration

External parameters:

- Orientation of real world coordinates
- Position of the camera

Internal parameters:

- Optical center
- Focal length
- Lens distortion parameters

https://docs.opencv.org/4.x/dc/dbb/tutorial_py_calibration.html

https://www.youtube.com/watch?v=3h7wgR5fYik

3. Possible use cases for semantic image segmentation:

- Autonomous driving: detection of other vehicles, road lanes, pedestrians, or sidewalks
- GeoSensing: analyze information about land usage, such as agricultural areas, forests, or areas of water from satellite images
- Pose estimation/motion capture: identifying and tracking of body parts, such as legs, arms, head, or eyes
- Medicine: detection of affected brain areas by tumors

LEARNING CONTROL QUESTIONS

1. What are the typical tasks in computer vision?

2. What is the purpose of camera calibration?

3. What are the most commonly used types of features in computer vision?

4. What is the purpose of semantic segmentation?

Solutions

1. Recognition, motion analysis, image restoration, geometry reconstruction
2. To estimate extrinsic and intrinsic parameters of a camera
3. Blobs, corners, and edges
4. Cluster parts of the image that belong to the same object class

How did you like the course?

HOW DID YOU
LIKE THE COURSE?