LECTURER: Nghia Duong-Trung

DATA UTILIZATION

TOPIC OUTLINE

TOPIC OUTLINE

UNIT 4

IMAGE RECOGNITION

STUDY GOALS

On completion of this unit, you will have learned …

- … different types of image representation.
- … image processing applications.
- … data compression techniques.
- … different approaches to image processing, including classic image processing and modern deep-learning-based approaches.
- … image transformation techniques and integral transforms.
- … the basics of convolutional neural networks.
- … the basics of augmented reality concepts and applications.

EXPLAIN SIMPLY

1. What are image processing and its applications?
2. What are well-known forms of image representation?
3. Which different levels of feature extraction exist in image processing approaches?
4. What is the role of neural networks in image processing?
5. What are AR and its applications?

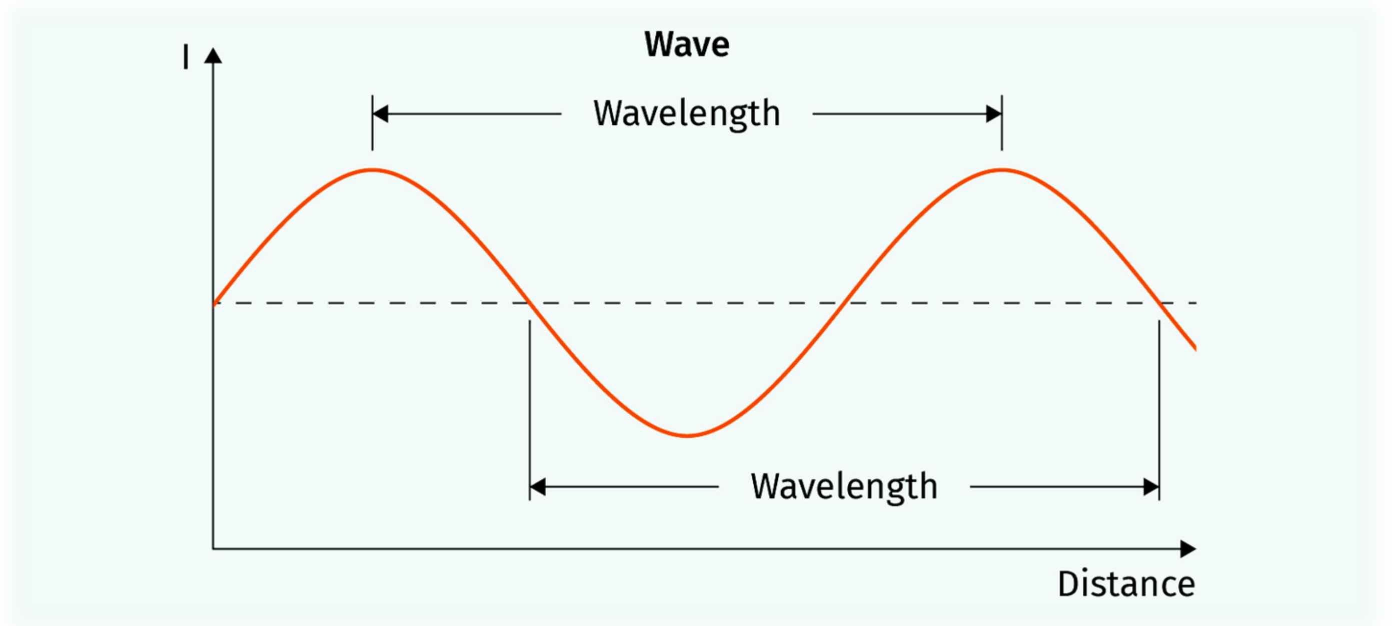
Image analysis:

process of extracting useful information from raw images

Image processing applications:

barcode reading, face detection, medical image analysis, surveillance, and movement detection

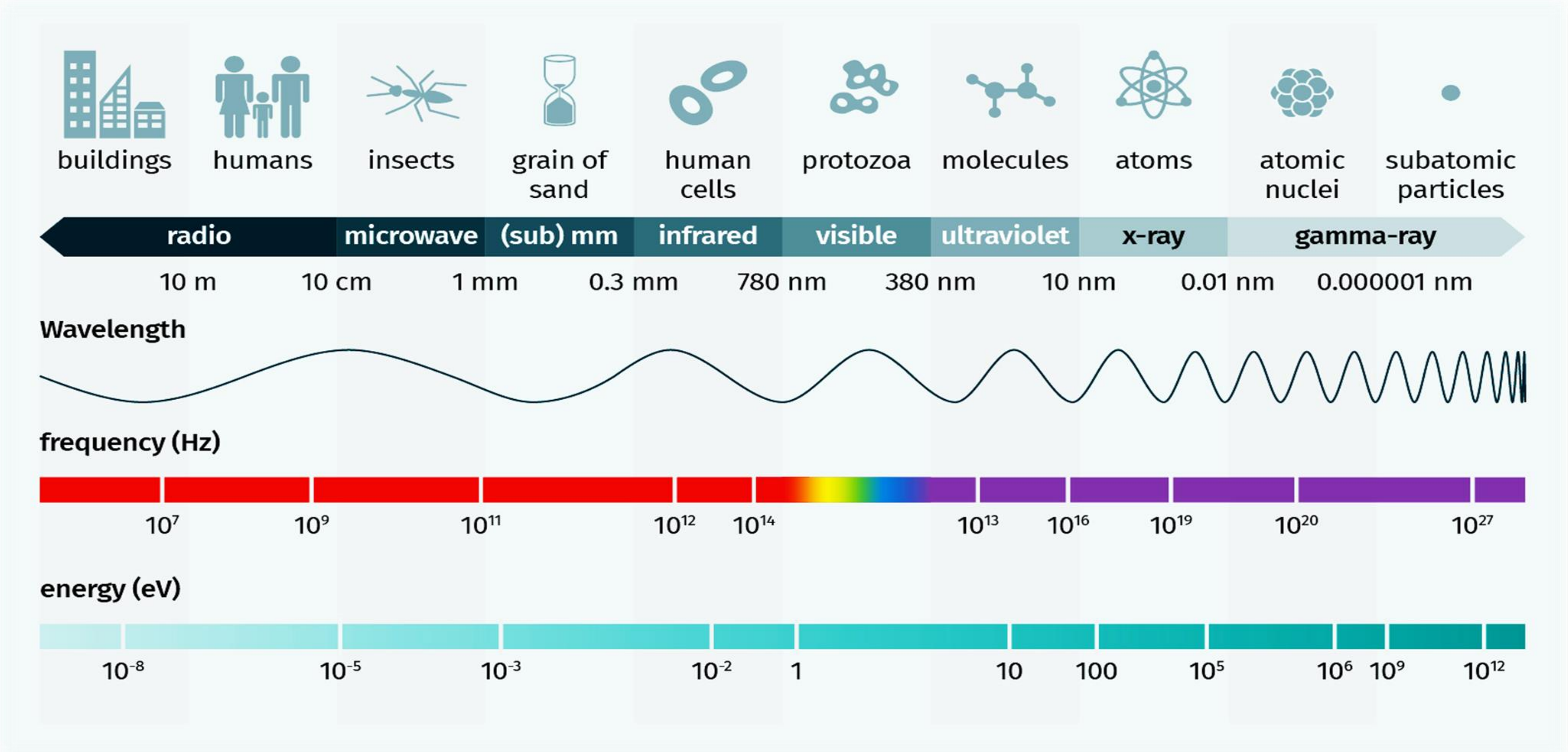
Electromagnetic wavelengths



- Human eye receptors are sensitive to **light wavelengths** in the interval between **400 (violet) and 800 nanometers (red)**.

Source of the graphic: Course Book DLMBBD01, p. 92

Electromagnetic Wavelets

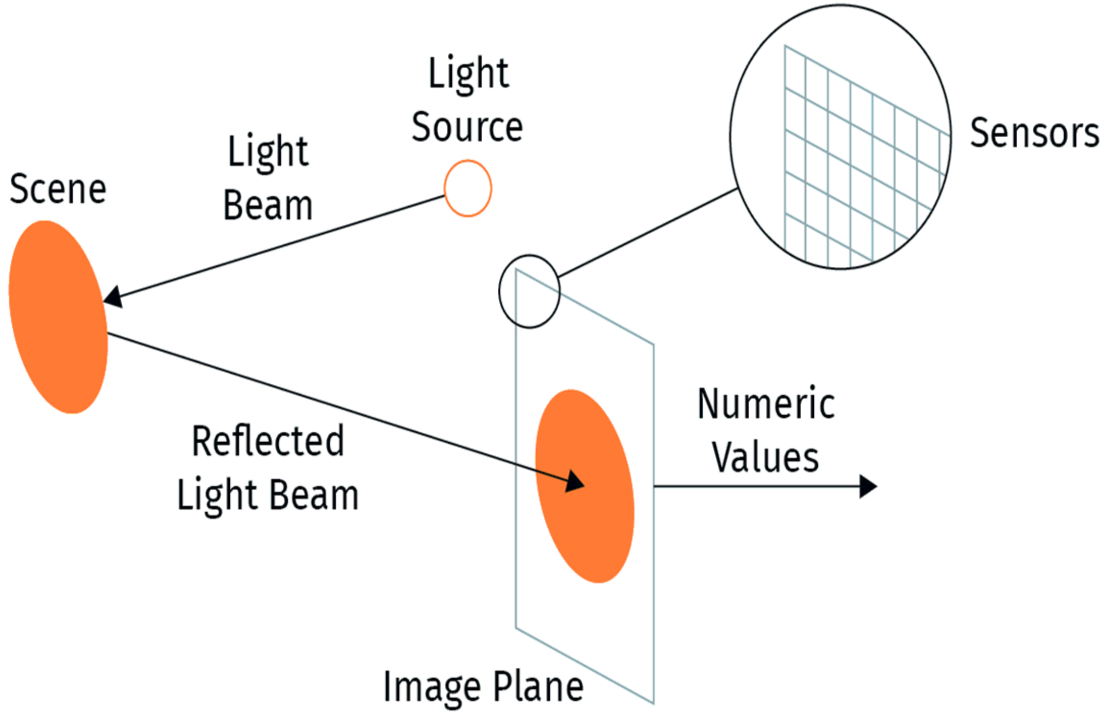


Source of the graphic: Stanford Solar Center, 2015

An image is either:

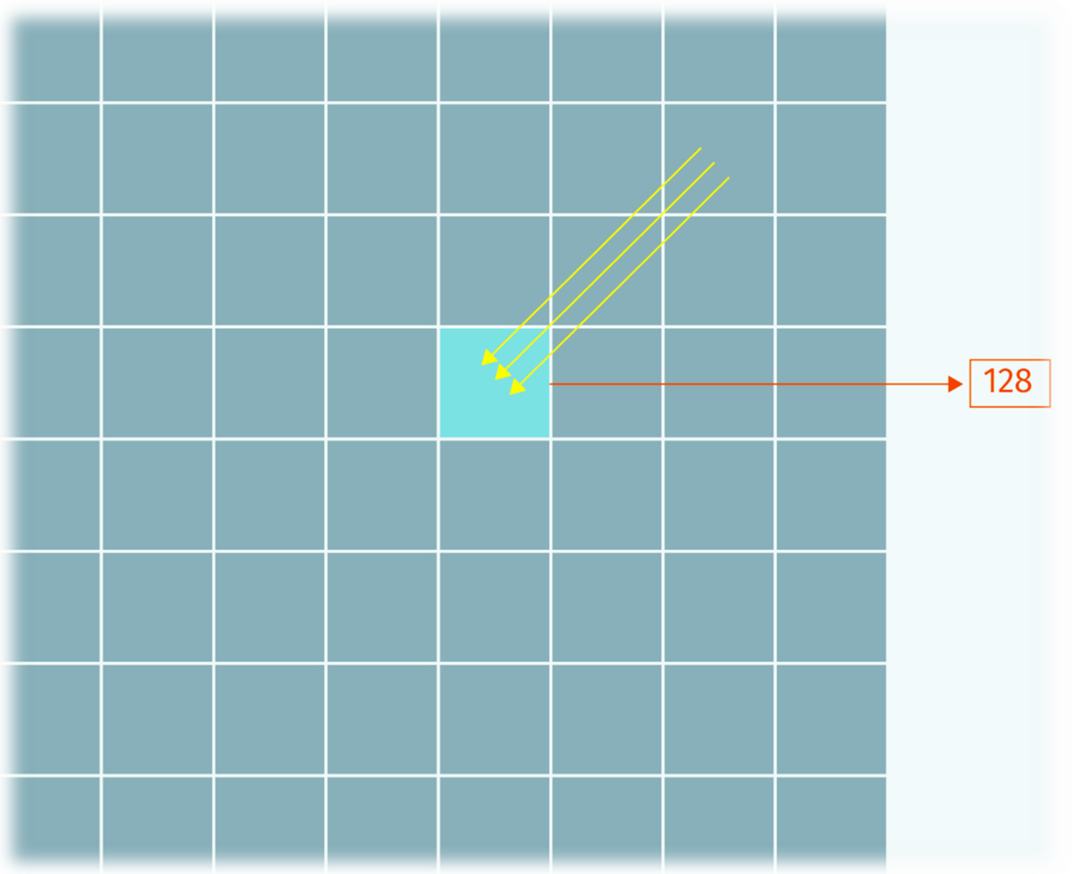
- the reflection of light from an object captured by a camera (e.g., **personal photographs**)
- the absorption of light by an object (e.g., **X-ray images**),
- the emission of light by an object (e.g., positron emission tomography (**PET**) images).

The camera is an **image creator** through which the information contained in the reflected light is extracted and stored.



Source of the graphic: Course Book DLMBBD01, p. 93

- An image plane can be viewed as a **matrix of discrete points** on a plane where each point represents a small area.
- The measured value returned by a sensor represents the **light intensity** of the whole sensor area.



The Sampling of Light by Sensors of the Image Plane

Source of the graphic: Course Book DLMBBD01, p. 94

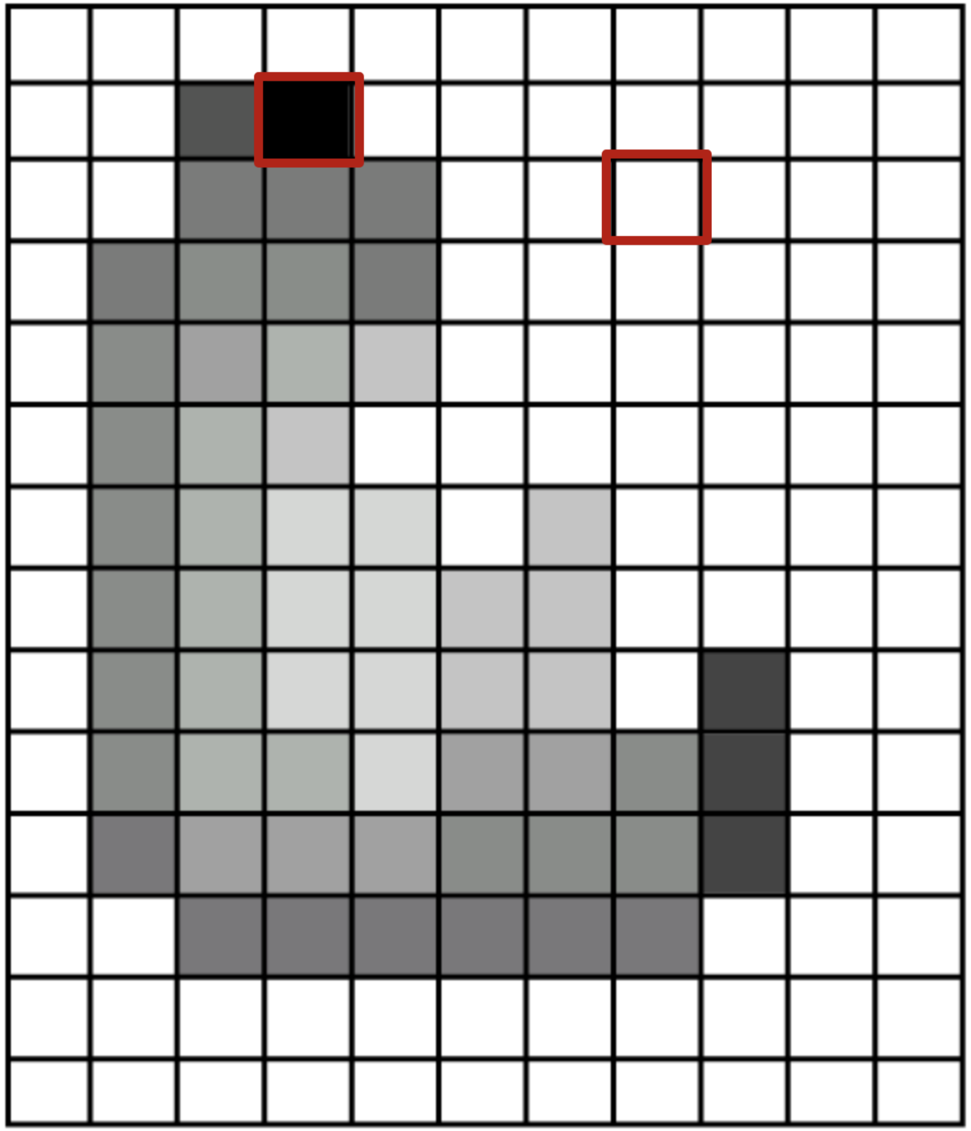
- A digital image can be represented as an M×N matrix of pixels which corresponds to a 2D rectangular representation of the image.
- In black and white images, each entry of the matrix stores a value between 0 and 255 which shows the intensity of the corresponding pixel.
- The intensity is called the **gray level (or gray value)** of that pixel.

- Each sensor returns a single value as the measured light intensity.
- The slanted edge of an object partly covers the area of a sensor, the corresponding pixel will be considered either part of the object or outside it.
- The slanted edges of objects are represented in stepped shapes.

- The measured values are normalized to a given range through a process called "**quantization**."
- The range of distinguished values between the brightest color (white) and darkest color (black) is called the **quantization level.**
- Typical quantization levels are 2 (black and white), 4, 16, 64, and 256.

Image Represented with 64(a), 16(b), and 2(c) Quantization Levels

Source of the image: Smarandache, n.d.



=

255	255	255	255	255	255	255	255	255	255	255
255	255	20	0	255	255	255	255	255	255	255
255	255	75	75	255	255	255	255	255	255	255
255	75	95	95	75	255	255	255	255	255	255
255	96	127	145	175	255	255	255	255	255	255
255	127	145	175	175	175	255	255	255	255	255
255	127	145	200	200	175	175	95	255	255	255
255	127	145	200	200	175	175	95	47	255	255
255	127	145	145	175	127	127	95	47	255	255
255	74	127	127	127	95	95	95	47	255	255
255	255	74	74	74	74	74	74	255	255	255
255	255	255	255	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255	255	255	255

0 = black; 255 = white

- most common format in digital photography
- Typically processed line by line, with each horizontal line called a **scan line.**
- The value of each cell of the matrix indicates the brightness of the corresponding pixel.

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	1	0.7	0	0.7	1	1	1	1
1	1	1	0.7	0	0	0	0.7	1	1	1
1	1	0.7	0	1	1	1	0	0.7	1	1
1	1	0	1	1	1	1	0	0	1	1
1	1	1	1	1	1	1	0	0	1	1
1	1	1	1	1	1	1	0	0.7	1	1
1	1	1	1	1	1	0	0.7	1	1	1
1	1	1	1	1	0	0.7	1	1	1	1
1	1	1	1	0	0.7	1	1	1	1	1
1	1	1	0	0.7	1	1	1	1	0.7	1
1	1	0	0	1	1	1	1	0	1	1
1	0	0	0	0	0	0	0	1	1	1
1	1	1	1	1	1	1	1	1	1	1

Source of the graphic: Course Book DLMBBD01, p. 96

- A vector image is represented as a collection of elementary shapes (e.g., **points, lines, and curves**) where a point has a (x, y) coordinate (i.e., its position on the x-axis and y), with each line or curve having its own properties.
- One of the advantages of vector images compared to bitmap images is that the quality of the image does not suffer when zoomed in. Because of this property, vector images are ideal for printing.

APPLICATIONS

Medical Imaging

Computer Games

Image Processing Applications

Visual Arts

Smart Railways

Compression methods are necessary to **reduce size** and **the amount of memory** needed to store images.

Data Compression Methods

Lossless Image Compression

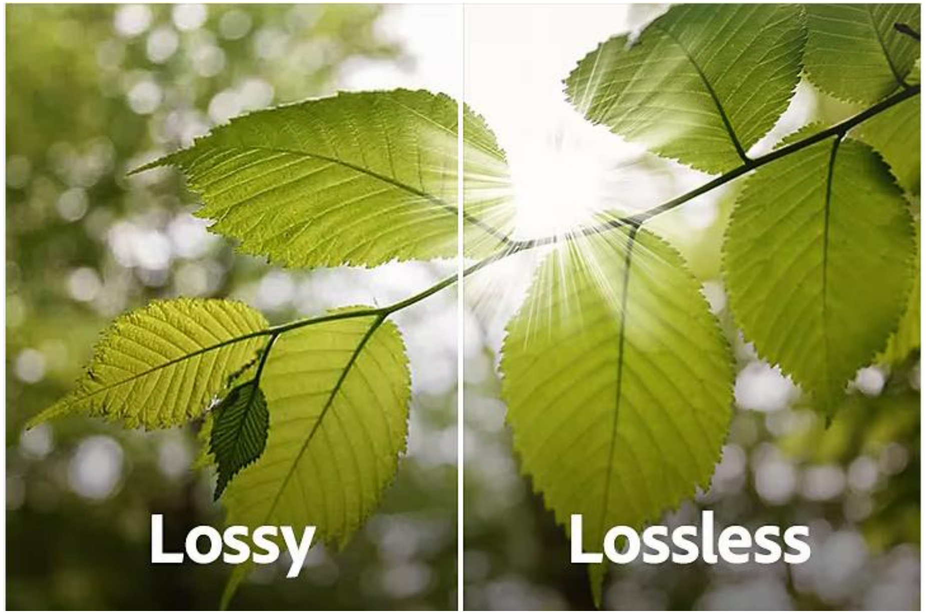
pixels in a tiny neighborhood often share the same values (i.e., same color), so identifying repeated values in the original image can yield a more compact image.

Lossy Image Compression

Lossy compression algorithms are only used in the final step of an image processing task when a compressed copy of an image needs to be created and stored. These images cannot be used as a working format during the intermediate steps of an image processing task.

LOSSY VS LOSSLESS

- Lossy image compression is a process that removes some of the data from your image file, reducing the overall file size. This process is irreversible, meaning that the file information will be removed permanently.
    - .JPEG
- Lossless image compression will not reduce image quality. That is because lossless compression only removes additional, non-essential data automatically added by the device used to take the photo.
    - .RAW, . PNG, .BMP



Source image: Adobe

IMAGE FORMATS

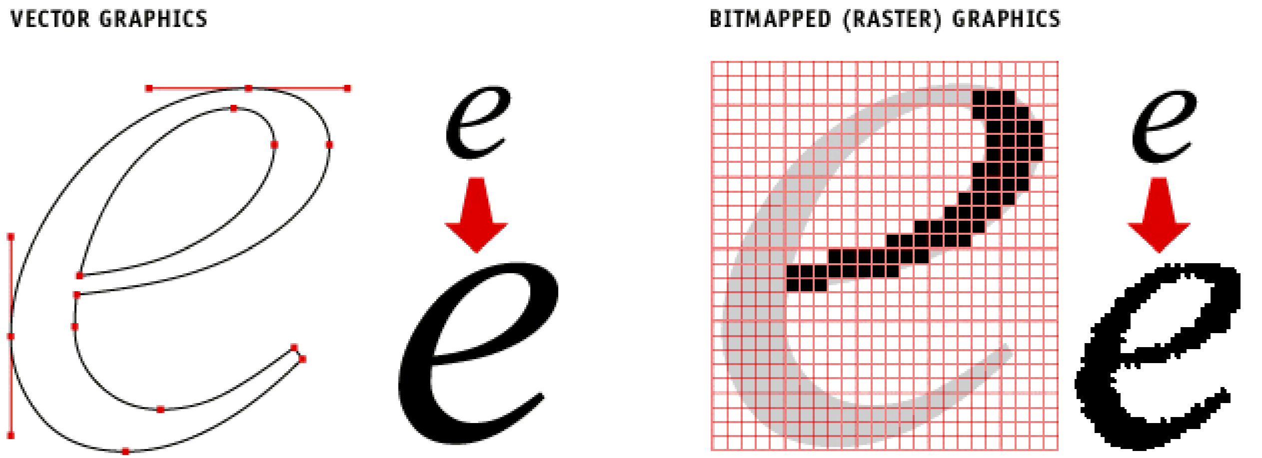
Extension	Name	Raster/ Vector	Description
.bmp		Raster	
.jpg	JPEG (Joint Photographic Experts Group)	Raster	Lossy compression format well-suited for photographic images
.png	Portable Network Graphic	Raster	Lossless compression image supporting 16-bit sample depth andA lpha channel
.gif	Graphic Interchange Format	Raster	8-bit indexed bitmap format superseded by PNG on all accounts except animation
.tiff, .tif	Tagged Image File Format	Raster	

Source of the table: Course Book DLMBBD01, p. 92

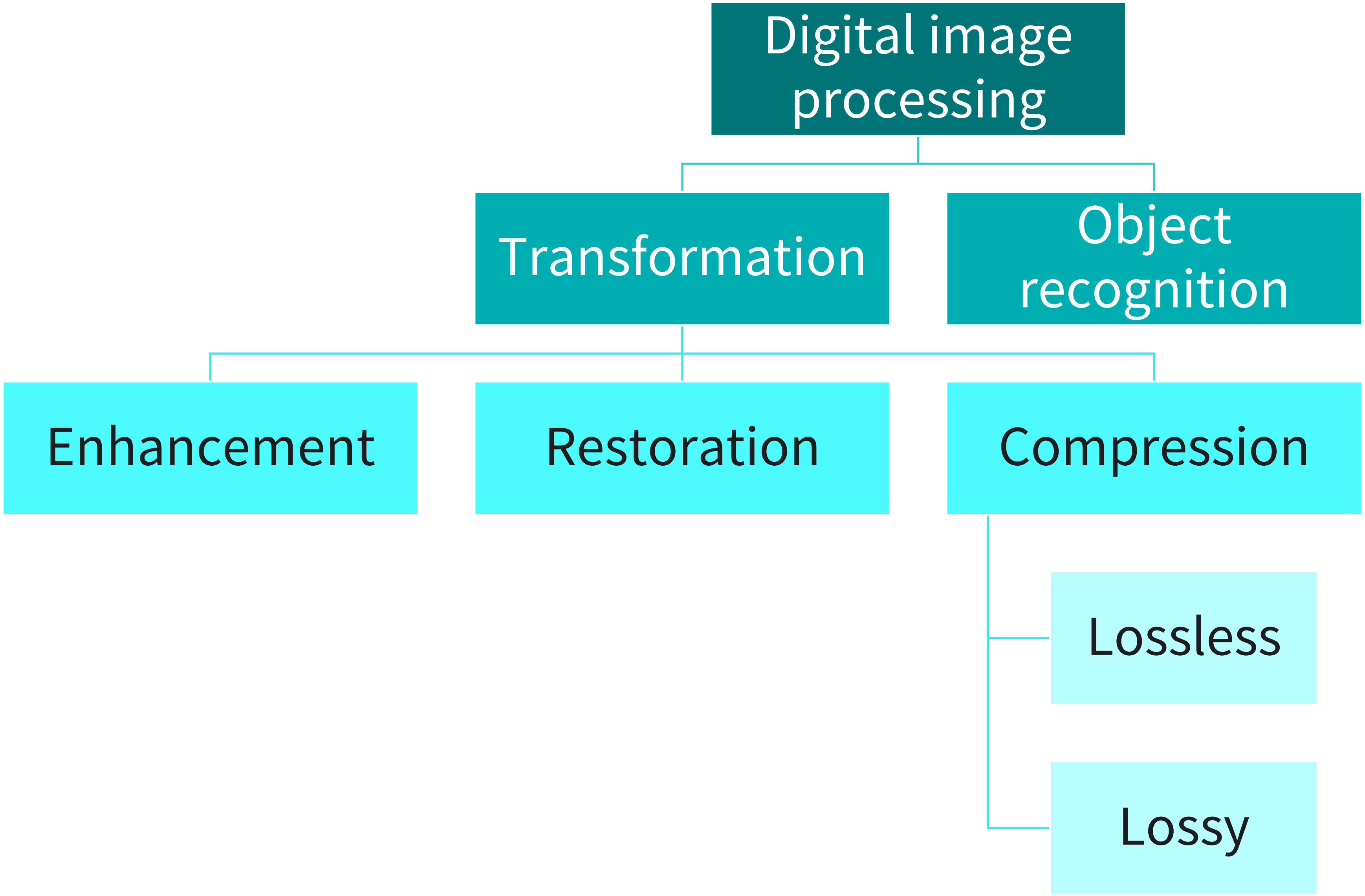
IMAGE FORMATS

.ai	Adobe Illustrator Document	Vector	Native format of Adobe Illustrator(based on .eps)
.eps	Encapsulated PostScript	Vector	The industry standard for including vector graphics in print
.ps	PostScript	Vector	Vector-based printing language used by many laser printers as electronic paper for scientific purposes
.pdf	Portable Document Format	Vector	A modernized version of .ps, adopted by the general public as the electronic print version
.svg	Scalable Vector Graphics	Vector	XML-based W3C standard incorporating animation, gaining adoption
.swf	Shockwave Flash	Vector	Binary vector format with animation and sound, supported by most major Web browsers

Source of the table: Course Book DLMBBD01, p. 92


VECTOR GRAPHICS
BITMAPPED (RASTER) GRAPHICS


VECTOR
RASTER



Source of the graphic: Course Book DLMBBD01, p. 103

Low-level features (edge detection)

- The value of pixels reflects the visual properties of the surface from which the light is reflected, the ambient light reflected from other objects in the scene, the type, color, and intensity of the light source, and the characteristics of the camera.
- Edge detection: Pixels on the border of an object should have neighbors from the object with the same color as they have and neighbors with different colors that belong to other objects or the background.

Mid-level features (segmentation)

- Unifying information from local processing of pixels yields more descriptive evidence about the objects in the images.
- Segmentation: process of dividing images into disjointed areas based on an image property such as color, intensity, or texture, where each area is uniform in terms of that property.
- Generally, each area defines an object, but this is not a necessity.

Mid-level features (template matching)

- Identifying objects in images still requires integrating a priori information about shapes and objects.
- A priori information can be modeled as generic templates to define specific objects.
- Template matching is useful to find a shape in the picture.

FEATURE EXTRACTION (SEGMENTATION)

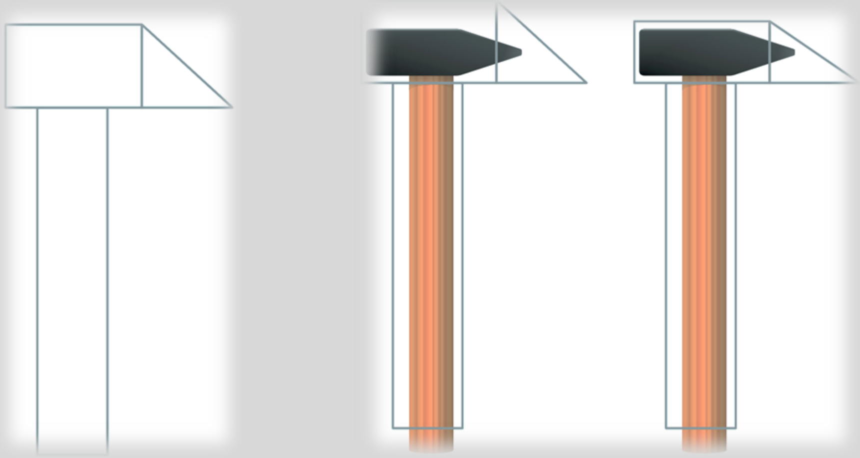
Source of the images: University of Central Florida Center for Research in Computer Vision, 2015; Khan & Leonardi, 2015

TEMPLATE MATCHING

The simplest and most obvious way to verify the presence of instances of the template within an image is placing the template at different locations on the image and comparing its values with the corresponding image pixels.

DEFORMABLE TEMPLATE

Deformable templates are parametric templates that can adapt with the input data. The figure shows a template of a hammer.



Source of the graphics: Course Book DLMBBD01, pp. 108-109

Oldest methods used in image processing:

– Discrete Fourier Transform (DFT),

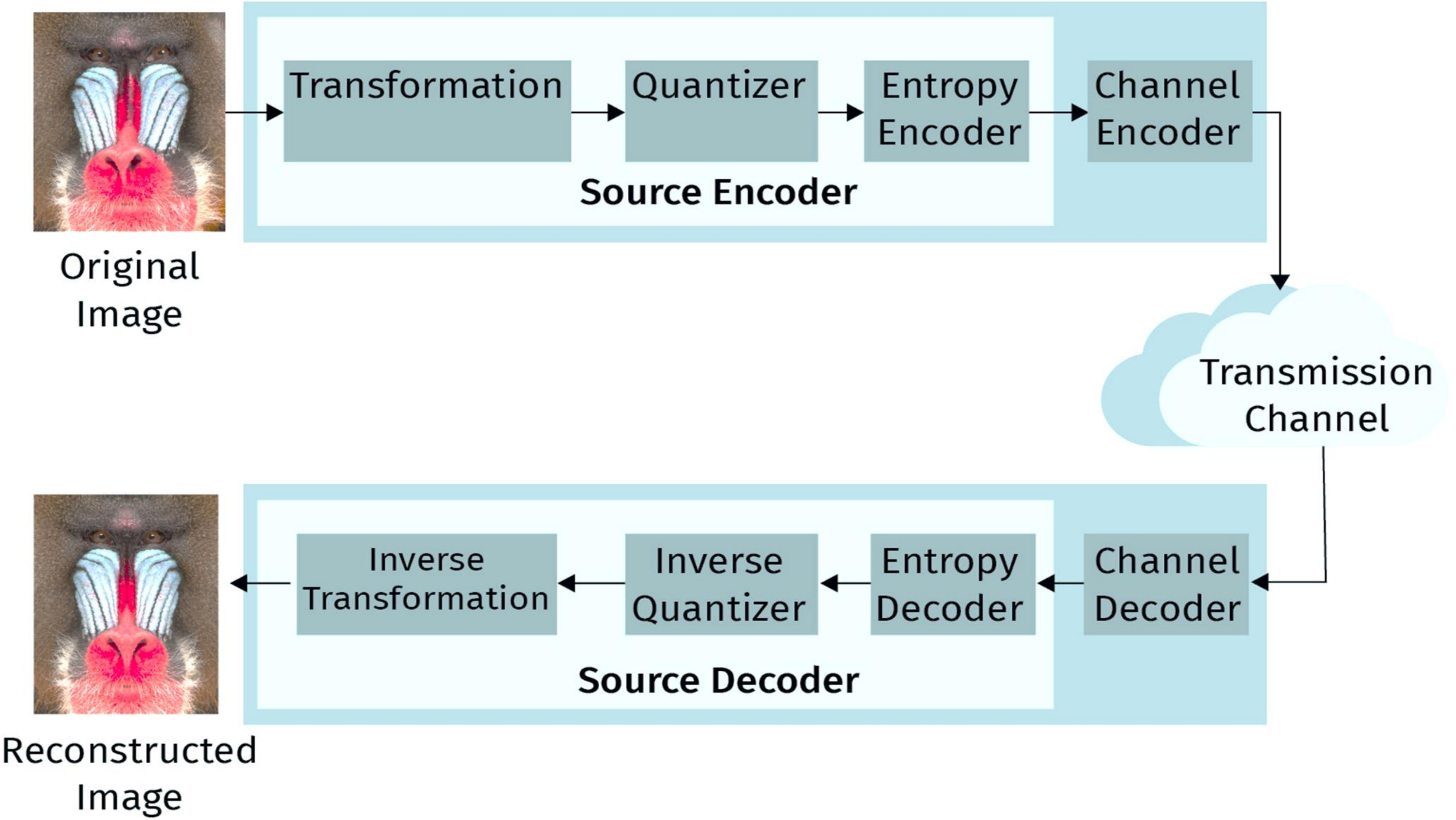
– Haar Transform

– Walsh-Hadamard Transforms

– Discrete Cosine Transform (DCT).

With the help of integral transforms, problems (or equations) in a certain domain get mapped onto another domain where they are easier to solve. After solving the problem in the second domain, the inverse transform must be applied to find the solution in the original domain.

COMPONENTS OF TRANSMISSION SYSTEM



Source of the graphic: Khayam, 2003

Image recognition:

– important application of machine learning and data analytics in the context of image processing

Image recognition is the automated recognition of objects (e.g., faces, places, digits, actions) in an image.

Training sets or training data:

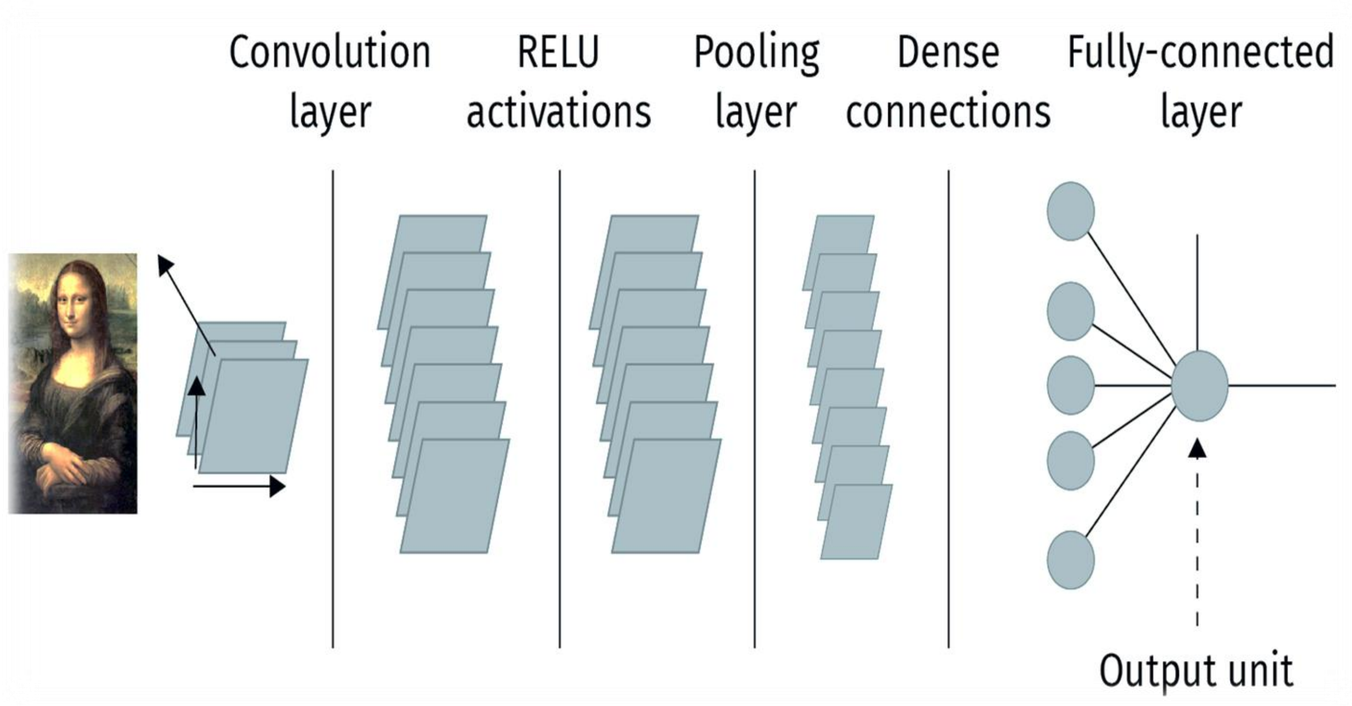
– pre-labeled images to train a smart recognition system

Deep learning models and convolutional neural networks (CNNs) outperform other machine learning methods.

- **Idea:** Network itself uses this convolution approach to extract features automatically.
- **Process**: Layers are put on one after another, implementing the convolution operator, and during the training process the network automatically learns which filters (kernels) would be most relevant to processing the image best.

Typical convolutional neural network parts:

- An input layer
- A series of successive convolution
- Pooling layers
- Neurons in each layer are connected via links to the next layer.



Source of the graphic: Pattanayak, 2017

https://poloclub.github.io/cnn-explainer/

https://distill.pub/2017/feature-visualization/

- **Definition:** Augmented reality (AR) is a technology that alters, in an interactive way, a user's view of their surrounding environment by adding extra information to physical objects to make them more informative.
- **Goal:** AR aims to integrate the real-world physical environment with virtual information in real-time. This virtual (or "augmented") information is generated by a computer and may give a wide range of perceptual information.

Examples:

- [ ] health
- [ ] manufacturing
- [ ] education
- [ ] advertising

EXPLOITING THE VISUAL: IMAGE RECOGNITION FOR AUGMENTED REALITY

Source of the graphic: Hiren, 2019

On completion of this unit, you will have learned …

- … different types of image representation.
- … image processing applications.
- … data compression techniques.
- … different approaches to image processing, including classic image processing and modern deep-learning-based approaches.
- … image transformation techniques and integral transforms.
- … the basics of convolutional neural networks.
- … the basics of augmented reality concepts and applications.

SESSION 4

TRANSFER TASK

TRANSFER TASK

Image processing has a wide range of applications in different industries. Describe and present one of its applications in transportation and how it could help to improve processes.

Explore https://roboflow.com/

Channel: https://www.youtube.com/c/roboflow

Notebooks: https://github.com/roboflow/notebooks

TRANSFER TASK
PRESENTATION OF THE RESULTS

Please present your results.
The results will be discussed in plenary.

1. The process of dividing an image into disjointed areas based on an image property such as color, intensity, texture, etc., is…
   a) edge detection
   b) segmentation
   c) feature extraction
   d) template matching

2. Which is the correct sequence of layers in a typical convolutional neural network?

   a) convolution, activation, pooling, dense connection, fully connected, output

   b) convolution, pooling, activation, dense connection, fully connected, output

   c) convolution, activation, dense connection, pooling, fully connected, output

   d) convolution, activation, dense connection, fully connected, output

3. For which image type will the quality not suffer when zoomed in?
    a) raster image
    b) bitmap image
    c) digital image
    d) vector image

LIST OF SOURCES

Hiren, S. (2019, March 8). *Mobile BI market: Global analysis, sales revenue, cost structure, forecasting 2019–2024* [article]. https://ourcryptojournal.com/mobile-bi-market-global-analysis-sales-revenue-cost-structure-forecast-ing-2019-2024/167457/

Khan, K., Mauro, M, & Leonardi, R. (2015). Multi-class semantic segmentation of faces. In: *Proceedings of the International Conference on Image Processing*, Quebec, 27 August. IEEE.

Khayam, S. A. (2003). *The discrete cosine transform (DCT): Theory and application.* Michigan State University.

Pattanayak, S. (2017). *Pro Deep Learning with TensorFlow: A mathematical approach to advanced artificial intelligence in Python.* Apress.

Smarandache, F. (n.d.). *Neutrosophic mathematical morphology for medical image* [powerpoint]. http://fs.unm.edu/ScPr/NeutrosophicMathematicalMorphology.pdf

Stanford Solar Center. (2015). *UV Light* [article]. http://solar-center.stanford.edu/about/uvlight.html

University of Central Florida Center for Research in Computer Vision. (2015). *Geosementic segmentation* [paper]. https://www.crcv.ucf.edu/projects/Geoseman-tic/