Computer Vision

Tushar B. Kute,
http://tusharkute.com

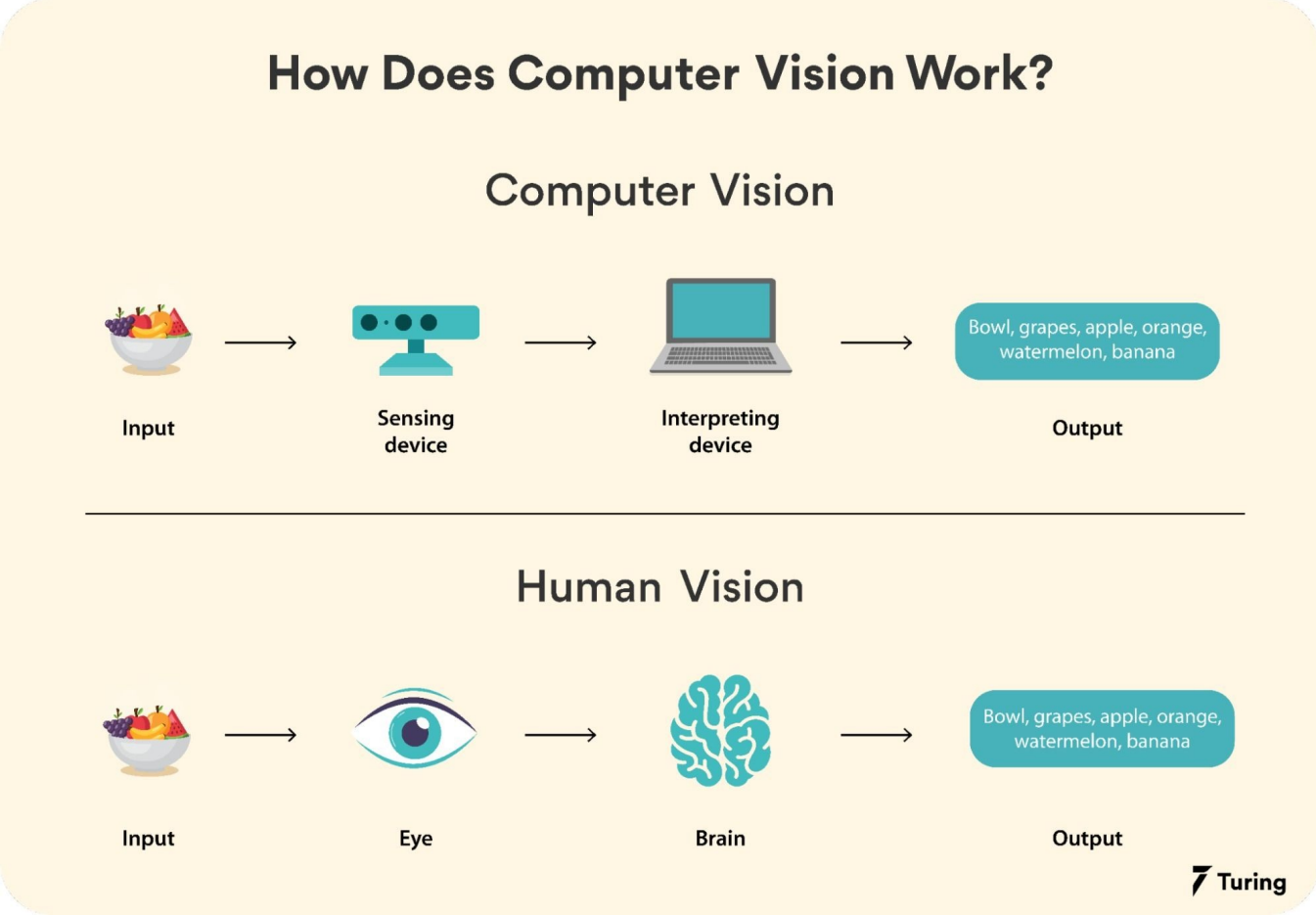
Computer Vision

- Computer vision is one of the fields of artificial intelligence that trains and enables computers to understand the visual world.
- Computers can use digital images and deep learning models to accurately identify and classify objects and react to them.
- Computer vision in AI is dedicated to the development of automated systems that can interpret visual data (such as photographs or motion pictures) in the same manner as people do.

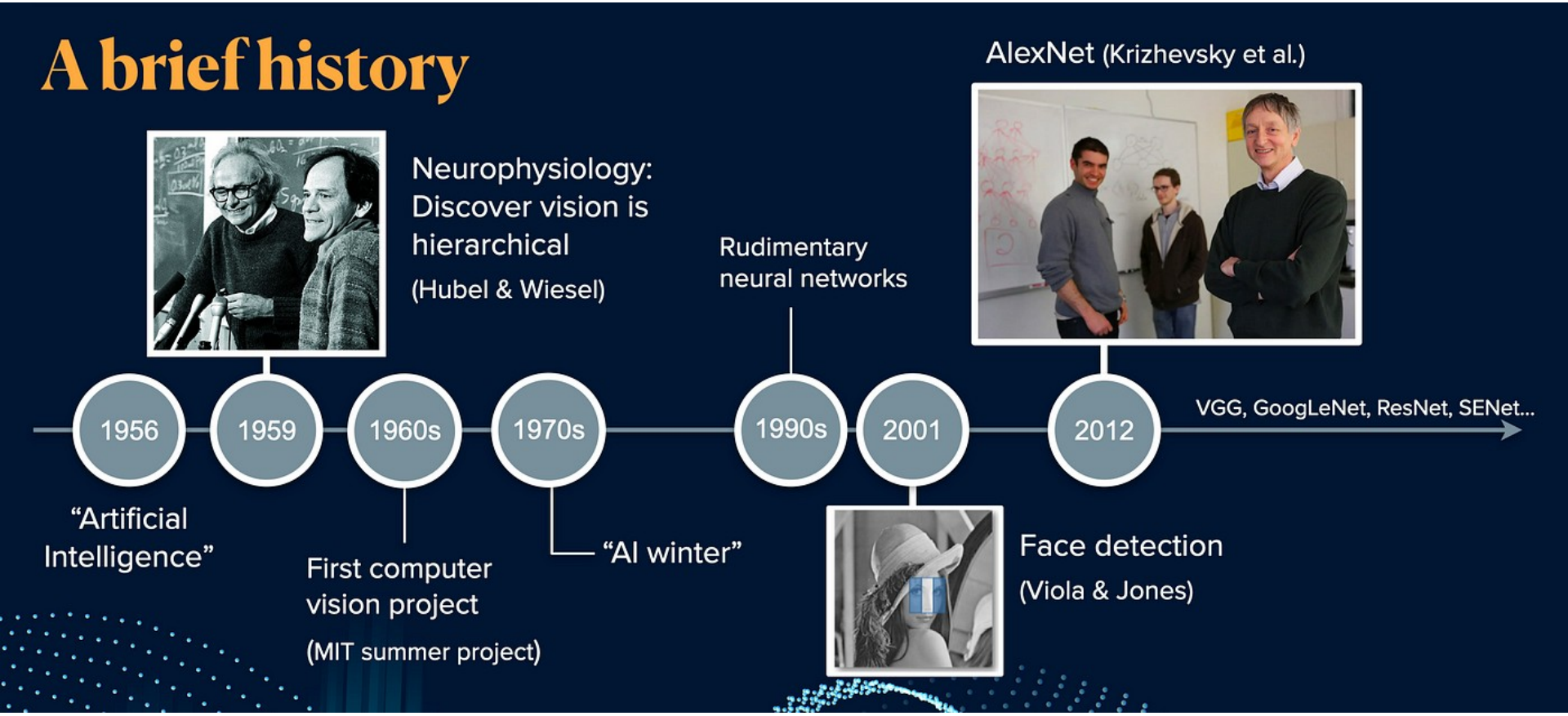
Computer Vision

- The idea behind computer vision is to instruct computers to interpret and comprehend images on a pixel-by-pixel basis.
- This is the foundation of the computer vision field. Regarding the technical side of things, computers will seek to extract visual data, manage it, and analyze the outcomes using sophisticated software programs.
- The amount of data that we generate today is tremendous i.e. 2.5 quintillion bytes of data every single day.
- This growth in data has proven to be one of the driving factors behind the growth of computer vision.

Computer Vision: How?

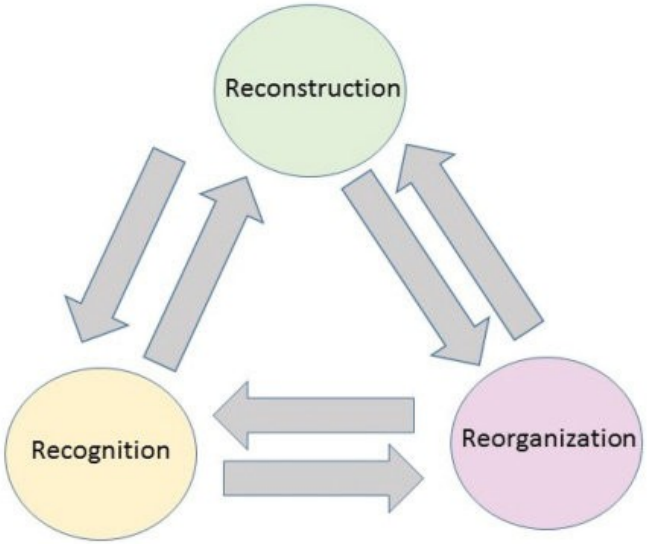


Computer Vision: History



Computer Vision and NLP

- The combination of natural language processing and computer vision involves three key interrelated processes: recognition, reconstruction, and reorganization.



Computer Vision and NLP

- Recognition:
  - This process involves assigning digital labels to objects within the image.
  - Examples of recognition are handwriting or facial recognition for 2D objects, and 3D assignments handle challenges such as moving object recognition which helps in automatic robotic manipulation.

Computer Vision and NLP

- Reconstruction:
  - This process refers to 3D scene rendering given inputs from particular visual images by incorporating multiple viewpoints, digital shading, and sensory depth data.
  - The outcome results in a 3D digital model that is then used for further processing.

Computer Vision and NLP

- Reorganization:
  - This process refers to raw pixel segmentation into data groups that represent the design of a pre-determined configuration.
  - Low-level vision tasks include corner detection, edges, and contours; while high-level tasks involve semantic segmentation, which can partly overlap with recognition processes.

Computer Vision and NLP: Relations

- Natural language processing tasks are deemed more technically diverse when compared to computer vision procedures.
- This diversification ranges from variable syntax identification, morphology and segmentation capabilities, and semantics to study abstract meaning.
- Complex tasks within natural language processing include direct machine translation, dialogue interface learning, digital information extraction, and prompt key summarisation.

Computer Vision Levels

- Computer vision is divided into three basic categories that are as following:
- Low-level vision:
  - includes process image for feature extraction.
- Intermediate-level vision:
  - includes object recognition and 3D scene Interpretation
- High-level vision:
  - includes conceptual description of a scene like activity, intention and behavior.

Applications

- Autonomous Vehicles:
  - Self-driving automobiles use CV systems to gather information regarding their surroundings and interpret that data to determine their next actions and behavior.
- Robotic Applications:
  - Manufacturing robotic machines using CV, 'view' and 'comprehend' their surroundings to perform their scheduled tasks.
  - In manufacturing, such systems inspect assembly items to determine faults and tolerance limits - simply by 'looking' at them as they traverse the production line.

Applications

- Image Search and Object Recognition:
  - Applications use CV data vision theory to identify specific objects within digital images, search through catalogs of product images, and extract information from photos.
- Facial Recognition:
  - Businesses and Government departments use facial recognition technology (that have adopted CV) to 'see' precisely what an individual is trying to gain access to.

Future

- Designing:
  - Within the area of home design, designer clothes, jewelry making, etc., customer systems can understand verbal or written requirements and thereby automatically convert these instructions to digital images for enhanced visualization.
- Describing Medical Images:
  - computer vision systems can be trained to identify more modest human ailments and use digital imagery in finer detail than human medical specialists.

Future

- Converting Sign Language: to speech or written text to assist the deaf and hard of hearing individuals in interacting with their surroundings. This enhanced capability can ensure their better integration within society.
- Surrounding Cognition: Constructing an intelligent system that 'sees' its surroundings and delivers a (recorded) spoken narrative. This outcome will be of use for visually impaired individuals.
- Converting Words to Images: Producing intelligent systems that convert spoken content to a digital image may assist people who do not talk and hear.

Thank you

@mitu_skillologies

@mITuSkillologies

@mitu_group

@mitu-skillologies

@MITUSkillologies

kaggle

@mituskillologies

Web Resources
https://mitu.co.in
http://tusharkute.com

@mituskillologies

contact@mitu.co.in

tushar@tusharkute.com