18CSE390T

Computer Vision

Topics	Hours
Introduction to Computer Vision - Image formation - Geometric primitives - 2D,3D Transformations - 3D to 2D Projection - Lighting, Reflectance and shading - Sampling and aliasing - Image processing Point operators - Pixel transforms - Color transforms - Histogram equalization - Linear filtering - Non Linear filtering - Fourier transforms - Two-dimensional Fourier transforms, Wiener filtering	9

Image Processing

- **Image:** An image is a two-dimensional function f(x,y), where x and y are the spatial (plane) coordinates, and the amplitude of f at any pair of coordinates (x,y) is called the intensity of the image at that level.
- If x,y and the amplitude values of f are finite and discrete quantities, we call the image a digital image. A digital image is composed of a finite number of elements called pixels, each of which has a particular location and value.
- The **digital image processing** deals with developing a digital system that performs operations on an digital image.
- Field of enhancing the images by fine tuning many parameters and features of the image.
- Image processing basically includes the following three steps:
  - Importing the image via image acquisition tools;
  - Analyzing and manipulating the image;
  - Output in which result can be altered image or report that is based on image analysis.

Pixel intensity value

$f(1,1) = 103$

Pixel location

rows columns

$f(645{:}650, 1323{:}1328) =$

83	82	82	82	82	82
82	82	82	81	81	81
82	82	81	81	80	80
82	82	81	80	80	79
80	79	78	77	77	77
80	79	78	78	77	77

$f(2724,2336) = 88$

Consider the following image (2724x2336 pixels) to be 2D function or a **matrix** with **rows** and **columns**

In 8-bit representation Pixel intensity values change between 0 (Black) and 255 (White)

Introduction to Computer Vision

- Providing vision or eyes to a machine
- A field of study that seeks to develop techniques to help computers “see” and understand the content of digital images such as photographs and videos.
- Purpose of computer vision is to program a computer to "understand" a scene or features in an image.
- Focus on extracting information from the input images or videos to have proper understanding of them to predict the visual input from human brain.
- *At an abstract level, the goal of computer vision problems is to use the observed image data to infer something about the world.*

Goals of Computer Vision

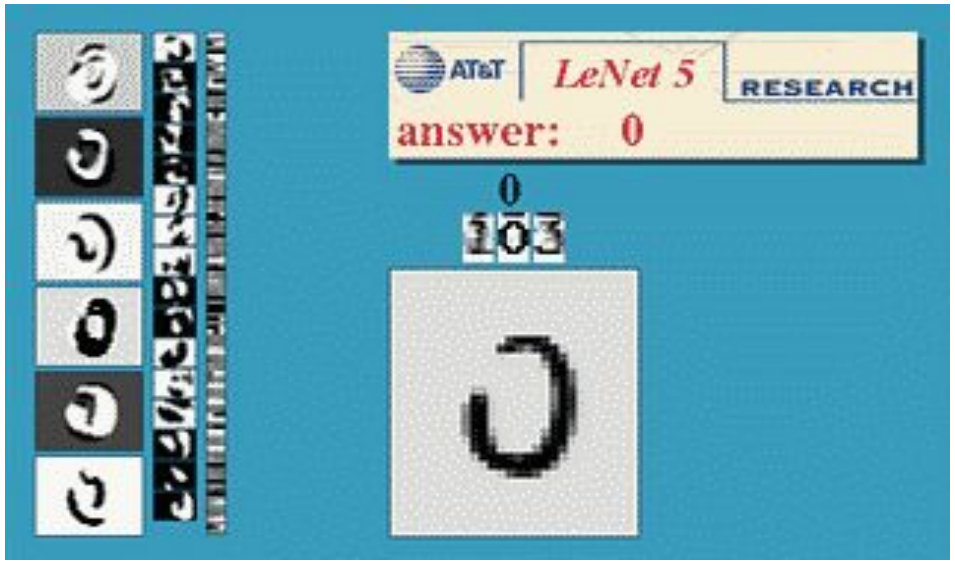
- The **detection, segmentation, localization, and recognition of certain objects** in images (e.g., human faces)
- The **evaluation of results** (e.g., segmentation, registration)
- **Registration of different views** of the same scene or object
- **Tracking an object** through an image sequence
- **Mapping a scene to a three-dimensional model of the scene**; such a model might be used by **a robot to navigate the imaged scene**
- **Estimation of the three-dimensional pose of humans and their limbs**
- **Content-based image retrieval** - Searching for digital images by their content

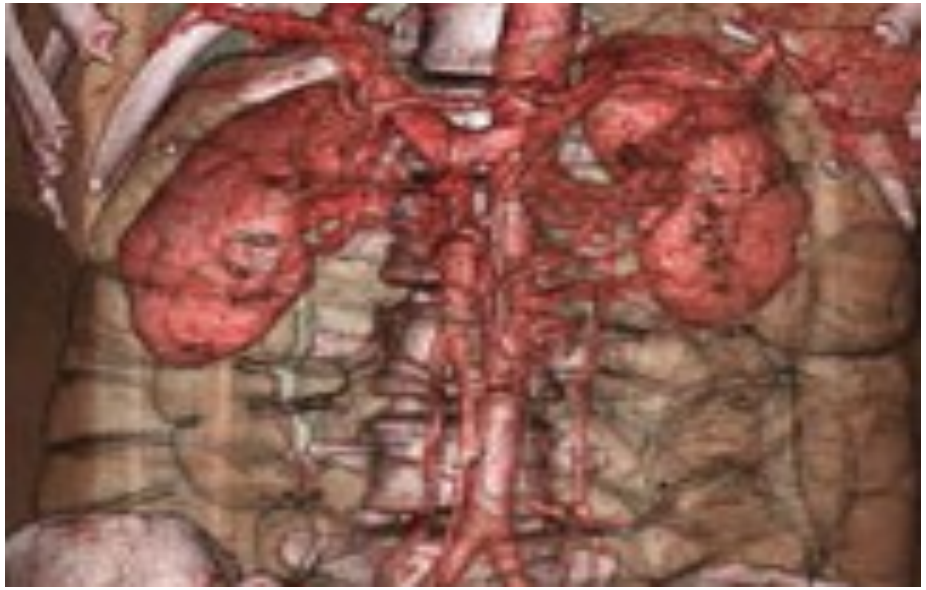
Applications of computer vision

- **Facial recognition:** Computer vision has enabled machines to detect face images of people to verify their identity. Initially, the machines are given input data images in which computer vision algorithms detect facial features and compare them with databases of fake profiles. Popular social media platforms like Facebook also use facial recognition to detect and tag users. Further, various government spy agencies are employing this feature to identify criminals in video feeds.
- **Healthcare and Medicine:** Computer vision has played an important role in the healthcare and medicine industry. Traditional approaches for evaluating cancerous tumors are time-consuming and have less accurate predictions, whereas computer vision technology provides faster and more accurate chemotherapy response assessments; doctors can identify cancer patients who need faster surgery with life-saving precision.
- **Self-driving vehicles:** Computer vision technology has also contributed to its role in self-driving vehicles to make sense of their surroundings by capturing video from different angles around the car and then introducing it into the software. This helps to detect other cars and objects, read traffic signals, pedestrian paths, etc., and safely drive its passengers to their destination.
- **Optical character recognition (OCR) :** Optical character recognition helps us extract printed or handwritten text from visual data such as images. Further, it also enables us to extract text from documents like invoices, bills, articles, etc.
- **Machine inspection:** Computer vision is vital in providing an image-based automatic inspection. It detects a machine's defects, features, and functional flaws, determines inspection goals, chooses lighting and material-handling techniques, and other irregularities in manufactured products.
- Retail (e.g., automated checkouts): Computer vision is also being implemented in the retail industries to track

- **3D model building:** 3D model building or 3D modeling is a technique to generate a 3D digital representation of any object or surface using the software. In this field also, computer vision plays its role in constructing 3D computer models from existing objects. Furthermore, 3D modeling has a variety of applications in various places, such as Robotics, Autonomous driving, 3D tracking, 3D scene reconstruction, and AR/VR.
- **Medical imaging:** Computer vision helps medical professionals make better decisions regarding treating patients by developing visualization of specific body parts such as organs and tissues. It helps them get more accurate diagnoses and a better patient care system. E.g., Computed Tomography (CT) or Magnetic Resonance Imaging (MRI) scanner to diagnose pathologies or guide medical interventions such as surgical planning or for research purposes.
- **Automotive safety:** Computer vision has added an important safety feature in automotive industries. E.g., if a vehicle is taught to detect objects and dangers, it could prevent an accident and save thousands of lives and property.
- **Surveillance:** It is one of computer vision technology's most important and beneficial use cases. Nowadays, CCTV cameras are almost fitted in every place, such as streets, roads, highways, shops, stores, etc., to spot various doubtful or criminal activities. It helps provide live footage of public places to identify suspicious behavior, identify dangerous objects, and prevent crimes by maintaining law and order.
- **Fingerprint recognition and biometrics:** Computer vision technology detects fingerprints and biometrics to validate a user's identity. Biometrics deals with recognizing persons based on physiological characteristics, such as the face, fingerprint, vascular pattern, or iris, and behavioral traits, such as gait or speech. It combines

Variety of real-world applications

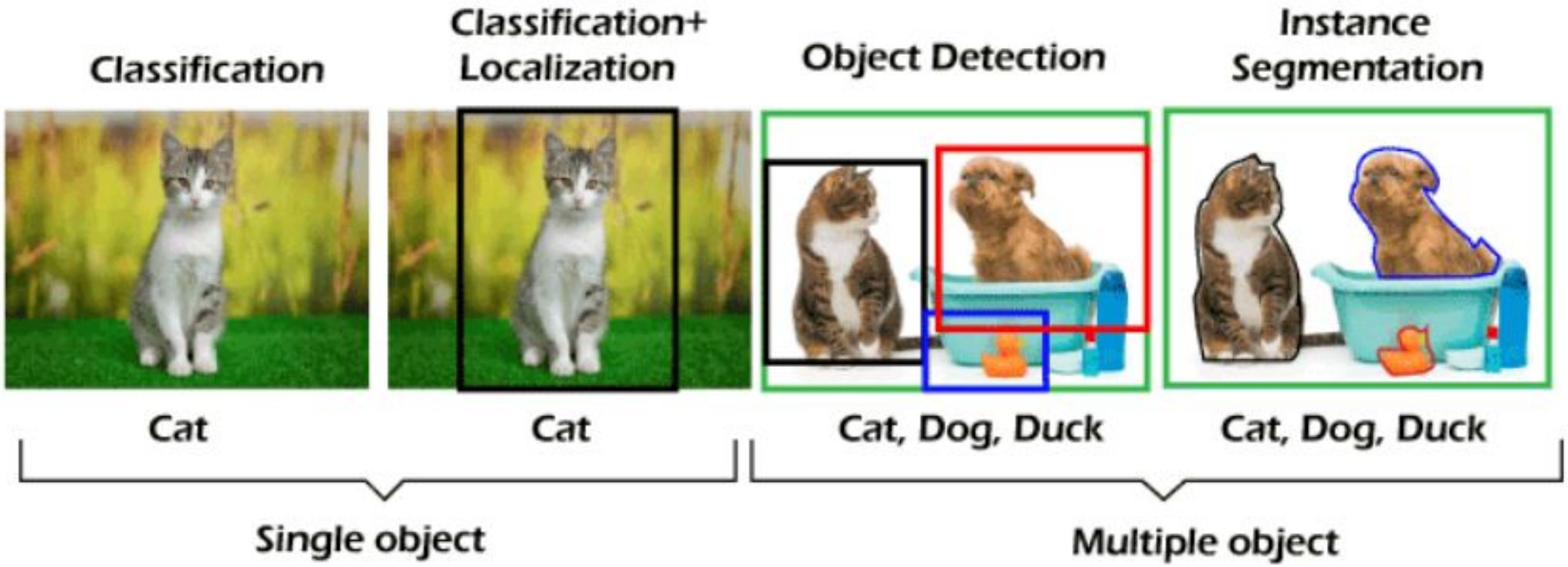




Tasks for which computer vision can be used

- **Image Classification:** Image classification is a computer vision technique used to classify an image, such as whether an image contains a dog, a person's face, or a banana. It means that with image classification, we can accurately predict the class of an object present in an image.
- **Object Detection:** Object detection uses image classification to identify and locate the objects in an image or video. With such detection and identification technique, the system can count objects in a given image or scene and determine their accurate location, along with their labelling. For example, in a given image, there is one person and one cat, which can be easily detected and classified using the object detection technique.
- **Object Tracking:** Object tracking is a computer vision technique used to follow a particular object or multiple items. Generally, object tracking has applications in videos and real-world interactions, where objects are firstly detected and then tracked to get observation. Object tracking is used in applications such as Autonomous vehicles, where apart from object classification and detection such as pedestrians, other vehicles, etc., tracking of real-time motion is also required to avoid accidents and follow the traffic rules.
- **Semantic Segmentation:** Image segmentation is not only about detecting the classes in an image as image classification. Instead, it classifies each pixel of an image to specify what objects it has. It tries to determine the role of each pixel in the image.
- **Instance Segmentation:** Instance segmentation can classify the objects in an image at pixel level as similar to semantic segmentation but with a more advanced level. It means Instance Segmentation can classify similar types of objects into different categories. For example, if visual consists of various cars, then with semantic segmentation, we can tell that there are multiple cars, but with instance segmentation, we can label them according to their colour, shape, etc.

Computer Vision Tasks



Computer Vision Process

1. Capturing an Image

A computer vision software or application always includes a digital camera or CCTV to capture the image. So, firstly it captures the image and puts it as a digital file that consists of Zero and one's.

2. Processing the image

In the next step, different CV algorithms are used to process the digital data stored in a file. These algorithms determine the basic geometric elements and generate the image using the stored digital data.

3. Analyzing and taking required action

Finally, the CV analyses the data, and according to this analysis, the system takes the required action for which it is designed.

Image Formation

- Before analyzing and manipulating images, we need to establish a **vocabulary for describing the geometry of a scene**.

- Need to understand the **image formation process** that produced a particular image given a set of
  - **Lighting conditions,**
  - **Scene geometry,**
  - **Surface properties, and**
  - **Camera optics.**