

Computer Vision Basics Interview Q&A

Question 1: What is computer vision?

Answer: Computer vision is a multidisciplinary field that empowers machines to comprehend and interpret visual information, mimicking human visual perception. It involves the development of algorithms and systems to automatically analyze, understand, and make decisions based on images or videos.

Question 2: Explain the difference between image processing and computer vision.

Answer:

- **Image Processing:** Involves manipulating image data to enhance certain features, such as filtering or contrast adjustments, primarily focusing on improving visual quality.
 - **Computer Vision:** Goes beyond by interpreting visual data, enabling machines to recognize patterns, objects, and scenes, and make decisions based on the extracted information.
-

Question 3: What are pixels, and how do they relate to image resolution?

Answer:

- **Pixels:** Are the smallest units in a digital image, representing dots of color.
 - **Image Resolution:** Is determined by the total number of pixels, where higher resolution means more pixels, providing greater detail and clarity.
-

Question 4: Describe the RGB color model.

Answer:

- **RGB Color Model:** Represents colors using combinations of red, green, and blue intensities.

- **Representation:** Each pixel is defined by three values (0-255), indicating the intensity of red, green, and blue. For example, (255, 0, 0) represents pure red.
-

Question 5: What is image segmentation?

Answer:

- **Image Segmentation:** Involves dividing an image into segments based on characteristics like color or intensity.
 - **Purpose:** Facilitates object isolation within an image for further analysis, enhancing the interpretation of complex scenes.
-

Question 6: Explain edge detection and its importance.

Answer:

- **Edge Detection:** Identifies boundaries between objects in an image based on intensity changes.
 - **Importance:** Essential for recognizing object structures, aiding in tasks like object recognition and feature extraction.
-

Question 7: What are feature detectors and descriptors?

Answer:

- **Feature Detectors:** Algorithms identifying key points (corners, edges) in an image.
 - **Descriptors:** Quantify local information around key points, crucial for matching and comparing features between images.
-

Question 8: Define convolution in the context of image processing.

Answer:

- **Convolution:** Mathematical operation using a kernel to combine pixel values, essential for tasks like blurring or edge detection.
 - **Kernel:** A small matrix determining the convolution operation's impact on neighboring pixels.
-

Question 9: What is a Convolutional Neural Network (CNN)?

Answer:

- **CNN:** A specialized neural network for visual data processing, employing convolutional layers to automatically learn hierarchical features from images.
 - **Layers:** Includes convolutional and pooling layers for feature extraction and downsampling.
-

Question 10: How does image classification work?

Answer:

- **Training:** Involves feeding labeled data to a CNN to learn associations between features and class labels.
 - **Testing:** The trained CNN predicts class labels for new, unseen images based on learned features.
-

Question 11: Discuss object detection and its challenges.

Answer:

- **Object Detection:** Identifies and locates multiple objects within an image, providing spatial information.
 - **Challenges:** Include scale variation, occlusion, viewpoint variation, and cluttered backgrounds.
-

Question 12: Explain the concept of image augmentation.

Answer:

- **Image Augmentation:** Involves artificially diversifying training datasets by applying transformations like rotation, scaling, or brightness changes.
 - **Purpose:** Enhances model generalization and robustness by exposing it to a variety of data variations.
-

Question 13: What are some common image preprocessing techniques?

Answer:

- **Normalization:** Scaling pixel values to a standard range.
 - **Resizing, Cropping:** Adjusting image dimensions and removing unnecessary portions.
 - **Grayscale Conversion:** Reducing images to a single channel.
 - **Noise Reduction:** Applying filters to reduce image noise.
-

Question 14: Describe how face detection works.

Answer:

- **Face Detection:** Involves Haar cascades, ROI selection, feature extraction, and classification to locate and identify faces within an image.
 - **Applications:** Crucial in facial recognition, security, and photography.
-

Question 15: What is the role of pooling in CNNs?

Answer:

- **Pooling:** A downsampling operation in CNNs to reduce spatial dimensions while retaining important features.
 - **Types:** Max pooling (retaining maximum values) and average pooling (calculating average values).
-

Question 16: Explain the concept of transfer learning in computer vision.

Answer:

- **Transfer Learning:** Adapting a pretrained model to a new task by leveraging learned features.
 - **Steps:** Pretraining on a large dataset, using features for a new model, and fine-tuning on task-specific data.
-

Question 17: Discuss the applications of computer vision in autonomous vehicles.

Answer:

- **Object Detection:** Identifying and tracking pedestrians, vehicles, and obstacles.
 - **Lane Detection:** Recognizing road lanes for proper vehicle positioning.
 - **Traffic Sign Recognition:** Understanding and obeying traffic signs.
-

Question 18: What are GANs, and what is their role in image generation?

Answer:

- **GANs (Generative Adversarial Networks):** Neural networks with a generator and discriminator, trained simultaneously.
 - **Role:** Used for image generation, style transfer, and data augmentation.
-

Question 19: Explain the concept of optical character recognition (OCR).

Answer:

- **OCR (Optical Character Recognition):** Technology converting scanned documents or images into editable and searchable text.
 - **Steps:** Text detection, character recognition, and text extraction.
-

Question 20: What is a depth map, and how is it created?

Answer:

- **Depth Map:** An image containing information about the distance or depth of objects in a scene.
- **Creation:** Often generated through stereo vision, comparing disparities between images captured from different perspectives.