Below are **the MOST LIKELY and COMMON viva questions** for each of the 12 CVIP syllabus topics — **simple, exam-ready, crisp answers** that will help you score full marks in your external lab.

✅ 1. Image Preprocessing (Resizing, Filtering, Thresholding)

Q1. What is image preprocessing?

A: Image preprocessing is the process of improving the raw image quality to make it suitable for further processing like detection and recognition.

Q2. Why do we resize images?

A: To reduce computational cost, maintain uniform input size for ML models, and speed up processing.

Q3. What is filtering?

A: Filtering smooths or sharpens images by modifying pixel values.
Examples: Gaussian blur (smooth), Median filter (remove salt & pepper noise).

Q4. What is thresholding?

A: Thresholding converts a grayscale image to binary by assigning pixel = 0 or 255 based on a threshold value.

Q5. Difference between global and adaptive thresholding?

- **Global:** One threshold for the entire image.
- **Adaptive:** Threshold varies for different regions.

✅ 2. Edge Detection (Sobel, Canny, Laplacian)

Q1. What is edge detection?

A: Technique to detect boundaries in an image where there is sudden intensity change.

Q2. What is Sobel operator?

A: Uses two kernels (Gx, Gy) to detect edges in horizontal and vertical directions.

Q3. Why is Canny considered the best?

A: Because it uses noise removal, gradient calculation, non-max suppression, and hysteresis thresholding → strong & accurate edges.

Q4. What is Laplacian?

A: Second-order derivative operator that detects all edges without direction.

Q5. Difference between Sobel and Canny?

- **Sobel:** Basic gradient method.
- **Canny:** Multi-stage, more accurate, less noise sensitive.

✅ 3. Object Detection Using Contours

Q1. What is a contour?

A: A curve joining continuous points of the same intensity.

Q2. Why do we use contours?

A: For shape detection, object counting, segmentation, and boundary extraction.

Q3. What is a bounding box?

A: A rectangular box drawn around the detected contour.

Q4. How do we find contours in OpenCV?

A: Using cv2.findContours() after thresholding or edge detection.

Q5. Difference between contour detection and edge detection?

- **Edges:** Finds sharp intensity changes.
- **Contours:** Track shape based on edges or binary image.

✅ 4. Feature Extraction (HOG, SIFT, ORB)

Q1. What is feature extraction?

A: Conversion of image into numerical features to detect patterns.

Q2. What is HOG?

A: Histogram of Oriented Gradients — captures edge direction patterns, used for human detection.

Q3. What is SIFT?

A: Scale-Invariant Feature Transform — detects keypoints invariant to scale, rotation, illumination.

Q4. What is ORB?

A: Oriented FAST and Rotated BRIEF — FAST keypoints + BRIEF descriptors, faster & free alternative to SIFT.

Q5. Which is fastest: SIFT, SURF, or ORB?

A: ORB.

✅ 5. Face Detection (Haar Cascades / DNN)

Q1. What is Haar Cascade?

A: A machine learning–based classifier that uses Haar-like features + Adaboost + Cascade structure.

Q2. Why is Haar Cascade fast?

A: Because it rejects non-face regions early using cascaded stages.

Q3. What is DNN face detection?

A: Deep learning model (e.g., OpenCV's ResNet SSD) that detects faces more accurately than Haar.

Q4. Difference between Haar and DNN?

- **Haar:** Fast, but less accurate.
- **DNN:** Slower but highly accurate and robust.

Q5. What input image format do Haar cascades require?

A: Grayscale.

✅ 6. Train SVM / KNN for Image Classification

Q1. What is SVM?

A: Support Vector Machine — finds the best hyperplane that separates classes.

Q2. What is KNN?

A: A non-parametric classifier that predicts label based on nearest k neighbors.

Q3. Which is faster in training: SVM or KNN?

A: SVM (KNN is slow at prediction).

Q4. Why do we flatten images for SVM?

A: SVM requires 1D feature vectors.

Q5. What is hyperparameter 'k'?

A: The number of neighbors used in KNN classification.

✅ 7. Decision Tree Classifier

Q1. What is a decision tree?

A: A tree-like model where nodes represent decisions based on features.

Q2. What is Gini impurity?

A: A measure of how mixed the classes are at a node.

Q3. What is entropy?

A: A measure of uncertainty; used in information gain splitting.

Q4. What is overfitting?

A: When the model memorizes training data; fixed using pruning.

Q5. What is pruning?

A: Removing unnecessary branches to reduce complexity.

✅ 8. Logistic Regression (Binary Classification)

Q1. Is logistic regression linear or nonlinear?

A: Linear classifier, but outputs probability using a sigmoid.

Q2. What is the sigmoid function?

A: Converts values into probabilities (0–1).

Q3. Why is it called "regression"?

A: Because it predicts probability, then converts to class.

Q4. What is threshold value?

A: Default 0.5 to classify positive or negative.

Q5. What is cost function?

A: Log-loss.

✅ 9. PCA (Principal Component Analysis)

Q1. What is PCA?

A: A technique to reduce dimensionality while preserving maximum variance.

Q2. Why use PCA?

A: Removes redundancy, speeds up training, improves visualization.

Q3. What does PCA return?

A: Principal components (new axes with maximum variance).

Q4. What is eigenvalue?

A: Amount of variance captured by its eigenvector.

Q5. What is the maximum number of principal components?

A: Equal to the number of features.

✅ 10. Simple Neural Network (Keras)

Q1. What is a neural network?

A: A model inspired by the human brain with layers of neurons.

Q2. What are the types of layers?

- Input

- Hidden
- Output

Q3. What is activation function?

A: Function applied to neuron output (ReLU, Sigmoid, Softmax).

Q4. What is loss function?

A: Measures error between prediction and actual value.

Q5. What is an epoch?

A: One complete pass through the training dataset.

✅ 11. CNN for MNIST Digit Recognition

Q1. Why use CNN instead of ANN?

A: CNN automatically extracts features using filters.

Q2. What is a convolution layer?

A: Applies filters to extract edges, textures, shapes.

Q3. What is pooling?

A: Downsampling to reduce dimensions (max-pooling is common).

Q4. Why MNIST is used?

A: It's a standard benchmark for digit recognition.

Q5. What accuracy does CNN normally achieve on MNIST?

A: 98–99%.

✅ 12. Real-Time Emotion Recognition (Webcam + Pretrained Model)

Q1. What model is used?

A: CNN-based FER models (FER2013, DeepFace, emotion-FER).

Q2. What preprocessing is needed?

A: Face detection → grayscale → resize → normalization.

Q3. How does the system work?

1. Capture frame
2. Detect face
3. Crop + preprocess
4. Predict emotion
5. Display label

Q4. Why use pretrained models?

A: Training emotion datasets requires huge data & GPU.

Q5. Common output emotions?

A: Happy, Sad, Angry, Fear, Neutral, Surprise, Disgust.

✅ **Want ultra-short one-line answers for last-minute revision?**

Just tell me — I can generate a **one-page cheat sheet** for your viva.