Assignment 4: Continuous Skip-Gram Model, Image Segmentation, and Continuous Bag-of-Words Model

Question 1: Define Continuous Skip-Gram Model

* Definition: Neural network-based word embedding model in NLP, part of the Word2Vec framework by Mikolov et al. (2013).
* Objective: Predicts surrounding context words for a target word, capturing semantic relationships.
* Key Features:
* Goal: Maximize probability of context words given a target word.
* Architecture: Input layer (target word), hidden layer (word embeddings), output layer (context probability distribution).
* Training: Adjusts word vectors to capture semantic similarities (e.g., "king" and "queen").
* Applications:

* Text classification
* Information retrieval
* Sentiment analysis

Question 2: Explain Image Segmentation and Object Detection

* Image Segmentation: Partitions an image into segments to identify objects.
* Types:
* Semantic Segmentation: Classifies each pixel (e.g., road, car).
* Instance Segmentation: Differentiates between instances of the same class.
* Techniques:
* Thresholding
* Clustering (e.g., K-means)
* Deep Learning (e.g., Fully Convolutional Networks)
* Object Detection: Identifies and localizes objects with bounding boxes and class labels.
* Techniques:
* Traditional: HOG, Haar Cascades

* Deep Learning: YOLO (real-time detection), Faster R-CNN (efficient detection)
* Applications:
* Autonomous vehicles
* Video surveillance
* Robotics
* Augmented reality

Question 3: Discuss Continuous Bag-of-Words (CBOW) Model
* Definition: Word embedding model in Word2Vec, predicts target word based on context words.
* Key Features:
* Objective: Maximize probability of target word given context words.
* Architecture: Input (context words), hidden layer (word embeddings), output (target word probability).
* Training: Adjusts embeddings to improve target word prediction.
* Applications: Similar to Skip-Gram-used

in sentiment analysis, text classification, recommendation systems.

* Comparison with Skip-Gram:
* Input: CBOW uses context words; Skip-Gram uses target word.
* Performance: CBOW is faster and better for smaller datasets; Skip-Gram performs well on larger datasets with rare words.