Dataset Construction - Annotation Guidelines

- Each figure and its corresponding caption must have a separate bounding box.
  - Figures should be assigned to exactly 1 of the 10 predefined classes .
  - Captions are always assigned the class "Caption".
  - Ensure no overlap between figure and caption annotations.
- Bounding Box Rules
  - Draw tight bounding boxes around each figure and its caption.
  - The caption box should cover only the text of the caption, not surrounding text.
  - The figure box should include only the visual content of the figure, avoiding page borders or surrounding text.
- Classifying Figures
  - Carefully examine the content and concept behind each figure.
  - Assign the most appropriate class from the predefined categories listed below.
  - If a figure could belong to multiple categories, choose the most dominant or relevant one.
- Subfigures and Complex Figures
  - If a figure consists of multiple subfigures labeled as (a), (b), (c), etc., annotate the entire figure as one bounding box.
  - If subfigures have separate captions, annotate them individually with their respective captions.
- For the Algorithms Code or Flowchart class, ensure that only the code or flowchart is included in the bounding box, excluding body text explanations.
- Do not include surrounding text or unrelated parts of the page in the bounding box.
- Do not annotate tables or equations; this task is only for figures and captions.
- There should be no overlapping or duplicate annotations.
- **Class Definitions**: Each figure must be assigned exactly one of the following classes:
  - Caption: Text that describes a figure.
    - Example: "Figure 3: Architecture of the proposed model."
  - **Pipeline**: Figures representing entire processing workflows, often spanning multiple steps and modules. Figures illustrating a high level overview of multi-step processes, workflows, or methodologies without detailed representations of each component.
    - Examples: End-to-end ML pipeline diagrams
    - Examples: Overview of the object detection process using YOLO
  - **Model Architecture**: Figures depicting the structural design of machine learning or deep learning models.
    - Examples: Transformer model, LSTM or YOLO architectural diagram
  - **Auxiliary Diagrams**: Schematic representations, flowcharts, or conceptual illustrations.

- Examples: System design diagrams, logic flow representations.
- **Real Images**: Photographic images or realistic visual content extracted from real-world sources.
  - Examples: Images from datasets, captured photographs, images of people, animals, places or objects.
- **Illustrations**: Figures providing explanatory visual aids for a concept or process.
  - Examples: Illustrative sketches, educational examples, artistic depictions.
- **Graph Plots**: Graphs, charts and mathematical plots.
  - Examples: Line graphs, bar charts, scatter plots, histograms.
- **Model Performance and Metrics**: Figures showing model evaluation results, benchmarking, and performance graphs.
  - Examples: Precision-recall curves, accuracy vs. epochs graphs, performance tables.
- **Statistics and Analysis**: Figures containing statistical results, experimental comparisons, or analytical visualizations.
  - Examples: Performance comparison graphs, confusion matrices, regression analysis plots.
- **Algorithms/Code/Flowchart**: Figures containing algorithmic representations, such as code snippets, pseudo code, or flowcharts.
  - Examples: Code blocks (e.g., Python, C++, pseudo code), Flowcharts detailing algorithmic steps, Structured representations of an algorithm's execution flow.