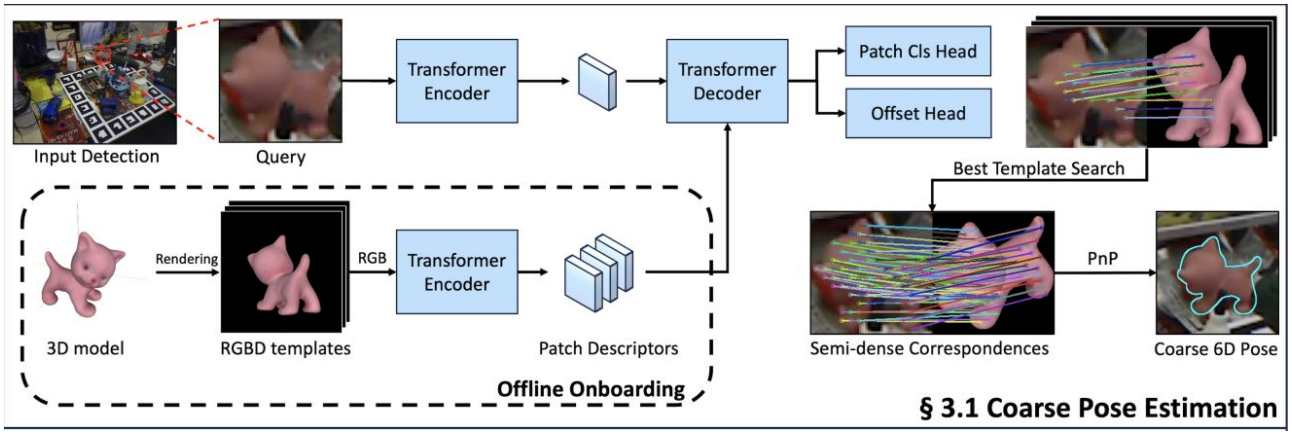
Moon, S., Son, H., Hur, D., & Kim, S. (2025). Co-op: Correspondence-based novel object pose estimation. In CVPR 2025.

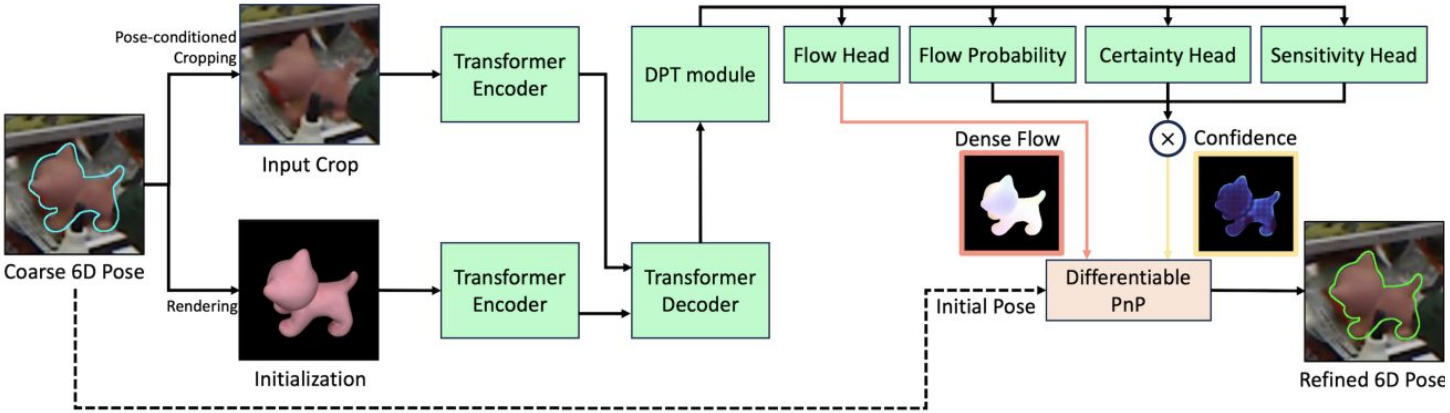
Methodology Summary

- Coarse Pose Estimation
  - Estimate semi-dense correspondences between input and small amount of rendered templates.
  - Apply PnP algorithm on correspondences to obtain an initial (coarse) pose estimation.
  - Uses a hybrid of classification+regression to improve generalization and reduce template count.



Methodology Summary

- Pose Refinement
  - Use a render-and-compare strategy with probabilistic dense correspondences (article refers to it as ‘flow’).
  - Train the model from end-to-end with a differentiable PnP layer to refine the coarse pose.
  - Helps deal with objects being occluded/obstructed from view and texture-less objects.



Technical Problems

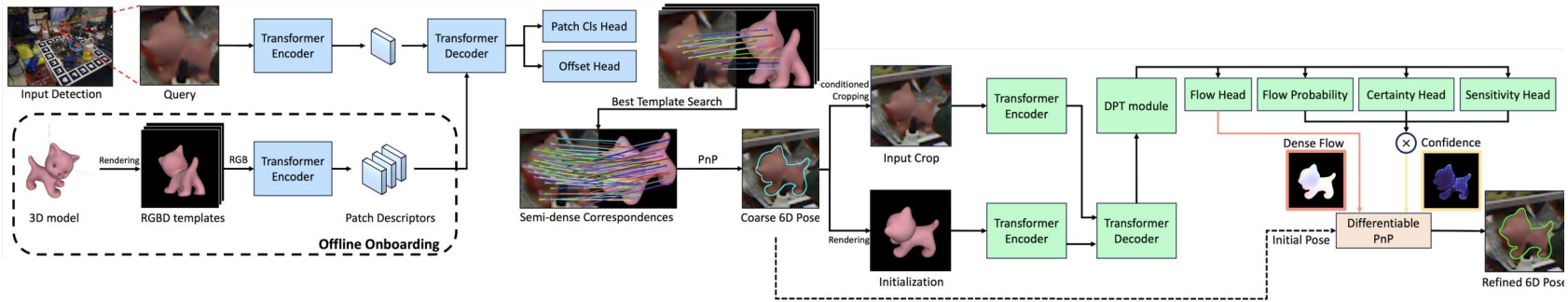
- Data and Training Overhead
  - Traditional deep learning needs large, object-specific datasets and retraining for each new object.
- Inefficiency of Template Matching
  - Model-based methods compare against many rendered templates, slowing computation.
- Dependency on Segmentation
  - Other methods rely on clear, accurate segmentation masks, making them error-prone if an imperfect mask is provided.
- Challenging Cases
  - Texture-less and occluded objects often cause failures.

Solutions to Technical Problems

- Data and Training Overhead
  - Co-op requires no fine-tuning or extra training.
- Inefficiency of Template Matching
  - Co-op uses far fewer templates while maintaining high accuracy.
- Dependency on Segmentation
  - Co-op directly predicts the object’s coarse pose from the entire image using a hybrid representation (patch-level classification + offset regression).
- Challenging Cases
  - Co-op’s pose refinement handles these cases via probabilistic correspondences and confidence learning.

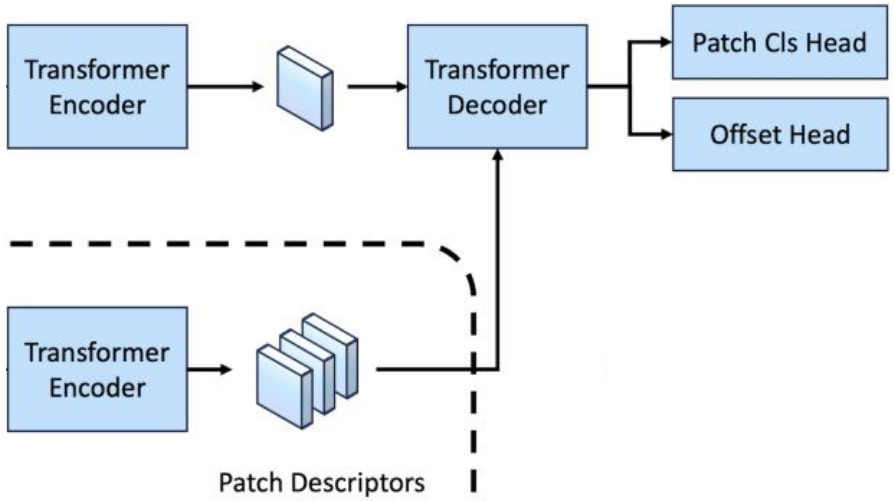
Novel Contributions

- Novel Unseen Object Pose Estimation Framework
  - Two-stage system achieving superior accuracy on all seven BOP Challenge datasets without retraining for new objects.



Novel Contributions

- Hybrid Representation for Coarse Estimation
  - Combines patch-level classification and offset regression for fast, accurate, and generalizable pose estimation with few templates.



Novel Contributions

- Probabilistic Flow for Pose Refinement
  - Refines poses via dense probabilistic correspondences, learning match reliability to handle occlusion and texture-less objects.