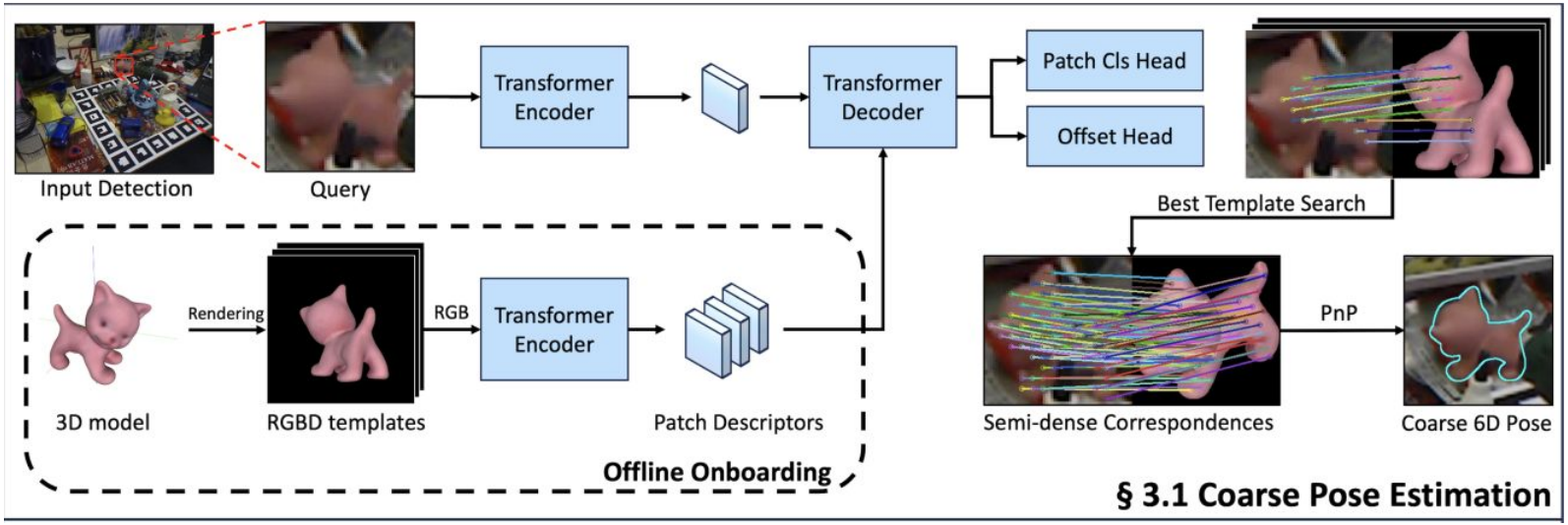
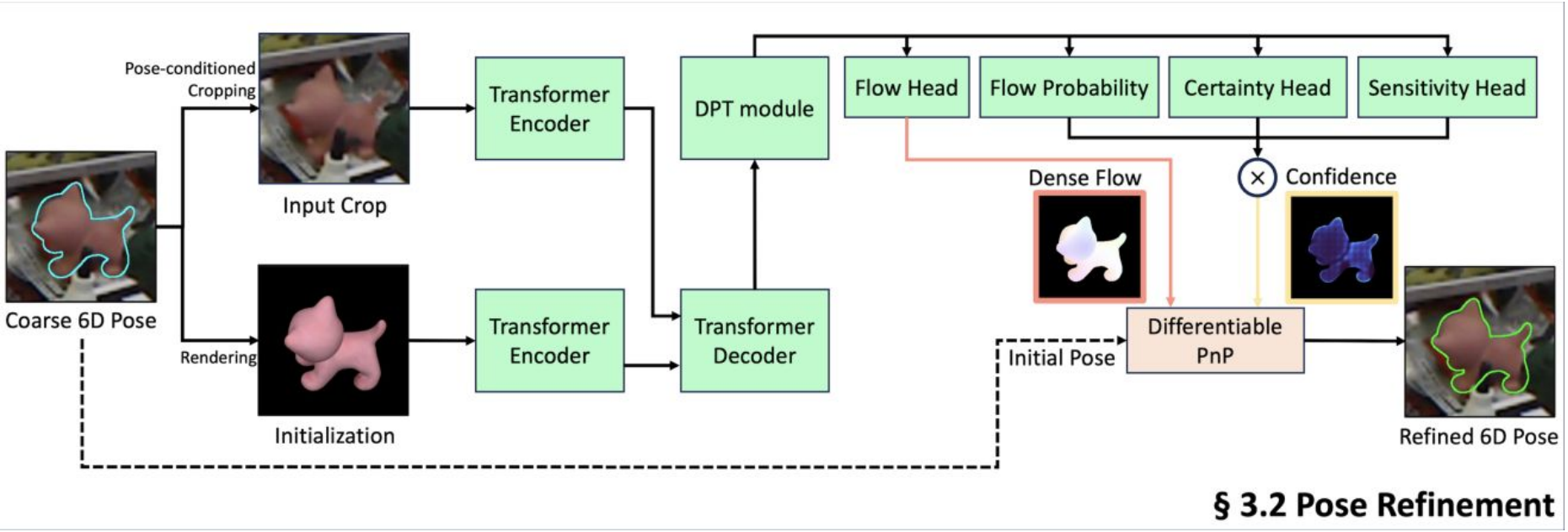
Moon, S., Son, H., Hur, D., & Kim, S. (2025). Co-op: Correspondence-based novel object pose estimation. In CVPR 2025.


Input Detection
Query
Transformer Encoder
Transformer Decoder
Patch Cls Head
Offset Head
Best Template Search
3D model
Rendering
RGBD templates
RGB
Transformer Encoder
Patch Descriptors
Offline Onboarding
Semi-dense Correspondences
PnP
Coarse 6D Pose
§ 3.1 Coarse Pose Estimation


Pose-conditioned Cropping
Input Crop
Coarse 6D Pose
Rendering
Initialization
Transformer Encoder
Transformer Encoder
Transformer Decoder
DPT module
Flow Head
Flow Probability
Certainty Head
Sensitivity Head
Dense Flow
Confidence
Initial Pose
Differentiable PnP
Refined 6D Pose
§ 3.2 Pose Refinement

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.
- 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.