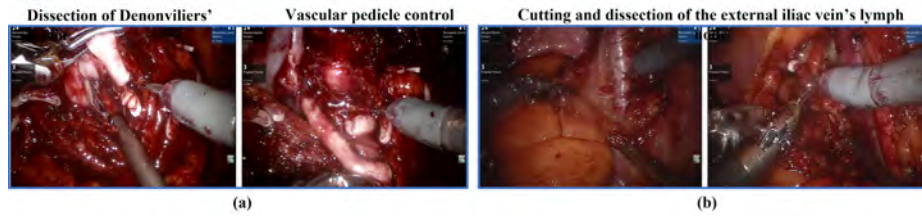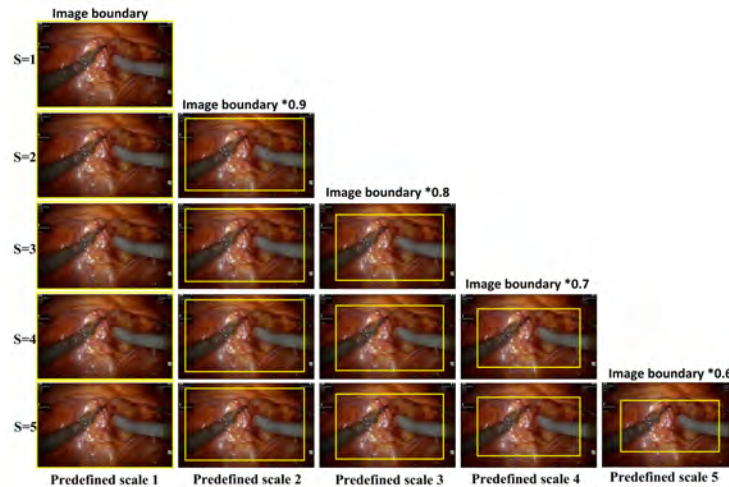# Supplementary Material for Paper: See, Predict, Plan: Diffusion for Procedure Planning in Robotic Surgical Videos

Paper ID 2373

Fig. 1. Samples exhibit minimal variance across different steps in (a), while images within the same step display significant variance in (b).

Fig. 2. Examples of predefined scales: The first scale corresponds to the image boundary, while the subsequent scales denote shrinkages of 10%, 20%, 30%, and 40% respectively, maintaining the center of the cropping boundary fixed.
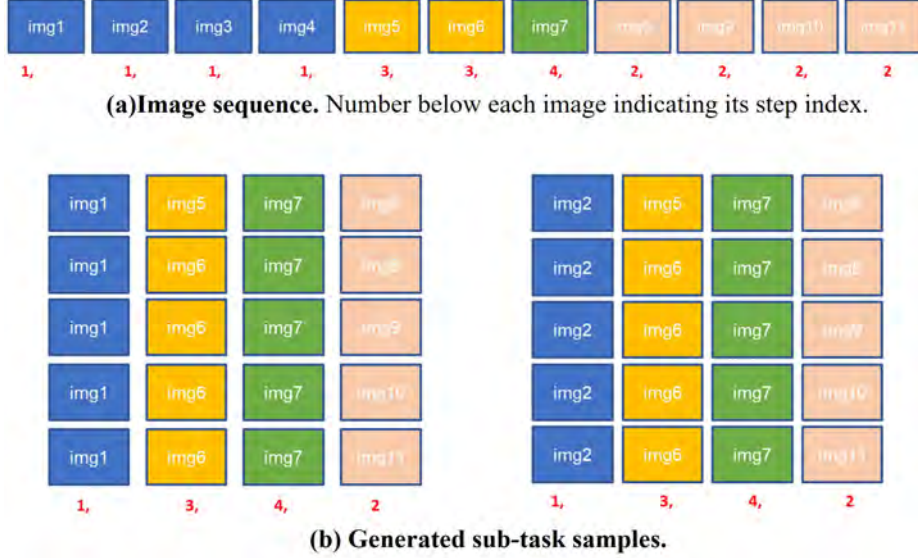
(a)Image sequence. Number below each image indicating its step index.

(b) Generated sub-task samples.

Fig. 3. Dataset construction illustration: we divided the data into sub-tasks, each comprising four different steps belonging to the same phase using sliding window approach.
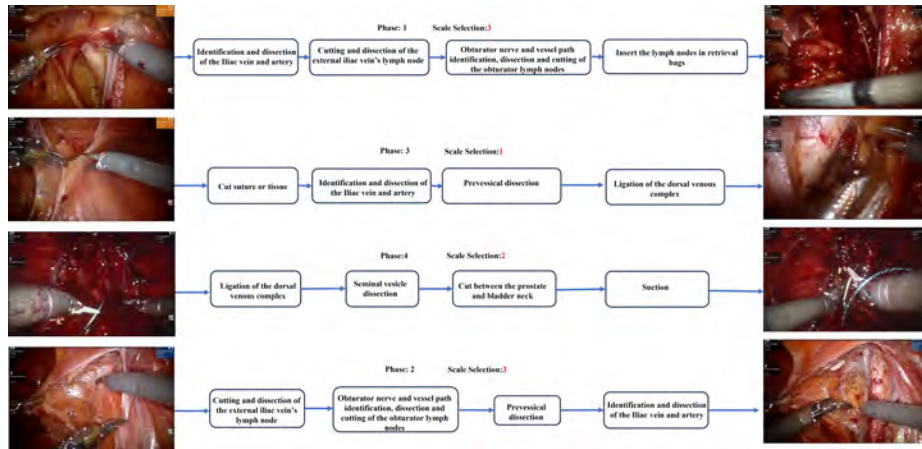

Fig. 4. More visualization examples.