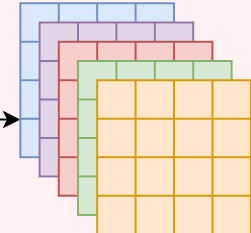
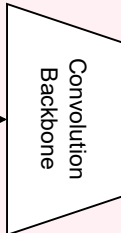


## Keypoint Feature Extraction



Cropped Input Image



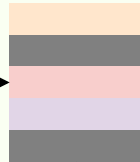
Keypoint Feature Map

Flatten

## Visibility Prediction and Keypoint Masking

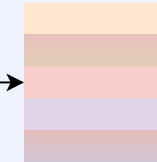
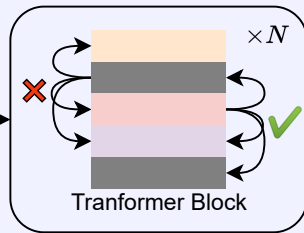


Visibility State Prediction



Visibility-Guided Keypoint Masking

## Keypoint Reasoning with Transformers



Output Keypoint Feature