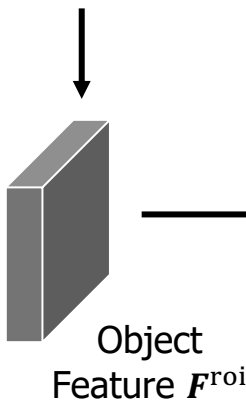
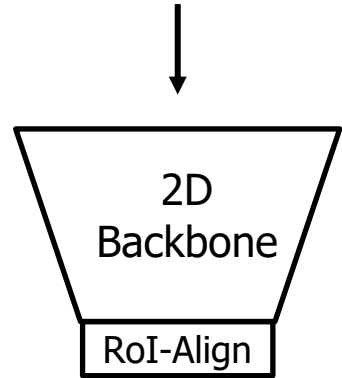
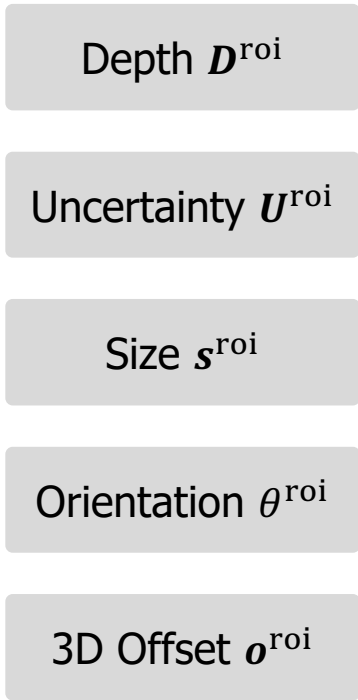




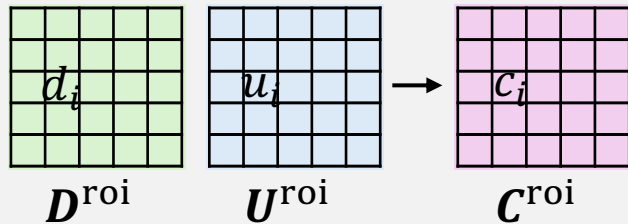
Input image  $I$



### 3D Heads

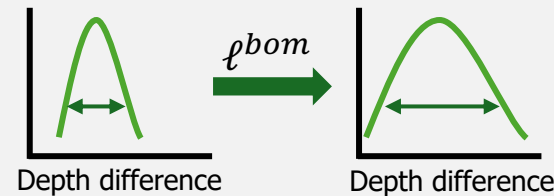


### Training

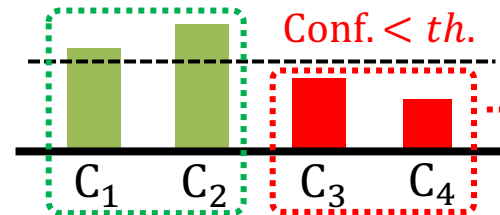


### Soft Best-of-Many (BoM) Loss

$$\ell^{\text{bom}} = \frac{\sqrt{2}}{u_{k^*}} |d_{k^*} - d^{\text{gt}}| + \log u_{k^*}$$



### Hypothesis Filtering



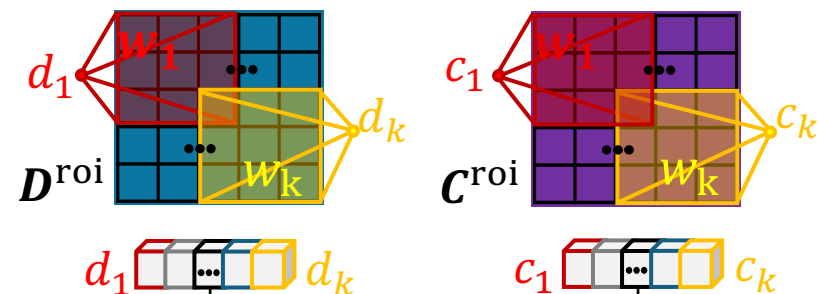
Confidence

3D Boxes

### Final Predictions



### RoI Division



Multi-hypothesis 3D Boxes