

MT-DETR Early Fusion across modalities: Complexity and Performance

