Semantic **Granularity** The image shows a person riding a red bicycle on a road with a red car in the Fine-grained background. The road is lined with semantic trees on both sides and there is another person riding another bicycle in front of her. The date "9/22/2023" is visible in the bottom visual grounding detailed caption & object detection FLD-5B (Comprehensive Annotations) A woman riding a bike red vintage car down a on street street next semantic Coarse to a red car. phrase segmentation visual grounding person caption Florence-2 car & object detection road (Unified Architecture) classification semantic None segmentation regional proposal **Spatial Hierarchy** Pixel-level Image-level Region-level