

Table 1: Demonstration of generated scenes under different multimodal layout instructions.

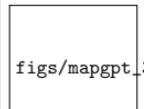
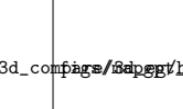
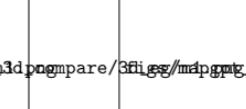
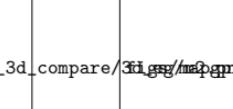
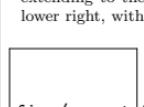
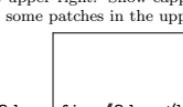
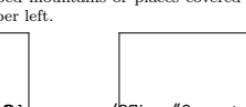
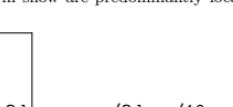
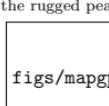
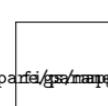
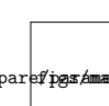
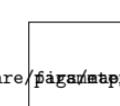
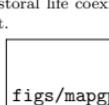
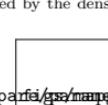
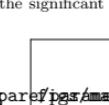
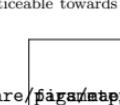
Demo 1	<p>Layout Instruction: The map displays a mix of land cover with buildings concentrated towards the center and edges of the upper regions. Roads traverse the map horizontally and vertically, intersecting near the left edge of the central area. Water bodies are scattered throughout, with a more prominent presence in the lower and left areas. The surrounding areas are predominantly covered in farmland, with patches of grassland in the lower left section.</p> 			
Demo 2	<p>Layout Instruction: The map shows a large area of low bushes and grasslands covering the entire left side and extending to the upper right. Snow-capped mountains or places covered in snow are predominantly located in the lower right, with some patches in the upper left.</p> 			

Table 2: Generated scenes under different environmental configuration instructions.

Layout <p>Layout Instruction: The map displays a substantial presence of forested regions, primarily concentrated in the central area and extending towards the edges. Grasslands or low bushes are scattered throughout. A significant body of water is noticeable towards the lower left.</p>				
	<p>Environmental Configuration Instruction: In a cartoon-style landscape, the terrain is predominantly mountainous with a sunny autumn day. In the upper left, two eagles are patrolling the skies, their sharp eyes scanning the forested regions below. Meanwhile, in the middle left, an ancient warrior stands idle, perhaps guarding the path through the forest. The central area is lush with forest extending towards the edges, providing a rich tapestry of greens against the backdrop of mountains. To the lower left, a significant body of water glistens under the sun, offering a serene contrast to the rugged peaks.</p>			
Demo 1	 Overview	 Local Scene (i)	 Local Scene (ii)	 Local Scene (iii)
	<p>Environmental Configuration Instruction: In the upper left of the mountainous terrain, amidst the rain and mist of an autumn afternoon, three ancient warriors patrol the rugged landscape. Their vigilant presence contrasts with the serene middle left, where five sheep are grazing peacefully. The scene captures the realism of a world where ancient and pastoral life coexist, surrounded by the dense forests and the significant body of water noticeable towards the lower left.</p>			
Demo 2	 Overview	 Local Scene (i)	 Local Scene (ii)	 Local Scene (iii)