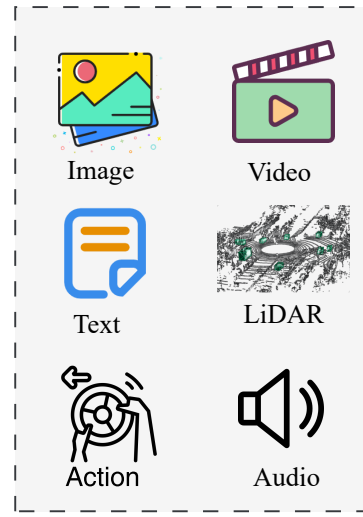


Multi-modal Multi-task Foundation Models for Road Scene Understanding



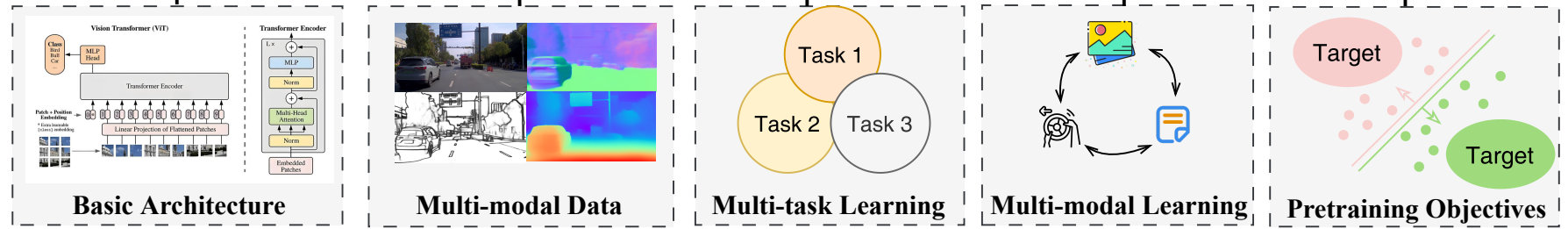
Multi-modal Input



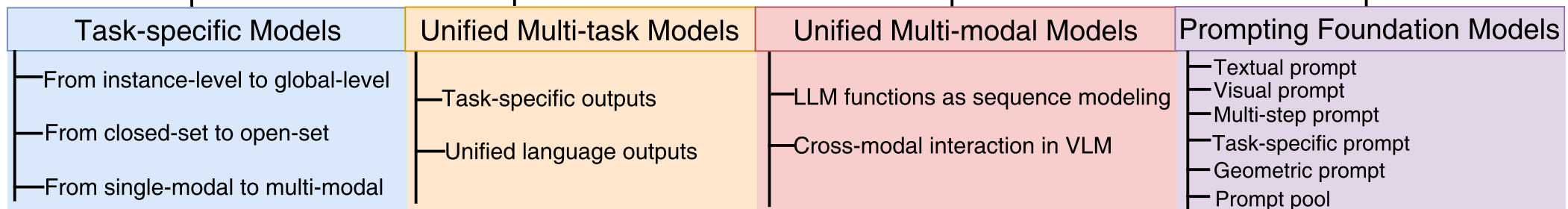
Multi-task Output



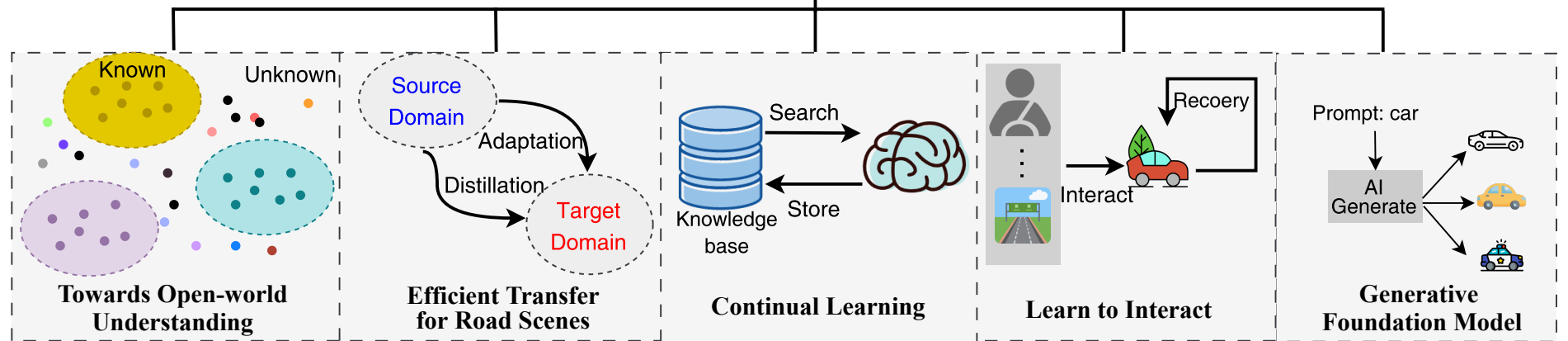
Prerequisites



Common Practices



Advanced Models



Challenges & Trends

