

**Short description:** High-frequency mmWave communication enables ultra-high data rates and low latency but faces considerable challenges due to severe path loss, especially in non-line-of-sight (NLoS) scenarios. Augmenting radios with visual sensing has recently proven effective, as cameras provide rich environmental context that helps predict obstructions and guide proactive network actions. In this CONVERGE Challenge, we invite participants to develop machine learning models that integrate visual and radio data to address key communication tasks in high-frequency wireless systems. The challenge consists of four independent tracks—blockage prediction, UE localization and position prediction, channel prediction, and beam prediction—based on a rich, real-world multimodal dataset collected in a controlled indoor mmWave testbed. This challenge offers an opportunity to benchmark cross-modal learning approaches and promotes interdisciplinary collaboration among the wireless communications, signal processing, computer vision, and AI communities.