

Multilingual VLMs

- Most current VLMs are trained in English, limiting access for non-English speakers.
- **Challenges:**
 - Lack of high-quality multilingual vision-language datasets
 - Tokenization issues (e.g., BPE, WordPiece) optimized for English
 - Fragmentation of common words in non-English scripts
 - Loss of semantics and struggles with languages like Thai or Japanese
- **Real-World Need:** Image-based learning, customer service, and healthcare in multilingual regions.