OCR Noise / Poor
Table Parsing

Fluent but Incorrect
Numeric Generation

Weak Numeric
Grounding

Vision → Language
Mismatch