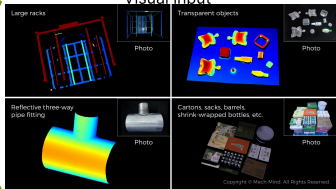


## Visual input



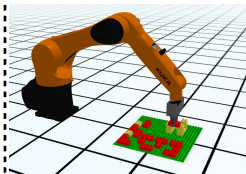
**Objects  
features**

**Human  
instructions**

**Robotics  
language  
model**

**Generated  
codes**

## Validate simulation



**NOT OK** ❌

**OK** ✅

**Output  
codes for  
robot**