Distributed ATL Architecture

Visual Input
(56×56 RGB)

Language Input
"red circle above
blue square"

Visual Cortex
(Conv Encoder)

Language Cortex
(Embedding)

f_vis (64-dim)

f_lang (64-dim)

Project

Project

Distributed ATL

200 Prototypes

$\alpha_i = \text{softmax}(\mathbf{p}_i \cdot \mathbf{f}/\tau)$

α_vis

α_lang

Binding Quality = cos(α_vis, α_lang)

Key: Composition emerges from PATTERN similarity, not single-winner matching