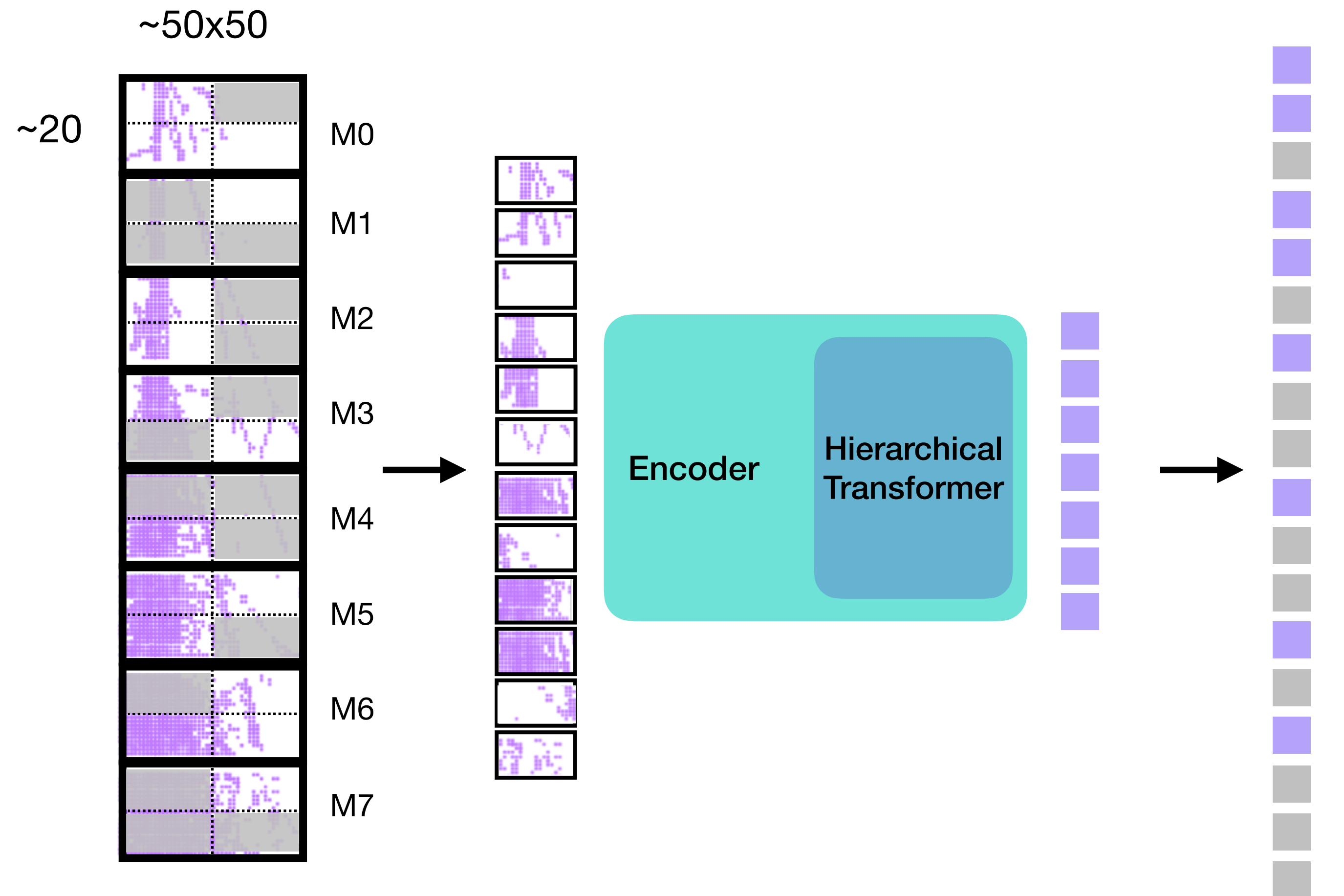
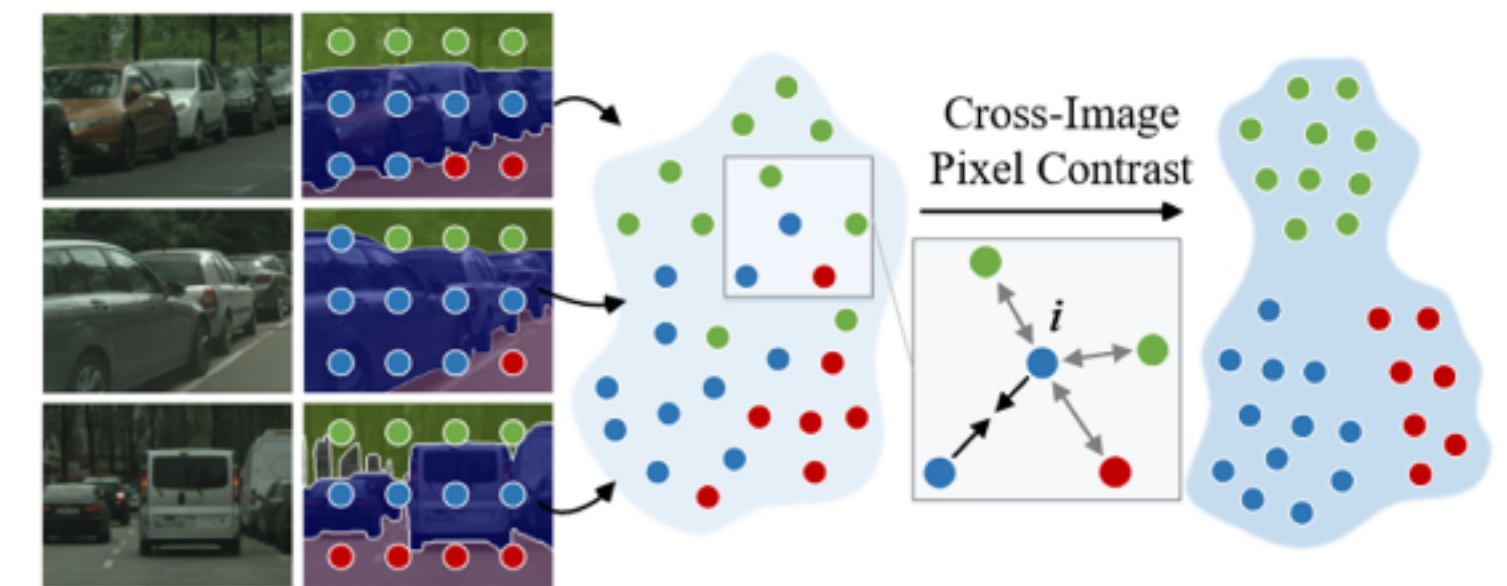


Pre-Training

Hierarchical Transformer



Contrastive Learning

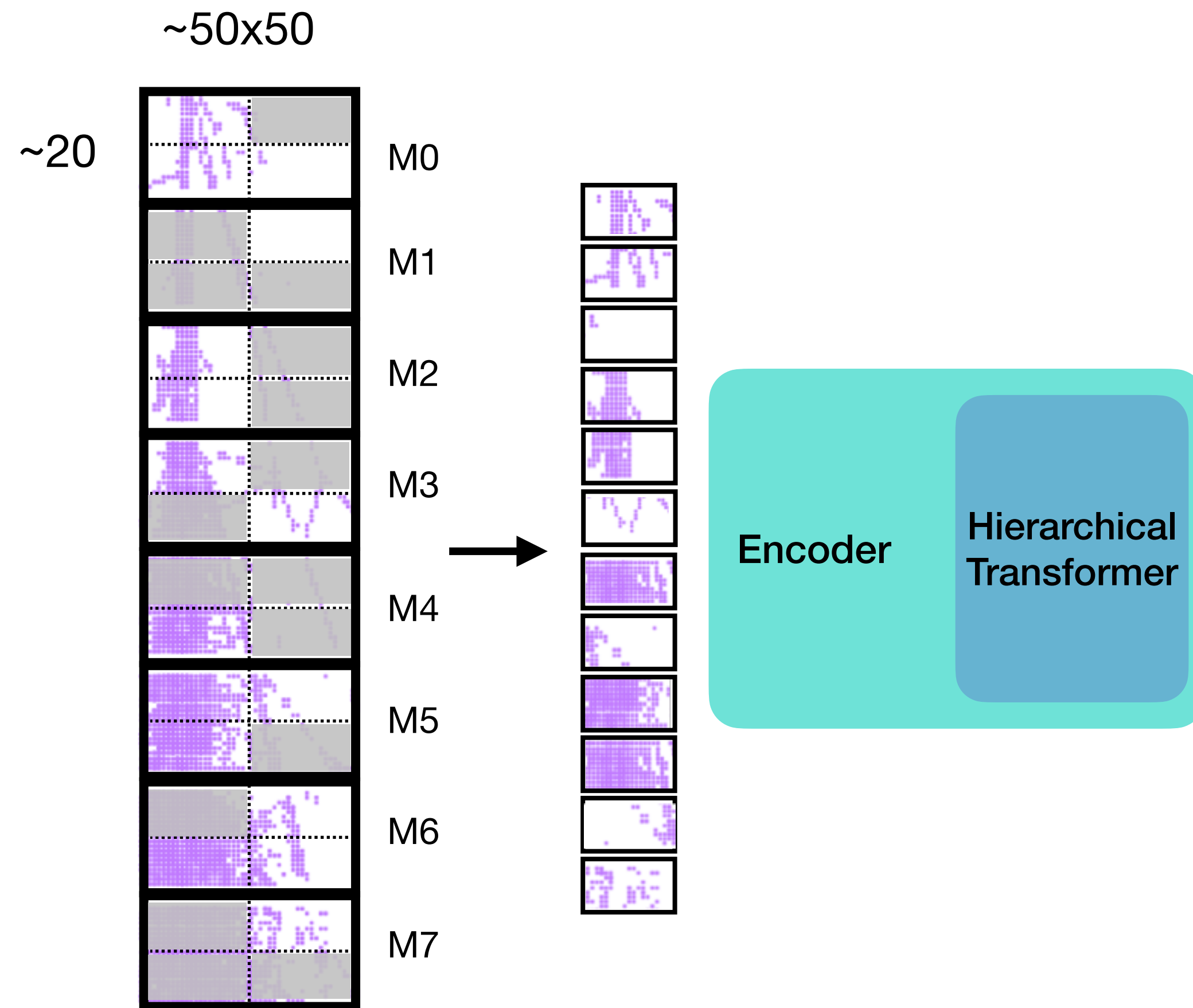


Pixel-wise contrastive learning:

Each pixel (embedding) i is pulled closer to pixels of the same *PDG class*, but at the same time pushed far from pixels from other *PDG classes*

Pre-Training

Hierarchical Transformer



- **Intra-Module Attention (The Local View):**
Summarize hits *within* each detector module independently
- **Inter-Module Attention (The Global View):**
Combine module summaries → learn long-range relations & event topology. Also uses global information from other sub detectors (RealCal, HCal, MuTag)

