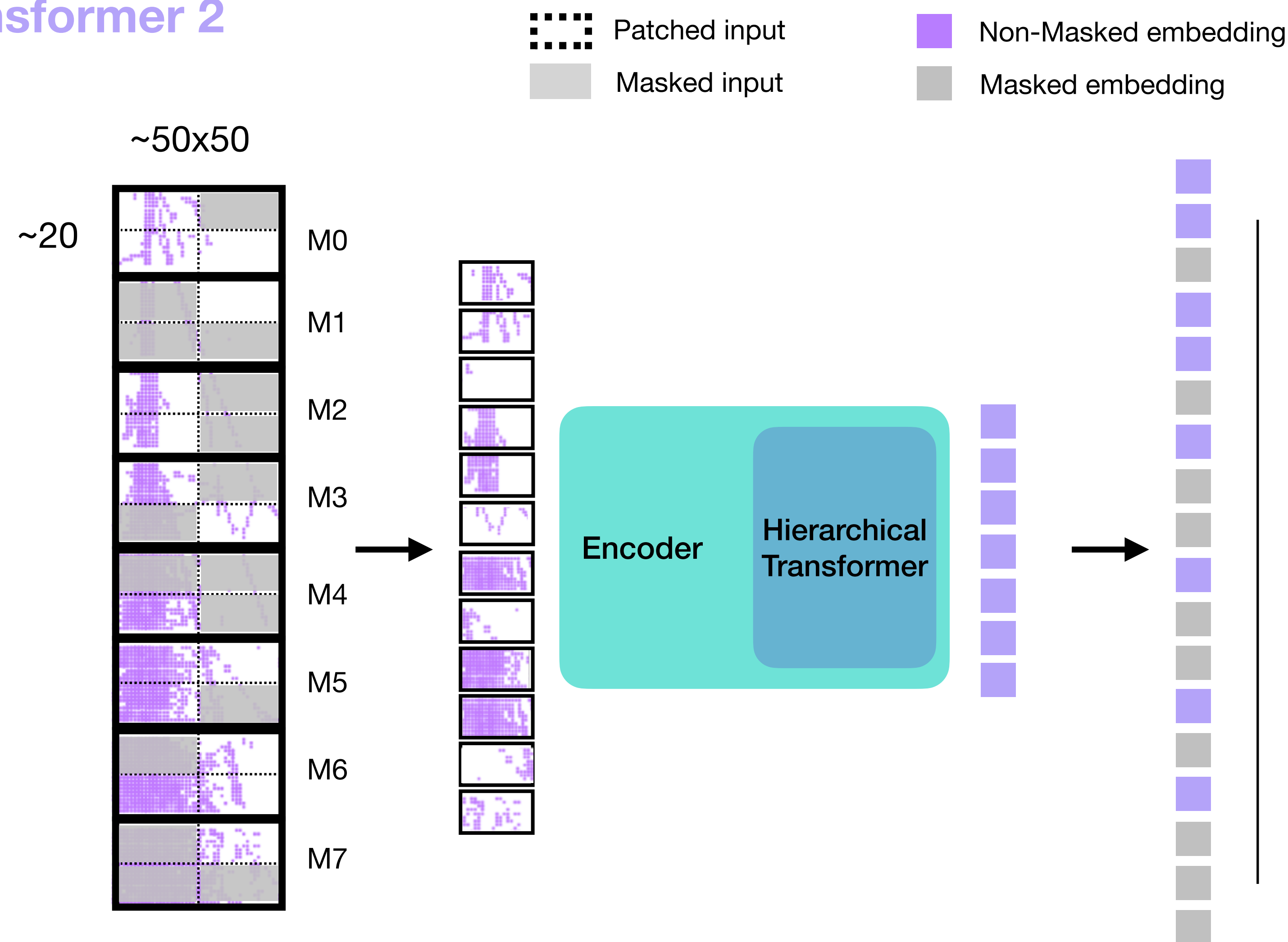


Pre Training

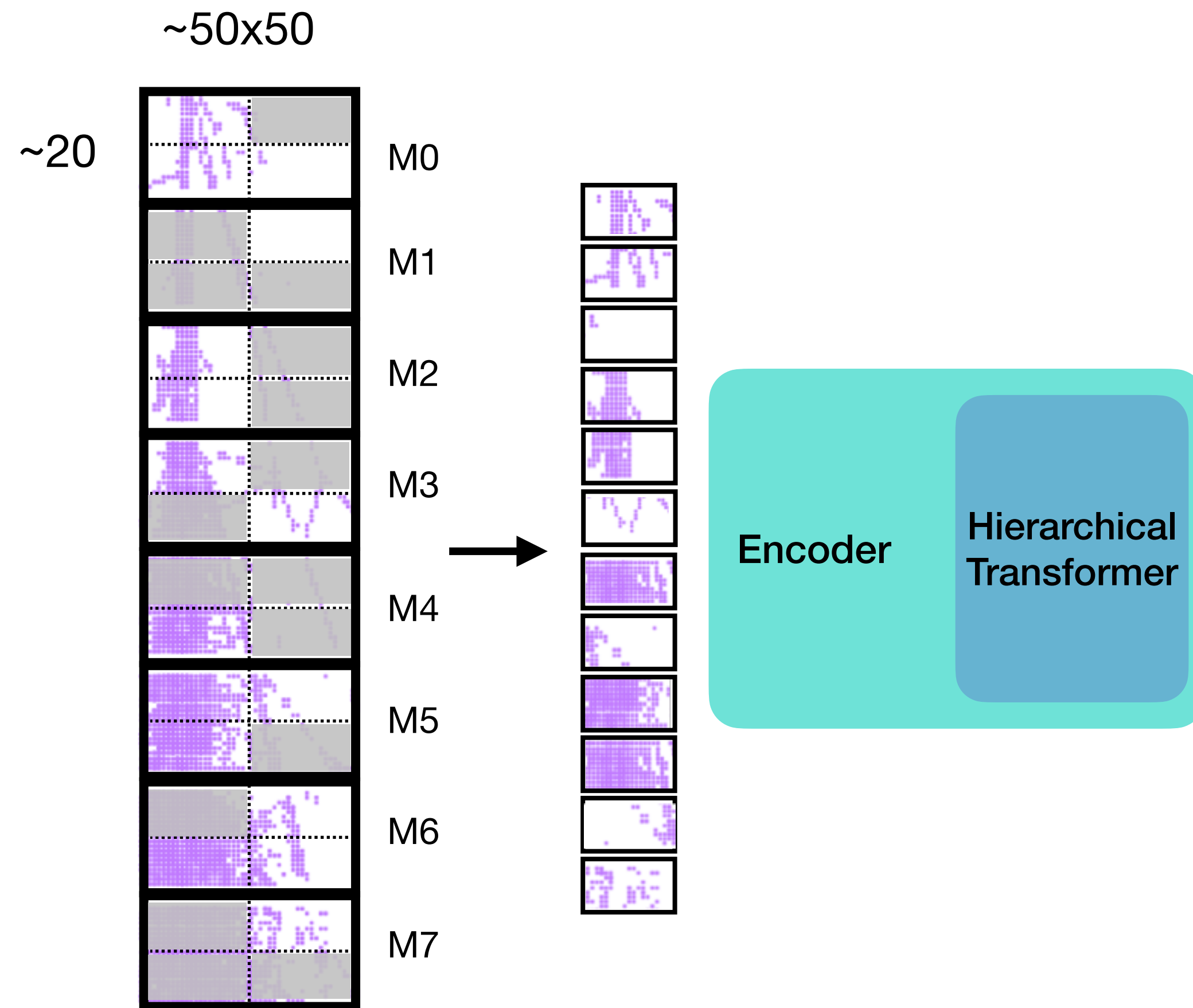
Transformer 2



Contrastive Learning??? Bkup

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)

