Sparse Submanifold Neural Network

~20

~50x50

MO

M1

M2

M3

M4

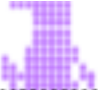
M5

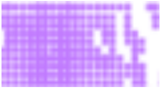
M6

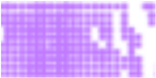
M7

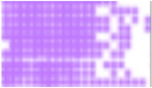
……

Encoder

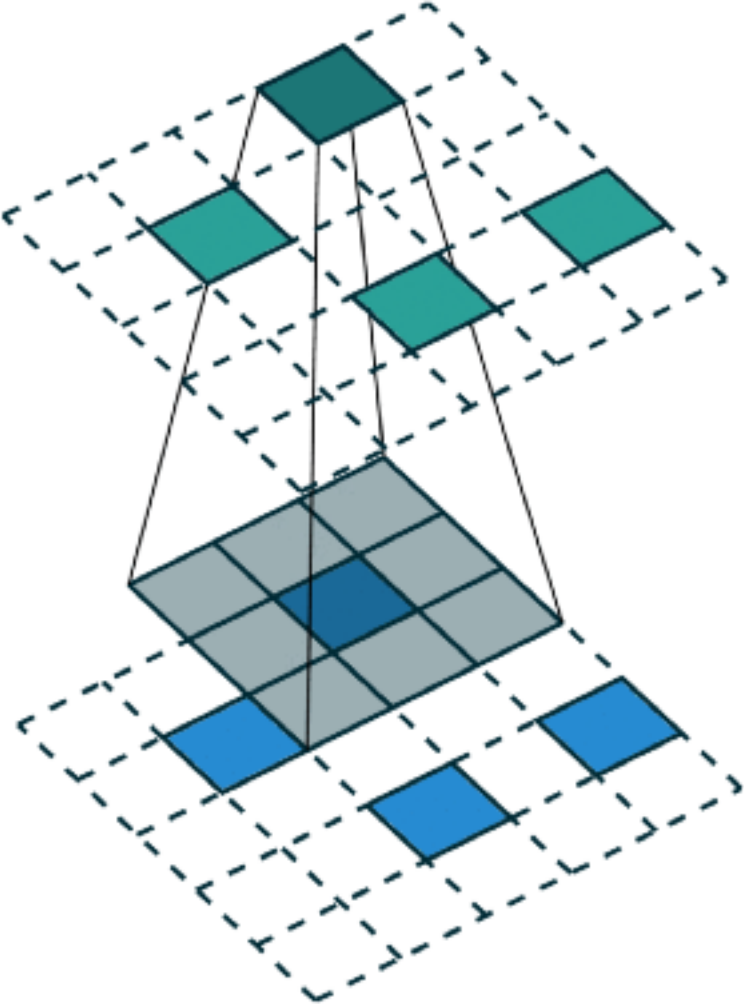








Pre-Training



1 feature

768 features

16x16

1 vector embedding

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

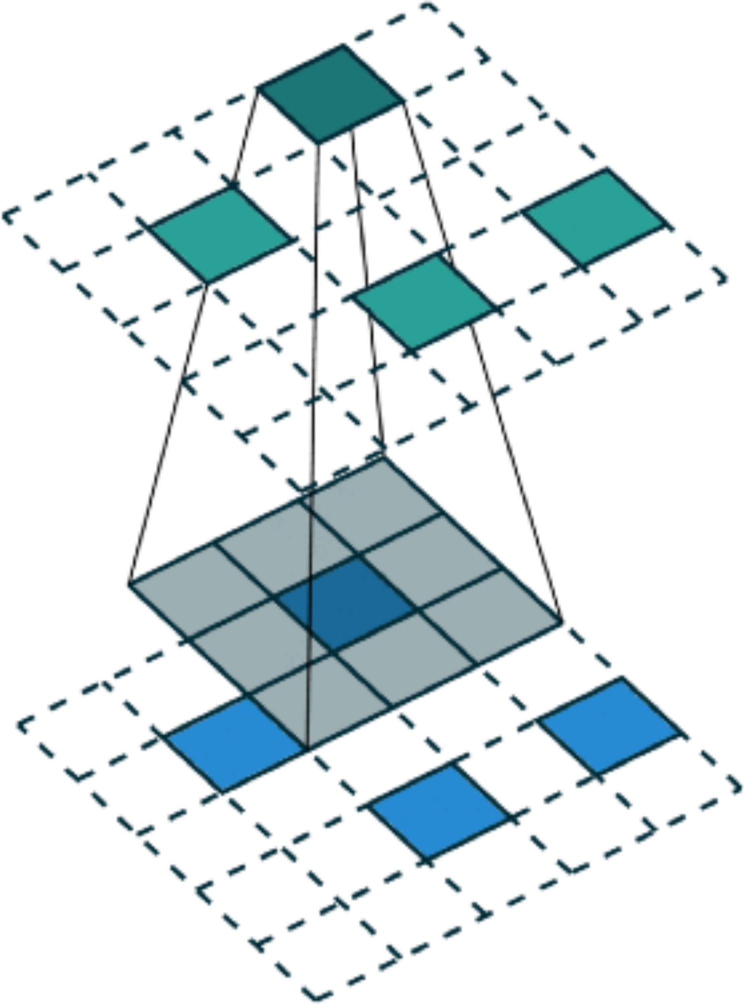
4

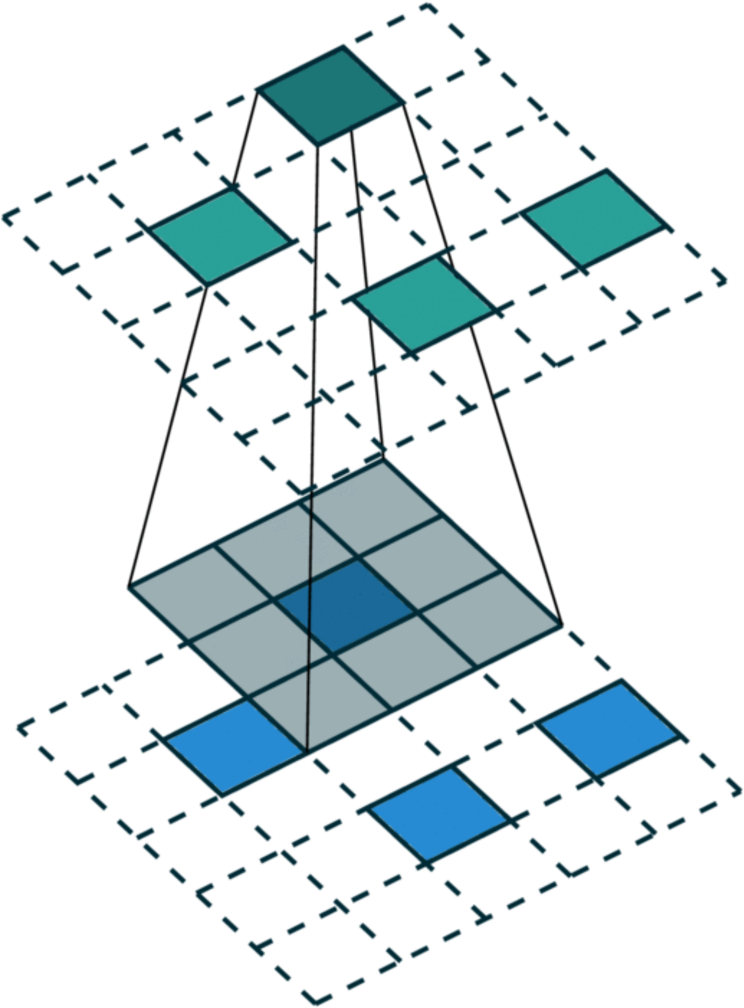
SCNN reduces the spatial coordinates while increasing the feature dimensions.

It performs convolution only on active voxels, effectively ignoring the vast empty regions typical in detector data.

SCNN

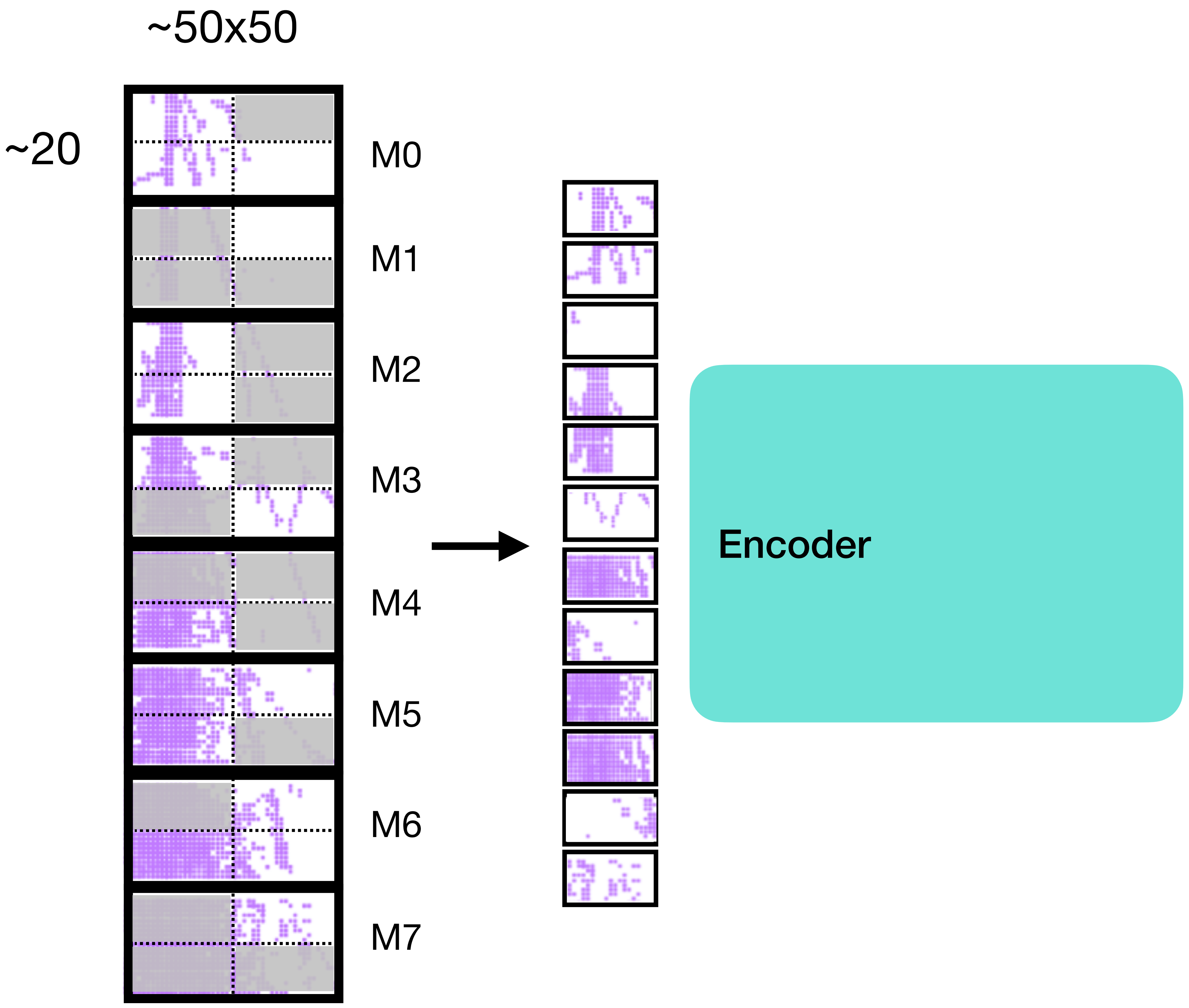
SCNN: Creates a vector embedding for every patch in the input



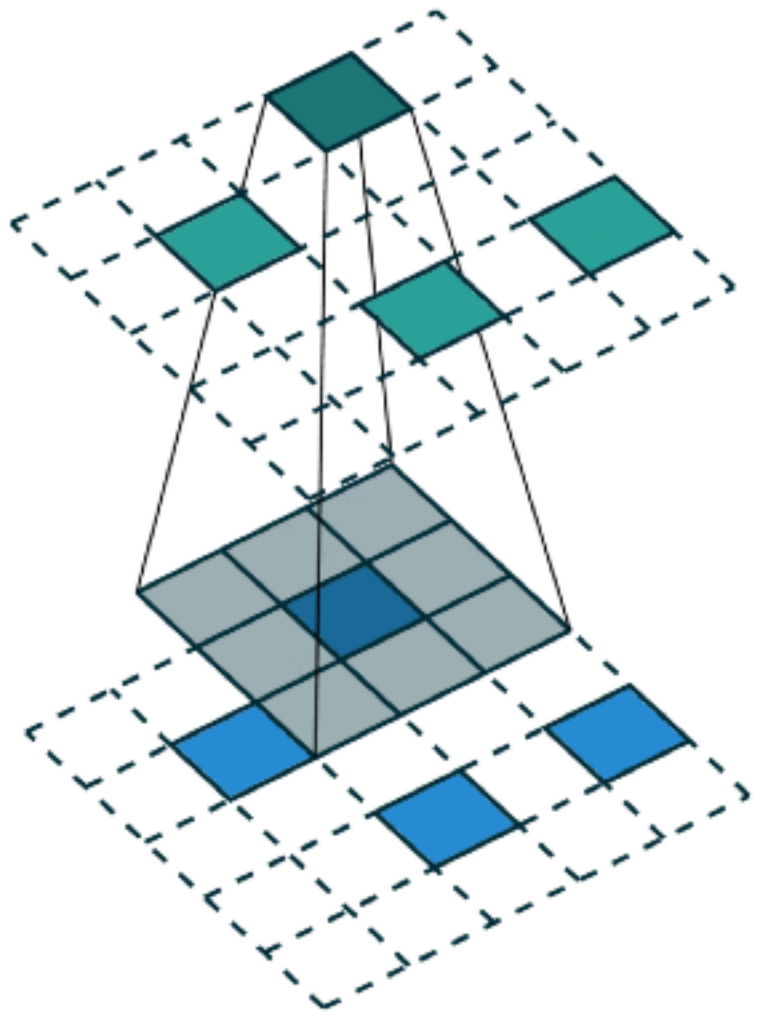


Pre-Training

Sparse Submanifold Neural Network

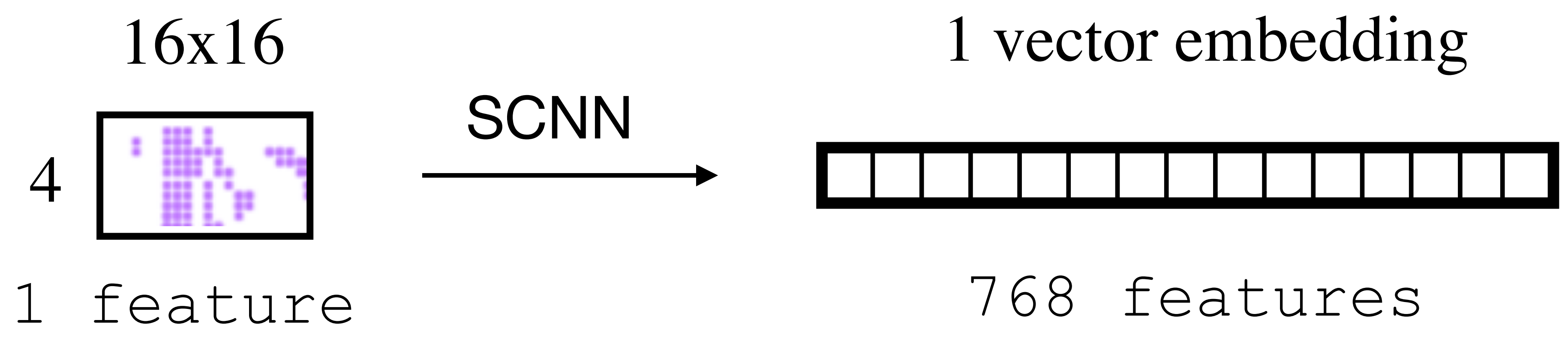


SCNN: Creates a vector embedding for every patch in the input



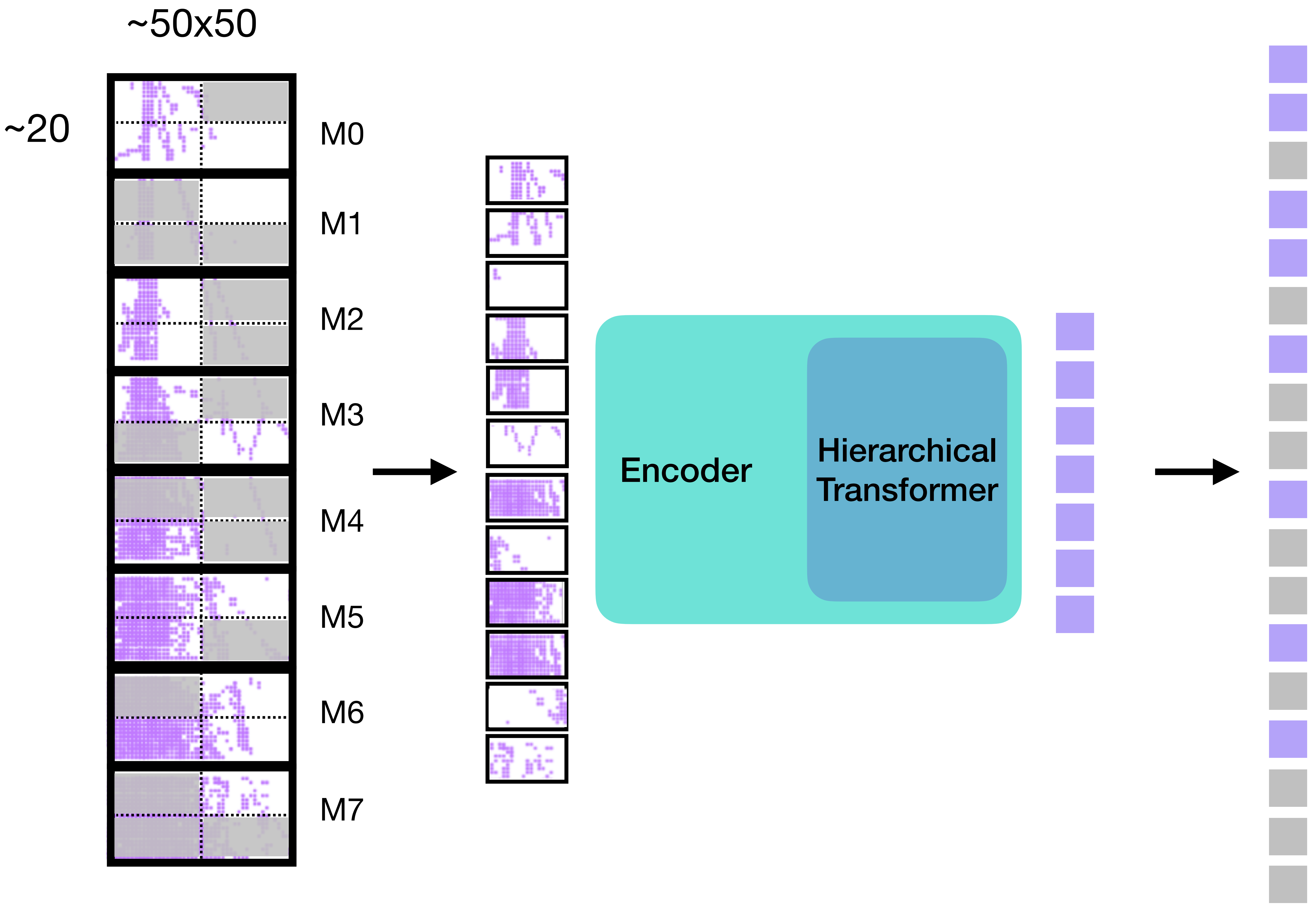
SCNN reduces the spatial coordinates while increasing the feature dimensions.

It performs convolution only on active voxels, effectively ignoring the vast empty regions typical in detector data.



Pre-Training

Hierarchical Transformer



~50x50

~20

M0

M1

M2

M3

M4

M5

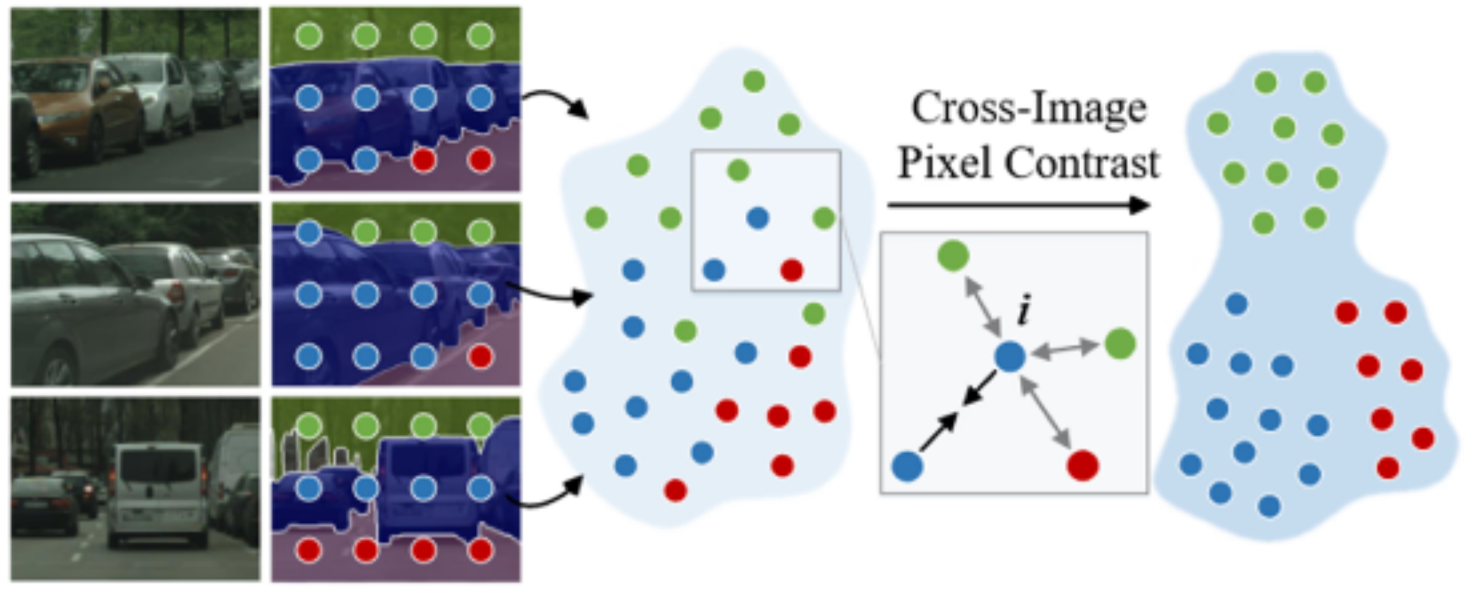
M6

M7

Encoder

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*