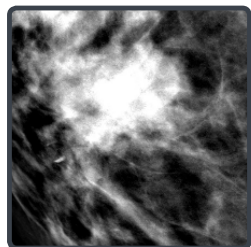


EHRs tabular data

Lesion Image



Multi-hot
Encoding

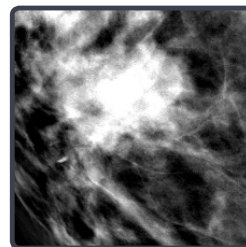
Vision Transformer
Encoder

Shallow Neural
Network

Fusion & Classification
Module

EHRs tabular data

Lesion Image



Tab2Text
Module

Vision Transformer
Encoder (CLIP)

Text Transformer
Encoder (CLIP)

Fusion & Classification
Module

