

STEP 1

Medical Book

UMLS

Cobblestone Pattern

Figure

Crystalline Structures

Patient

Blue-gray Ovoid Nest

Image

⋮

Fried-egg Appearance

Color

LLM

Prompt which requires LLM to filter the concepts unrelated to the disease diagnosis task and garbled text.

Cobblestone Pattern

Crystalline Structures

Blue-gray Ovoid Nest

⋮

Fried-egg Appearance

Discovered Concept Bank

Atypical Pigment Network

Irregular Streaks

Irregular Dots and Globules

⋮

Blue Whitish Veil

Predefined Concepts

Frozen Text Encoder

C_1^M C_1^K
 C_2^M C_2^K
 C_3^M C_3^K
 C_4^M C_4^K
 \vdots
 C_m^M C_k^K

Concept Embeddings

STEP 2

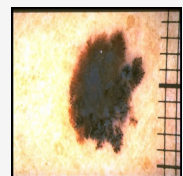


Image Encoder

Predefined Concepts

C_1^K C_2^K C_3^K C_4^K ... C_k^K

K

V

A

Q

A

Q

⋮

A

Q

A

Q

Multi-head cross-Attention (MA)

B

B

⋮

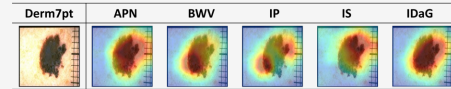
B

Decision Layer

Explain

Lesion Diagnosis: **Melanoma**.

☐ Global Predefined Concepts:



☐ Local Discovered Top-5 Concepts:

1. Annular-granular Pattern
2. Pigmented Lesions
3. Basal Cell Carcinoma
4. Milky-red Areas
5. Parallel Furrow Pattern

Textual and Visual Explanations for Diagnosis

A

Concept Adapters

B

Bottleneck Layers

P

Projector and Similarity Score

C_i^K

Concept Embeddings

C_j^M

T

Pseudo Local Textual Embedding

C^L

Local Concept Embedding

SM

SoftMax Layer

Q

K

V

Query, Key and Value for MA

S_3

S_7

S_4

S_{10}

...

S_{22}

S_1

S_2

S_3

S_4

...

S_m

Sorting

P

T

S_1

S_2

S_3

S_4

...

S_m

SM

\odot

C^L

Concept Complement Module

K

V

C_3^M

C_7^M

C_4^M

C_{10}^M

C_{20}^M

...

C_3^M

C_7^M

C_4^M

C_{10}^M

C_{20}^M

...

C_3^M

C_7^M

C_4^M

C_{10}^M

C_{20}^M

...

C_3^M

C_7^M

C_4^M

C_{10}^M

C_{20}^M

...

C_3^M

C_7^M

C_4^M

C_{10}^M

C_{20}^M

...

C_3^M

C_7^M

C_4^M

C_{10}^M

C_{20}^M

...

C_3^M

C_7^M

C_4^M

C_{10}^M

C_{20}^M

...

C_3^M

C_7^M

C_4^M

C_{10}^M

C_{20}^M

...

C_3^M

C_7^M

C_4^M

C_{10}^M

C_{20}^M

...

C_3^M

C_7^M

C_4^M

C_{10}^M

C_{20}^M

...

C_3^M

C_7^M

C_4^M

C_{10}^M

C_{20}^M

...

C_3^M

C_7^M

C_4^M

C_{10}^M

C_{20}^M

...

C_3^M

C_7^M

C_4^M

C_{10}^M

C_{20}^M

...

C_3^M

C_7^M

C_4^M

C_{10}^M

C_{20}^M

...

C_3^M

C_7^M

C_4^M

C_{10}^M

C_{20}^M

...

C_3^M

C_7^M

C_4^M

C_{10}^M

C_{20}^M

...

C_3^M

C_7^M

C_4^M

C_{10}^M

C_{20}^M

...

C_3^M

C_7^M

C_4^M

C_{10}^M

C_{20}^M

...

C_3^M

C_7^M

C_4^M

C_{10}^M

C_{20}^M

...

C_3^M

C_7^M

C_4^M

C_{10}^M

C_{20}^M

...

C_3^M

C_7^M

C_4^M

C_{10}^M

C_{20}^M

...

C_3^M

C_7^M

C_4^M

C_{10}^M

C_{20}^M

...

C_3^M

C_7^M

C_4^M

C_{10}^M

C_{20}^M

...

C_3^M

C_7^M

C_4^M

C_{10}^M

C_{20}^M

...

C_3^M

C_7^M

C_4^M

C_{10}^M

C_{20}^M

...

C_3^M

C_7^M

C_4^M

C_{10}^M

C_{20}^M

...

C_3^M

C_7^M

C_4^M

C_{10}^M

C_{20}^M

...

C_3^M

C_7^M

C_4^M