Attributes
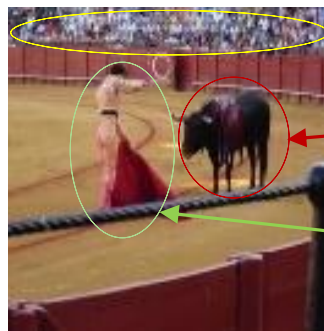
- spectating/ being in an audience
- dark bull and matador
- ......
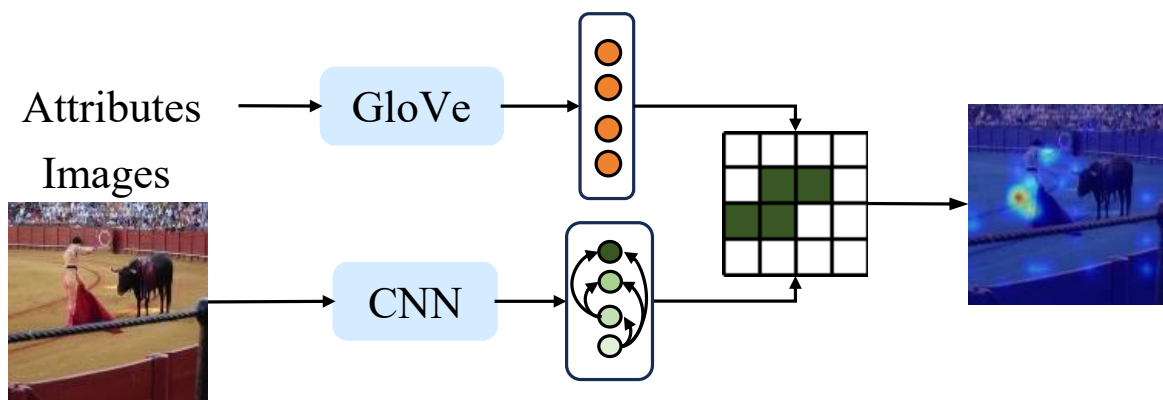
spectators
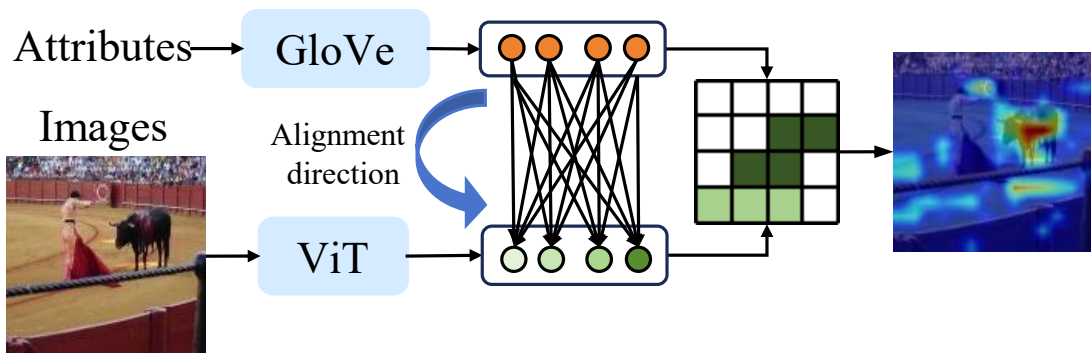
Bull

matador

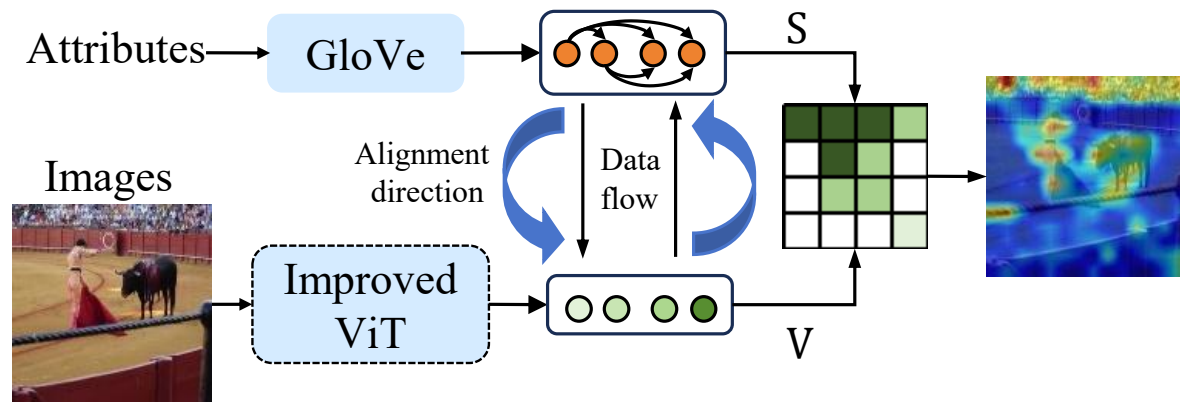(a) Key Local Visual Features and Semantic Attributes for a Bullfighting Scene

Attributes

Images

GloVe

CNN

(b) CNN-based Self-attention Method

Attributes

Images

GloVe

Alignment direction

ViT

(c) ViT-based Attribute-guided Method

Attributes

Images

GloVe

Alignment direction

Data flow

Improved ViT

S

V

(d) Our Method