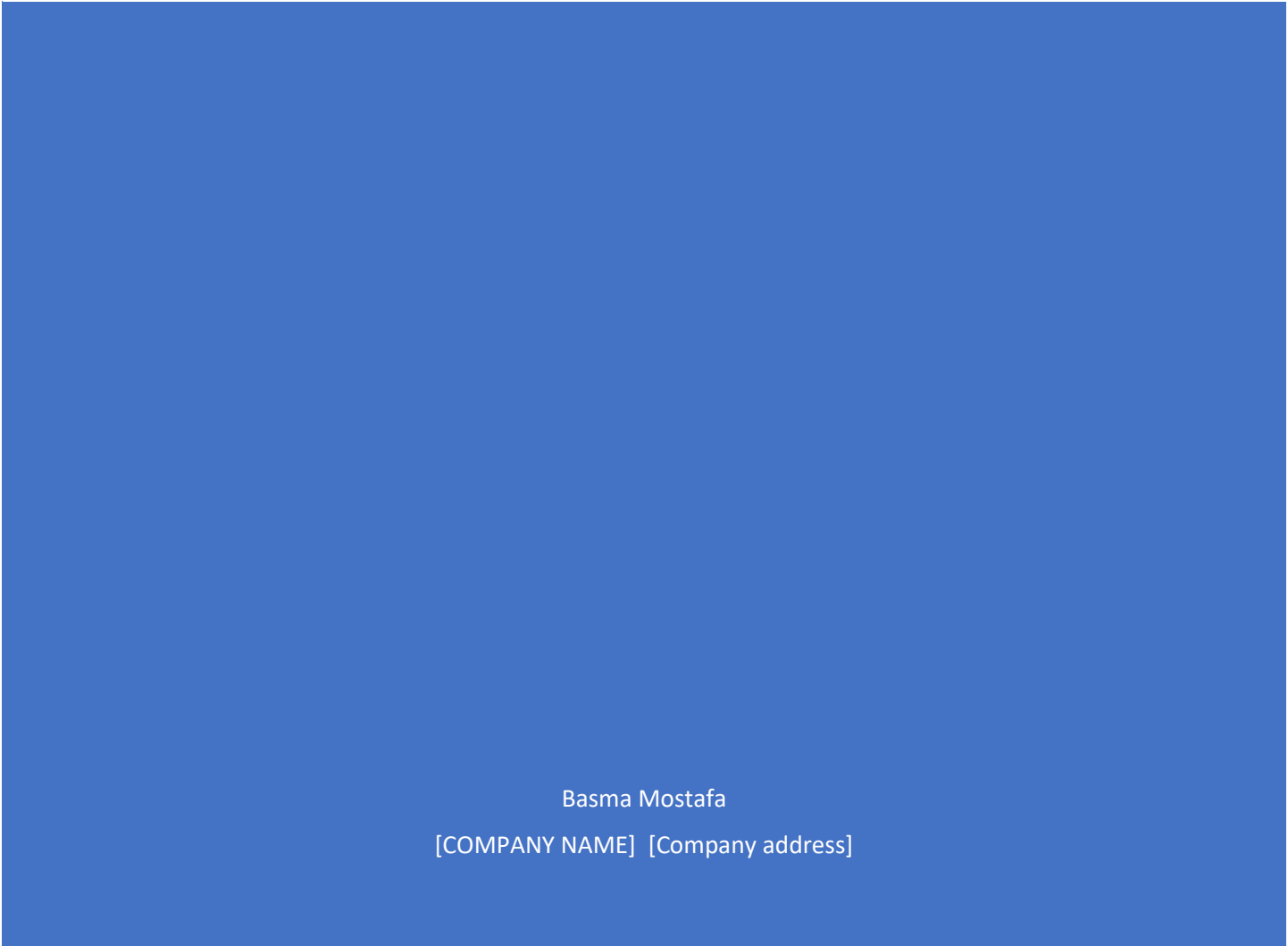




# BASMA MOSTAFA



Basma Mostafa  
[COMPANY NAME] [Company address]

# Attention Mechanisms in Computer Vision

## attention mechanisms:

methods for diverting attention to the most important regions of an image and disregarding irrelevant parts are

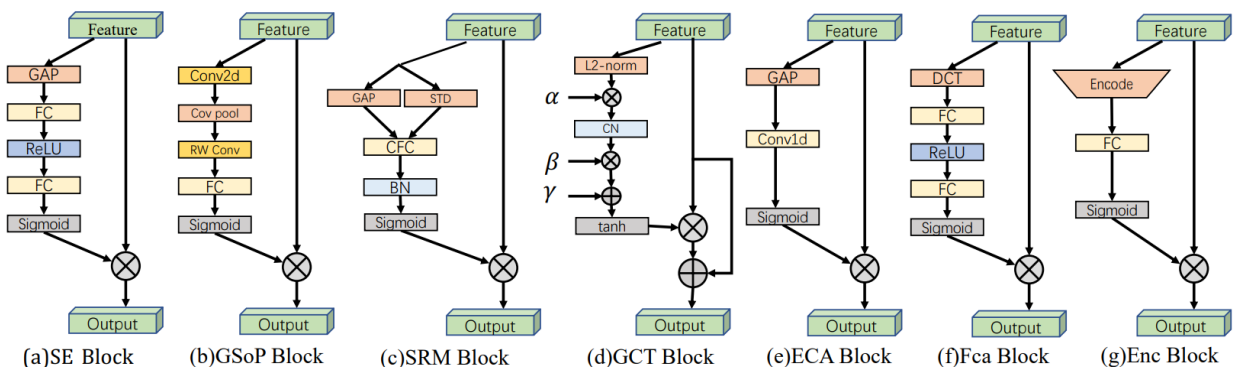
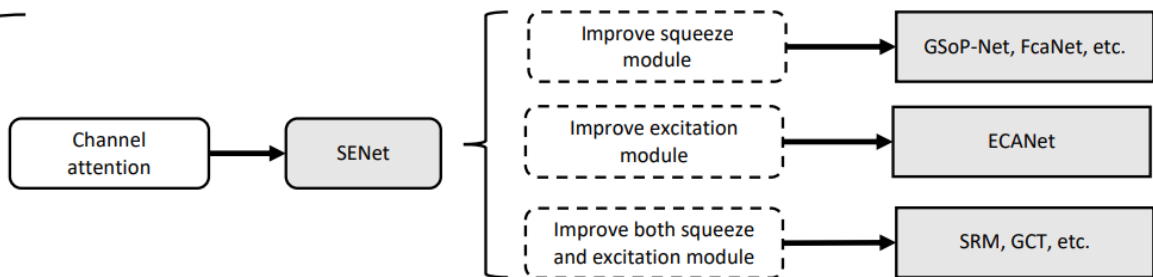
## Attention tasks:

1. image classification
2. face recognition
3. medical image processing
4. object detection
5. semantic segmentation
6. Video understanding
7. image generation

## Attention modules:

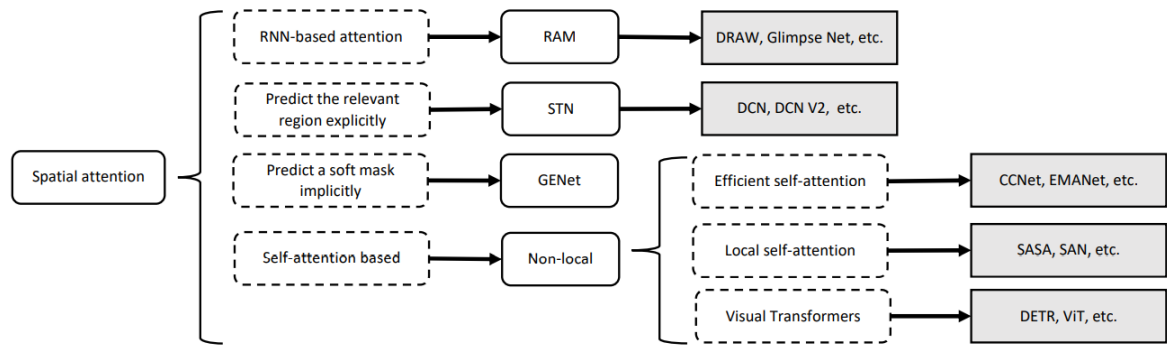
### 1. channel attention

Generate attention mask across the channel domain and use it to select important channels.



## 2. spatial attention

Generate attention mask across spatial domains and use it to select important spatial regions

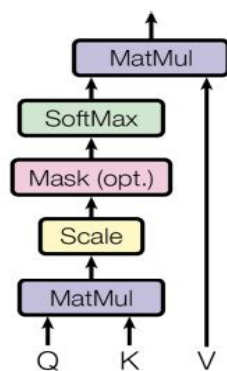


- Attention Gate: focus on targeted regions while suppressing feature activations in irrelevant regions instead of operating on particular regions of interest (ROI), which requires excessive and wasteful use of computational resources and model parameters
- self-attention: is used as a spatial attention mechanism to capture global information.

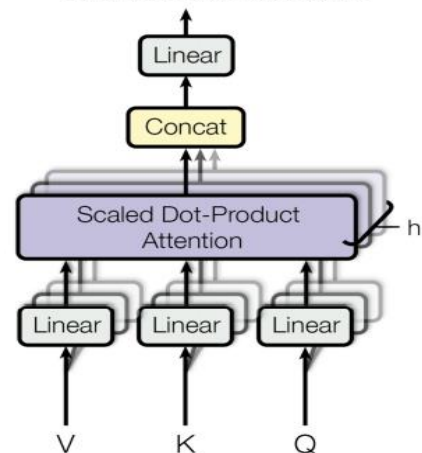
Taking a 2D image as an example, given a feature map, self-attention first computes the queries, keys and values Q, K, V and reshaping operations. Then self-attention can be formulated as:

$$A = \text{SoftMax} (QK^T)$$

Scaled Dot-Product Attention

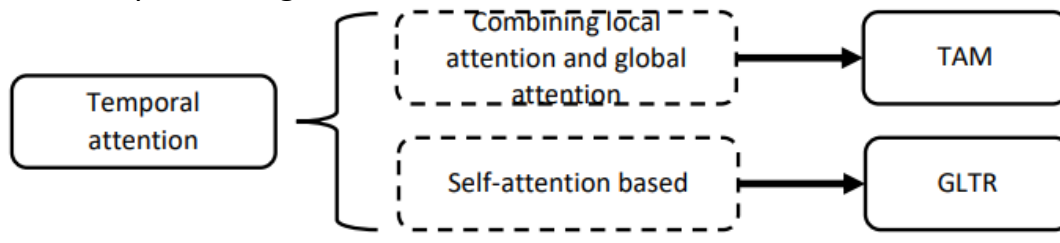


Multi-Head Attention



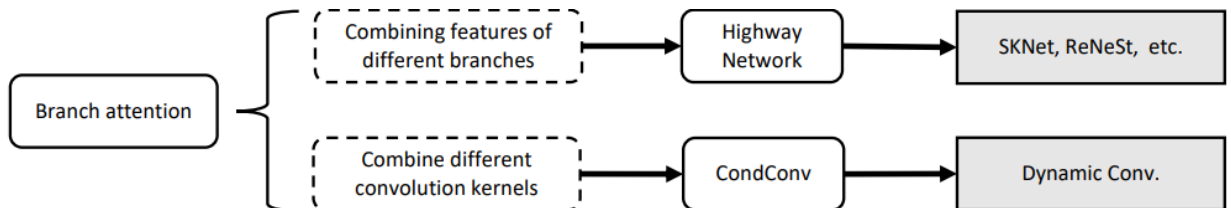
### 3. temporal attention

Generate attention mask in time and use it to select key frames. usually used for video processing



### 4. branch attention

Generate attention mask across the different branches and use it to select important branches



### 5. channel & spatial attention

Predict channel and spatial attention masks separately or generate a joint 3-D channel, height, width attention mask directly and use it to select important features.

Channel & spatial attention combines the advantages of channel attention and spatial attention.

### 6. Spatial & temporal attention

Compute temporal and spatial attention masks separately or produce a joint spatiotemporal attention mask to focus on informative regions.

Spatial & temporal attention combines the advantages of spatial attention and temporal attention as it adaptively selects both important regions and key frames.