**Convolutional Attention Module (CAM)**

Convolutional neural networks (CNNs) can perform better by focusing on the most important regions of the input feature maps thanks to a process called the Convolutional Attention Module (CAM). This concentration enhances the network's capacity to identify significant characteristics and filter out unimportant or irrelevant data.

Attention Mechanism: The attention mechanism reduces the weight of some input data points while emphasizing others. By giving distinct weights to various input components, it enables the network to concentrate on key traits.

Spatial Attention: Concentrates on "where" the key features are situated within the feature map's spatial dimensions (height and breadth).

Channel Attention: Determines which feature maps (or channels) are more relevant by concentrating on "what" the essential characteristics are across the channel dimension.

**Different Convolutional Attention Module**

**Convolutional Block Attention Module (CBAM)**

Traditional CNN feature mapping is enhanced by the Convolutional Block Attention Module (CBAM), which concentrates on key spatial and channel regions. Convolutional Block Attention Modules function in a sequential manner, highlighting important spatial regions and defining critical channels.

Two separate features are obtained by applying average and max pooling to the input feature map F.A Multi-Layer Perceptron (MLP) with ReLU activation processes these features and generates a channel attention map (fc). After that, the SAM concentrates on F′ and creates unique feature representations. A convolution kernel and sigmoid activation function analyze the unified feature map to produce a spatial attention map (fs).

**Residual Attention Network (RAN)**

Remaining units and attention mechanisms are used by a convolutional neural network (CNN) called a residual attention network (RAN) to produce attention-aware features. The attention-aware properties of RANs vary with the depth of layers since they are composed of several attention modules stacked on top of one another. The residual units traverse 2-4 levels with batch normalizations and nonlinearities using skip connections.

The latest feed forward network design allows for end-to-end training of RANs. They have been applied to image classification and to differentiate between optic disc drusen (ODD) and healthy optic discs.In one study, for instance, RANs were able to remove background from photographs of hot air balloons by matching blue color characteristics from the bottom layer with a sky mask. Additionally, a balloon instance mask was used to refine part characteristics from the top layer.

**Self-Attention Generative Adversarial Networks (SAGAN)**

For tasks involving the creation of images, attention-driven, long-range dependency modeling is made possible by the Self-Attention Generative Adversarial Network, or SAGAN. Using just spatially localized points from lower-resolution feature maps, conventional convolutional GANs produce high-resolution details. With SAGAN, cues from every feature position may be used to produce information. Additionally, the discriminator may verify the consistency of highly detailed features located in distant regions of the image with one another. The self-attention mechanism produces outputs that are more detailed and cohesive by calculating relationships between far-off regions in a picture.

**Dual Attention Network (DANet)**

The Dual Attention Network (DANet) is an architecture for neural networks that uses Position attention and Channel attention to improve feature representation. The network is able to capture both spatial and channel-wise dependencies thanks to this dual attention strategy, which is essential for a number of computer vision applications including object identification and picture segmentation. Identify spatial relationships between features at various points throughout the feature map using the Position Attention Module (PAM). The Channel Attention Module (CAM) is designed to highlight significant feature channels by capturing inter-channel relationships.

**Squeeze-and-Excitation Networks (SENet)**

The squeeze-and-excitation process is employed by SENet. After generating channel-wise data by globally averaging feature maps, it runs them through a tiny MLP to determine channel attention weights. The original feature maps are scaled using these weights.