In humans, attention is a core attribute of all perceptual and cognitive operations. The idea of simulating human attention, which first emerged in the field of computer vision, sought to reduce the computational complexity of image processing while improving performance by introducing a model that focused only on specific areas of an image rather than the entire image.

However, the real starting point for the mechanisms of attention as we know them today is usually in the field of natural language processing.Attention was initially implemented in the machine translation model to solve some problems in the structure of recurrent neural networks.After scholars emphasized the benefits of attention, the technique improved and quickly became popular for tasks as diverse as text classification, image captioning, sentiment analysis, and speech recognition.The structural model based on attention mechanism can not only record the location relationship between information, but also measure the importance of different information features according to the weight of information.Dynamic weight parameters are established by making relevant and irrelevant choices on information features to strengthen key information and weaken useless information, thus improving the efficiency of deep learning algorithm and improving some defects of traditional deep learning.