近些年来，深度学习技术得到了长足的发展，为图像分类、机器翻译、语音识别等多媒体问题提供了高精度高性能的解决方案。自高精度的图像分类CNN网络AlexNet【来源请求】问世以来，卷积神经网络凭借其精度优势，逐渐成为物体识别、图像分类、图像语义分析等图像、视频处理领域问题的主流选择。在AlexNet被提出之后，学界发展了大量精度更高、结构更复杂的卷积神经网络模型，如R-CNN【来源请求】、Fast R-CNN【来源请求】、Faster R-CNN【来源请求】、YOLO【来源请求】。基于卷积神经网络的图像处理技术在业界也得到了广泛应用，并渗透到了每个人的日常生活中。

如，在自动驾驶领域，Google基于图像物体检测技术启动了智能自动驾驶汽车计划【https://www.google.com/selfdrivingcar/】，截止至2012年，其推出的Waymo自动驾驶系统已经装上Lexus RX450h试验车已经完成了30万英里的无人驾驶旅程。2017年，Waymo项目已经推出了其第一款基于量产车型的自动驾驶汽车，并且在亚利桑那州开始召集体验用户为他们提供回馈；

在医疗方面【http://www.nvidia.com/object/deep-learning-in-medicine.html】，深度学习技术也可以通过预训练的神经网络对CT影像、组织切片影像等图像进行分析，为医护人员做出诊断提供可靠的参考。日前即将落幕的Kaggle大数据竞赛Data Science Bowl 2017大赛的主题【https://www.kaggle.com/c/data-science-bowl-2017】即为根据CT图像等数据集准确预测肺癌。

标榜“智能”的智能手机厂商们也在利用深度学习技术，为手机用户开发更加智能的应用、提供更加智能的用户体验。从较为简单基础的相机识别人脸、自动聚焦，到复杂的云相册服务（识别并分类相册中的人脸，自动根据照片中含有“谁”来进行整理归类；识别各照片中的物体并标记，以便用户检索相册中的特定照片，等）。而随着更新的卷积神经网络模型在深度、广度上进一步变复杂，这类任务的计算量对服务器的压力也进一步增大。一些智能手机厂商为了减少服务器的负担，决定采取直接在手机上运行卷积神经网络模型的做法。