



王飞跃，1990年获美国伦塞利尔理工学院（RPI）博士学位，1990年起在美国亚利桑那大学先后任助教授、副教授和教授。曾任中国科学院自动化研究所副所长，现为中科院自动化研究所研究员、复杂系统管理与控制国家重点实验室主任。主要研究机器人、复杂系统、智能控制、人工智能、智能交通、社会计算、知识自动化等领域。

卷首语

Foreword

计算知识视觉：迈向智能视觉推理的新范式

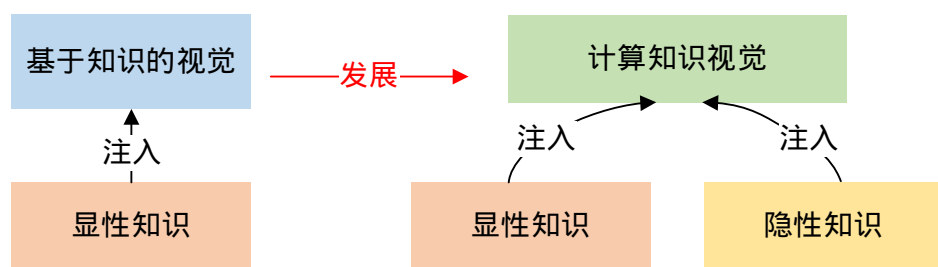
王飞跃

大数据时代的信息急速膨胀和计算硬件的快速发展，使得计算机视觉开始崭露头角。尤其是，在人脸识别、目标检测、图像分割等任务中，近年来基于深度学习的计算机视觉模型取得了优秀的研究成果，并且大部分成果已经从实验室阶段迈入了我们的日常生活。此外，计算机视觉正在与智慧安防、智能机器人、无人驾驶、智能家居等诸多前沿领域互相结合与渗透，以期更好地服务大众、造福社会。

尽管许多计算机视觉模型在一些方面已经取得了极大的成功，但是现有的这类模型却容易犯一些在正常成年人眼里低级而又不可思议的错误。例如，特斯拉生产的自动驾驶汽车就曾将白色货车的货箱错认成了天空，进而引发交通事故，造成了人员伤亡。显然，我们人类的视觉系统在正常情况下不可能将白色的货箱错认成天空。

人类不仅擅长处理各种直观的视觉任务，还善于利用已有的“知识”进行视觉推理和理解。例如，当驾驶员观察路况意图变道时，驾驶者可能首先观察到前后左右有几辆车和这些车所在的位置，随后会根据这些车辆的指示灯和相对位置来决定自己何时变道。而现有模型却很难获得这些“知识”，从而导致其在进行处理视觉推理与视觉理解时常受限制。那么，我们何不将人类知识与现有的视觉模型相结合，从而弥补现有视觉模型的不足呢？

其实，人类知识和人工智能的交汇由来已久。在20世纪70年代，爱德华·费根鲍姆（Edward A. Feigenbaum）倡议要将知识引入到人工智能中，并提出了“知识工程”的概念，后来又在此基础上，提出“知识表征”。随着知识工程和知识表征的发展，我们现在耳熟能详的语义网、知识图谱以及新兴的知识范畴等技术逐渐出现并得到蓬勃发展。



前事不忘后事之师。启发于知识工程，一种迈向智能视觉推理的新范式——计算知识视觉（computational knowledge vision, CKV）就此诞生。计算知识视觉的关键在于将人类知识融入到现有的计算机视觉模型之中。为了达到这一点，通常需要用结构化的方法来表征人类知识。换句话说，要将人类知识转化为计算机可以接受输入的结构化知识。由于人类知识存在显式和隐式两种类型，那么对应的结构化知识模型也是两种。另外，计算知识视觉不仅考虑了视觉模型和结构化知识的表征，还考虑了结构化知识作用下的视觉模型的推理和理解。一般来说，可以通过构建结构化知识来指导视觉模型的训练，视觉模型也可以从具体任务中获得知识，从而增强结构化知识的效用。相信这一新的研究方向，可将智能视觉推理方法推向新的范式 and 高度。

未来，计算机视觉应用将更多地出现在我们的日常生活中，人们对智能视觉系统的有效性和稳定性方面的要求会越来越高，而结合人类知识的计算知识视觉无疑会为计算机视觉领域的发展带来新的可能。随着智能技术、物联网等的进一步发展，特定知识在不同的智能视觉系统中必将发挥更大作用。