



第6讲

人工智能技术应用

主讲：丛润民

01

智慧之眼——视觉感知



02

智能之躯——具身智能



03

AI卫士——智慧医疗



04

AI助手——智慧生活



目
CONTENT
录

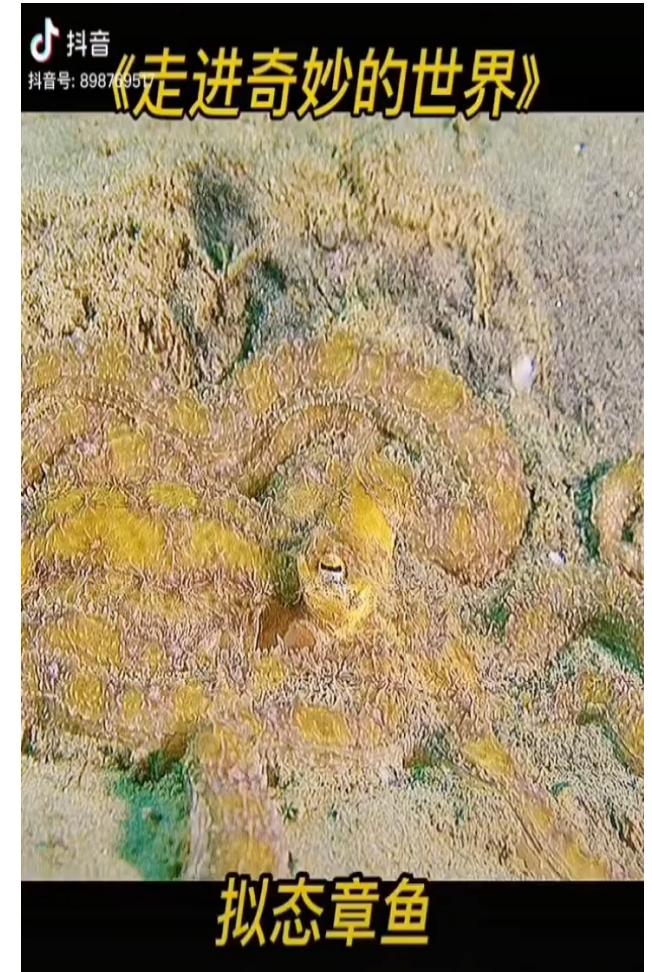




01

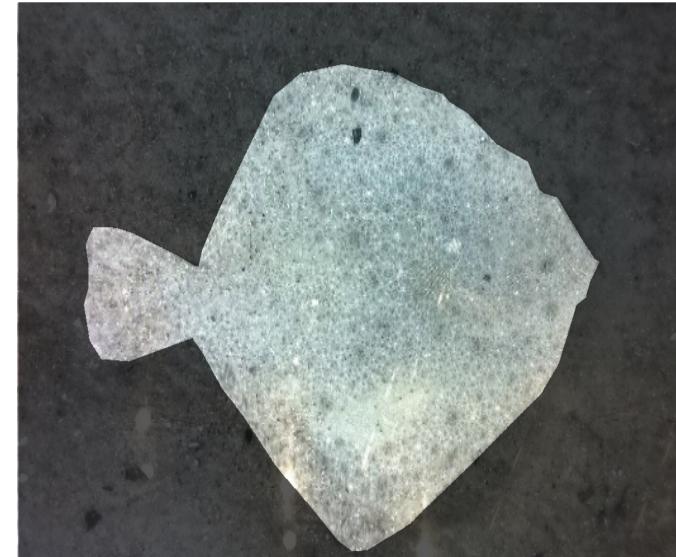
智慧之眼——视觉感知

- 在自然界中，动物试图通过使身体的结构适应周围环境的结构来隐藏自己，这有助于它们避免被捕食者识别。
- 伪装目标检测任务旨在让计算机能够自动识别场景中的伪装目标，是一个非常具有挑战性的任务，应用场景包括珍稀物种发现和军事场景中伪装敌人的识别等等……



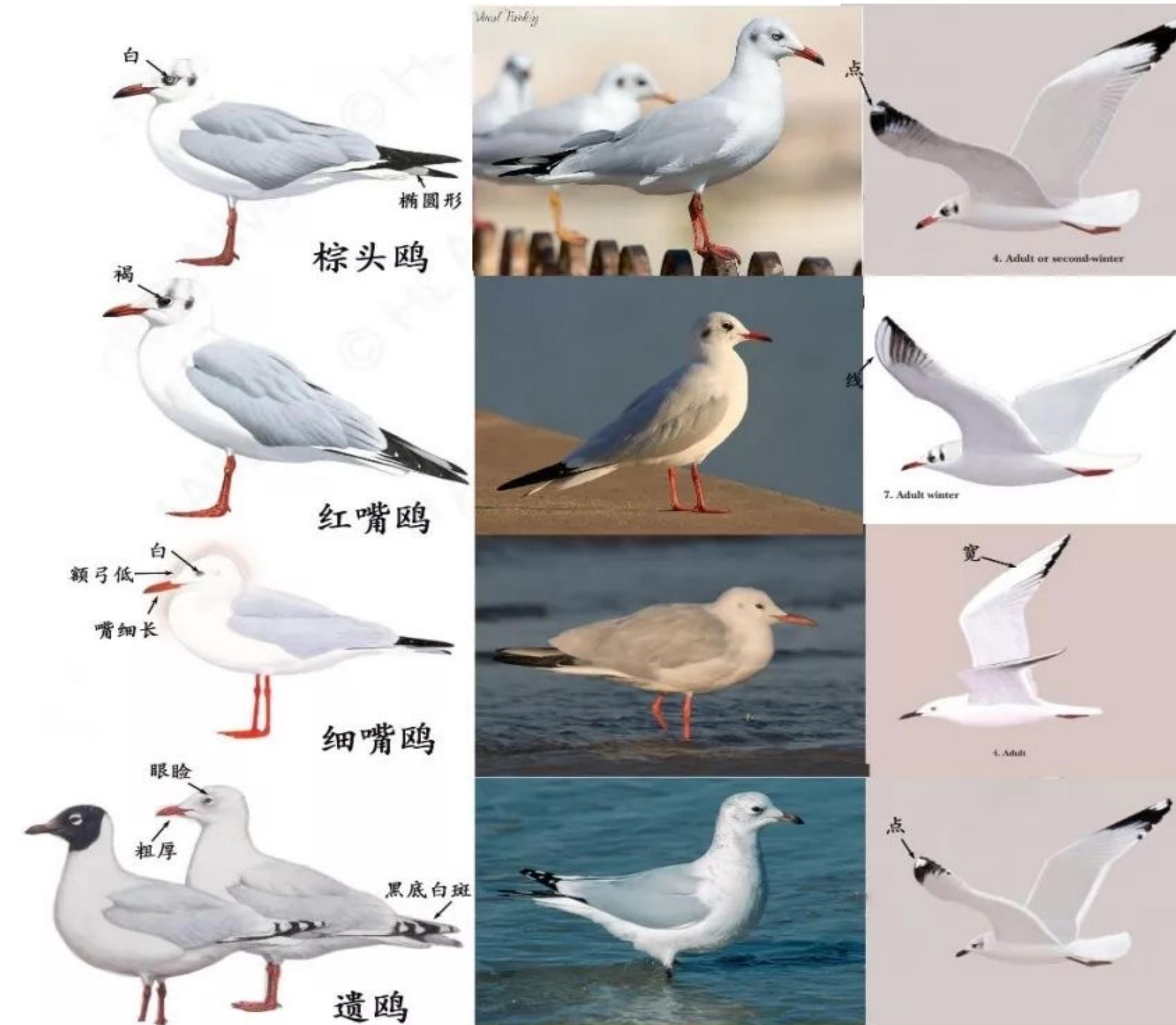
开始测试 •

请说出图像中伪装的目标在哪里？



1.0

你的眼睛真的是尺吗？

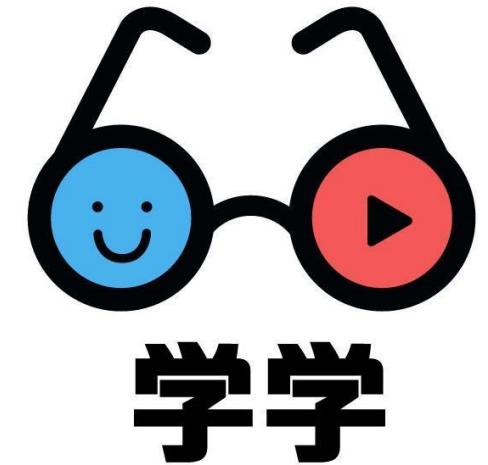


棕头鸥：虹膜白色，头白色，嘴红，尖端黑色，眼后有一黑褐色斑，头顶有两道不明显斑纹，上体浅灰色，翼尖黑色，下体白色，停歇时候可以观察到初级飞羽下部有一椭圆形白斑。

红嘴鸥：虹膜暗褐色，头白色，嘴红，尖端黑色，眼后有一黑褐色斑，头顶有两道不明显斑纹上体浅灰色，翼尖黑色，下体白色，有些个体泛粉红色，停歇时候可以观察到初级飞羽翼下全白，飞行时表现为初级飞羽外侧两根羽毛为白色，仅先端黑色。

细嘴鸥：虹膜白色，其额弓明显较低，嘴细长为暗红色，眼后黑色斑模且较小，上体浅灰色，翼尖黑色，下体白色或粉红色，停歇时候可以观察到初级飞羽翼下全白，飞行时候初级飞羽外侧四根羽毛为白色，仅尖端黑色，观感为翅上白色羽毛较宽。

遗鸥：虹膜褐色，具有白色眼睑，头污白色，枕部及颈部一般具有灰褐色斑，嘴红色较粗厚，因脖子较短，整体看起来较矮胖，停歇时候可以观察到初级飞羽上部白色斑块较明显、下部有一椭圆形白斑。



1.0

你的眼睛真的是尺吗？

开始测试 ··:

请将图片与其正确的名字连起来



棕头鸥

红嘴鸥

细嘴鸥

遗鸥



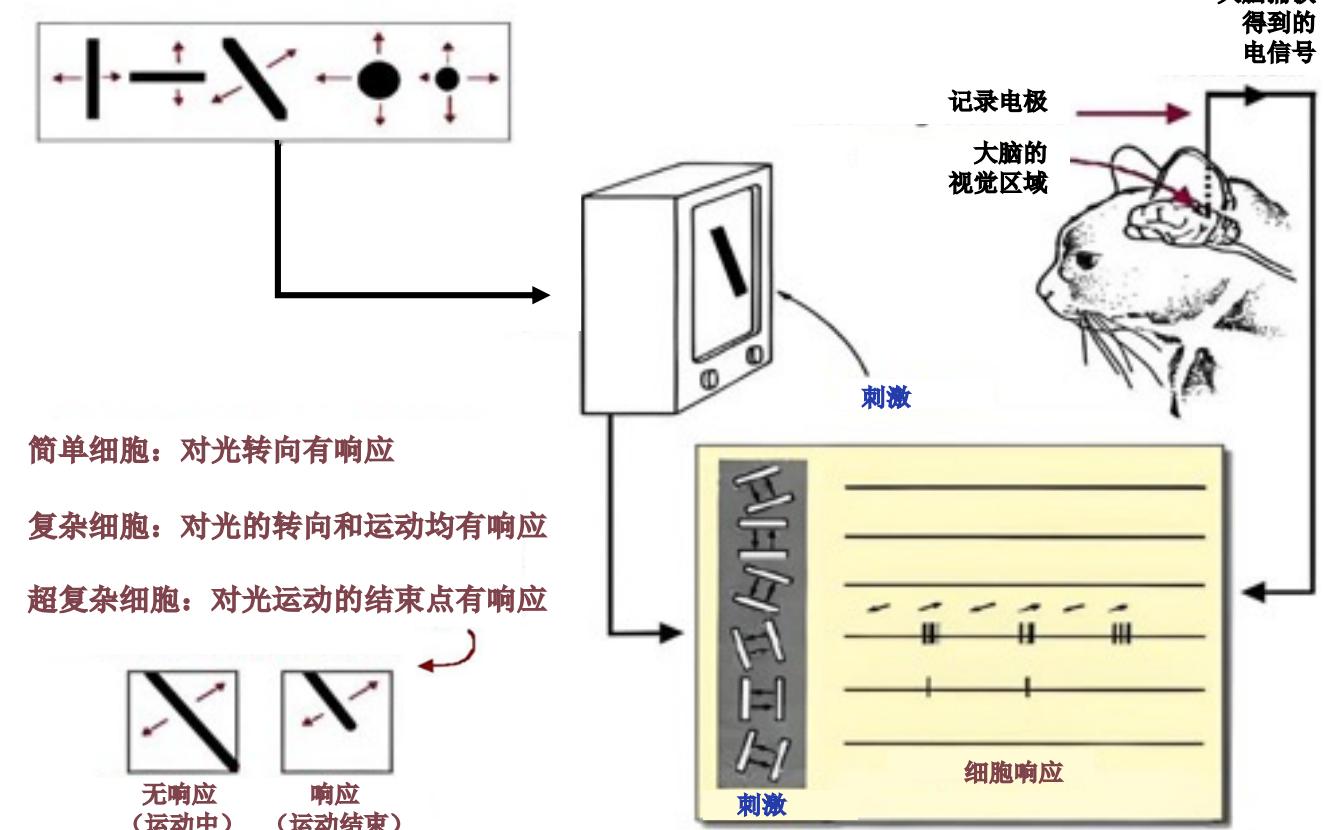
- 计算机视觉是一门让计算机“看见”并理解视觉信息的技术科学。其核心目标是让计算机像人类一样，能够**感知、分析和理解图像或视频**。
- 简言之，计算机视觉就是**让机器“看懂”世界**，从图像或视频中提取有用的信息，并做出相应的判断或决策。



- 图像增强 ✓ 自动驾驶
- 图像分类 ✓ 智慧医疗
- 目标检测 ✓ 无人安防
- 语义分割 ✓ 目标检测
- 三维视觉 ✓ 工业检测
- 图像生成 ✓ 虚拟现实

20世纪50年代——二维图像的分析与识别

- 1959年，大卫·休伯尔和托斯登·威塞尔通过猫的视觉实验，揭示了大脑皮层细胞对不同视觉特征的感知机制，奠定了生物视觉和计算机视觉的基础。
- 1959年，拉塞尔·基尔希等人发明了首台数字图像扫描仪，推动了数字图像处理技术的发展。



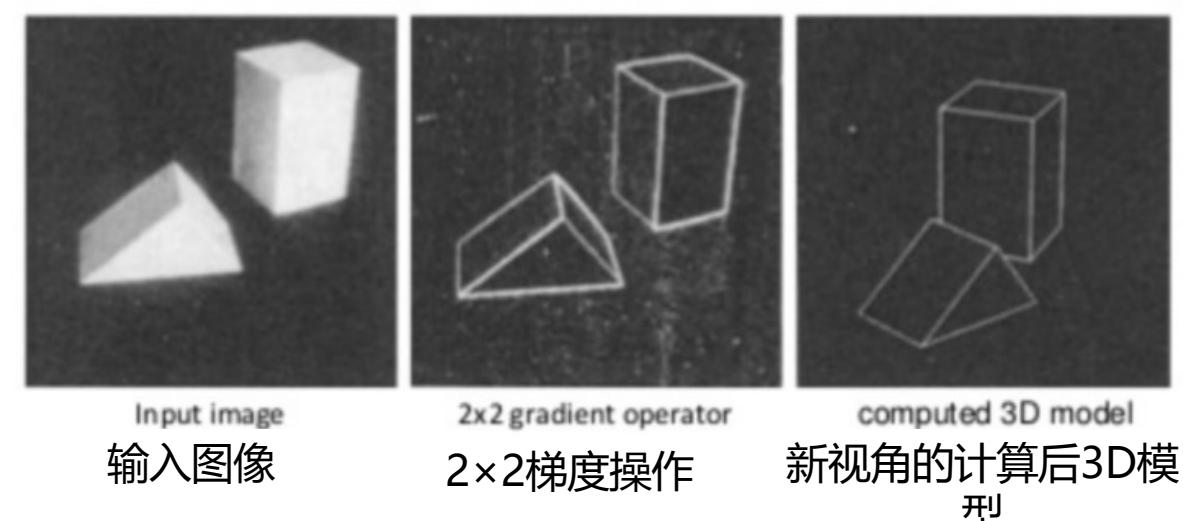
猫视觉实验

20世纪60年代——聚焦三维视觉理解

- 1965年，拉里·罗伯茨在《三维固体的机器感知》中提出了从二维图像提取三维信息的方法，开启了计算机视觉研究的新纪元。
- 1969年，威拉德·博伊尔和乔治·史密斯研发了电荷耦合器件（CCD），极大提高了数字图像采集的效率。



Larry Roberts
1963, 1st thesis of Computer Vision



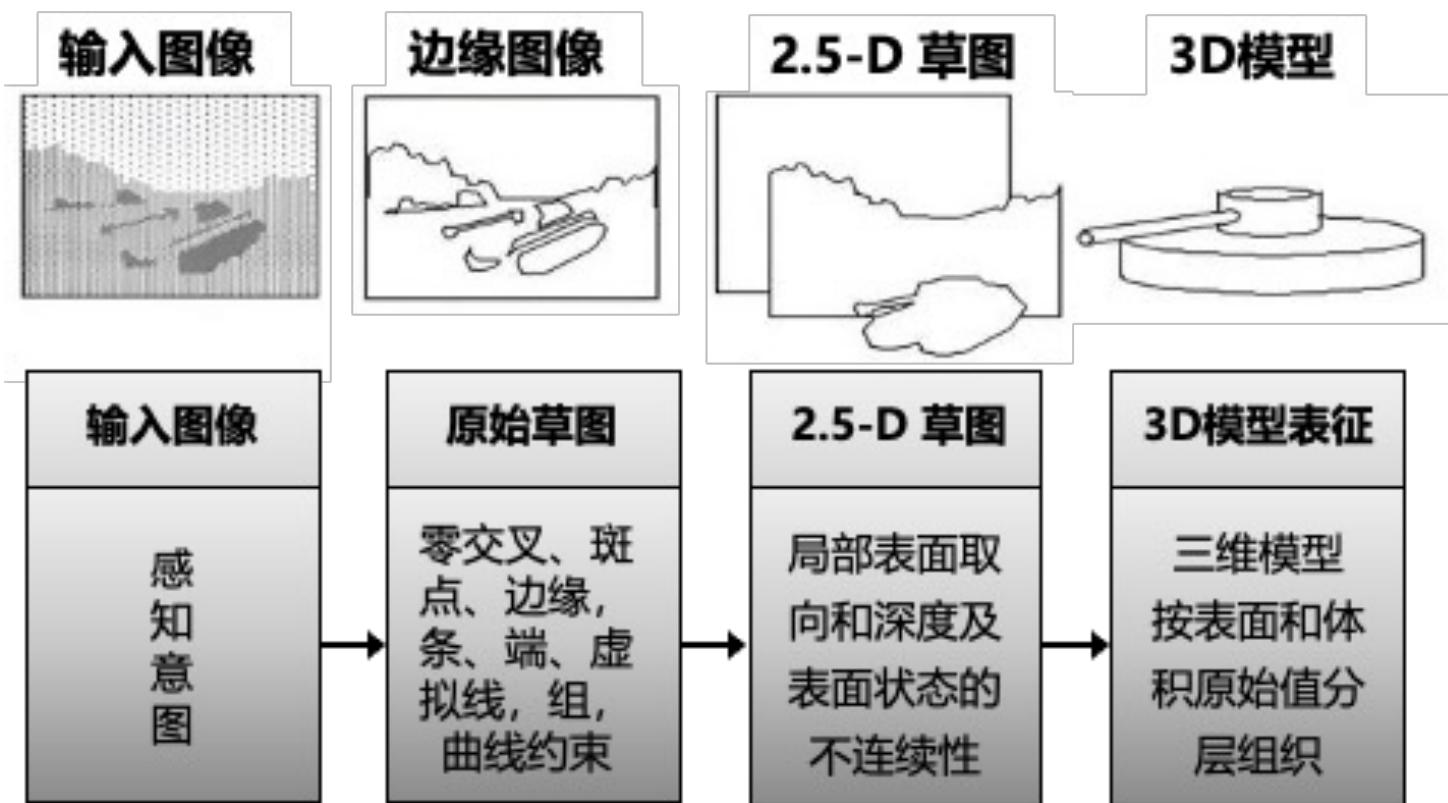
拉里·罗伯茨与《三维固体的机器感知》

20世纪70年代——理论体系的系统化构建时期

- 1977年，大卫·马尔提出“计算视觉”理论，创新性地描述了视觉信息处理的阶段性模型。

20世纪80年代——独立学科的确立与理论向应用的飞跃

- 1982年，大卫·马尔的《Vision》一书推动了计算机视觉理论的成熟，标志着该领域成为独立学科。
- 1982年，日本COGEX公司推出了首套工业光学字符识别系统，标志着计算机视觉开始进入工业应用。



“计算视觉”理论



David Marr被认为是计算机视觉之父，他在视觉信息处理方面的理论对后来的研究产生了深远的影响。为了纪念他，IEEE国际计算机视觉大会设立了马尔奖，每两年评选一次，授予在计算机视觉领域做出杰出贡献的论文。这个奖项代表了计算机视觉研究方面的最高荣誉之一。

20世纪90年代——特征对象识别开始成为研究重点

- 1992年，支持向量机提出，在分类、回归和异常检测等任务中表现出色，成为当时机器学习的主流方法之一。
- 1998年，杨立昆改进了卷积神经网络，提出了LeNet-5模型，奠定了现代卷积神经网络的基础。
- 1999年，NVIDIA推出GPU，开启了并行计算和图像处理的新纪元。

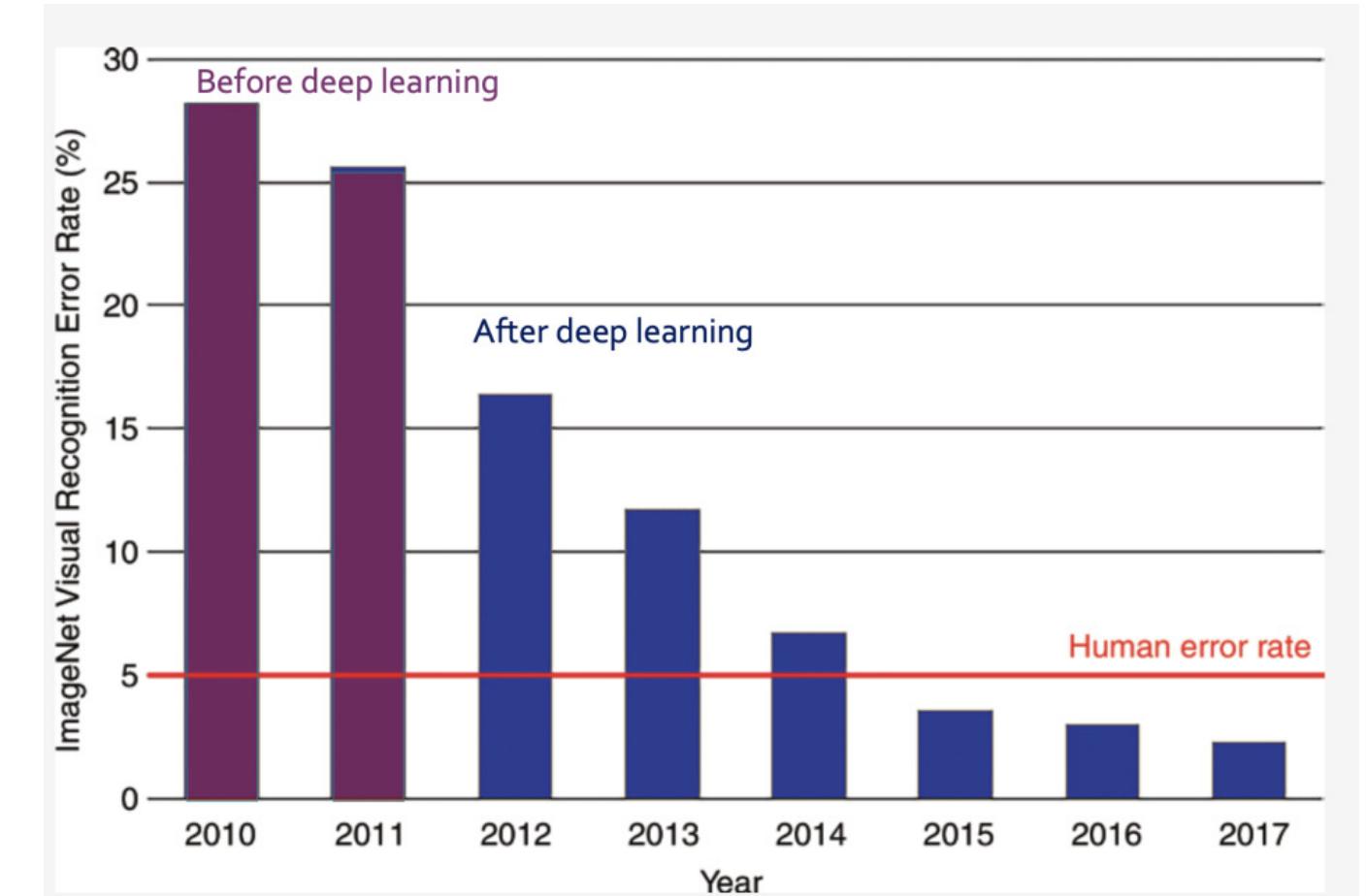
21世纪初，从机器学习迈向深度学习

- 2001年，提出了第一个实时人脸检测框架，首次实现了在普通硬件上实时检测人脸的目标。
- 2005年，提出了基于方向梯度直方图（HOG）的行人检测方法，成为重要的视觉工具。
- 2006年，Pascal VOC数据集推出，极大推动了对象分类技术的发展。
- 2006年，辛顿等人提出深度信念网络，标志着深度学习时代的开始，为后续的卷积神经网络等技术奠定基础。
- 2009年，李飞飞教授及其团队发布ImageNet数据集，极大丰富了计算机视觉领域的研究数据资源。



2010年至今——深度学习引领的全面爆发

- **2010-2017年，ImageNet数据集与挑战赛的推动：**从2010年到2017年，ImageNet挑战赛 (ILSVRC) 成为推动深度学习技术快速发展的重要平台，极大提升了目标检测等视觉任务的性能。
- **2012年，AlexNet的突破：**2012年，亚历克斯·克里切夫斯基、伊尔亚·苏茨克维和杰弗里·辛顿提出的 AlexNet深度卷积神经网络赢得ImageNet竞赛，标志着深度学习正式成为计算机视觉的核心技术，推动了卷积神经网络的广泛应用。



深度学习给ImageNet挑战赛带来的性能突破

杰弗里·辛顿



A close-up portrait of a woman with dark, shoulder-length hair. She is looking slightly upwards and to her right with a neutral expression. The background is blurred, showing warm autumn colors.

AI女神李飞飞

希

一个底层移民的逆袭

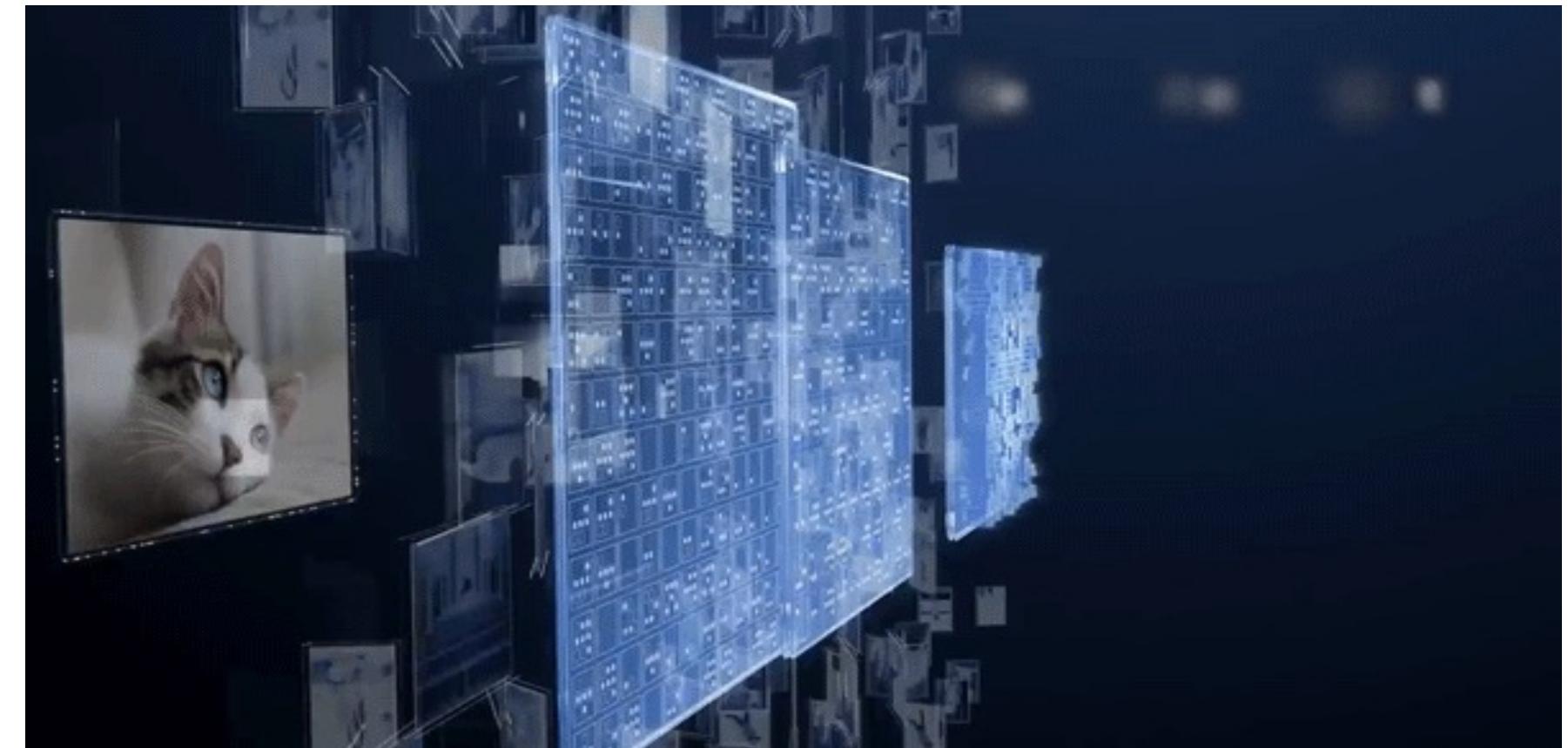
- 2014年，伊恩·古德费洛提出**生成对抗网络（GAN）**，通过生成器与判别器竞争优化，革新了图像生成技术，成为计算机视觉领域的重大突破之一。
- 2016年，林宗毅等人提出**特征金字塔网络**，实现多尺度特征融合。
- 2017年，**PyTorch和TensorFlow的普及**：PyTorch和TensorFlow两个深度学习框架的崛起，迅速成为研究人员的主流工具，为包括图像分类在内的多项任务提供了强大的支持。
- 2017年，何恺明等人提出的**Mask R-CNN**在实例分割任务上实现了重大突破，其创新的RoIAlign技术显著提升了分割精度，成为该领域的里程碑式工作。
- 2020年，**Transformer模型进军计算机视觉**：Vision Transformer (ViT) 的提出标志着Transformer模型正式应用于视觉任务，并迅速成为目标检测、语义分割和图像分类等任务的基础架构。
- 2020年，**扩散模型**作为生成模型的新范式，它不仅能够生成高质量、高分辨率的图像，还支持灵活的图像编辑功能，为图像修复、风格迁移、超分辨率重建等任务提供了新的解决方案。
- 2023年，**Segment Anything Model (SAM)** 模型凭借其零样本学习能力，在视觉分割领域取得突破，为下游应用提供了广泛的可能性，预示着计算机视觉向智能化和自动化迈进。

■ 图像分类目的

图像分类是指将输入的图像分配到一个预定义的类别集合中，类别可以是“猫”、“狗”或“车”等特定物体。

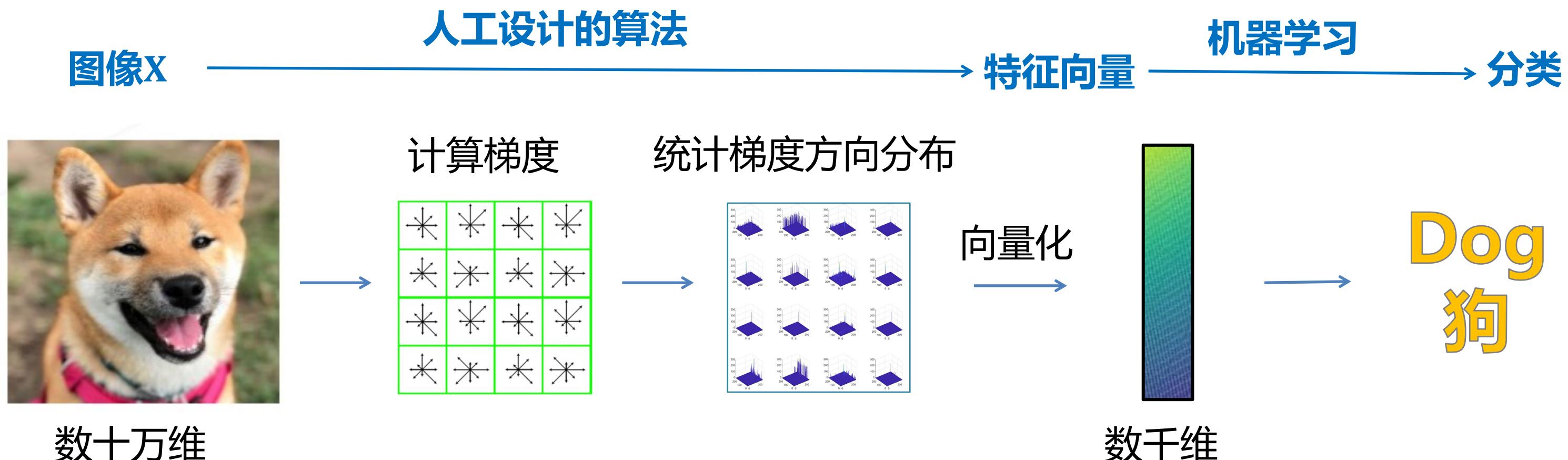
■ 广泛的应用场景

在多个领域都有广泛应用，例如自动驾驶中的行人识别、医学图像中的疾病检测、农业中的病虫害识别等。它是许多视觉任务的基础，例如面部识别、物体检测、场景理解等。随着深度学习的发展，图像分类已成为人工智能的重要应用之一。



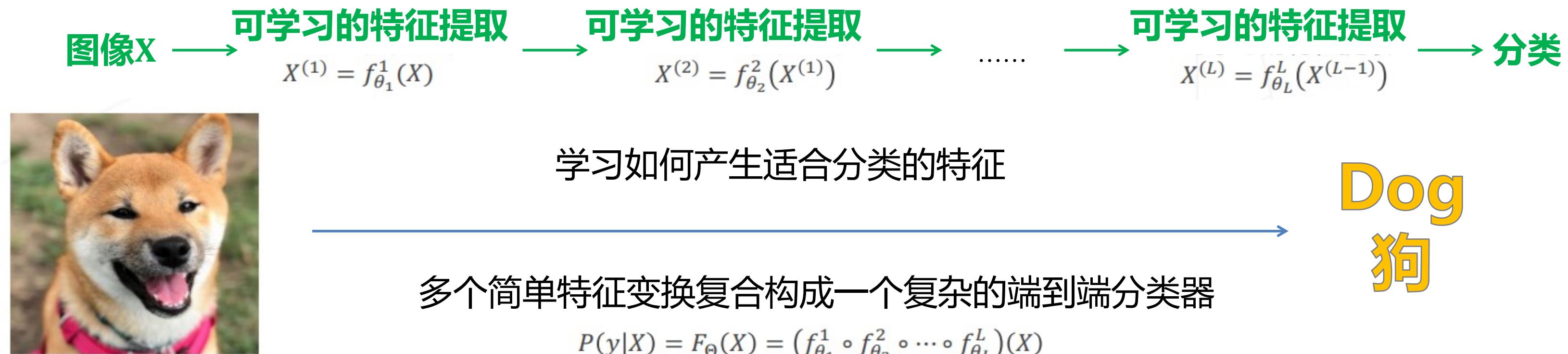
■ 传统图像分类方法

- 传统方法通常包括数据预处理、特征提取（如SIFT、HOG等）和分类器的训练（如支持向量机、K-近邻等）。
- 虽然这些方法在某些情况下能取得较好的效果，但其依赖于手工设计的特征提取过程，难以应对复杂任务。



■ 深度学习的崛起

- 随着计算能力和数据规模的提升，深度学习模型（如AlexNet、VGGNet、ResNet等）可以自动从图像中学习层次化的特征表示，极大提高了图像分类的准确性和效率。
- 深度学习通过卷积神经网络（CNN）等架构，减少了对手工特征的依赖，并且能够在大规模数据集上取得优异表现。



■ 图像分类的实际应用

- **人脸识别：**杭州萧山国际机场通过阿里云ET航空大脑的人脸识别技术，大幅提升了安检和食堂支付效率，旅客通过安检时仅需3秒，人脸识别还能有效应对整容、照片更新不及时等复杂情况，方便特殊人群使用。
- **医疗影像诊断：**针对青光眼筛查，腾讯觅影团队研发了基于深度学习的青光眼分类模型，准确率超过95%，显著提高了筛查效率和早期青光眼的诊断准确性，帮助缓解医疗资源不均问题。



■ 定义

细粒度图像分类是指在大类别（如“狗”或“鸟”）中，进一步细分到子类别（如不同品种的狗或鸟），从而**识别具有细微差异的物体**。

■ 挑战

由于细粒度图像分类**需要对局部细节做出精确判断**，环境因素（如光照、角度、背景等）容易干扰分类结果。此外，如何让模型学会区分细微差别并保持高准确性是一个难点。

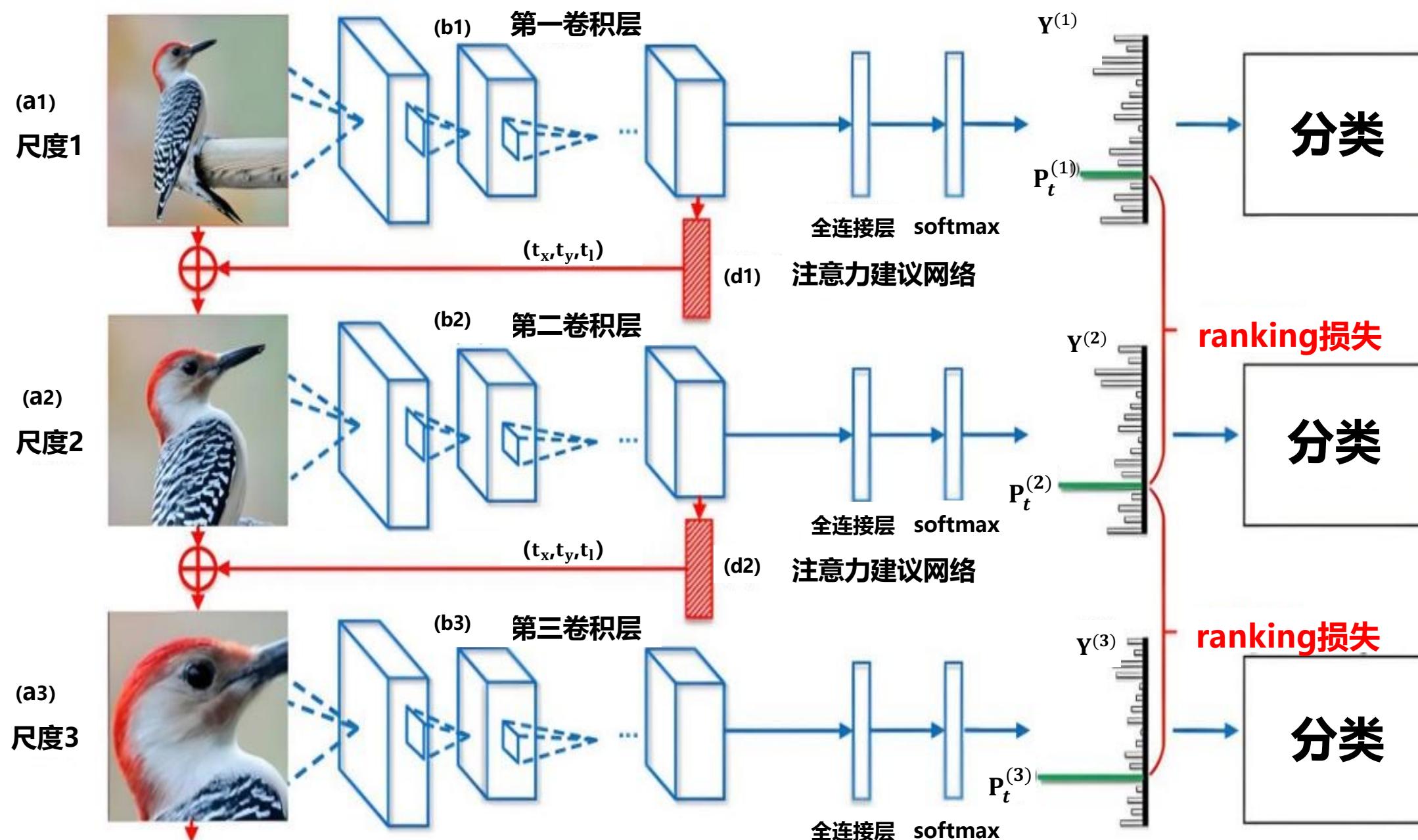
■ 应用

生物多样性保护与研究、医疗与生命科学、艺术品与文化保护、电子商务与零售等。



1.2 图像分类

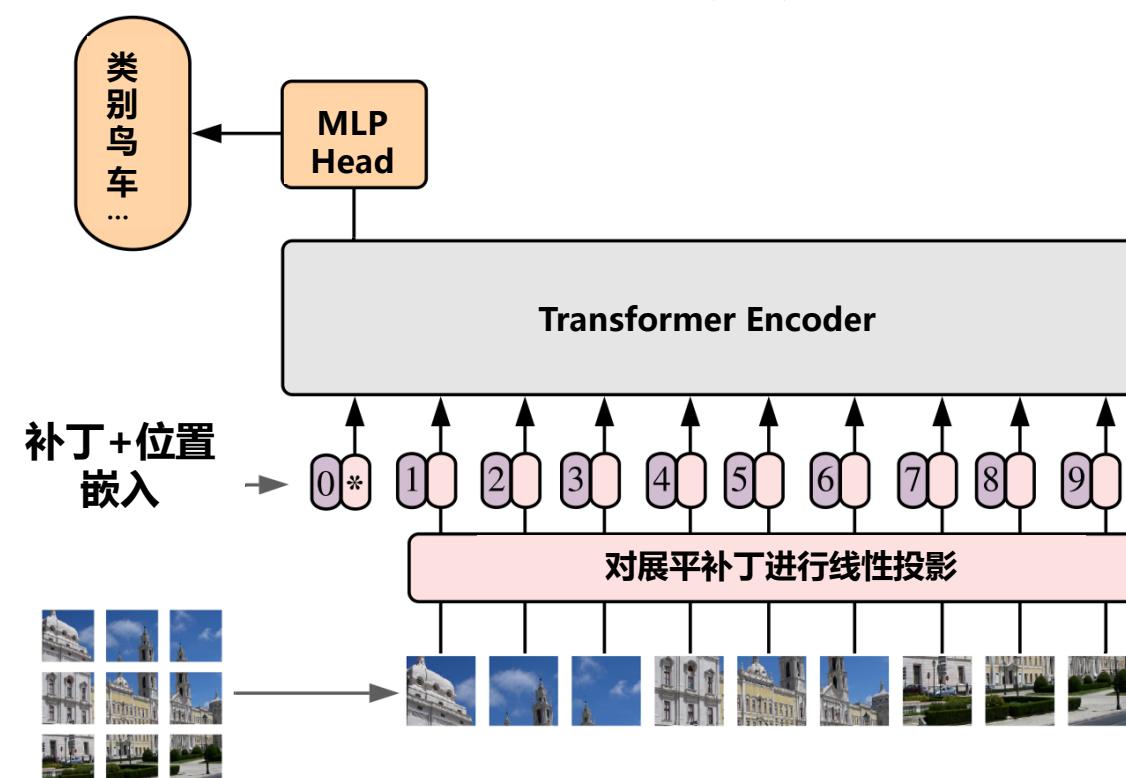
■ 经典方法——循环注意力卷积神经网络 (RA-CNN)



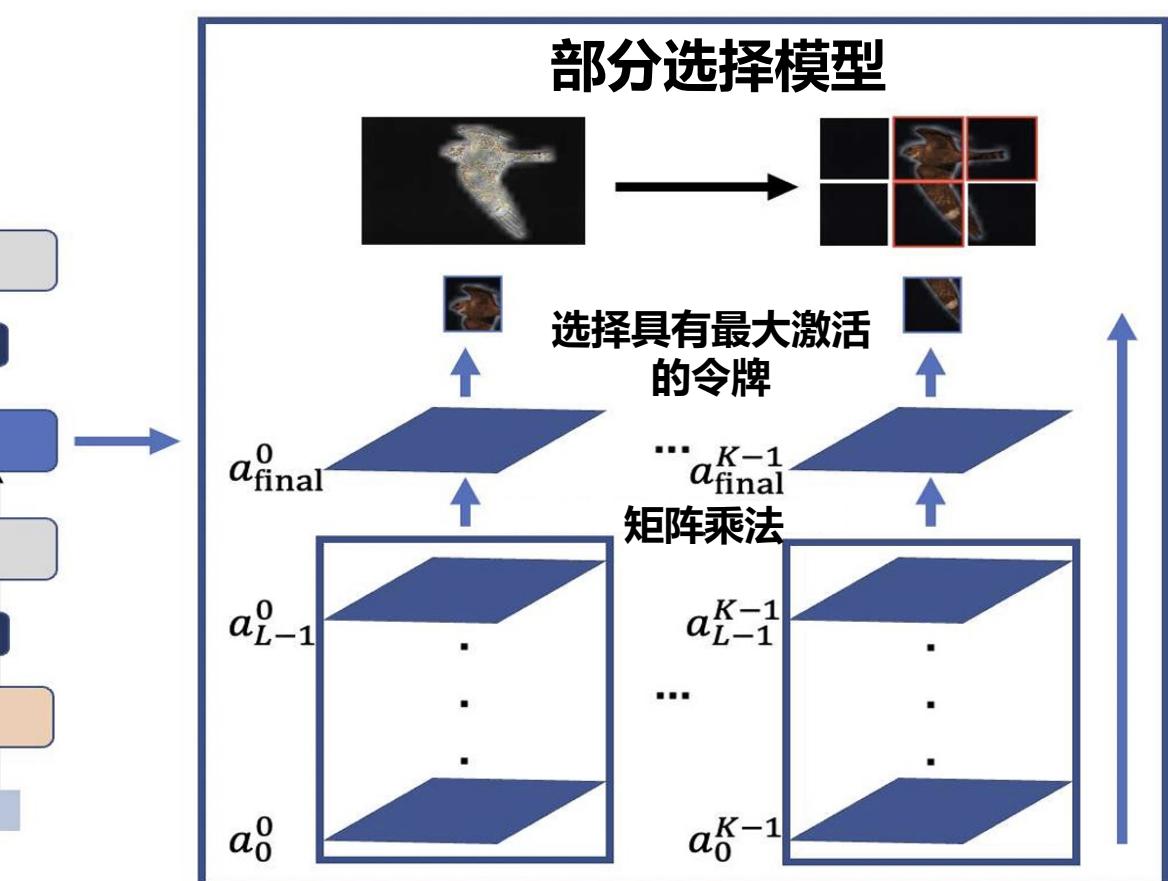
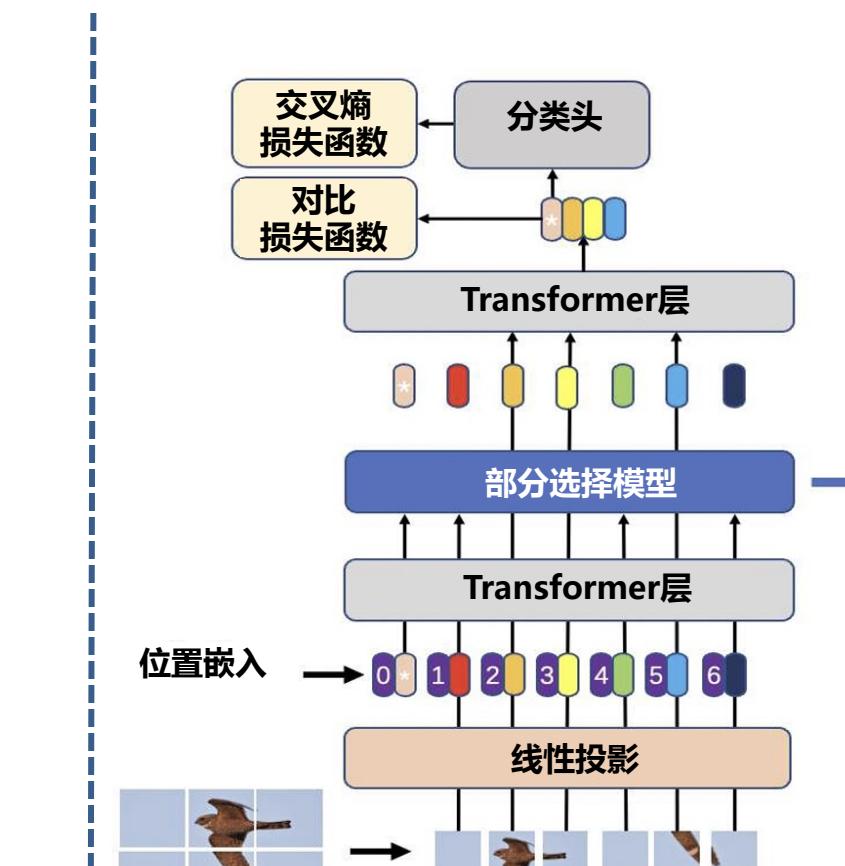
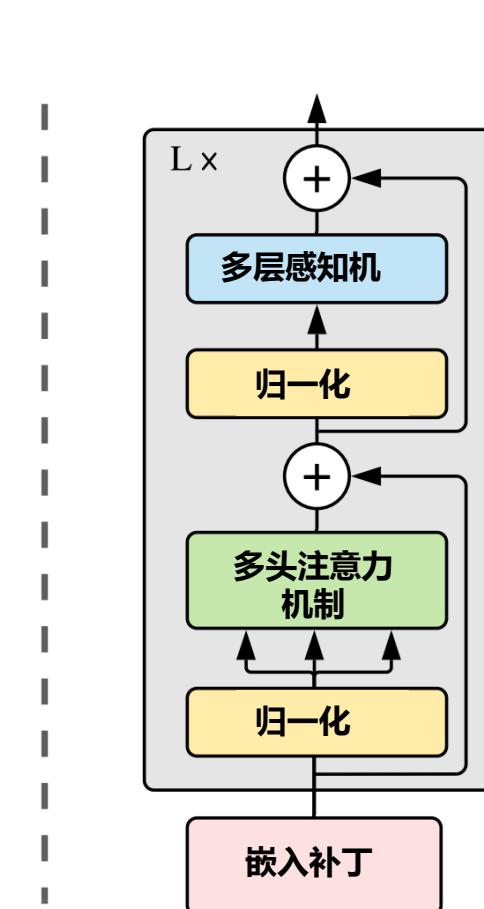
- 结合CNN与注意力机制：RA-CNN 将卷积神经网络与注意力机制相结合，使模型能够**自动聚焦于图像中的关键信息区域**，避免了对冗余信息的过度处理，提升了计算效率和精度。
- 递归结构：RA-CNN 引入了递归结构，采用递归计算，**逐步调整关注区域**，增强记忆与动态调整能力。
- 自适应特征提取：RA-CNN 能够通过反复的**迭代调整对图像特征的关注**，从而实现更精准的特征提取。

■ ViT模型的引入

➤ **Vision Transformer (ViT)** 模型将Transformer架构引入图像分类，通过自注意力机制捕捉图像中的全局信息。不同于传统CNN，ViT将图像分割成若干小块，并将这些块作为序列输入Transformer。通过自注意力，ViT能够建模图像的全局空间关系，尤其适用于大规模数据集。



Vision Transformer (ViT) 模型



TransFG模型

■ 挑战及未来的研究方向

- 细粒度图像分类的未来发展方向将围绕精度提升、效率优化、鲁棒性增强、数据依赖减少、应用场景扩展等方面展开。
- 例如，通过结合Transformer的全局信息捕捉能力与CNN的局部特征提取能力，可以进一步提升细粒度分类的性能。开发更加适合细粒度分类的大规模、高质量数据集将成为重点，数据集的提升对于模型的泛化能力至关重要。

特性	CNN(卷积神经网络)	Transformer
主要应用领域	图像处理、物体检测、图像分类	自然语言处理(NLP)、图像处理、序列任务
核心机制	卷积操作、池化操作	注意力机制(Self-Attention)
上下文捕获能力	局部上下文(通过感受野逐层扩大)	全局上下文(直接捕获任意位置间依赖关系)
并行化处理	卷积操作并行，但需逐层处理	自注意力机制完全并行处理
计算复杂度	通常较低，尤其是浅层网络	随着输入序列长度增加，计算复杂度增加
特征提取	固定卷积核，擅长局部特征提取	自适应特征提取，根据输入动态调整权重
模型参数量	较少(依赖卷积核数量和大小)	较多(依赖注意力头和层数)
训练速度	通常较快，适合硬件加速	并行性好，但因参数量大训练时间可能更长
处理长距离依赖	较弱(需多层卷积扩展感受野)	很强(直接通过自注意力机制处理)
数据结构假设	强(适用于有局部相关性的图像数据)	弱(适用于各种类型的数据)



更大规模的细粒度分类数据集



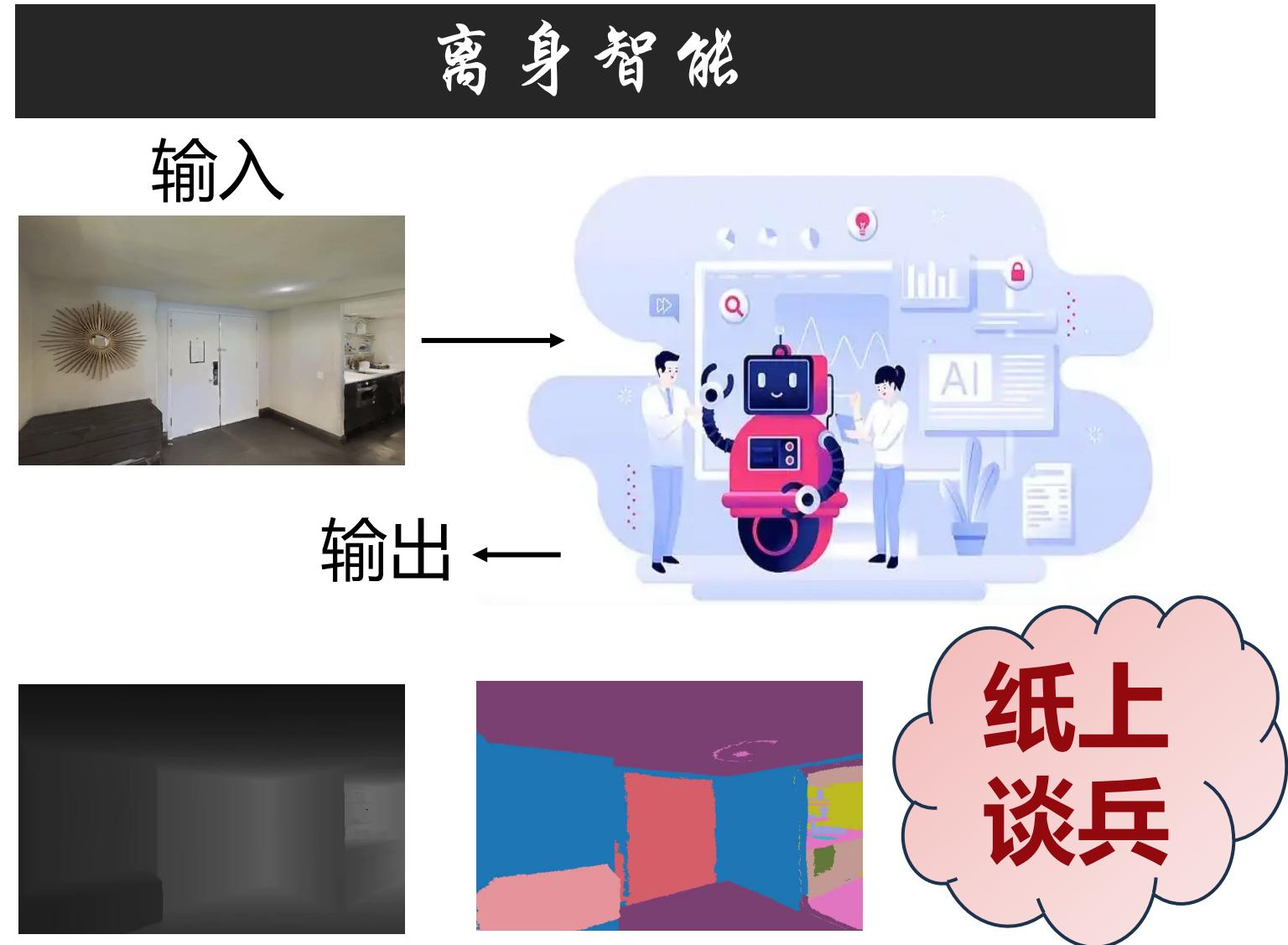
02

智能之躯——具身智能

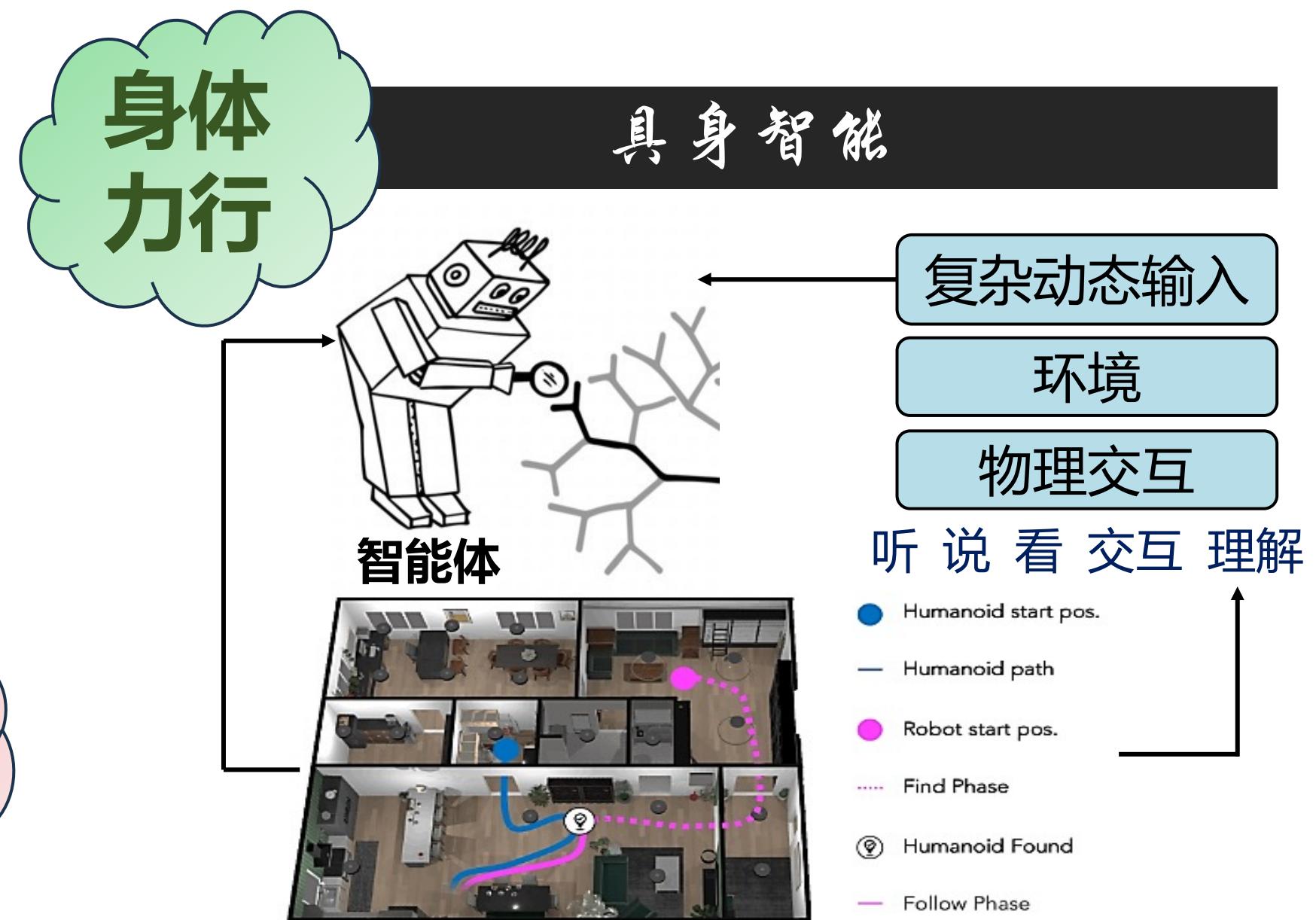
什么是具身智能？

2.1

人工智能的机械战甲——具身智能



单一的符号智能往往与真实世界相脱节，
认知与身体解耦

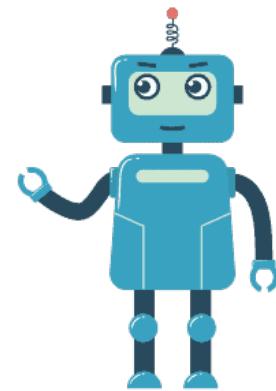


智能是具身化和情景化的，具身智能可通过
与真实世界的交互完成任务

- 具身智能是以智能体作为本体支撑，不再局限于被动响应，而是能够像生物体一样，主动适应环境变化，应对噪声干扰，并适时调整自身行为。

2.1

人工智能的机械战甲——具身智能



AI智能体

机器人等人工智能系统。



感知输入

搭载各类传感器收集环境信息，以支持智能系统分析与学习。



达到解决复杂问题的目标，
例如运动规划、导航等。



交互学习

在与环境交互中学习，
直至达成目标。



世界模型

能对世界时空维度的信息进行抽象表达和理解。

- **具身化 (Embodiment)** : 是一个跨学科概念，强调身体（本体）在认知、情感、社会互动中的核心作用，反对传统的身心二元论。
- **具身智能 (Embodied AI)** : 特指那些**拥有物理形态**，并能**直接参与物理世界交互**的智能系统，如服务型机器人、智能无人驾驶车辆等。它们通过“身体力行”的方式，展现出高度的环境适应性与任务执行能力。
- **具身任务**: **像人类一样**通过观察、移动、对话以及与世界互动从而完成的一系列任务。
- **多模态**: 一个模型或系统**能够处理多种不同类型的输入数据**并融合它们生成输出，如文本、图像、音频和视频等。这种能力对于提升智能系统的环境感知与决策能力至关重要。
- **主动交互**: 机器人或智能体**与环境的实时交互**过程，从而提高智能体的学习、交流与处理问题的能力，是具身智能实现高效任务执行的关键。

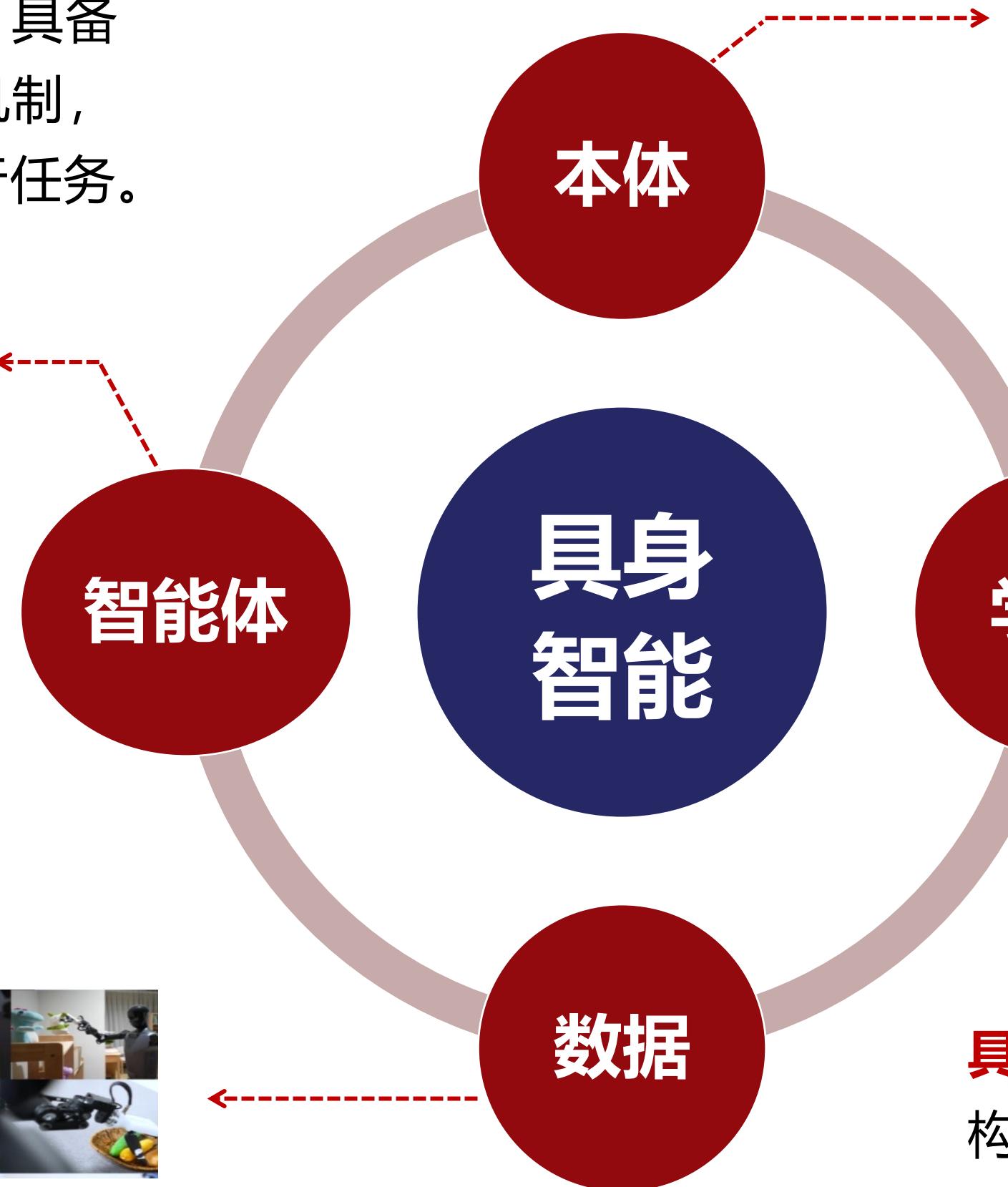
2.1

人工智能的机械战甲——具身智能

智能体作为本体的智能核心，具备敏锐的感知能力和动态决策机制，能够解析复杂环境并高效执行任务。

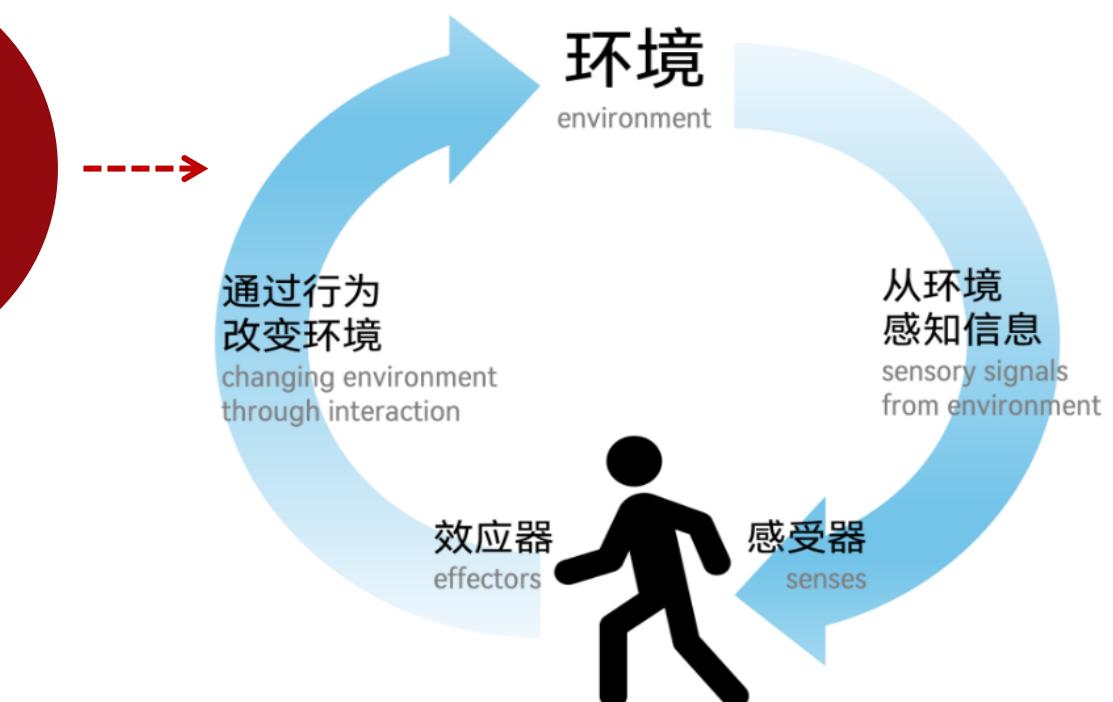


数据在机器学习与具身智能系统中至关重要，通过整合大量多样的具身数据，提升了智能体的任务执行成功率。



固定式机器人 自动驾驶汽车 四足机器人

本体在物理与虚拟空间中承担环境感知和任务执行。



具身学习通过智能体与环境及人类的互动，构建“感知-决策-行动(控制)”闭环。

➤ 核心技术1：具身感知

- ✓ **主动视觉感知：**智能体能够自主控制感知设备，如选择最佳视角和运用注意力机制。这种能力允许**智能体主动探索环境**，优化信息获取，从而提高任务执行效率。
- ✓ **三维视觉定位与物体感知：**智能体**需具备在三维空间中定位自身及周围物体的能力**，这对导航和物体操作至关重要。现代视觉编码器预训练技术增强了对物体类别、姿态和几何形状的精确估计，使智能体能在复杂动态环境中全面感知。
- ✓ **多模态感知融合：**除了视觉之外，触觉和听觉等感知模态同样重要，它们为智能体提供额外的环境信息。触觉传感帮助智能体感知物体的质地、重量和形状，支持精确的物体操作。**整合多模态感知数据**，能够显著提升智能体对环境的整体理解能力，使其在执行任务时更加灵活和高效。

➤ 核心技术1：具身感知的发展方向

感知大模型

- SAM: 视觉分割大模型
- DINO-X: 目标检测大模型



静态环境识别精度与人类相当

被动感知

第三人称

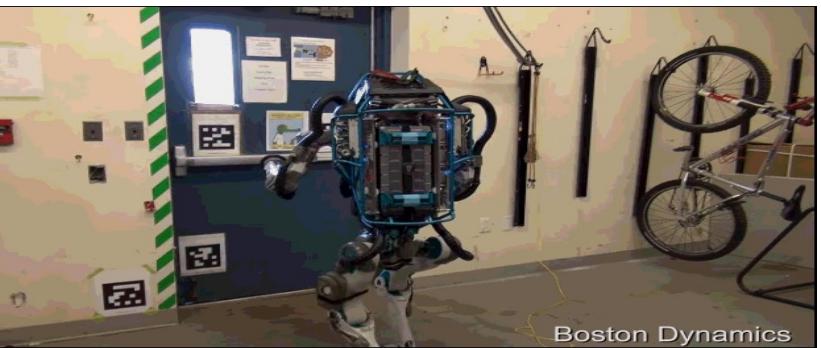
具身主动感知

通过主动获取图像，相比现有大模型的目标检测性能有显著提升



主动感知

第一人称



被门挡住了视线？

交互后，推开门看看是什么，能干什么？

具身交互感知

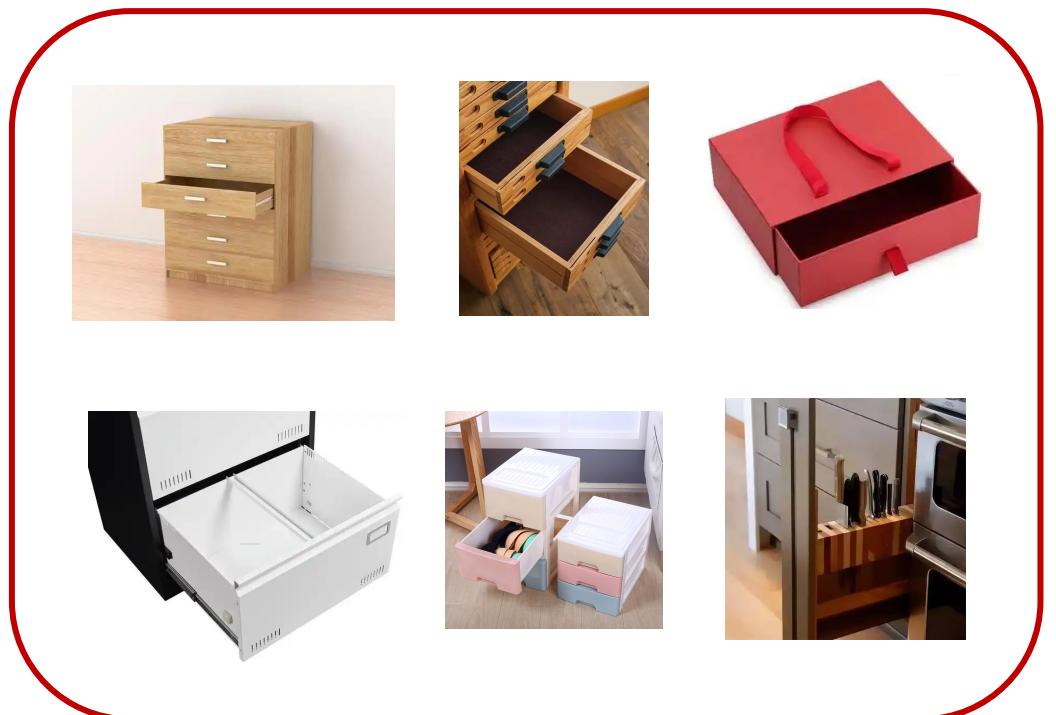
第一人称

行为交互+感知

具身感知模式从被动感知向主动交互感知方向发展

➤ 核心技术1：具身感知与非具身感知的区别

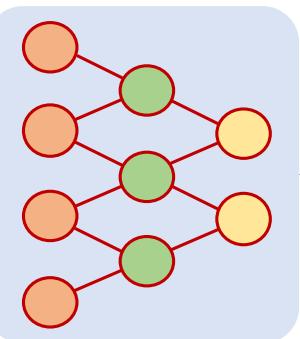
第三人称智能



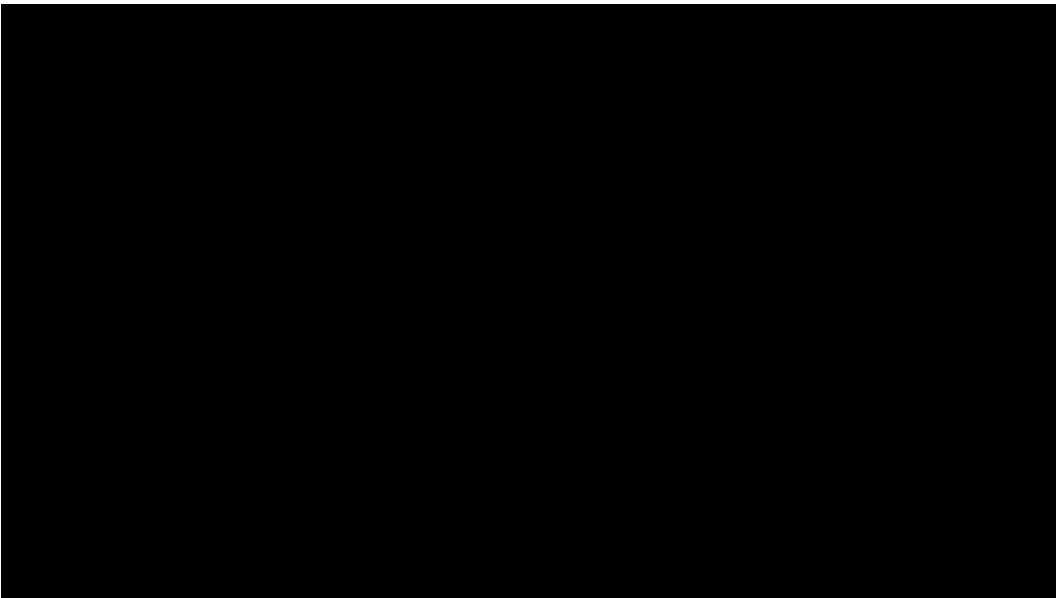
别人告诉我这就是抽屉

(a) 非具身感知

标签：抽屉



第一人称智能

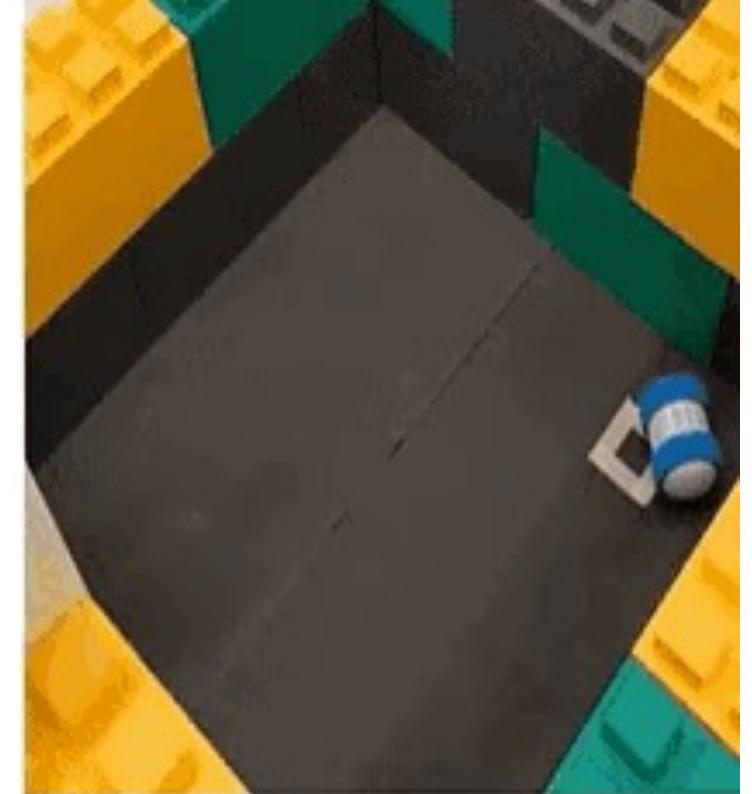
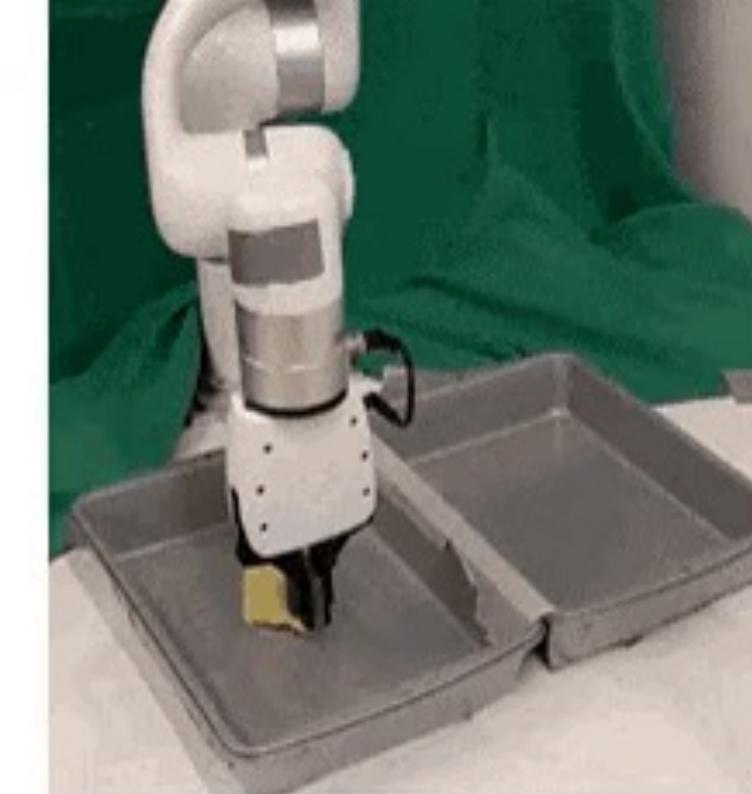
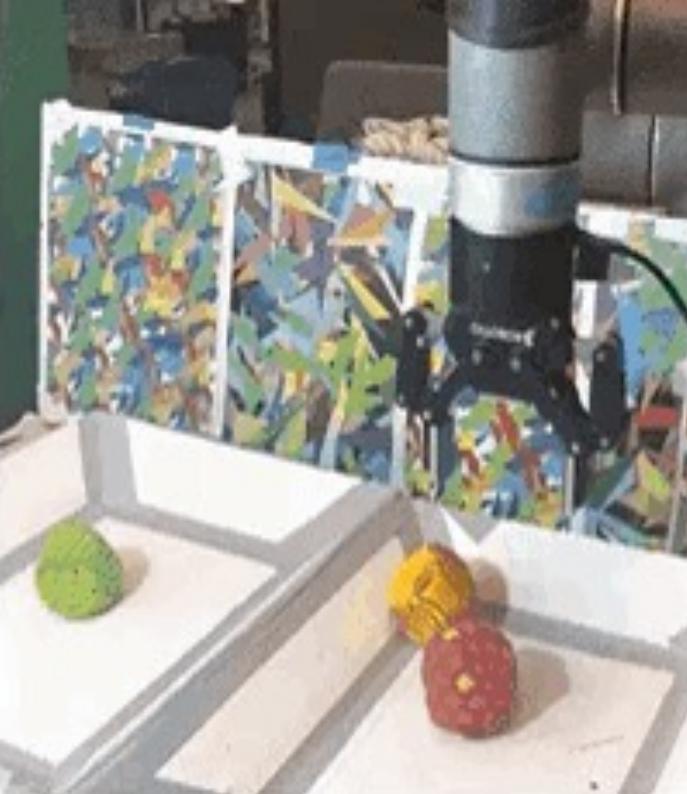


可以打开可以装东西
我亲身体验抽屉是什么

(b) 具身感知

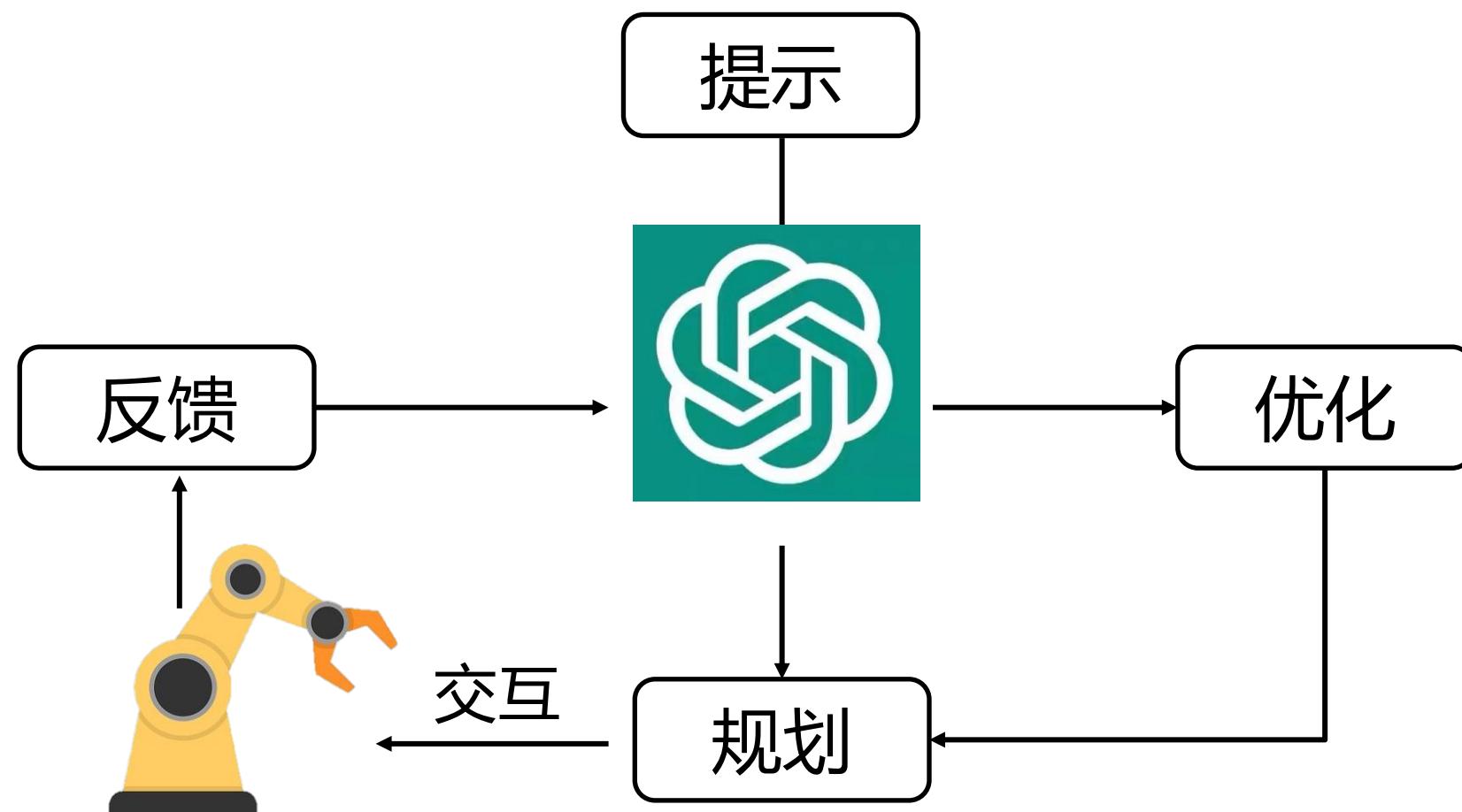
➤ 核心技术2：行为模块的应用示例

✓ 行为模块是**连接感知与行动的纽带**，它基于丰富的感知数据或人类指令，操纵智能体执行复杂的物体操作任务。这一过程融合了语义理解、场景感知、决策制定与稳健的控制规划。

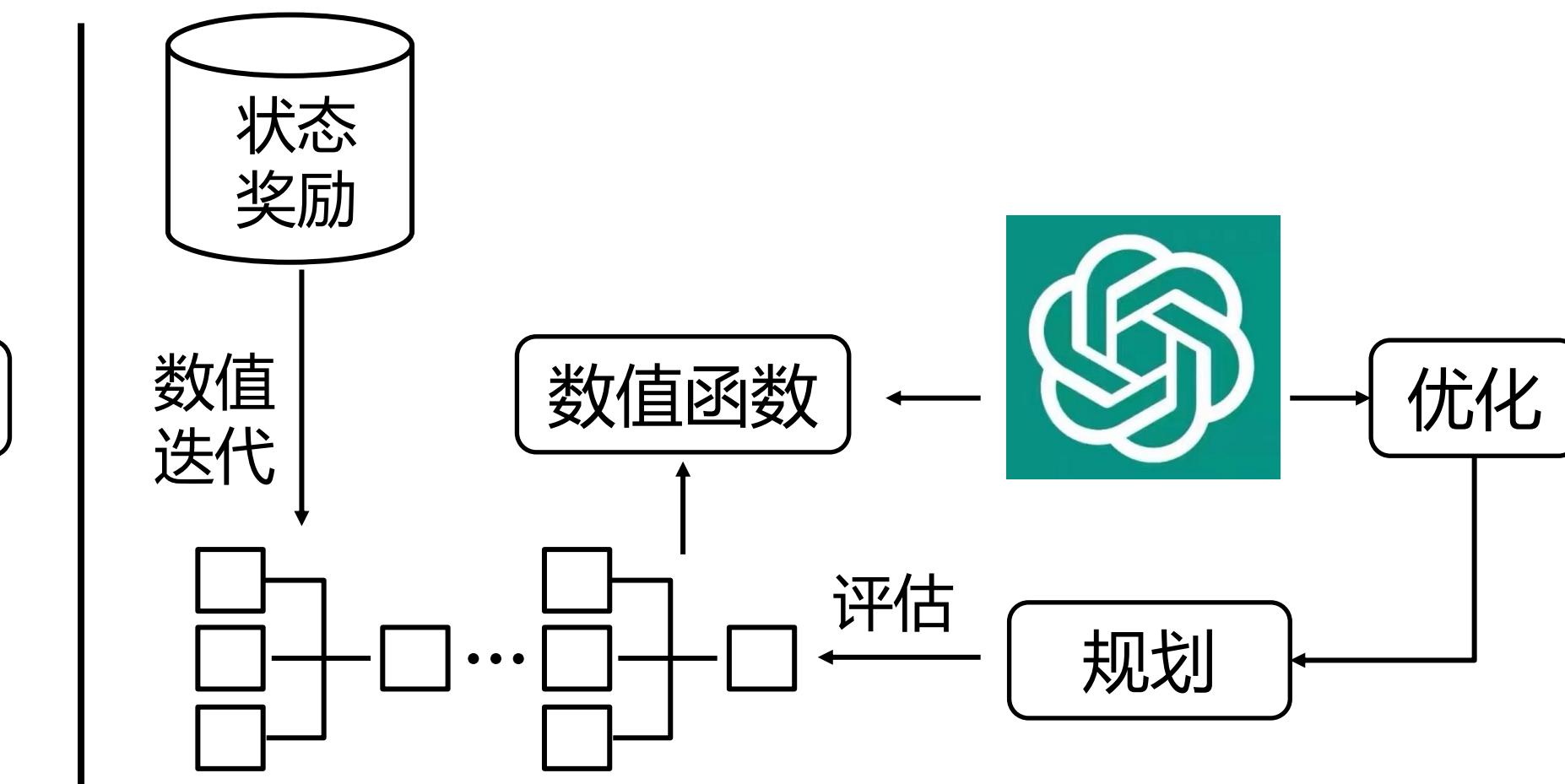


➤ 核心技术2：行为模块的规划策略

- ✓ 行为模块规划策略包括基于**物理反馈**和**强化学习**的规划。前者通过人类提示和实时反馈确保任务准确性，适用于高精度控制；后者通过交互和奖励机制优化策略，适应复杂动态环境，实现自主学习和快速适应。



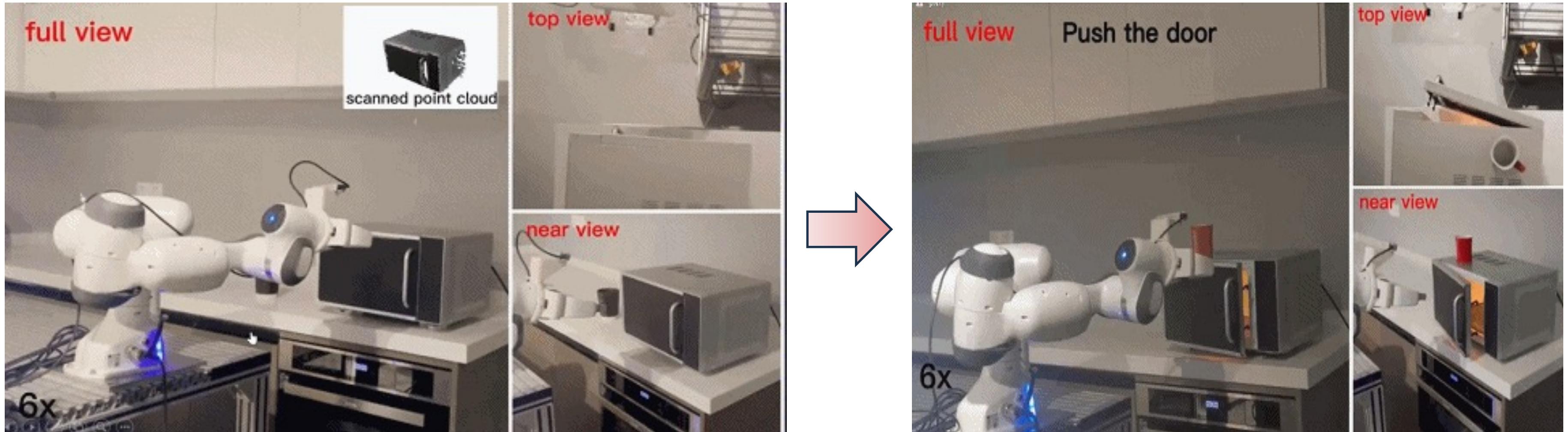
(a) 基于物理反馈的规划



➤ 核心技术3：具身交互概念

- ✓ **人类监督与反馈的重要性：**人类在监督智能体行为轨迹的同时，确保其行动符合需求，并保障交互的安全、合法及道德边界。尤其在医学诊断等敏感领域，人类监督能有效弥补数据局限性与算法能力的不足。
- ✓ **从被动感知到主动交互的转变：**智能体通过在线互动实现模型发展与进化，从第三人称的被动感知转向第一人称的主动交互感知。如智能体能够通过行为交互主动适应实际场景，如“被门挡住视线”的情况。
- ✓ **人类与智能体交互的两种范式：**具身交互分为“不平等互动”模式，即“指导者-执行者”范式，人类发布指令，智能体辅助完成任务；以及“平等互动”模式，智能体与人类共同决策，预示更加协同的未来。

➤ 核心技术3：具身交互示例



初始状态下，机器人拉微波炉门的准确度较差

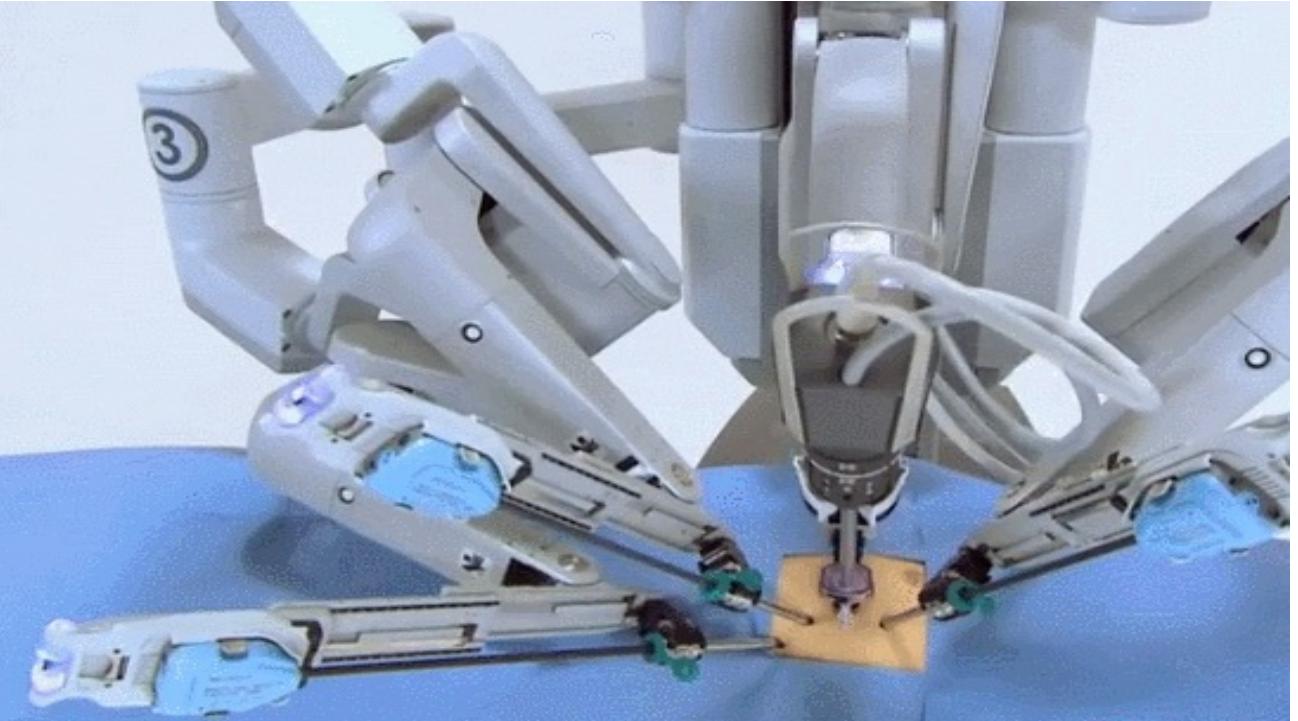
具身交互让机器人学会准确拉开微波炉的门，之后就可以在上面加技能了，比如把东西塞进去

2.2

具身智能应用——智能机器人



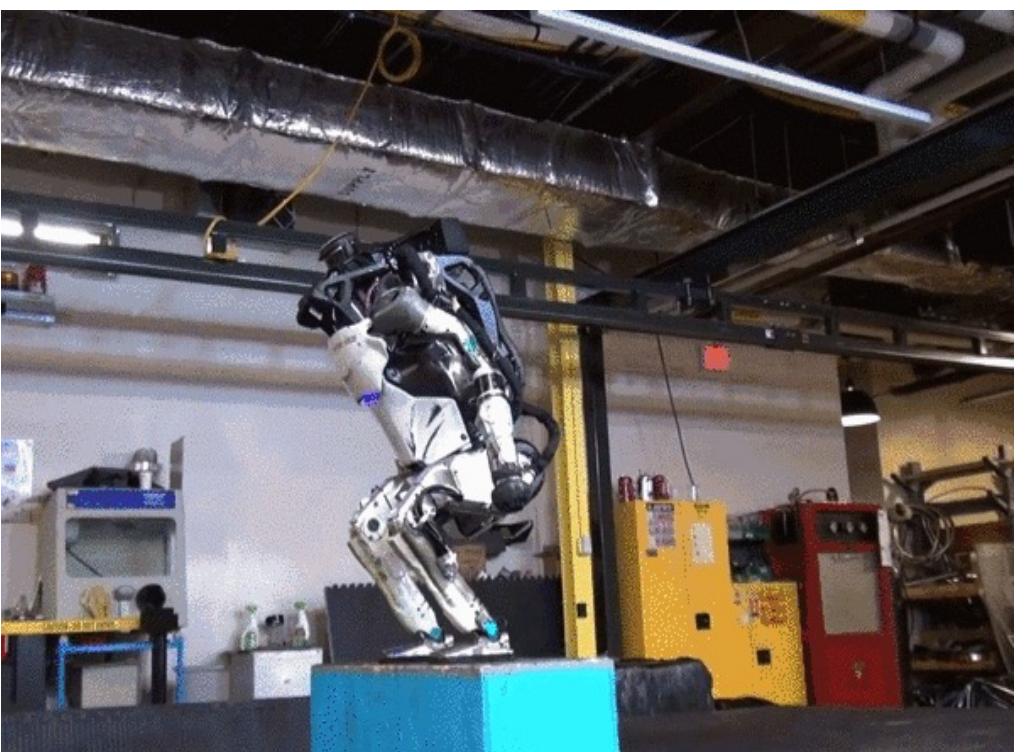
装配空调



手术机器人



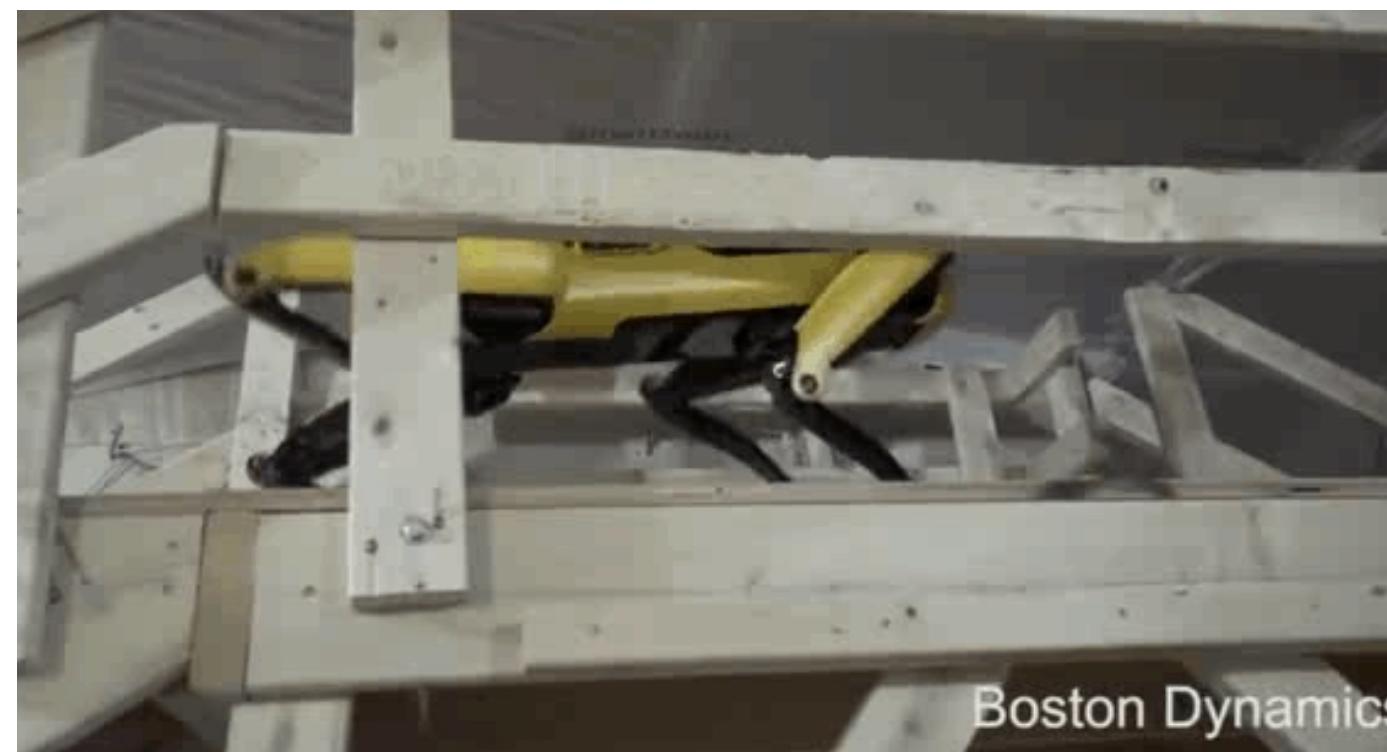
工业喷涂



后空翻动作



家庭服务



机器狗巡检

CCTV 4
中文国际

cctv.com
央视网

人形机器人

AI

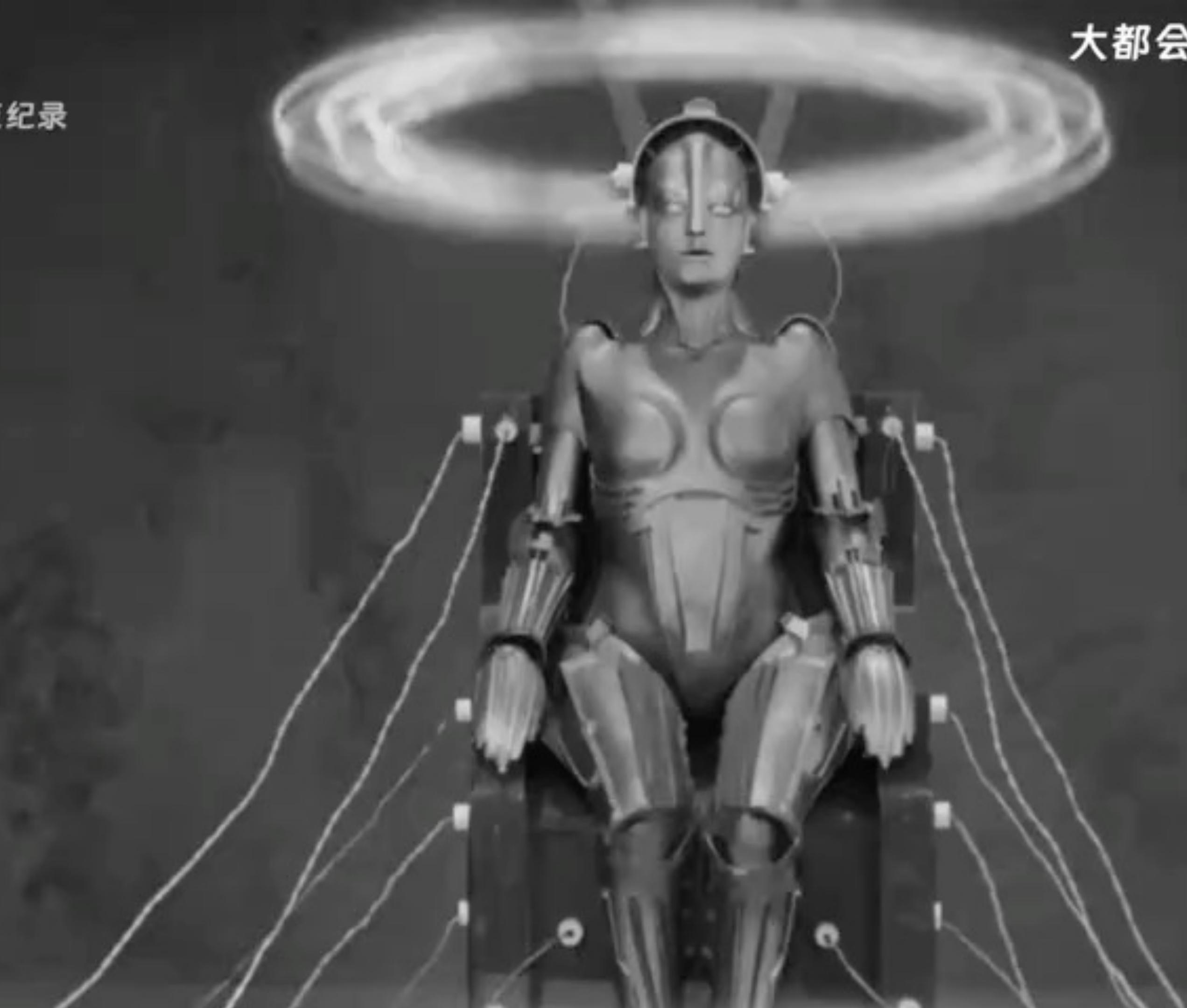
今日
亚洲

今日
亚洲



大都会Metropolis (1927)

Look
— 路克纪录





03

AI卫士——智慧医疗



3.1

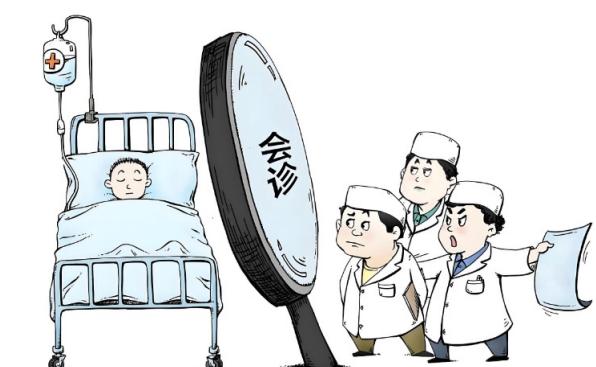
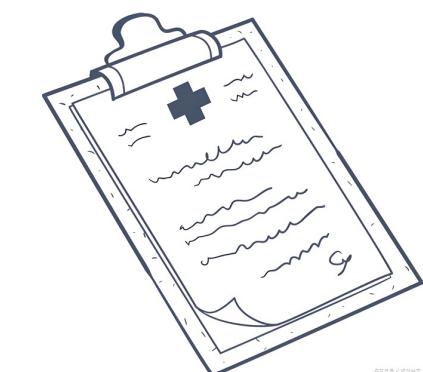
医生的智能助手：智慧医疗



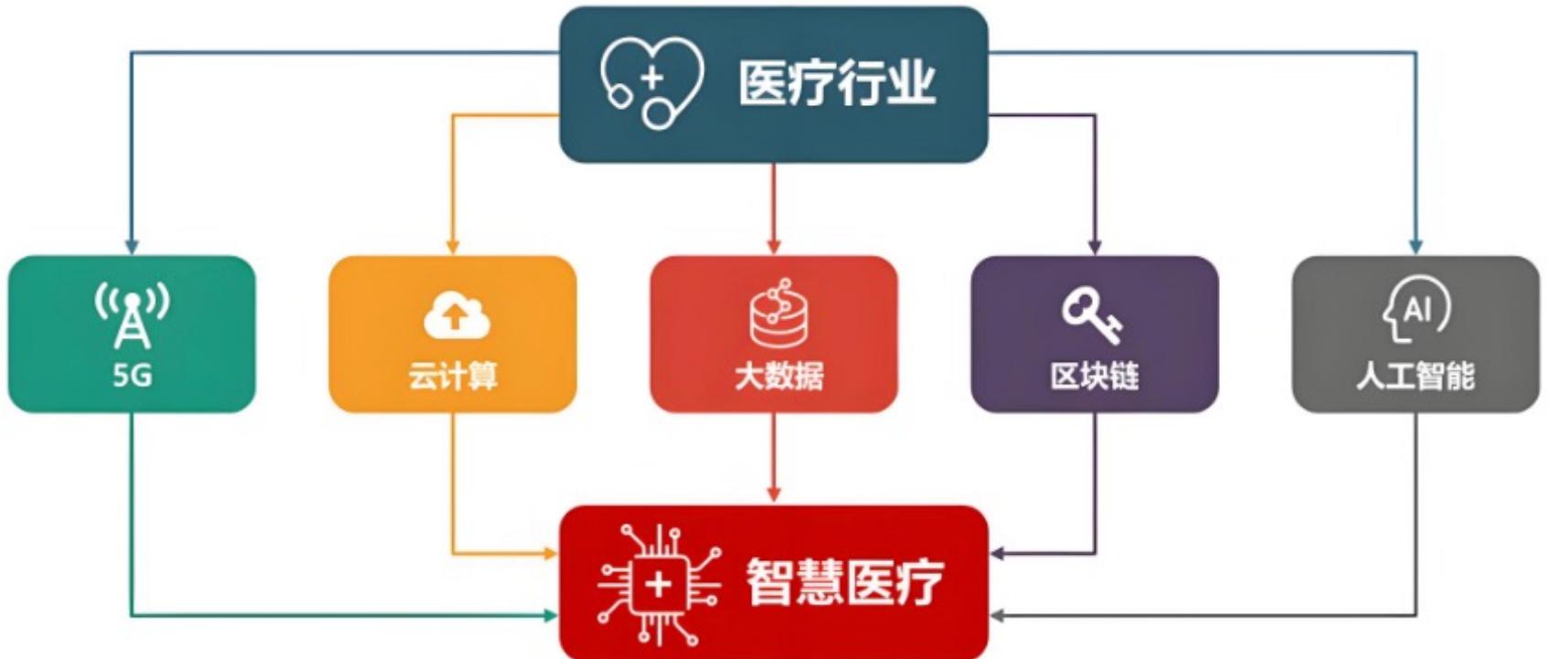
传统中医

- 医疗资源配置不均衡
- 医疗数据的记录和查询繁琐，不利于管理
- 依靠医生的经验和专业知识进行辨证论治
- 注重患者与医生面对面交流

传统西医

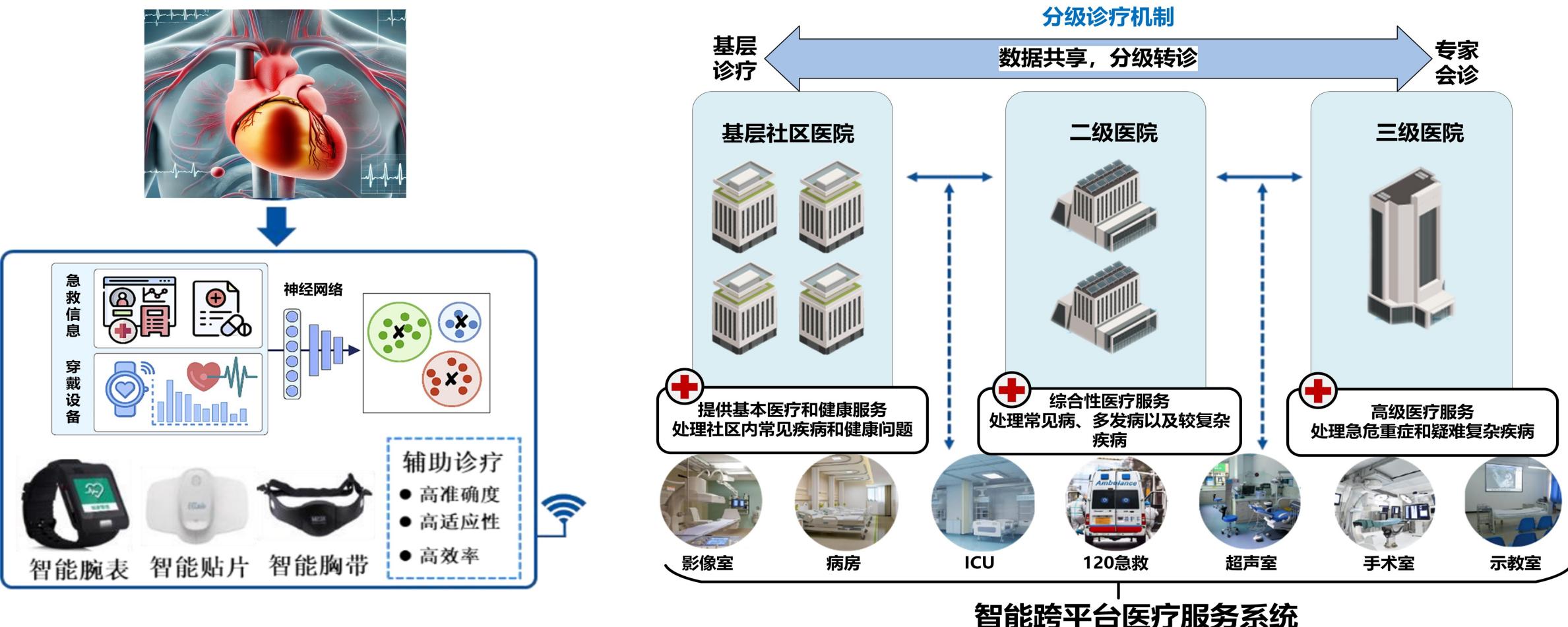


3.1 医生的智能助手：智慧医疗



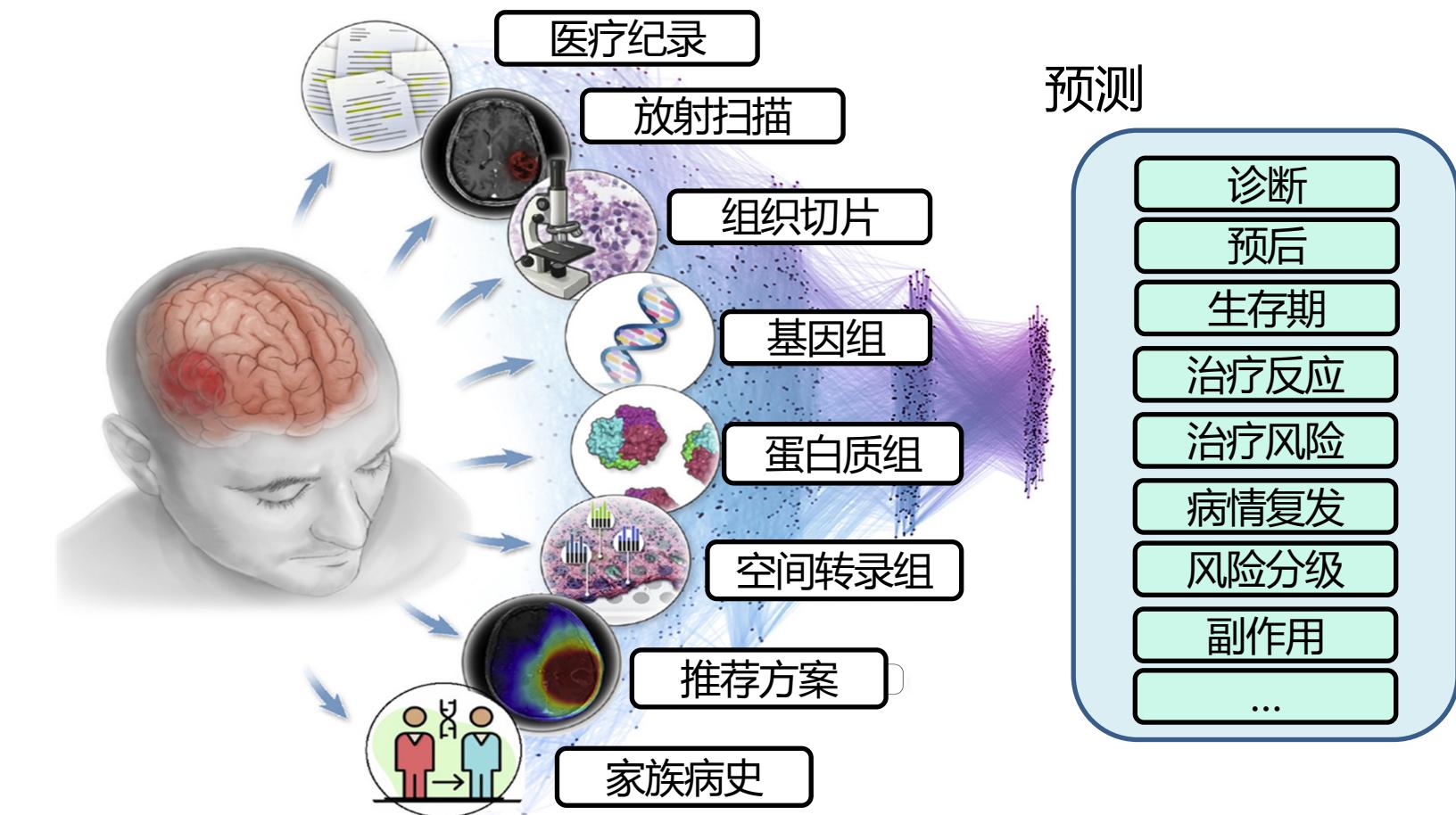
智慧医疗，又称为数字医疗或智能医疗，是指运用大数据、云计算、物联网、人工智能等现代信息技术的新型医疗服务模式。

智慧医疗**以患者数据为核心**，通过构建智能化医疗信息平台，实现**医疗机构、医务人员、医疗设备与患者之间**的互联互通，不断优化医疗资源配置，提高医疗服务效率和质量。



3.1 医生的智能助手：智慧医疗

■ **数据**在医疗行业中扮演着至关重要的角色，是智慧医疗生态中的核心要素，是推动医疗进步的关键驱动力。



医学数据包括但不限于患者的个人信息、病史记录、影像资料、检验结果、基因序列、生理参数等，以及医疗机构的运营数据、医疗资源分配数据等，它们都是构成医学大数据的重要组成部分。

✓ 海量性
✓ 多样性

✓ 高价值
✓ 隐私性

- **云计算技术**：构建医疗云平台，提供海量数据存储与强大算力，打破信息孤岛，实现跨机构协同诊疗与资源共享。
- **物联网技术**：通过各类传感器与智能设备实时采集患者生命体征数据，实现远程持续监控和个性化健康管理。
- **5G技术**：凭借超高带宽与超低延时特性，确保远程手术、高清视频会诊的实时稳定传输，大幅提升急救效率。
- **人工智能技术**：基于深度学习赋能疾病早期筛查、辅助诊断、医学影像分析及药物挖掘，显著提升精准医疗水平。
- **大模型技术**：深入理解与分析海量医学知识库和临床数据，赋能智能问诊、诊疗方案生成与前沿科研突破。



医疗大模型破晓

医疗惠及普惠众生

医疗手术机器人



新闻综合



上海早晨

■ 《金融资产管理公司不良资产业务管理办法》发布。

■ 第37届中国电影金鸡奖

15°C
18°C 08:01

3.2

AlphaFold：开启药物研发新纪元



蛋白质 被低估的营养素

[人体成分占比] → 蛋白质

NOTES

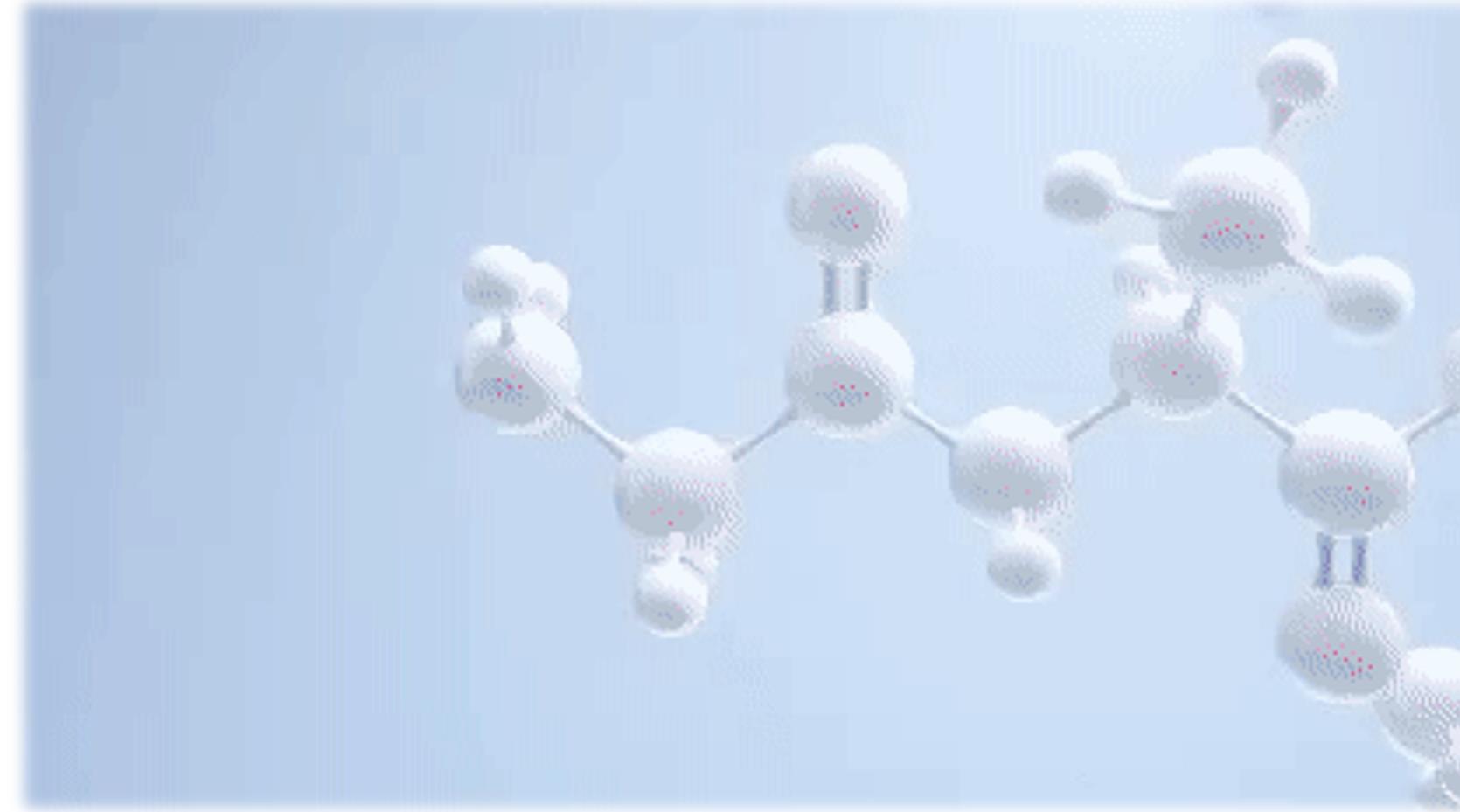
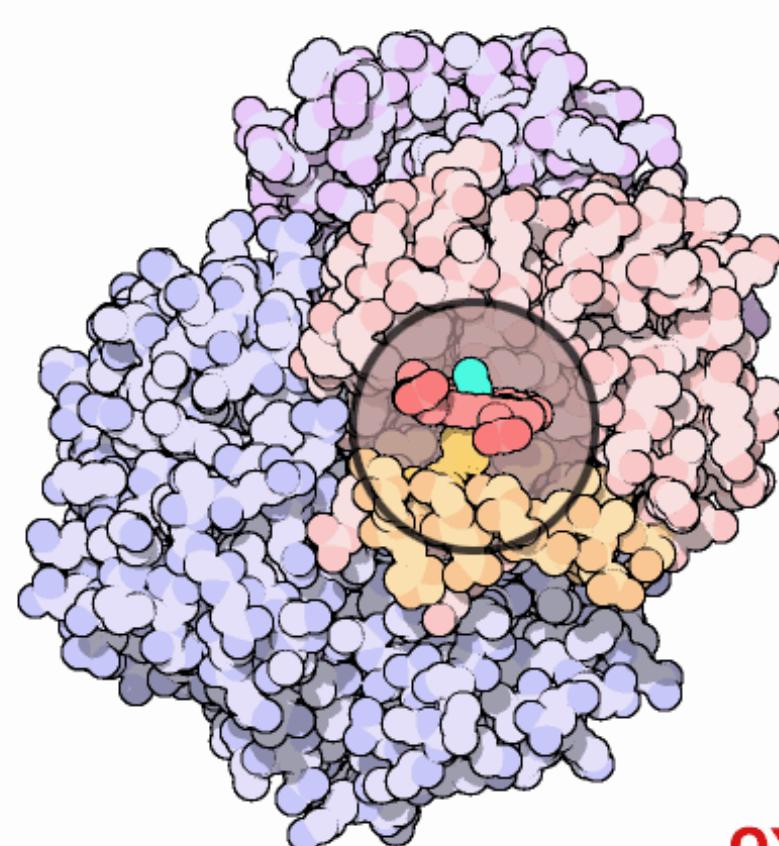
蛋白质的作用

- 构成各种组织、器官
- 免疫功能
- 催化作用（酶）
- 维持酸碱平衡
- 合成激素（生长激素、胰岛素）
- 维持体液渗透压
- 提供能量（三大供能物质之一）
- 运输功能（血红蛋白运输氧气，脂蛋白运输胆固醇）



3.2

AlphaFold：开启药物研发新纪元



蛋白质是由氨基酸通过肽键连接而成的一类高分子化合物，是生物体内最重要的生物大分子之一。它们的基本结构单位是氨基酸，而氨基酸的种类、数目和排列顺序决定了蛋白质的一级结构（氨基酸序列）。此外，蛋白质还具有复杂的二级、三级和四级结构，这些**蛋白质结构共同决定了它的功能和特性。**

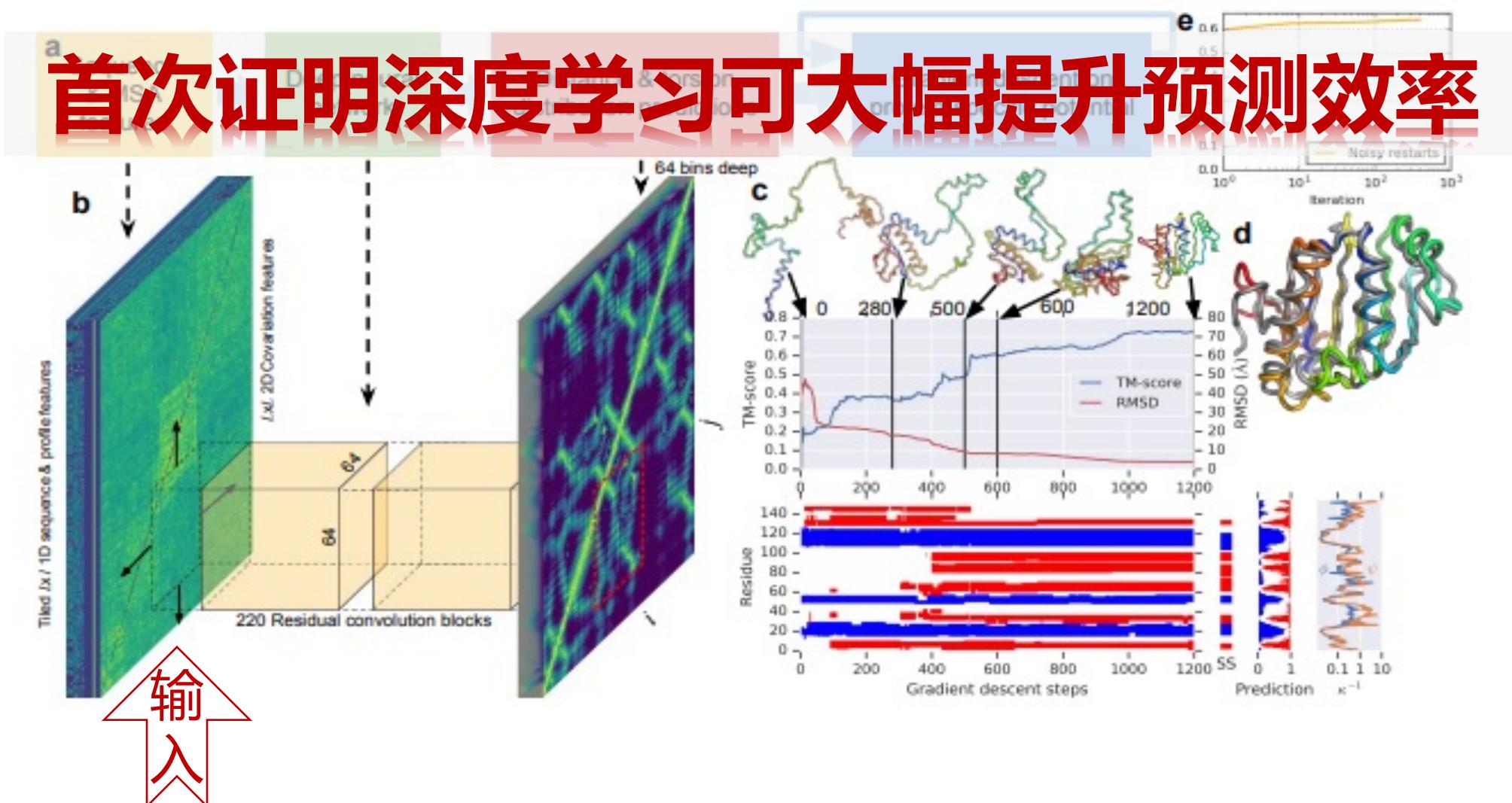
3.2

AlphaFold：开启药物研发新纪元



3.2 AlphaFold：开启药物研发新纪元

- 谷歌DeepMind自2018年起，利用生成式人工智能技术，相继推出了AlphaFold系列模型，引领蛋白质结构预测的革命。初代AlphaFold核心思想是通过**对蛋白质序列中两两残基距离的概率分布进行预测**，进而推断出蛋白质的三维结构。



氨基酸序列：目标蛋白质的氨基酸序列（一级结构）

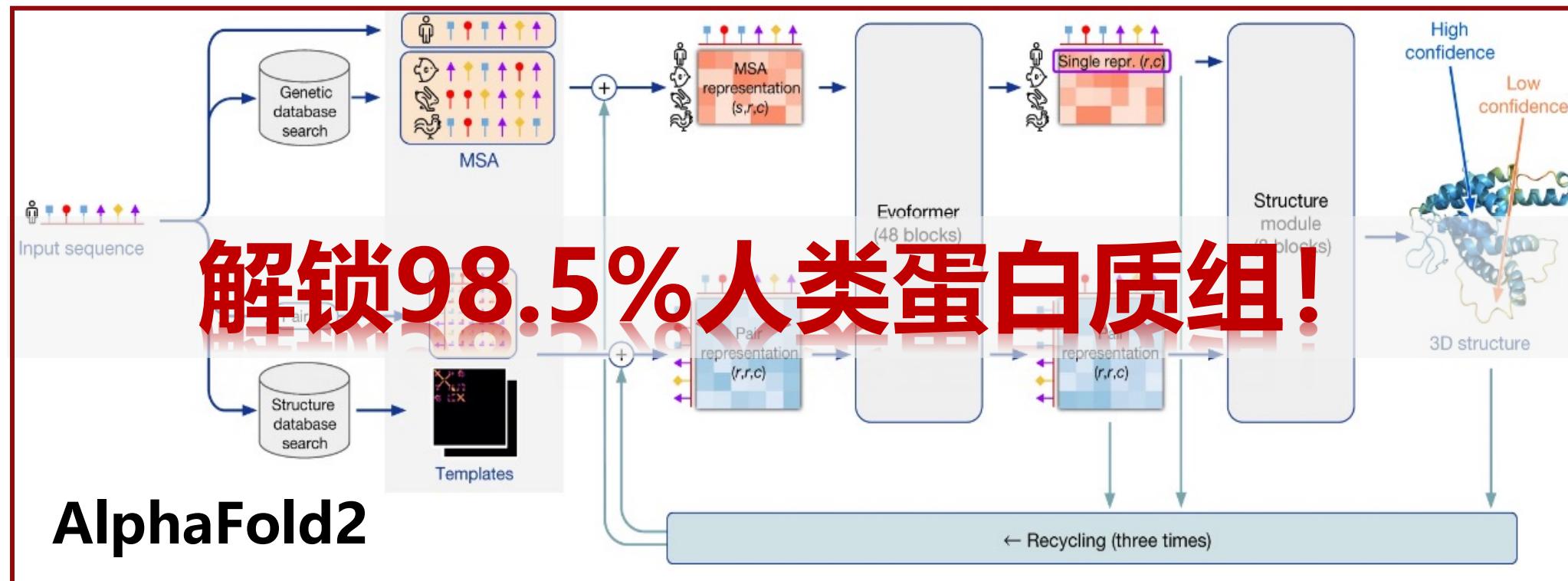
多序列比对 (Multiple Sequence Alignment) : 通过比对目标蛋白质与同源蛋白质的序列，捕捉进化信息。

多序列比对特征提取：利用多序列比对信息，从每个残基对的MSA中提取特征，并将其作为深度卷积神经网络的输入之一，帮助模型更好地理解蛋白质的进化信息。

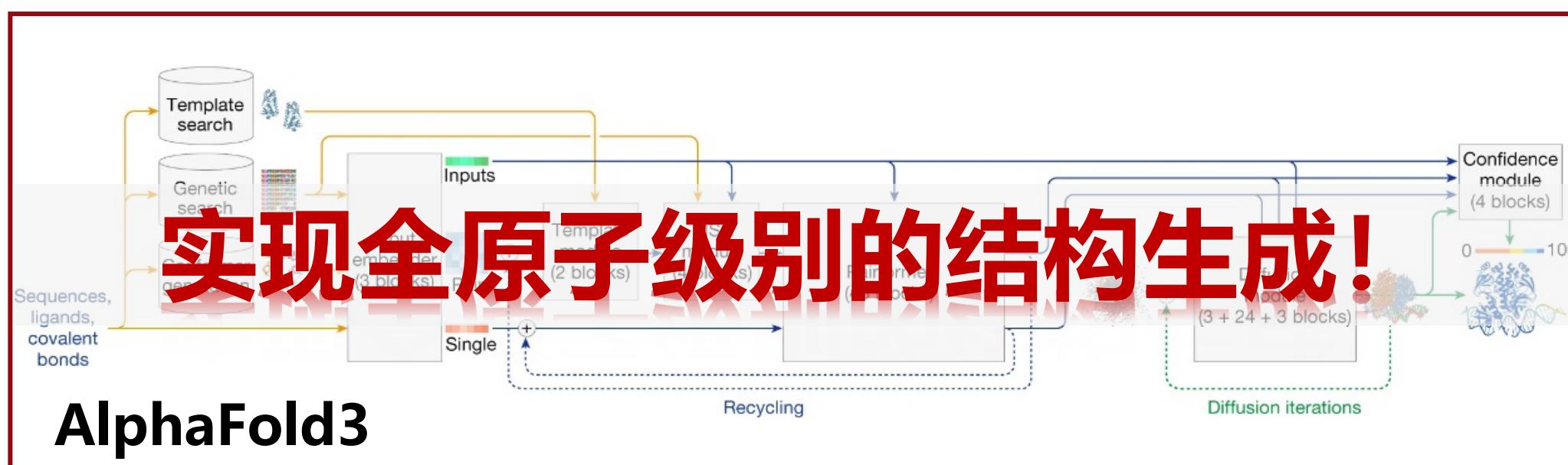
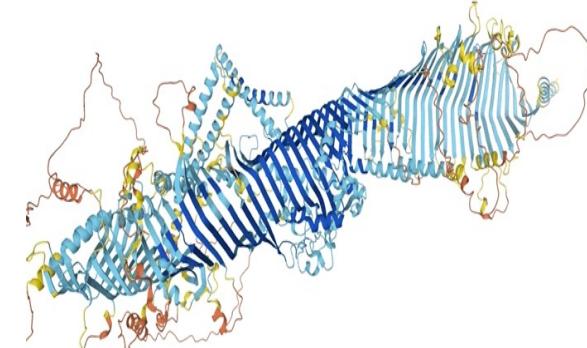
深度学习在蛋白质结构预测领域的初步尝试

3.2 AlphaFold：开启药物研发新纪元

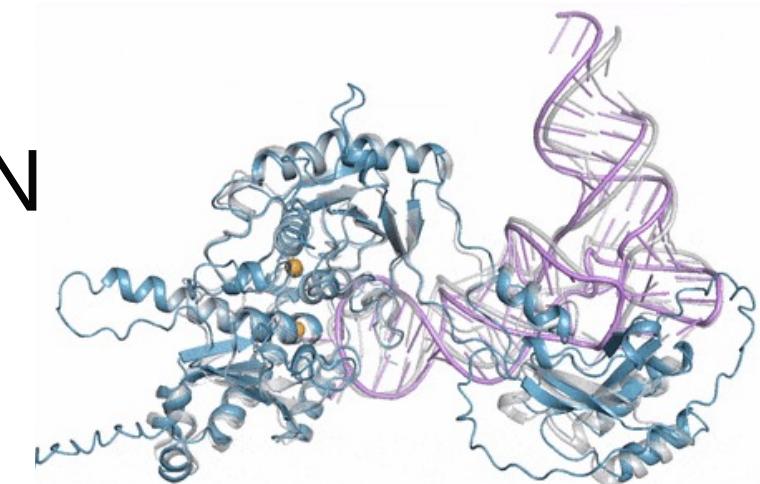
■ AlphaFold2在大部分蛋白质结构预测任务中达到了**接近实验**（如X射线晶体学、冷冻电镜）**的精度**，AlphaFold3实现了**多分子复合物和动态结构**的高精度预测。



- ✓ MSA数据利用更充分
- ✓ 统一的Transformer架构
- ✓ 更多样化的数据

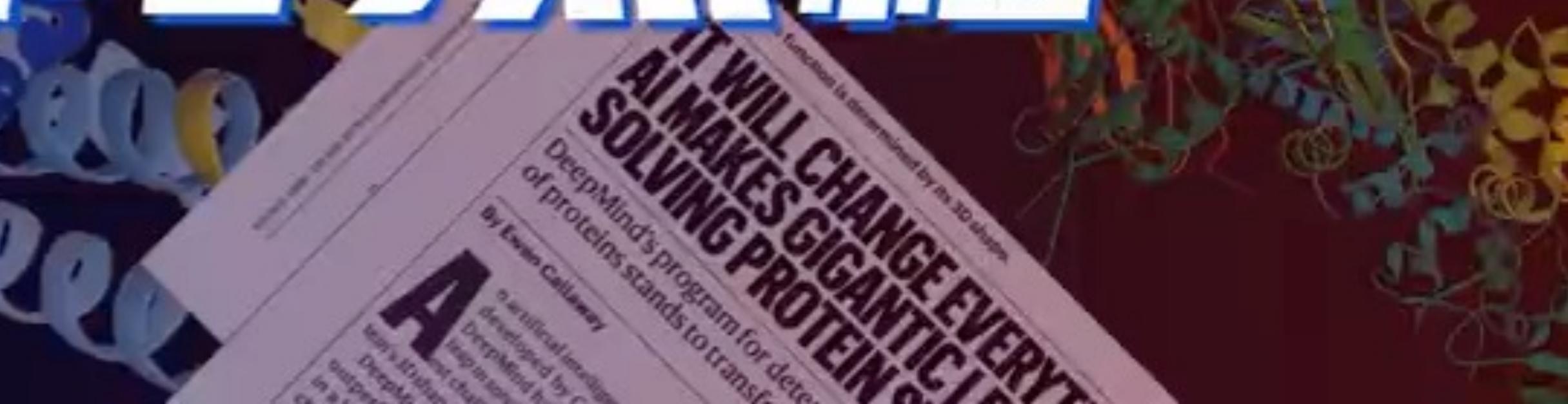


- + 减少MSA处理量
- + 引入扩散模块与GNN
- + 近亿级数据
- + 跨蒸馏技术



量子位 bili bili

AI 蛋白 折叠的黑相





04

AI助手——智慧生活

赛博李一哈 bilibili

朋友们!

朋友们

➤ 人工智能为时装业的零售商和顾客带来了诸多好处，接下来将从**硬件（即智能穿戴）与软件（即时尚助手）**两个方面介绍智能时尚的内涵与技术原理。



智能穿戴是指将智能技术集成到日常穿戴设备中，以增强功能、提升用户体验。常见的智能穿戴设备包括**智能手表、智能眼镜、智能手环、智能服装和智能鞋等**。



时尚助手是一种利用人工智能和大数据技术，帮助用户在时尚领域做出更好决策的工具或应用。它可以通过分析用户的喜好等为其**提供个性化的时尚建议**，包括服装搭配、购物推荐、潮流趋势分析等，还可以辅助设计师**进行服装设计**，激发创意。

你好帅杨 bili bili

智能穿戴设备

2024年 7月



4.1

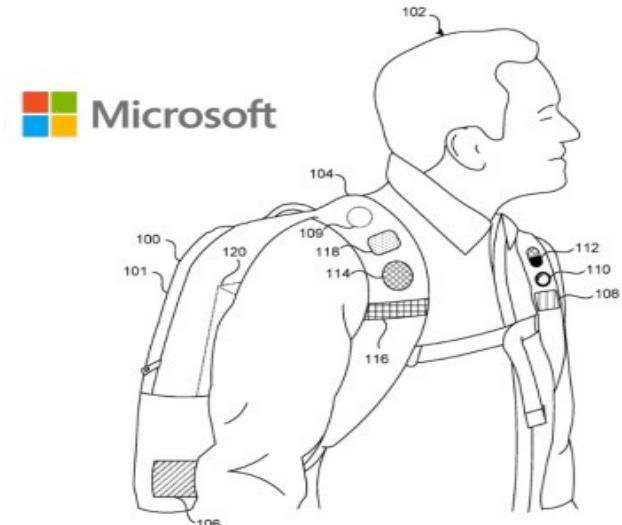
“AI” 上时尚



智能眼镜



Apple Watch



微软包



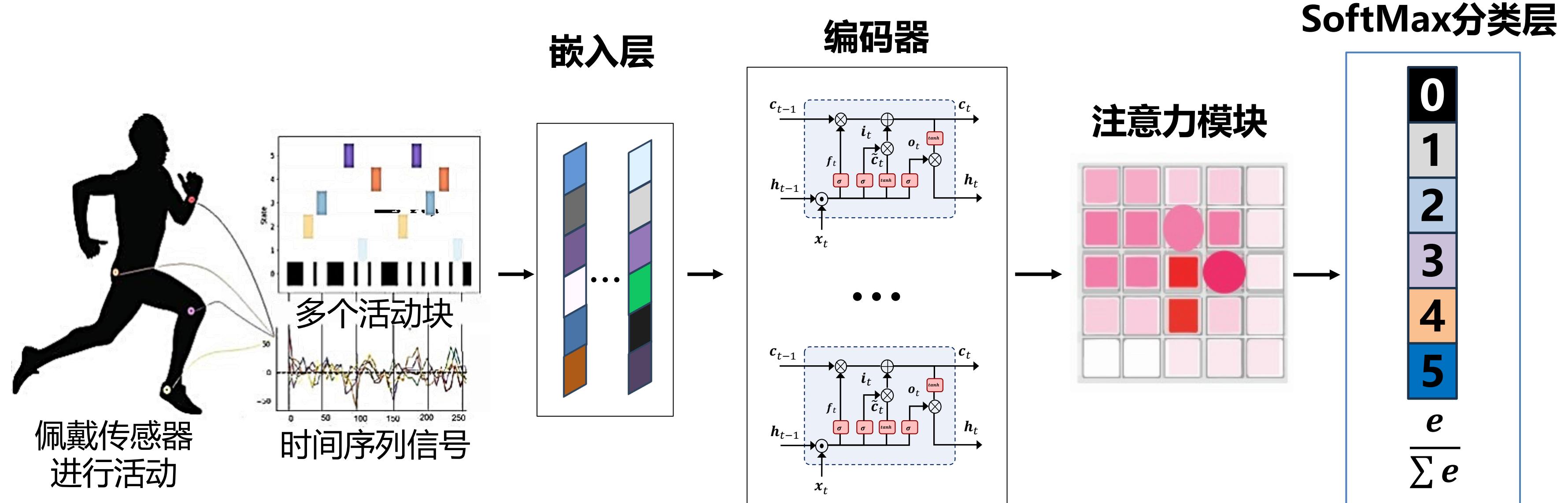
智能鞋

时尚单品



健康卫士

➤ 智能可穿戴开源框架——DeepConvLSTM人体活动识别框架



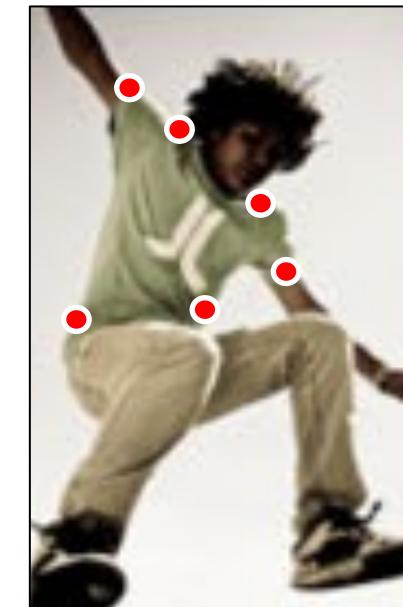
- **虚拟试衣技术**作为时尚与科技融合的产物，通过整合用户偏好、身体尺寸及风格取向，通过先进算法提供个性化时尚推荐，并使用户在不脱去身上衣物的情况下，完成变装查看效果的目的。
- **AI时尚设计**是指利用人工智能（AI）技术辅助或主导时尚设计的过程。通过机器学习、计算机视觉、自然语言处理等技术，AI可以帮助设计师进行创意生成、趋势预测、图案设计、面料选择等工作，从而提升设计效率、降低成本并推动创新。



➤ 关键点检测与要素解析



关键点检测



确定试衣对象的身体关键位置节点

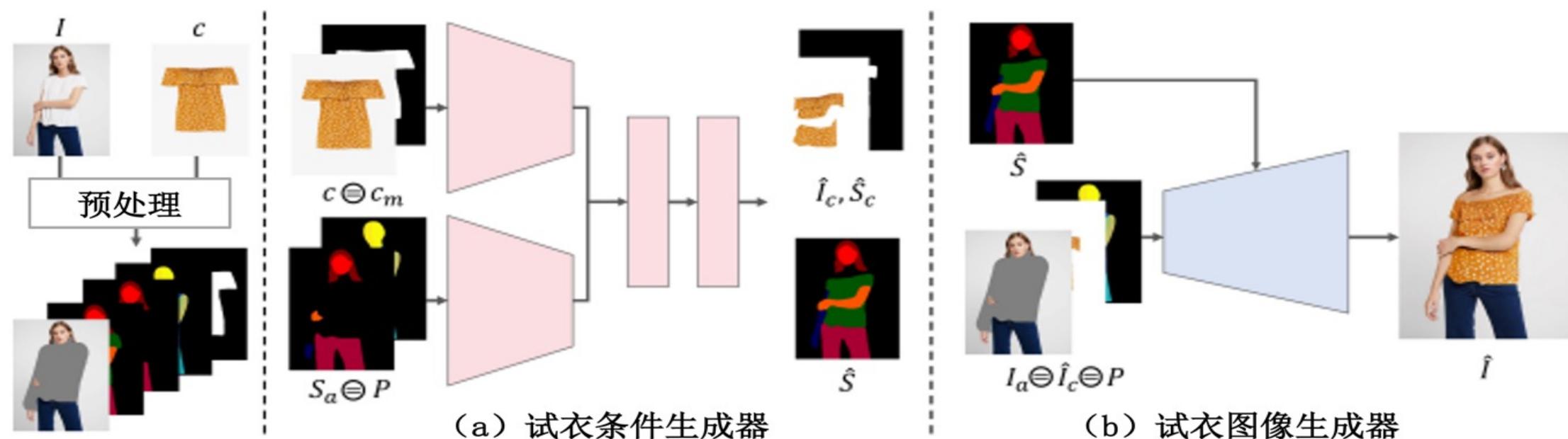
时尚要素解析



将试衣对象进行部位解析以便试衣

➤ AI虚拟换装合成

利用生成对抗网络、扩散模型等**生成类模型**将目标衣物与试衣对象进行合成。

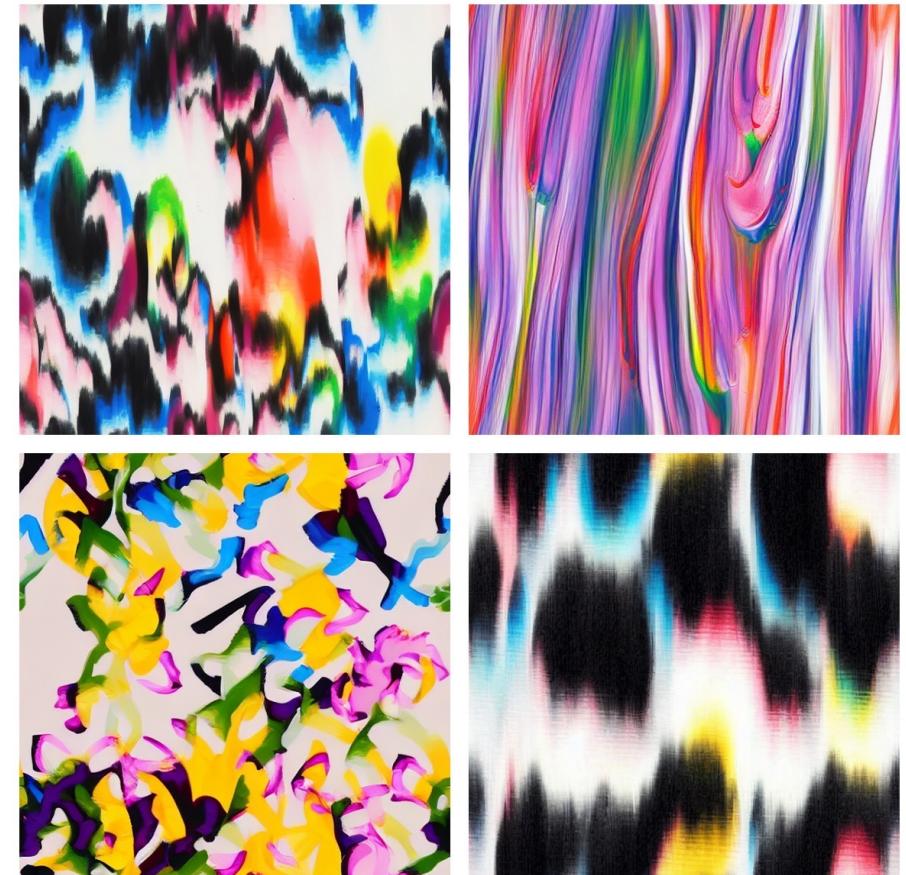


4.1

“AI”上时尚

去找找

下面哪款衣服是用AI设计出来的？



➤ 人工智能生成内容（AIGC）技术带来的突破

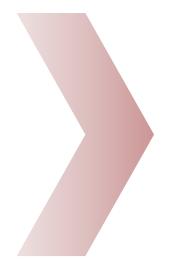
- AIGC (Artificial Intelligence Generated Content)，即人工智能通过学习大量的数据，来实现**自动生成各种内容**，如**文本、图像、音频、视频等**，是人工智能1.0时代进入2.0时代的重要标志。GAN、CLIP、Transformer、Diffusion、预训练模型、多模态算法等技术的累积融合，催生了AIGC的爆发。
- 早期AIGC在文本生成领域开启内容创作落地，后期逐渐向图像、音视频等多模态领域扩展，加之**大模型技术**的推动下，AIGC有了更多的可能性。
- 中国联通基于元景多模态大模型的图像生成能力打造AIGC产品——**“衣影”辅助设计大模型**，以人机对话互动方式提供服装领域多模态知识整合、分析、输出等服务，降低服装设计周期，提高新品开发效率。

4.2

“AI” 护盘中餐



- 人工智能与自动化技术已深刻变革了传统的农业模式，诞生了智慧农业这一新兴智慧经济体，减轻了劳动力需求，同时也减少了对环境的影响，提高了作物品质。



➤ 精准农业

- 精准农业旨在借助信息技术实现在**最佳时间和地点**种植**最适宜作物**，优化农业生产，提升生产效率并降低成本。
- **农作物分类**利用神经网络等技术，实现了作物的精准识别与生长参数的科学管理。**自主作业机器人**通过配备传感器来实时捕捉农田环境信息，并以此控制机器人做出灵活的决策。



农业除草机器人



农业车辆机器人



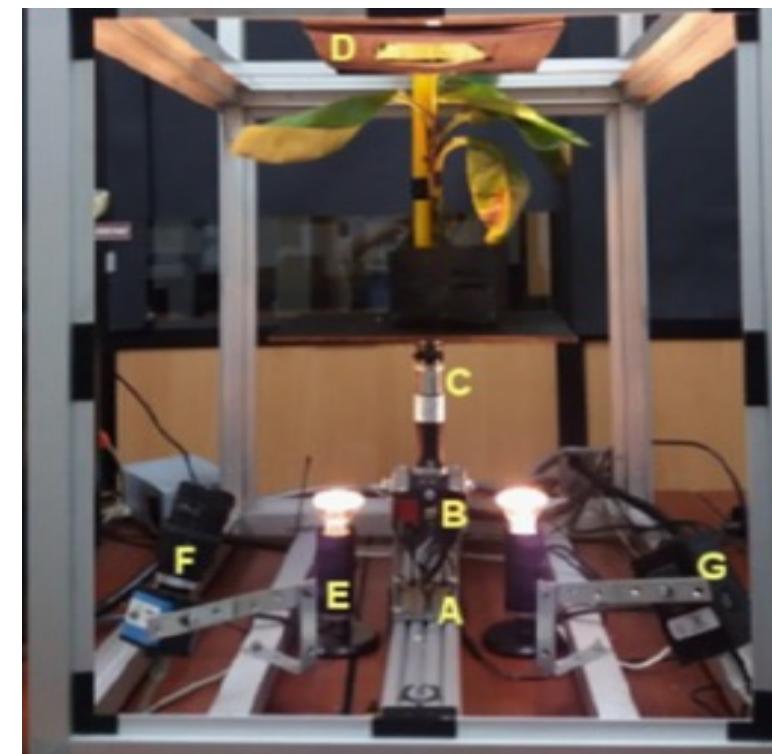
农业喷洒机器人

➤ 农作物病害检测

构建有效的**植物疾病预防与控制体系**对于保障作物健康生长和粮食安全至关重要。传统的依靠专家人工目检来识别植物疾病的方法因成本高且专业性强而显得不足，亟需现代化解决方案以实现及时准确的疾病识别。



高光谱和叶绿素荧光成像



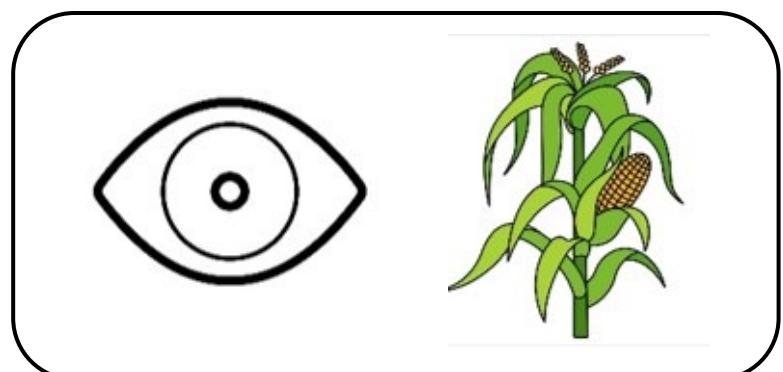
高光谱成像用于香蕉植株病害扫描



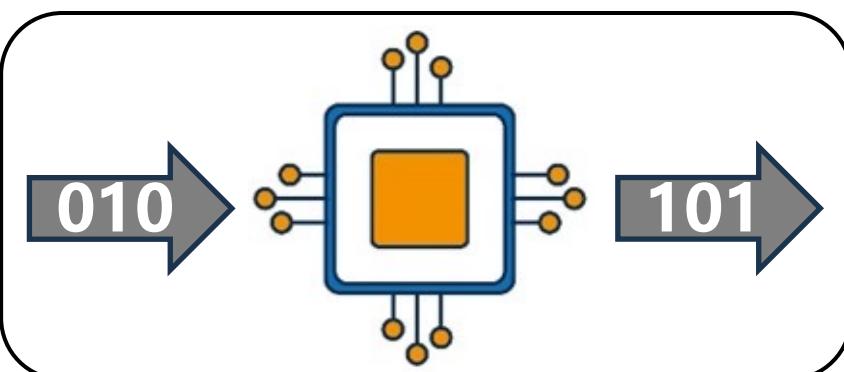
红外和热成像用于柑橘绿化检测

➤ 农作物表型分析

口 生物体的可观察特性称为**表型**，包括行为、生化属性、颜色、形状和大小等；在植物科学中，通过自动化计算机视觉系统进行的植物表型分析能够无须人工干预地**检测和测量作物特征**，以确定作物的生长潜力和抗病性。



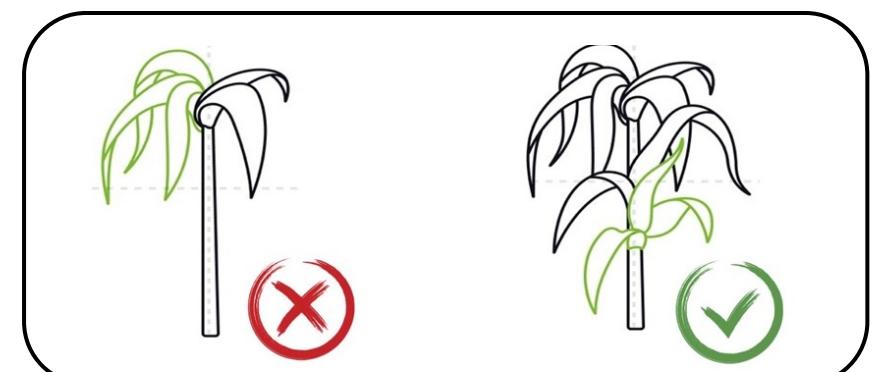
作物表型



智能计算



特征分析



育种选择

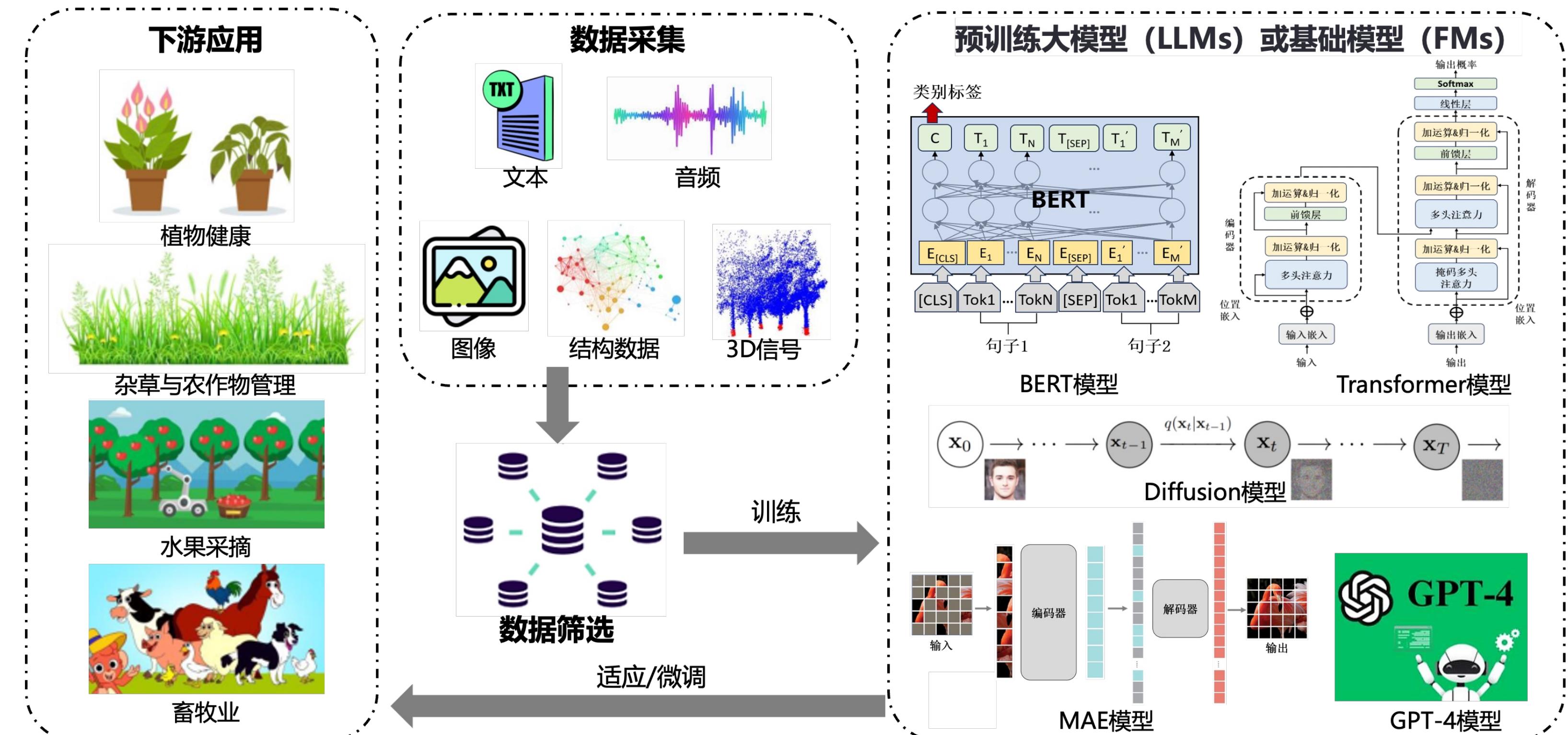
筛选形态特征、生理特征和病原体

计算机视觉/机器学习、地理信息系统、作物模型与算法

大数据管理策略、训练策略、育种平台

选择形态特征、生理特征和病原体

➤ 基于多模态大模型的农业应用概览





➤ 基于多模态大模型的农业应用概览

神农大模型2.0版本

- 多模态交互与智能化推理
 - 知识图谱与向量数据库
 - 并行加速推理算法
 - 育种、种植、养殖和气象
- 四个专业门类大模型



第一次体验小度控制智能家居

视频来源：好看视频-大涌路7号

4.3

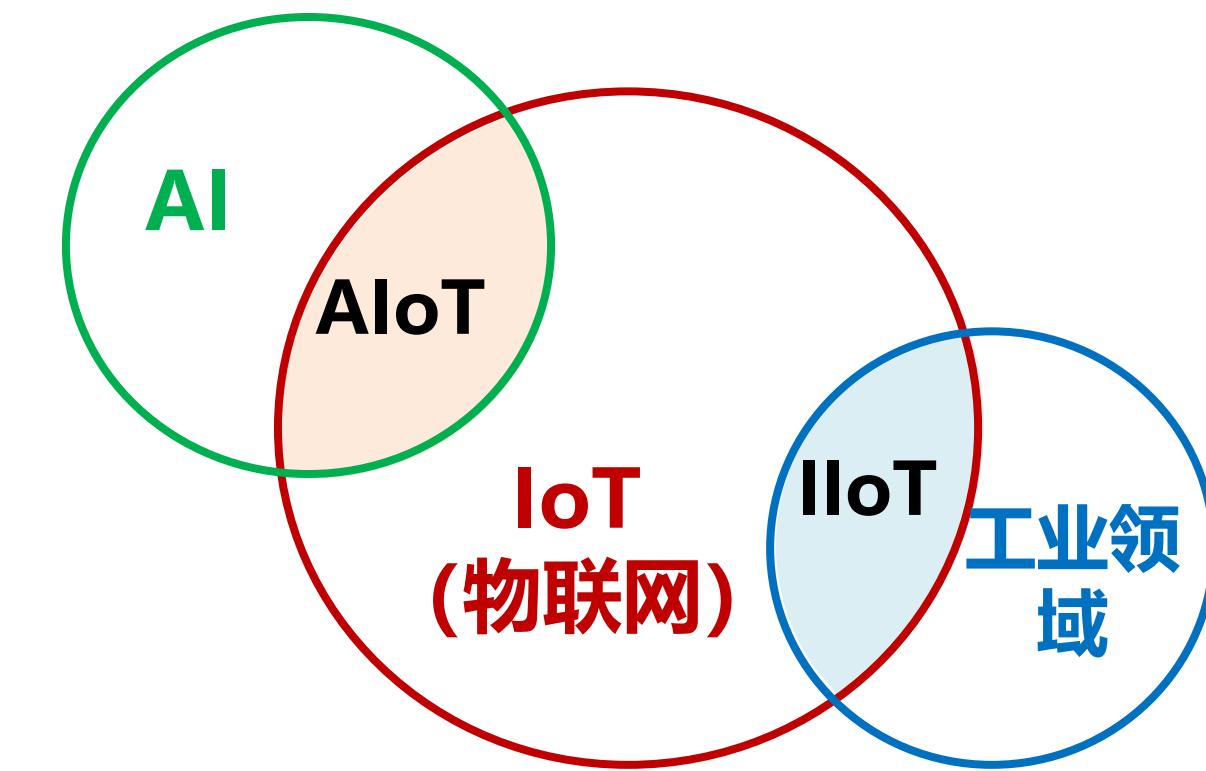
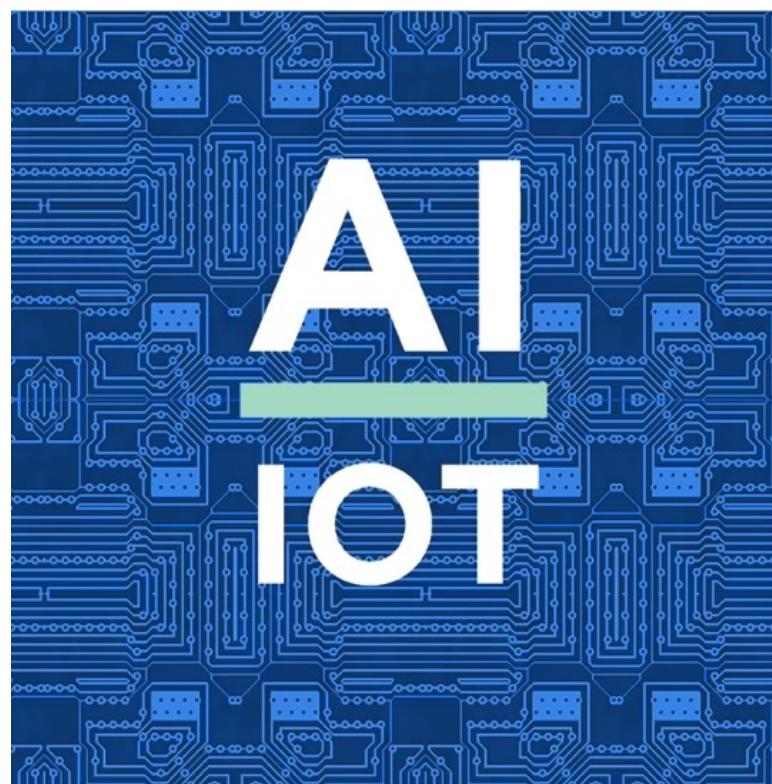
家中管家

➤ **智能家居**是指通过物联网、人工智能、云计算等技术，将家居设备、家电和系统连接起来，实现自动化控制、远程管理和智能化操作的家庭环境。



➤ 人工智能物联网AIoT

- AIoT（人工智能物联网）**融合了AI与物联网的优势**，通过智能计算能力和多元人机交互，实现了设备的精准控制与信息深度理解，推动了从“物联”到“智联”的升级。
- AIoT的研究体系涵盖感知层、操控层、应用场景及安全隐私保护，推动了智能家居等领域的智能化进程。



AIoT是AI与IoT的融合可使终端产品变得更加智能化

IoT、IIoT、AIoT的关系

AI与IoT安全及隐私保护融合

智能制造物联网

智慧农业

智能家居

智能交通控制系统

智慧医疗

智慧社区

AI与IoT应用场景融合

智能制造物联网

智慧负载均衡

智慧能耗管理

资源智慧调度

复杂事件智慧控制

云端边融合

AI与IoT操控层融合

数据智慧采集存储

数据智慧分析处理

感知设备智慧协同

AI与IoT感知层融合

数据与内容安全

应用服务安全

可信管理

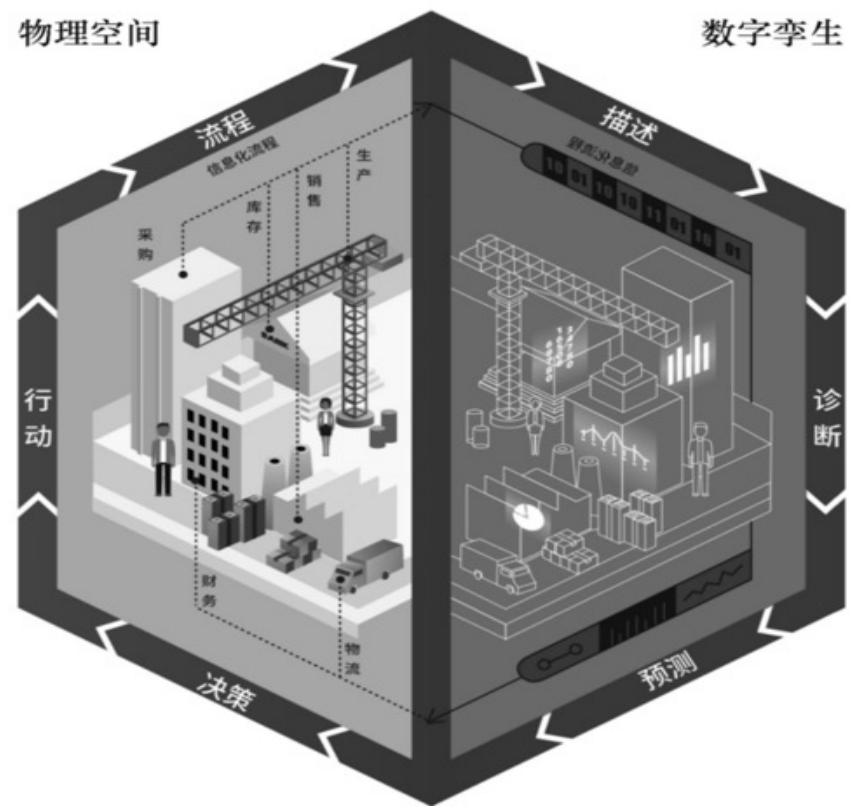
传输媒介安全

移动终端安全

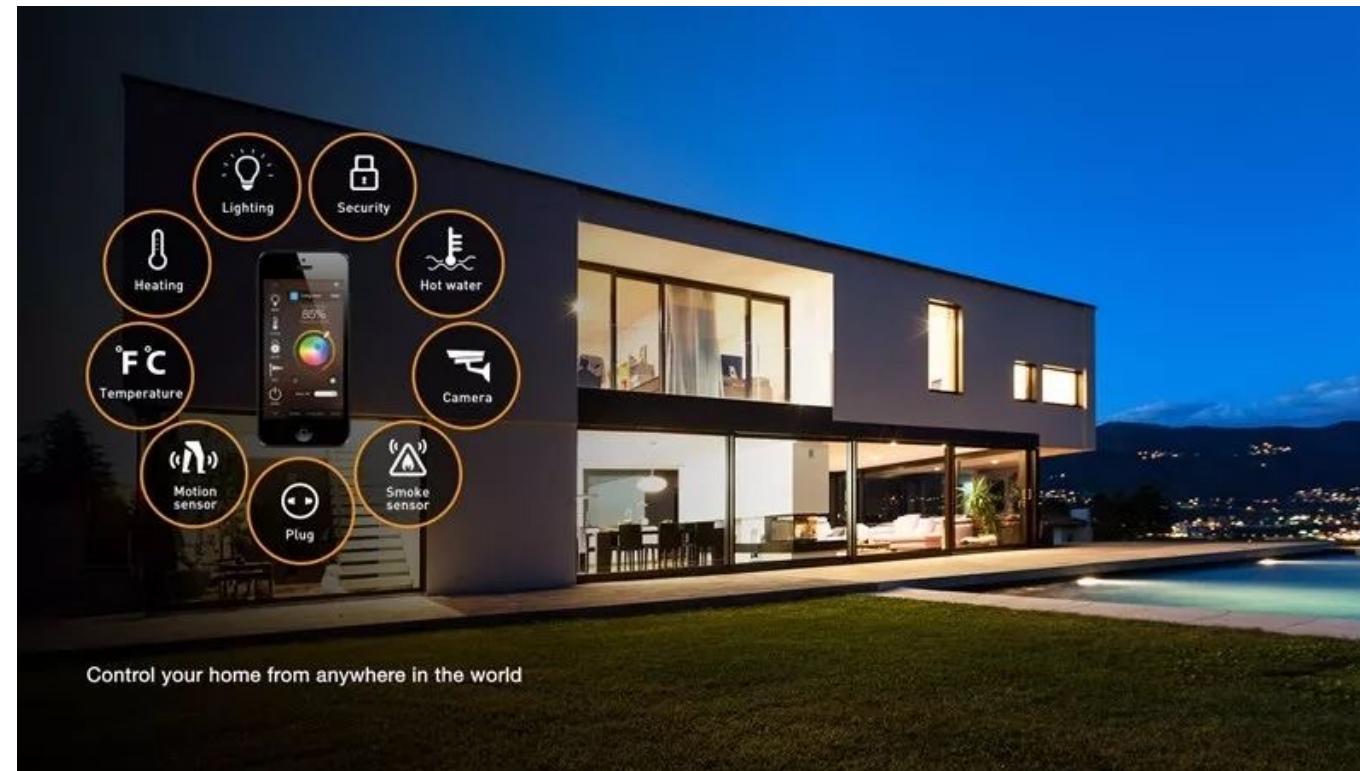
位置隐私保护

➤ 数字孪生技术

- 数字孪生技术从**远程控制、智能管理、安全防护和生活品质提升**四个方面为智能家居系统提供技术支持，使家庭管理更加便捷高效，同时增强了家庭安全性和居住舒适度。
- 通过这一技术，用户不仅可以跨越空间限制控制家居设备，还能实现设备间的智能联动，享受个性化规则带来的智能舒适生活，并通过安防监控获得更高的安全感。



数字孪生即现实世界的“数字副本”



数字孪生技术将成为未来家庭中央管理系统的重要组成部分



基于数字孪生技术的家居设计与虚拟看房服务

幻彩科技智慧孪生 bilibili

究竟什么是 数字孪生？

两分钟带你走近数字孪生

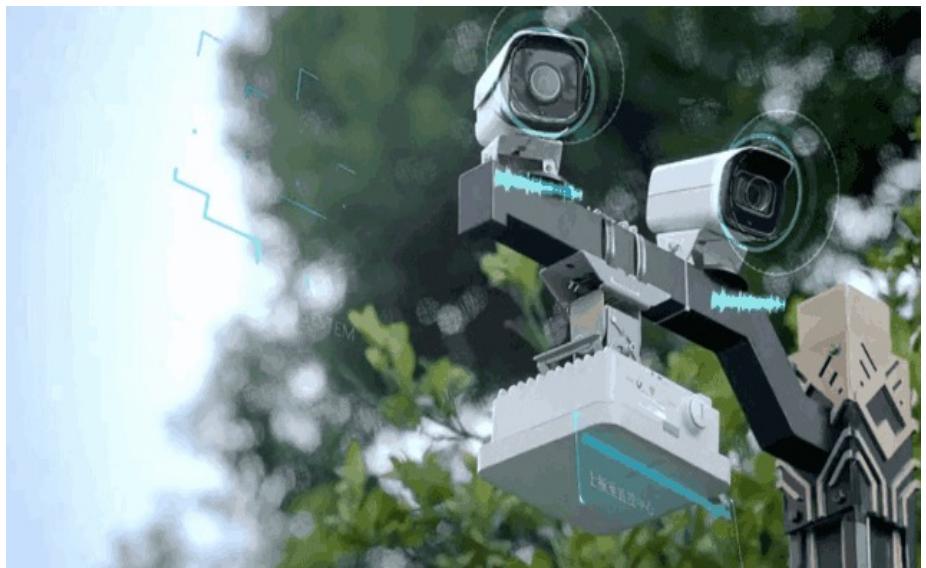
4.3

家中管家

➤ 数字孪生技术



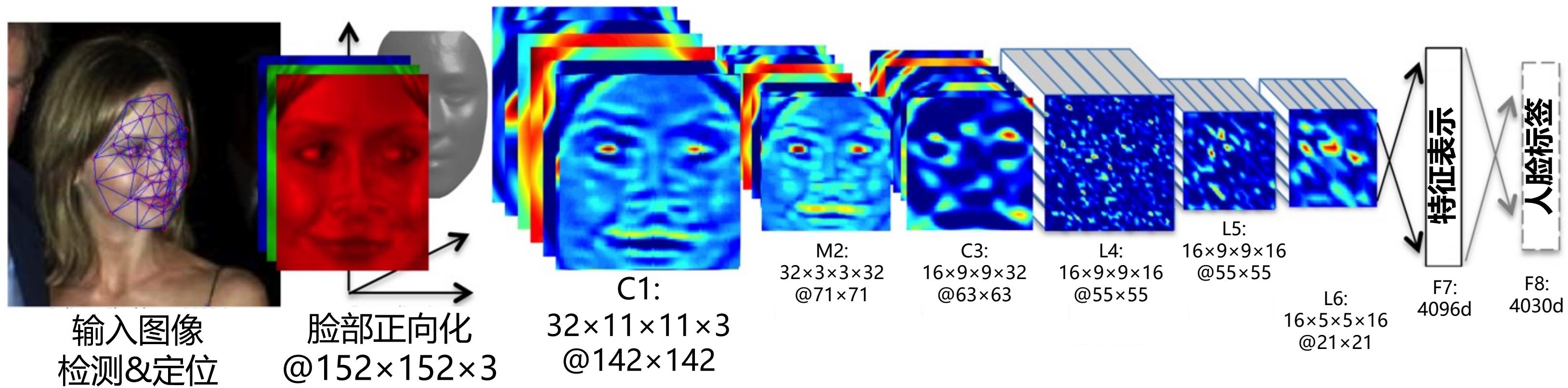
➤ 智慧社区以居民为中心，融合信息技术、物联网、大数据和云计算，全面提升社区管理和服务的智能化水平，增强居民的安全感与幸福感。





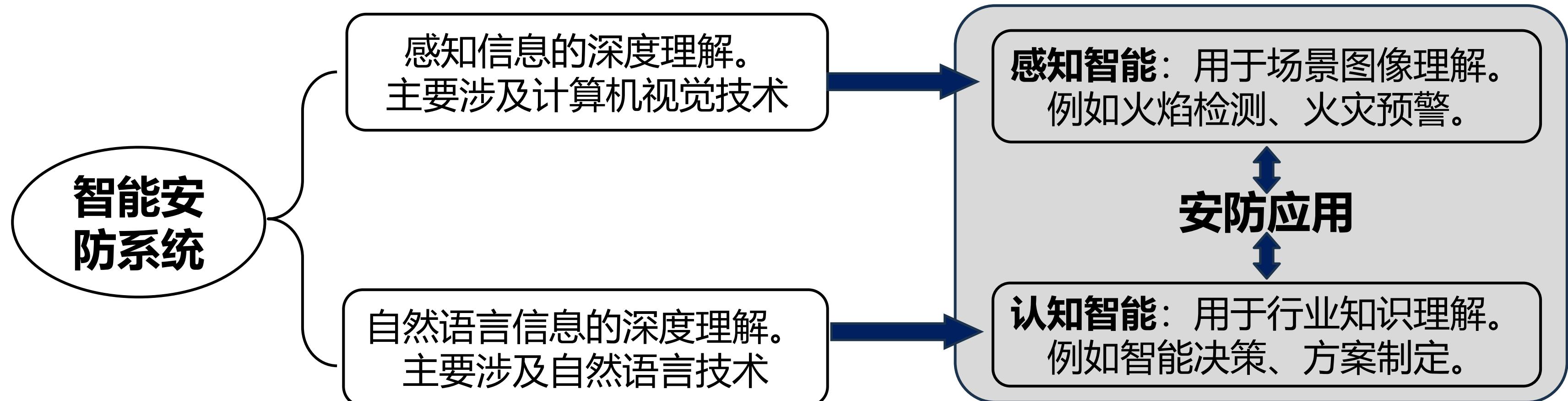
➤ 智能安防中的人脸识别技术

□ 深度学习算法在安防监控行业广泛应用于目标分析，如Facebook AI Research开发的DeepFace系统，通过3D对齐方法提高人脸识别准确性，并集成多种模型实现高精度识别与属性分析。DeepFace系统不仅在安全监控和身份验证中发挥作用，还适用于个性化娱乐互动、社交网络智能化及人机交互等多种场景。



➤ 智能安防中的异常检测技术

- 异常检测技术是保障家庭、社区及公共场所安全的核心手段之一，旨在识别和监测数据中存在的异常行为或事件，如火灾、斗殴、偷盗等。
- 异常行为或事件通常是指与普遍规律或典型模式不符合的数据点或样本。



体验

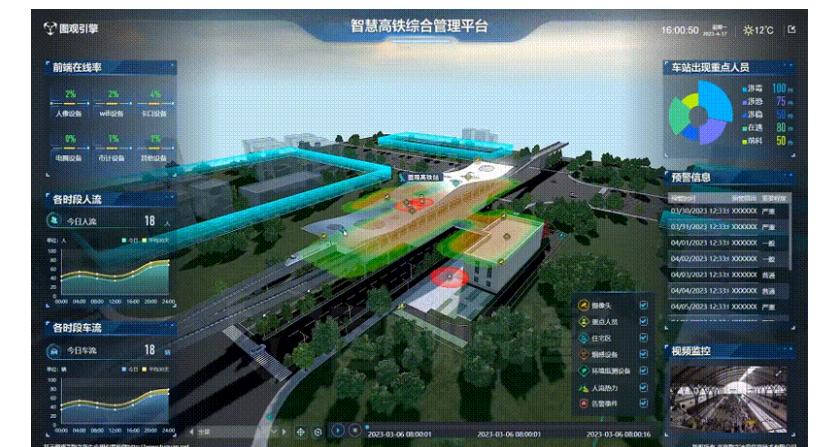
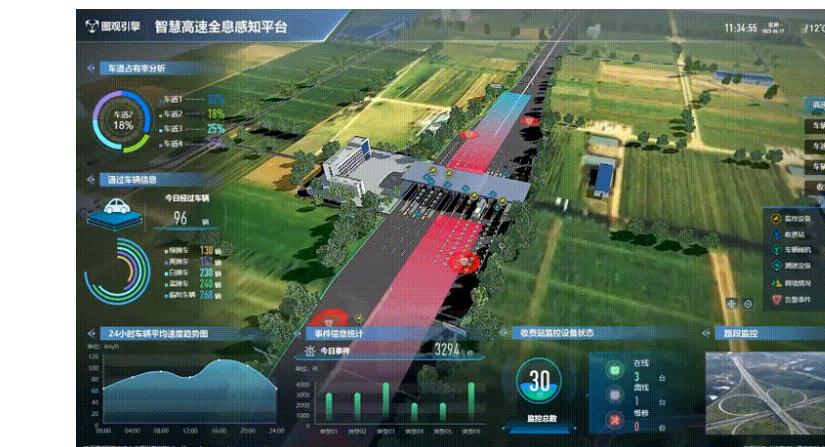
无人驾驶

开放道路

萝卜快跑



■ 智慧交通是以互联网、物联网等网络组合为基础，以智慧路网、智慧装备、智慧出行、智慧管理为重要内容的交通发展新模式，具有信息联通、实时监控、管理协同、人物合一的基本特征。



高速、铁路、公交、机场等综合管理平台

■ 智慧交通关键技术：

- 物联网技术可以实现对交通设备的智能化管理和控制；
- 云计算技术可以处理和分析海量的交通数据，提供强大的计算能力和存储能力；
- 大数据处理技术可以实时收集、处理和分析交通数据，提高交通管理的效率和准确性；
- 人工智能技术则可以实现交通系统的自我学习、自我优化和智能决策。

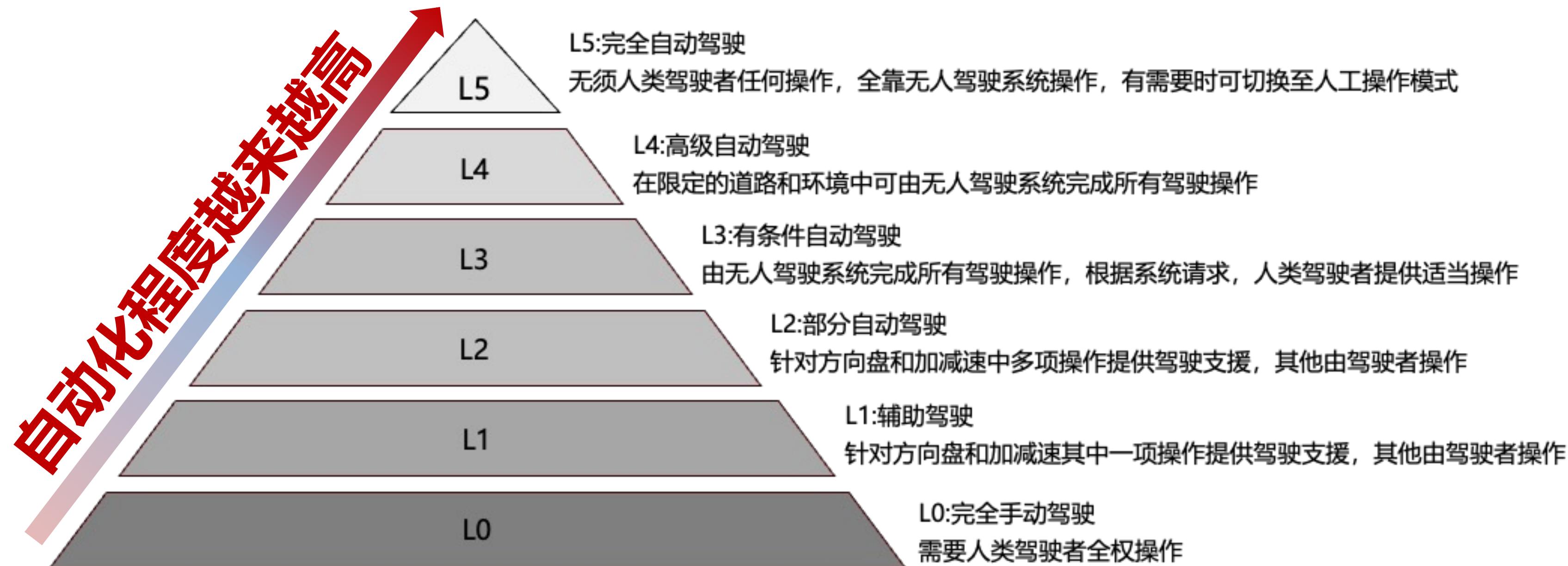
■ 智慧交通典型应用：

- 智能交通管理系统：交通信号控制、路况实时监测、流量预测、综合智慧管理等；
- 自动驾驶：实现车辆的自主驾驶，提高安全性；
- 智能公交系统：实现公交车精准调度和线路优化；
- 智能物流系统：实现智能车辆调度等；
- 智能公路系统：实现车辆安全预警、智能路网监控等；

4.4

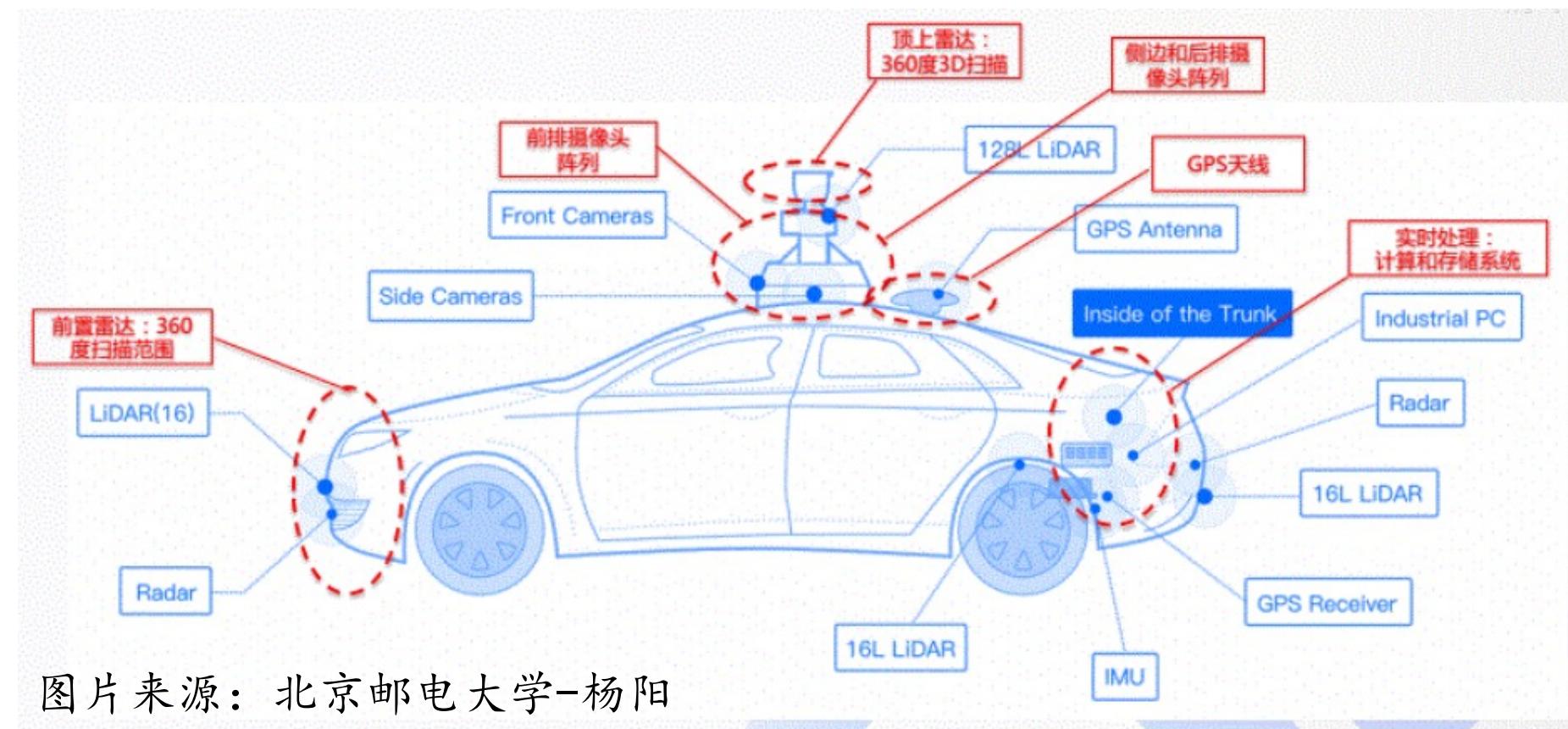
萝卜快跑

- 自动驾驶是指车辆在**没有人类干预**的情况下，能够**自主**感知环境、规划路径并执行驾驶操作的技术。它通过传感器、计算机视觉、人工智能和控制系统等技术的结合，实现车辆的自主导航和驾驶。

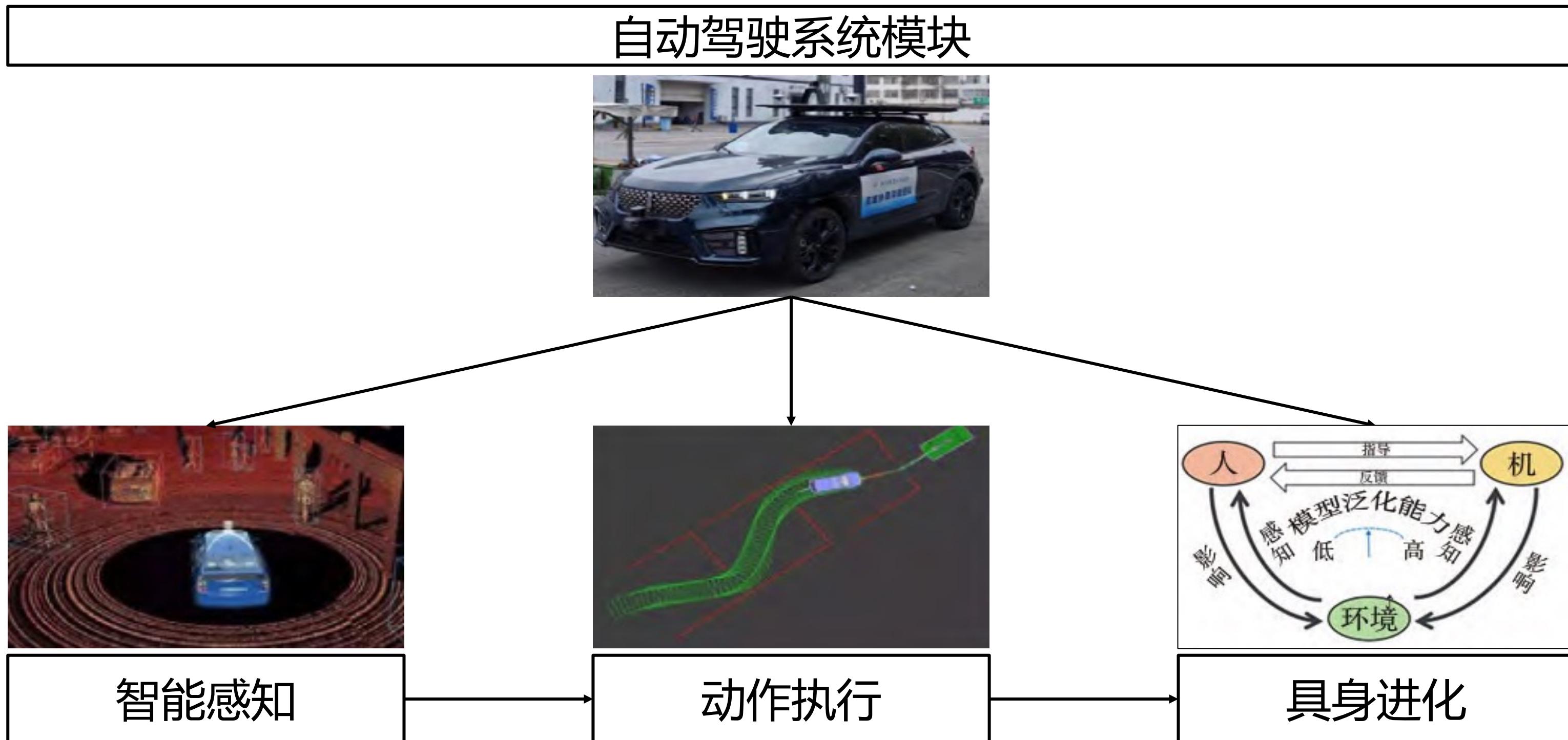


➤ 自动驾驶系统的整体框架

- 无人驾驶车由车、线控系统、传感器、计算单元等组成。
- 车作为载体，是无人驾驶实现的根本。线控系统通过电信号来控制汽车。传感器是无人驾驶车的眼睛，负责感受外部环境，如激光雷达、摄像头、毫米波雷达、超声波雷达、GPS、IMU等。计算单元则是无人驾驶车辆的大脑，传感器获取的信息经过计算单元的计算之后，输出一条可以供汽车安全行驶的轨迹，控制汽车行驶。



➤ 自动驾驶系统的整体框架



汉牛在中国率先开启了L2+辅助驾驶

所长Wy bilibili





第6讲

人工智能技术应用