



视觉中国供图

SAM 是一类处理图像分割任务的通用模型。与以往只能处理某种特定类型图片的图像分割模型不同，SAM 可以处理所有类型的图像。相比于以往的图像分割模型，SAM 可以识别各种输入提示，确定图像中需要分割的内容，还可以灵活集成到虚拟现实/增强现实等其他系统中，且目前对于一些它未见过或相对模糊的场景，也能实现较好的图像分割效果。

又一通用模型发布 SAM 是如何做到“分割一切”的

◎ 实习记者 裴宸纬

最近一段时间，人工智能通用模型领域频现“爆款”。4月，Meta 公司发布了一款名为“Segment Anything Model”(SAM)的通用模型，号称可以“零样本分割一切”。也就是说，SAM 能从照片或视频图像中对任意对象实现一键分割，并且能够零样本迁移到其他任务中。

在相关展示页面中，科技日报记者看到，在一张包含水果、案板、刀具、绿植、储物架等众多物体、背景杂乱的厨房照片中，该模型可迅速识别出不同的物体，以粗线条勾勒出物体轮廓，并用不同颜色对不同物体进行区分。“这就是 SAM 最重要的功能——图像分割。”中国科学院自动化研究所多模态人工智能系统全国重点实验室研究员、中国科学院大学人工智能学院教授杨戈向记者表示。

那么，SAM 的技术原理是什么？相比于此前的图像分割模型，该模型有何不同？未来又有可能在哪些方面应用？

图像分割通用模型泛用性强

杨戈向记者解释道，像 SAM 这样可以处理多种不同类型任务的人工智能模型，叫作通用模型。与之相对，那些专门处理一种类型任务的人工智能模型，叫作专有模型。

打个形象的比喻，通用模型就好比是一个“多面手”。它具有处理一般事务的能力，但是在精度等性能上往往会逊色于只处理一种类型任务的专有模型。

既然通用模型可能会在精度上低于专有模型，为什么还要费尽心力地开发通用模型？对此，杨戈表示，通用模型与专有模型定位不同。通用模型带来的，是解决分割问题的新范式，特别是帮助科研人员提升在解决专有任务时的效率，“以前，面对不同的任务需求，科研人员往往需要开发不同的专有模型来应对。这样开发出的模型精度确实会更高，但是往往也会付出较大的研发成本，而且研发的模型通用性不强。”杨戈说。

通用模型能够将所有任务都处理得“八九不离十”，因此科研人员往往只需在通用模型的基础上进行优化，使之更加符合任务需求即可，而不需要费尽心力地从零开始搭建专有模型。因此，通用模型的初始开发成本可能会高，但随着使用通用模型的次数越来越多，其应用成本也会越来越低。

SAM 就是一类处理图像分割任务的通用模型。与以往只能处理某种特定类型图片的图像分割模型不同，SAM 可以处理所有类型的图像。“在 SAM 出现前，基本上所有的图像分割模型都是专有模型。”杨戈补充

道，“打个比方，在医学领域，有专门分割核磁共振图像的人工智能模型，也有专门分割 CT 影像的人工智能模型。但这些模型往往只在分割专有领域内的图像时，才具有良好性能，而在分割其他领域的图像时往往性能不佳。”

有业内专家表示，相比于以往的图像分割模型，SAM 可以识别各种输入提示，确定图像中需要分割的内容，还可以灵活集成到虚拟现实/增强现实等其他系统中，且目前对于一些它未见过或相对模糊的场景，也能实现较好的图像分割效果；同时，SAM 建立了一套图像分割的通用模型，降低了对于特定场景建模知识、训练计算、数据标记的需求，有望在统一框架下完成图像分割任务。目前 Meta 公司已经开放共享 SAM 的代码和训练数据集。

利用海量数据实现准确分割

那么，SAM 是通过什么技术手段，实现对物体的识别与分割？尤其是在面对复杂环境、甚至没遇到过的物体和图像时，SAM 又是怎么做到的准确识别与分割的？

“根据 Meta 发布的相关论文，SAM 的模型结构其实并不是特别复杂。”杨戈告诉记者，“它用到了一个叫作‘编码解码器’的构架。”

记者了解到，SAM 先通过图像编码器为图像生成编码，同时用一个轻量级编码器将用户的文字提示转换为提示编码。然后，SAM 将图像编码和提示编码信息源组合在一起，输送到一个轻量级解码器中，用于预测分割掩码。这样一来，一旦使用者给出提示，则每个提示只需要几毫秒就能在浏览器中得到结果响应。

杨戈用了一个生动的例子解释 SAM 的运行原理。“比如说，给你一张带有猫和狗的图片，将照片中的猫标注出来，这就是提示；但是对于机器来说，它并不能直接‘明白’这种文字性提示，因此就需要将文字性提示转换为机器能够理解的提示编码。”同理，对于照片中的猫和狗，机器实际上并不能直接“明白”什么是猫、什么是狗，而是将照片中的猫和狗与图片编码对应起来。SAM 通过训练学习提示编码与图片编码的不同结合，理解人类在文字提示中表述的希望如何分割这张图片。一旦“将照片中的猫标注出来”这句提示被输入时，SAM 就能快速运行，得到人类想要的结果。

既然 SAM 并没有真正理解什么是猫、什么是狗，它又是如何准确地执行人类赋予的任务的呢？

“虽然 SAM 并没有完全理解人类的语言和视觉的能力，但是通过对海量数据的学习，SAM 仍然能够做到准确执行任务。”杨戈解释道，用于训练 SAM 的数据集的数据量，是以往最大数据集的 6 倍。在这个数据集中，包括 1100 万张图片，和 11 亿个标注（可简单理解为

11 亿个物体）。这是一个“走量”的过程，数据量越大，机器分割图像的能力就越准确；即使在某张图中出现了这 11 亿个物体之外的物体，机器也往往能够根据以往经验以较高的准确率“推断”出它是什么物体，并将其纳入自己的数据库，这就是为什么 SAM 对于从未见过的物体，也能有很好的识别与分割效果。

“需要注意的是，这 11 亿个标注也不是纯手工完成的，而是通过渐进的方式自动完成标注的。一开始，这个数据集中只有相对少量的标注数据。科研人员先用这些数据训练出一个模型，然后再让这个模型自动标注数据，并通过人工对标注结果进行改进，这就得到了比上一个数据集更大一些的数据集。如此循环往复，就能得到海量标注数据集。”杨戈补充道。

促进计算机视觉领域发展

功能如此强大的图像分割通用模型，将给计算机视觉领域带来哪些改变？

“我认为，SAM 的出现将为计算机视觉领域的科研人员带来工作范式上的变化。”杨戈对记者说道，“SAM 的出现确实会对目前一些与机器人视觉相关的研究领域造成冲击，但从总体上看，SAM 的出现会提升相关科研人员的效率。”

杨戈解释道，以往科研人员构建图像分割模型，是一个“从下到上、从零开始”的过程；而图像分割通用模型则将模型构建方式变成了“从上到下”，即在已有性能和泛化能力更强的模型基础上继续修改、优化，“这可能确实会取代某些专有模型，但从总体上看它有利于整个领域的发展。”

此外，在具体应用上，图像分割通用模型前景十分广阔。

工业中的机器视觉、自动驾驶、安防等一些原来采用计算机视觉技术的行业，因为长尾场景多，需要大量标签数据，因此训练成本较高。有了图像分割通用模型后，这些领域内定制化开发产品的成本可能会降低，由此带来毛利率的提升；还有一些领域，过去因为样本量少而难以应用深度学习等人工智能算法。现在，由于 SAM 在零样本或者少量样本上表现优异，一些新的应用领域将被拓展，比如从代码驱动变为视觉驱动的机器人、流程工业场景等。

同时，由于 SAM 可以接受来自其他系统的输入提示，因此科幻片中根据用户视觉焦点信息来识别并选择对应物体或将成为可能。

SAM 不仅将在上述这些前沿领域发挥作用，同样或将会用于人们的日常生活。“比如在医学影像诊断领域，SAM 可能会催生出生精度更高的医学影像模型，提升医疗水平；在拍照过程中，SAM 的加入或将实现更快更智能的人脸识别。”杨戈说道。

“知海图 AI”开启内测 为用户提供“热榜摘要”服务

科技日报讯（实习记者骆香茹）知乎创始人、董事长兼首席执行官周源在 2023 年知乎发现大会上宣布，知乎首个大语言模型“知海图 AI”开启内测。周源表示，在知乎的“讨论场”中，AI 正在扮演越来越重要的角色。

据悉，“知海图 AI”中文大语言模型由知乎和北京面壁智能科技有限公司（以下简称面壁智能）合作开发，拥有强大的逻辑推理能力和推理速度。

目前，“知海图 AI”已经运用到知乎热榜中，内测功能“热榜摘要”已正式上线，该功能会对知乎热榜问题中的回答进行要素抽取、观点梳理和内容聚合，将回答以梗概的形式呈现给用户，使用户更快、更全面地了解讨论的热点与焦点。

面壁智能联合创始人、首席技术官曾国洋介绍，“热榜摘要”页面聚合的回答并不是一成不变的，而是会随着话题的发酵和讨论的发展，变得越来越丰富。

“打个比方，就好像我们刚刚加入一个讨论场，大家七嘴八舌，非常热闹。假如想要参与到这个讨论中来，我们首先应该干什么呢？毋庸置疑，我们肯定应该拉着一个离我们最近的人，问他：‘你们在讨论啥？’”曾国洋解释，“现在，在知乎的热榜上，我们的大模型就成了这样的一个人。”

AI 会取代创作者的主体地位吗？面对人与 AI 的关系问题，周源认为，要做好两件事：坚持生态第一、承担社会责任。

他表示，当前，为了确保 AI 在逐步加入社区生态的过程中持续保持赋能状态，知乎在线上 AI 创作辅助功能的同时，呼吁创作者对自己的作品负责。此外，知乎也开始对 AI 产生的劣质内容进行识别。

“我相信，就如同过往的所有技术一样，AI 终究会服务于人，赋能于人。”周源表示，“我们不仅要夯实在应用层和数据层这两方面的独特优势。同时，我们也要在基础模型层成为新生产力的开发者。”周源说。

身兼厨师和服务员 AI 食堂持证“上岗”

◎ 本报记者 叶青

在微信小程序上下单，过了一会儿，AI 智能烹饪机器人——“盒小午”就为科技日报记者制作出一份热气腾腾的午餐。4月11日，记者在广州人力资源服务产业园一家科技服务公司的 AI 食堂体验了一番高科技的用餐服务。

这个 AI 食堂拥有广东省首张智能移动微食堂经营试点备案凭证，由广州市黄埔区市场监管局发放，填补了新业态食品经营许可证可办理的政策空白，也是该局“一类一策 3.0”服务的一个生动案例。

身兼厨师和服务员双重身份的“盒小午”，虽然体积不大，占地仅 2.5 平方米，却很“全能”。它可同时烹饪至少 36 种菜品，用餐高峰期同时做好 200 份热乎乎的“暖心餐”，还能同时实现生鲜现制、自动烹饪、智能控制、批量出餐、无油烟排放等。

“盒小午”由广州才和科技服务有限公司（以下简称才和公司）运营，广州一盒科技有限公司自主研发。

此前，由于智能餐饮机器人这一新兴餐饮服务领域暂无明确对应的食品经营许可证规范，智能餐饮机器人的合法化运营受阻，才和公司的业务拓展遭遇了瓶颈。

广州市黄埔区市场监管局得知这一情况之后马上响应，为其提供“一类一策 3.0”服务，量身定制备案指南，经安全评估后向才和公司发出了全省首张智能移动微食堂经营试点备案凭证。

“从申报到办证下来，只用了不到两周的时间！”才和公司负责人李希说。

目前，广州市黄埔区约有 15 万家小微企业，其中绝大部分小微企业无法自建合规食堂。智能移动微食堂经营试点备案凭证的发放，有助于推动智能移动微食堂的普及，为小微企业降低 50%—60% 的食堂建设投入成本，以及 40%—80% 的食堂综合运营成本。除企业外，智能移动微食堂应用场景还可拓展到学校、医院等。

架起供需桥梁 青岛打造人工智能应用新高地

◎ 本报记者 宋迎迎

4月11日，山东省青岛市人工智能场景开放暨创新产品对接会举行，在人工智能场景与企业之间搭建起“供给”与“需求”的桥梁纽带，推动人工智能企业全面参与数字青岛建设。

2022 年以来，青岛市工业和信息化局开发上线了“链万企”供需对接公共服务平台。“链万企”平台自上线以来，链接平台和链主企业 72 个，其中包括 4 家人工智能产业链链主企业，优化了人工智能发展资源配置，搭建起行业互通桥梁。“青岛市数字家庭产业与应用促进会副秘书长王爱玲介绍说。

“青岛有良好的产业基础和广阔的应用场景，是人工智能创新应用的天然试验场。近年来，青岛举办了‘人工智能精准赋能促进月’‘百企百景’对接洽谈会等活动，上线‘链万企’供需对接公共服务平台，持续开展‘工业赋能’‘未来城市’等场景征集工作，发布场景 4050 个，配对场景 931 个，在人工智能领域的工作取得了一系列成果。”青岛市委常委、副市长耿涛介绍说，下一步，青岛市将以促进人工智能与实体经济深度融合为主线，围绕场景创新应用，加快资本、人才、技术、数据、算力等要素汇聚，推动产业集聚和生态培育，打造人工智能场景应用新高地。

◎ 本报记者 王祝华 实习生 曲怡臻

AIGC(人工智能生成内容)、人工智能大模型、对话机器人……近年来，人工智能领域技术突破不断，创新成果不断融入社会各个领域，深刻改变着人们的日常工作、生活方式。博鳌亚洲论坛 2023 年年会上，多位专家学者、行业领袖围绕人工智能领域技术突破和对人类社会的影响等问题，进行了深度研讨交流。

理性看待人工智能颠覆性创新

华为云人工智能领域首席科学家田奇表示，最近人工智能领域前沿技术的巨大突破对社会的影响，或要远远超过当年智能手机的出现。

人工智能已经发展了 60 余年，为何在近几年突然获得了如此巨大的进步？田奇认为，这是几十年积累后的一次爆发，新出

现的人工智能技术，不仅改变着人机交互的方式，还极大地提高了生产力。

不过，专家也表示，当前人工智能还有一定的局限性，那就是还未出现拥有类情感智能的个体。人工智能虽然是个“文理通才”，但它仍建立在人类所有知识和理解的基础上，并没有超过人类整个群体的智慧，更不要说拥有自身情感和智慧了。“就比如最近大火的 ChatGPT，它的回答水平虽然已经超越很多个体乃至群体，但从机器的视角看，本质上它对于回答了什么一无所知。”小米集团董事长兼首席执行官雷军说。

平安银行行长特别助理蔡新发表示，人工智能技术的进步之所以能产生革命性意义，是因为其具有跨界、横向、打通的能力。

“科技的发明具有两面性。”清华大学公共管理学院院长朱旭峰表示，人工智能技术在提高各行各业生产效率的同时，又会对社会结构产生深刻影响，往往会伴随

着一些行业的消失与崛起。因此，如何利用新的工具，以适应行业新的发展，成为未来是否能够屹立潮头的重要因素。

“未来的 5 至 10 年是人工智能技术应用的红利变现期。”袁辉表示，人工智能的新产品已经快速融入各行各业，不断塑造新业态、新场景，带来巨大的商业价值。

大模型推动 AI 产业化再加速

AI 大模型是近年来最为热门的 AI 细分领域。相比于小模型，AI 大模型具备多个场景通用、泛化和规模化复制等诸多优势，被视为是实现通用人工智能的重要研究方向。

无论是最近大火的 ChatGPT，还是国内百度公司推出的文心一言，其实都是 AI 大模型的典型代表。

所谓 AI 大模型，就是经过大规模数据训练后，能够适应一系列任务的模型。深度学习作为人工智能的重要技术，完全

依赖模型自动从数据中学习知识，在显著提升性能的同时，也面临着通用数据激增与专用数据匮乏的矛盾。AI 大模型兼具“大规模”和“预训练”两种属性，面向实际任务建模前需在海量通用数据上进行预先训练，能大幅提升 AI 的泛化性、通用性、实用性。

此外，相对于传统的小型生成模式，AI 大模型能够大幅缩减特定模型训练所需要的算力和数据量，缩短模型的开发周期，还能得到更好的模型训练效果。可以说，AI 大模型的真实意义在于改变了模型的开发模式，将模型的生产由“小作坊”升级为工业化的“流水线”，而模型开发模式的转变，将使得 AI 技术能够更广泛地地下沉到一些长尾场景。

当然，AI 大模型的开发需要具备丰富的开发资源，以及庞大的数据、算力支撑。

田奇建议，我国要建构统一架构的多模态 AIGC，加大大模型和底层硬件的适配，从而达到降本增效的效果。