# Article information:

通用视觉大模型综述 - 知乎  
<https://zhuanlan.zhihu.com/p/547494670>

# Article summary:

1. 大型视觉模型的发展：过去两年中，大型视觉模型得到了快速发展，包括微软的swin-transformer系列、谷歌的vit系列和具有150亿参数的V-MOE模型。大型模型在各种任务上表现出更好的性能。

2. Baidu Wenxin UFO 2.0大型视觉模型：Baidu提出了统一特征优化技术（UFO），通过充分利用大数据和大模型，并考虑到大模型实施的成本和效率，解决了实施预训练大模型所面临的问题。该技术方案包括All in One和One for All两个主要内容，可以处理多个任务并支持不同类型硬件的灵活部署。

3. Huawei Pangu CV大型模型：华为发布了Pangu CV系列大模型，是当时最大的视觉预训练模型之一，包含超过30亿个参数。Pangu CV在ImageNet数据集上的小样本分类准确率达到行业最高水平。该模型通过收集和索引大规模数据进行预训练，旨在解决以往“车间式”AI开发模式并降低模型训练成本。

# Article rating:

Appears strongly imbalanced: The article is written in a biased or one-sided way, and the information it provides is not trustworthy enough to be considered a reliable source. You should consult other sources to find reliable information on the presented issues.

# Article analysis:

这篇文章主要介绍了几个通用视觉大模型，包括百度的Wenxin UFO 2.0、华为的Pangu CV和商汤的INTERN。然而，文章存在一些问题和偏见。

首先，文章没有提供足够的证据来支持其所述的大模型在各种任务上取得更好性能的主张。虽然提到了这些模型在一些公共数据集上达到了SOTA效果，但没有提供具体的实验结果或比较其他模型的数据。因此，读者无法确定这些大模型是否真的比其他方法更好。

其次，文章对于大模型带来的计算资源和部署问题只是简单地提及，并没有深入探讨这些问题可能带来的潜在风险和挑战。例如，在部署到边缘设备或终端设备时，大模型可能会面临存储和计算能力有限的硬件限制。然而，文章并未讨论如何解决这些问题或提出可能存在的解决方案。

此外，文章对于这些大模型背后的数据集和训练过程也没有进行充分考虑。它只是简单地提到了使用大量数据进行预训练，并将其迁移到下游任务中进行微调。然而，并未涉及到数据集选择的偏见或可能存在的数据集偏差问题。此外，文章也没有讨论大模型可能带来的隐私和安全风险。

最后，文章对于这些大模型的宣传内容较多，缺乏客观和平衡的报道。它只是简单地介绍了这些模型的优点和功能，而没有提及任何潜在的缺点或限制。这种片面报道可能会误导读者，并使他们对这些大模型持有过于乐观的态度。

总之，这篇文章存在一些问题和偏见，包括缺乏证据支持、忽视潜在风险、片面报道等。读者应该保持批判思维，并进一步研究和探索相关主题。

# Topics for further research:

* 大模型在各种任务上的性能表现是否真的比其他方法更好？
* 大模型在计算资源和部署方面可能面临的挑战和风险是什么？
* 大模型的训练过程中是否存在数据集选择的偏见或数据集偏差问题？
* 大模型可能带来的隐私和安全风险有哪些？
* 这些大模型的缺点和限制是什么？
* 这篇文章是否客观和平衡地报道了这些大模型的优点和缺点？

# Report location:

<https://www.fullpicture.app/item/db3a76fc85be7683f821dbe1a1693fba>