腾讯混元大模型

产品简介

腾讯云

版权所有：腾讯云计算（北京）有限责任公司

【版权声明】

©2013-2025 腾讯云版权所有

本文档（含所有文字、数据、图片等内容）完整的著作权归腾讯云计算（北京）有限责任公司单独所有，未经腾讯云事先明确书面许可，任何主体不得以任何形式复制、修改、使用、抄袭、传播本文档全部或部分内容。前述行为构成对腾讯云著作权的侵犯，腾讯云将依法采取措施追究法律责任。

【商标声明】

腾讯云

及其它腾讯云服务相关的商标均为腾讯云计算（北京）有限责任公司及其关联公司所有。本文档涉及的第三方主体的商标，依法由权利人所有。未经腾讯云及有关权利人书面许可，任何主体不得以任何方式对前述商标进行使用、复制、修改、传播、抄录等行为，否则将构成对腾讯云及有关权利人商标权的侵犯，腾讯云将依法采取措施追究法律责任。

【服务声明】

本文档意在向您介绍腾讯云全部或部分产品、服务的当时的相关概况，部分产品、服务的内容可能不时有所调整。您所购买的腾讯云产品、服务的种类、服务标准等应由您与腾讯云之间的商业合同约定，除非双方另有约定，否则，腾讯云对本文档内容不做任何明示或默示的承诺或保证。

【联系我们】

我们致力于为您提供个性化的售前购买咨询服务，及相应的技术售后服务，任何问题请联系 4009100100或95716。

文档目录

产品简介
- 产品概述
- 产品优势
- 应用场景

产品简介

产品概述

最近更新时间：2025-06-27 14:12:42

腾讯混元大模型（Tencent Hunyuan）由腾讯公司全链路自研，在高质量内容创作、数理逻辑、代码生成和多轮对话上性能表现卓越，处于业界领先水平。混元 API 支持 AI 搜索联网插件，通过整合腾讯优质的内容生态，提供强大的时新、深度内容获取和 AI 问答能力。

混元生文

混元文生文目前已上线16个版本，以下为模型特性的简要介绍，可作为业务接入时版本选择的参考依据。

模型类型	模型名称（api调用名）	版本更新时间	能力和特征	输入输出	相关文档
文生文-推理模型	hunyuan-t1-latest	2025-05-21	【官方推荐使用】【推理模型】 业内首个超大规模 Hybrid-Transformer-Mamba 推理模型，扩展推理能力，超强解码速度，进一步对齐人类偏好。	最大输入28k 最大输出64k	调用示例 立即体验 Blog地址
	hunyuan-t1-20250529	2025-05-29	【最新版本】 能力和特征： 优化文本创作、作文写作，优化代码前端、数学、逻辑推理等理科能力，提升指令遵循能力。	最大输入28k 最大输出64k	调用示例 立即体验 Blog地址
	hunyuan-t1-20250521	2025-05-21	【推荐】【热门】 1. 数学、代码、科学、知识问答能力显著增强。	最大输入28k 最大输出64k	调用示例 立即体验 Blog地址

			2. 优化文科数据，提升文创、文本理解能力。		
	hunyuan-t1-20250403	2025-04-03	提升项目级别代码生成能力；提升文本生成写作质量；提升文本理解 topic 的多轮、tob 指令遵循和字词理解能力；优化繁简混杂和中英混杂输出问题。	最大输入 28k 最大输出 64k	调用示例 立即体验 Blog地址
	hunyuan-t1-20250321	2025-03-21	全面搭建模型文理科能力，长文本信息捕捉能力强。 支持推理解答各种难度的数学/逻辑推理/科学/代码等科学问题。	最大输入 28k 最大输出 64k	调用示例 Blog地址
	hunyuan-a13b	2025-06-25	1. 适用场景：绝大部分场景，同时兼顾效果及推理性能。 2. 模型能力和特征：混元第一个混合推理模型，hunyuan-standard-256K 的升级版本，总参数 80B，激活 13B，默认是慢思考模式，支持通过参数或者指令进行快慢思考模式切换，慢快思考切换方式为 query 前加/	最大输入 224k 最大输出 32k	调用示例 立即体验

			no_think；整体能力相对上一代全面提升，特别数学、科学、长文理解和Agent 能力提升显著。 3. 模型结构：混元 MoE 结构。		
通用文生文	hunyuan-turbos-latest	2025-05-20	【最新版本】【效果最优】【官方推荐使用】hunyuan-TurboS 混元旗舰大模型最新版本，具备更强的思考能力，更优的体验效果，已更新至最新版本。	最大输入 28k 最大输出 16k	调用示例 立即体验 Blog地址
	hunyuan-turbos-20250604	2025-06-04	【最新版本】能力和特征：预训练底座升级，写作、阅读理解能力提升，较大幅度提升代码和理科能力，复杂指令遵循等持续提升。	最大输入 28k 最大输出 16k	调用示例 立即体验 Blog地址
	hunyuan-turbos-20250515	2025-05-20	一、发布特性： 1. 长短思维链自适应融合。 2. 全面提升数学、代码、逻辑和科学等理科效果，文科效果维持。 二、效果指标：	最大输入 28k 最大输出 16k	调用示例 立即体验 Blog地址

			理科：数学ood 测试集提升7%；逻辑难题提升5%；代码能力及体验提升5%；科学提升8%。		
	hunyuan-turbos-20250416	2025-04-16	【推荐】【热门】预训练底座升级，增强底座的指令理解及遵循能力；对齐阶段增强数学、代码、逻辑、科学等理科能力；提升文创写作质量、文本理解、翻译准确率、知识问答等文科能力；增强各领域Agent 能力，重点加强多轮对话理解能力等。	最大输入24k 最大输出8k	调用示例 立即体验 Blog地址
	hunyuan-turbos-20250313	2025-03-13	【推荐】【热门】预训练底座训练token 数升级；数学/逻辑/代码等思考能力提升；中英文通用体验效果提升，包括文本创作、文本理解、知识问答、闲聊等。	最大输入24k 最大输出8k	调用示例 立即体验 Blog地址
	hunyuan-turbos-longtext-128k-20250325	2025-03-25	擅长处理长文任务如文档摘要和文档问答等，同时也具备处理通用文本生成任务的能力。在长文本的分析和生成	最大输入128k 最大输出6k	调用示例 立即体验

			上表现优异，能有效应对复杂和详尽的长文内容处理需求。		
	hunyuan-large	2025-02-10	Hunyuan-large 模型总参数量约 389B，激活参数量约 52B，是当前业界参数规模最大、效果最好的 Transformer 架构的开源 MoE 模型。	最大输入 28k 最大输出 4k	调用示例 立即体验 腾讯混元-large
	hunyuan-standard-256K	2024-02-10	采用更优的路由策略，同时缓解了负载均衡和专家趋同的问题。长文方面，大海捞针指标达到 99.9%。MOE-256K 在长度和效果上进一步突破，极大的扩展了可输入长度。	最大输入 250k 最大输出 6k	调用示例 立即体验
	hunyuan-standard	2025-02-10	采用更优的路由策略，同时缓解了负载均衡和专家趋同的问题。MOE-32K 性价比相对更高，在平衡效果、价格的同时，可实现对长文本输入的处理。	最大输入 30k 最大输出 2k	调用示例 立即体验
	hunyuan-lite	2024-10-30	升级为 MOE 结构，上下文窗口为 256k ，在 NLP，代码，数学，行业等多项	最大输入 250k 最大输出 6k	调用示例 立即体验

			评测集上领先众多开源模型。		
机器翻译	hunyuan-translation	2024-10-25	支持中文和英语、日语、法语、葡萄牙语、西班牙语、土耳其语、俄语、阿拉伯语、韩语、意大利语、德语、越南语、马来语、印尼语15种语言互译，基于多场景翻译评测集自动化评估COMET评分，在十余种常用语种中外互译能力上整体优于市场同规模模型。	最大输入4k 最大输出4k	调用示例 立即体验
	hunyuan-translation-lite	2025-06-06	混元翻译专项模型，基于混元2B-Dense模型进行翻译能力专项优化，通过迭代高质量多语言SFT数据，强化多语种翻译能力。支持简体中文、繁体中文、粤语、印尼语、英语、日语、法语、葡萄牙语、西班牙语、土耳其语、俄语、阿拉伯语、韩语、意大利语、德语、越南语、马来语、印尼语等18种语言互译。	最大输入4k 最大输出4k	调用示例 立即体验
角色扮演	hunyuan-role	2024-07-04	混元最新版角色扮演模型，混元官方精调训练推出的角色扮演模	最大输入28k 最大输出4k	调用示例 立即体验

版权所有：腾讯云计算（北京）有限责任公司

			型，基于混元模型结合角色扮演场景数据集进行增训，在角色扮演场景具有更好的基础效果。		
Function Call	hunyuan-functioncall	2025-04-22	混元最新 MOE 架构 FunctionCall 模型，经过高质量的 FunctionCall 数据训练，上下文窗口达 32K，增强了复杂指令和多轮理解能力，增强了开源 BenchMark 的效果，在多个维度的评测指标上处于领先。	最大输入 28k 最大输出 4k	调用示例 立即体验
代码生成	hunyuan-code	2024-11-12	混元最新代码生成模型，经过 200B 高质量代码数据增训基座模型，迭代半年高质量 SFT 数据训练，上下文长窗口长度增大到 8K，五大语言代码生成自动评测指标上位居前列；五大语言10项考量各方面综合代码任务人工高质量评测上，性能处于第一梯队。	最大输入 4k 最大输出 4k	调用示例 立即体验

混元多模态模型

混元多模态模型重磅发布，支持图片生文，包括图片基础识别、图片内容创作、图片多轮对话、图片知识问答、图片分析推理、图片 OCR 等能力。以下为模型特性的简要介绍，可作为业务接入时版本选择的参考依据。

模型类型	模型名称	版本更新时间	能力和特征	输入输出	相关文档
混元图生文	hunyuan-t1-vision	2025-05-16	混元多模态理解深度思考模型，支持多模态原生长思维链，擅长处理各种图片推理场景，在理科难题上相比快思考模型全面提升。	最大输入8k 最大输出24k	调用示例 立即体验
	hunyuan-t1-vision-20250619	2025-06-19	混元最新版t1-vision多模态理解深度思考模型，支持多模态原生长思维链，相比上一代默认版本模型全面提升。	最大输入16k 最大输出24k	调用示例 立即体验
	hunyuan-turbos-vision	2025-05-23	此模型适用于图文理解场景，是基于混元最新 turbos 的新一代视觉语言旗舰大模型，聚焦图文理解相关任务，包括基于图片的实体识别、知识问答、文案创作、拍照解题等方面，相比前一代模型全面提升。	最大输入8k 最大输出24k	调用示例 立即体验
	hunyuan-turbos-vision-20250619	2025-06-19	混元最新版turbos-vision视觉语言旗舰大模型，在图文理解相关的任务上，包括基于图片的实体识别、知识问答、文案创作、拍照解题等上面相比上一代默认版本模型全面提升。	最大输入16k 最大输出16k	调用示例 立即体验

	hunyuan-large-vision	2025-05-26	此模型适用于图文理解场景，是基于混元Large训练的视觉语言大模型，支持任意分辨率多张图片+文本输入，生成文本内容，聚焦图文理解相关任务，在多语言图文理解能力上有显著提升。	最大输入8k 最大输出8k	调用示例 立即体验
	hunyuan-vision	2025-01-03	混元最新多模态模型，支持图片+文本输入生成文本内容。 图片基础识别：对图片中主体、元素、场景等进行识别 图片内容创作：对图片进行概述、创作广告文案、朋友圈、诗词等 图片多轮对话：输出单张图片进行多轮交互问答 图片分析推理：对图片中逻辑关系、数学题、代码、图表进行统计分析 图片知识问答：对图片包含的知识点进行问答，例如历史事件、电影海报 图片 OCR：对自然生活场景、非自然场景的图片识别文字。	最大输入6k 最大输出16k	调用示例 立即体验

混元生图

混元生图提供文生图 API 技术服务，基于混元文生图大模型，可结合输入的文本描述智能创作出与输入相关的图像内容，提供更高质量的图像创作、更强大的中文语义理解能力以及更广泛的应用场景。

模型类型	模型名称	版本更新时间	能力和特征	输入输出	相关文档
混元生图	混元生图	2024-09-11	输入 prompt 文本，生成图像。 prompt 扩写：支持 prompt 自动扩写优化。 绘画风格：支持18种预置风格选项 + prompt 自定义风格。 分辨率：支持8种分辨率。 生成数量：支持1 - 4张。 随机种子：支持随机种子生成或自定义固定种子生成。	最大输入1024字符文本 最大输出4张图片	调用示例 立即体验
	混元生图（多轮对话）	2024-08-28	输入 prompt 文本，生成图像，支持通过多轮对话的方式不断调整图像内容。	最大输入100轮对话文本 输出图片	调用示例
	文生图轻量版	2024-07-15	输入 prompt 文本，生成图像，支持更轻量的同步请求方式和更高的性能。	最大输入256字符文本 输出图片	调用示例

混元生3D

混元生3D 提供 API 技术服务，基于混元生3D 大模型，可结合输入的文本描述或者上传2D 图像快速生成高精度3D 模型，显著降低创作门槛，适用于游戏、影视、工业设计等领域，大幅提升效率。

模型类型	模型名称	版本更新时间	能力和特征	输入输出	相关文档
混元生3D	混元生3D	2025-03-11	支持根据文本描述或上传2D图像自动生成高精度3D 模型。	• 最大输入图分辨率5k • 最大输入200个 utf-8 字符	调用示例

				• 输出 3D 模型	

腾讯元器

腾讯元器平台是一款基于混元大模型开发的一站式智能体创作与分发平台，用户无需编程经验即可轻松创建属于自己的智能体。在这个平台上，用户可以通过描述来设定智能体的角色和流程，利用丰富的插件和知识库资源让智能体更加智能化。平台还提供基于 AI 大模型的智能对话功能，以及智能体商店，汇聚了众多富有创意和趣味性的智能体玩法，为用户带来全新的智能体验。

版权所有：腾讯云计算（北京）有限责任公司

产品优势

最近更新时间：2025-03-04 18:07:52

腾讯混元大模型是由腾讯全链路自研的通用大语言模型，拥有万亿参数规模，具有强大的中文理解与创作能力、逻辑推理能力，以及可靠的任务执行能力。

腾讯混元通过采用混合专家模型(MoE)结构，推动了性能提升和推理成本下降。在中文表现尤其是在文本生成、数理逻辑和多轮对话上性能表现卓越，整体处于业界领先水平。支持AI 搜索联网插件，通过整合腾讯优质的内容生态和全网搜索，使混元具备强大的时新、深度内容获取和AI问答能力。

同时，混元还采用了各种技术手段来提高模型的性能和效果，例如使用掩码策略、使用不同的优化算法、进行数据增强等。这些技术手段可帮助模型更好地处理文本数据，提高模型的泛化能力和生成效果。

腾讯混元大模型目前覆盖**四大核心能力：**

多轮对话

具备上下文理解和长文记忆能力，流畅完成各专业领域的多轮问答。

知识增强

有效解决事实性、时效性问题，快速检索相关内容，提升内容生成效果。

逻辑推理

准确理解用户意图，擅长数学推导，基于输入数据或信息进行推理、分析和规划。

内容创作

支持文学创作、文本概要、角色扮演、文生图和图生文等多模态创作能力，生成内容流畅、规范、中立、客观。

版权所有：腾讯云计算（北京）有限责任公司

应用场景

最近更新时间：2024-03-29 15:46:41

文档场景

可提供文档创作、文本润色、文本校阅、表格公式及图表生成等能力，提高创作效率，提升创作体验。

会议场景

可提供会中问答、会议总结、会议待办项整理等能力，简化会议操作并提高会议效率。

广告场景

可提供智能化的广告素材创作，提升营销内容创作工作效率。

营销场景

构建智能导购，帮助商家提升服务质量和服务效率。

版权所有：腾讯云计算（北京）有限责任公司