

更多AI工具可直接访问：<https://www.faxianai.com/>

视频模型：可灵

文档可编辑，欢迎贴上你的测试效果

- 专用语言模型，提升生成效果

2024年6月6日快手发布“可灵”视频大模型：

[🇺🇸 快手【可灵】大模型介绍PPT](#)

- 生成超过120秒1080P视频
- 模拟真实物理特性（重力、光影反射、液体流动等）
- 准确建模复杂运动场景（高速奔跑的动物、月球行走的宇航员等）
- 画面连贯，动作流畅，细节真实
- 支持用户输入控制信息，丰富内容控制能力
- 类Sora的DiT结构，用Transformer代替卷积网络
- 自研3D VAE网络，提升视频重建质量

 官网链接<https://kling.kuaishou.com/mobile>

小互报导：

[🔗 https://xiaohu.ai/p/9119](https://xiaohu.ai/p/9119)

[🔗 https://x.om/imxiaohu/status/18040](https://x.om/imxiaohu/status/18040)

大聪明的技术详解

中学生能看懂：快手「可灵」和「Sora」背后 DiT 技4术

https://mp.weixin.qq.com/s/2PrMgNAL0Er_vNjhqSbTHQ

6月21日发布图生视频功能

群公告

一、可灵产品更新📢

🎉 “图生视频”和“视频续写”功能上线啦！欢迎大家前往快影APP体验！

1. 图生视频：上传任意图片，生成5秒精彩视频
2. 视频续写：对生成视频一键续写4~5秒，支持多次续写（最长3分钟）
3. 文生视频多比例选择：新增9:16和1:1视频尺寸选择，提供更多创作可能

🔗 更多介绍详见官网：

二、产品体验权限申请流程📝

1. 第一步：快影申请👉快影APP - AI创作 - AI生视频 - 填写信息
2. 第二步：问卷填写👉<https://ppg.viviv.com/s/Xku4zYHe>

每日根据以上填写信息优先开放权限，如申请超过3日仍未开通，可私聊 @可灵社群小助理

三、创作者激励活动🏆

1. 长期激励：带话题标签 #可灵 #快手可灵# 可灵AI，在快手APP发布有额外流量激励！
2. 话题投稿：参与快手定向投稿任务，获得专属流量激励，带2个话题 #可灵AI创意大赏 #可灵，关注并@可灵AI @快手热点 官号，详细规则：<https://a.d4t.cn/DpFGhZ>
3. 可灵官号投稿：<https://ppg.viviv.com/s/kiqAWHrd>

欢迎大家踊跃投稿！👏

7月6日功能更新发布首尾帧功能

1. 上线网页版：<https://klingai.kuaishou.com/>
2. 增加首尾帧功能
3. 文生视频模式下增加运镜控制功能

7月24日发布会员体系

1. 结束免费内测，使用可灵无需再申请

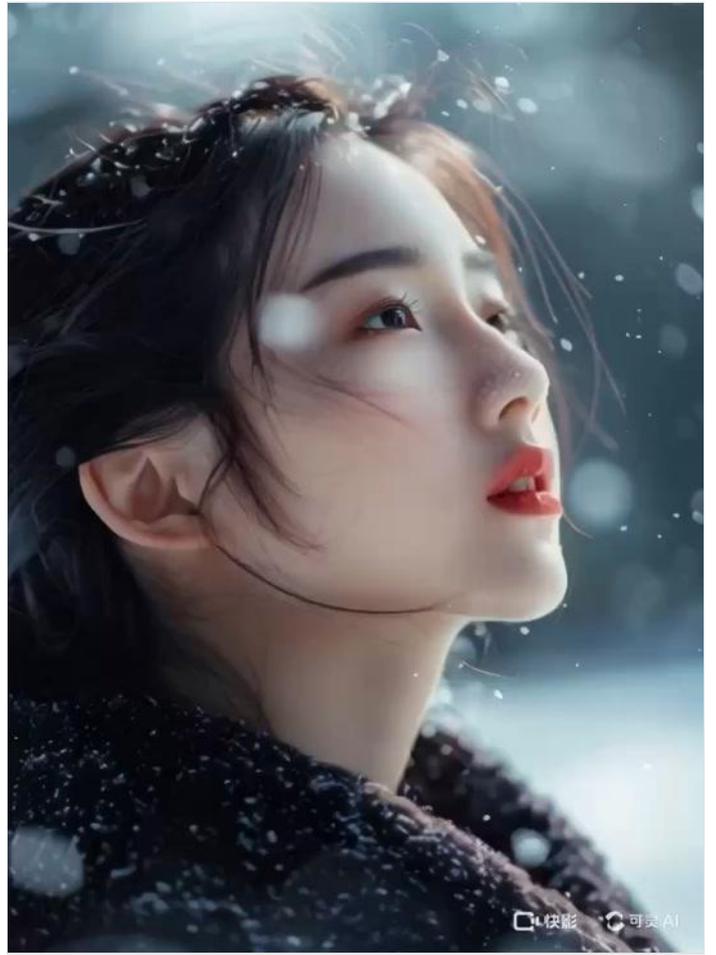
2. 引入灵感值，生成视频消耗灵感值
3. 使用高级功能（视频延长、高表现模式、大师运镜等）和获取更多灵感值，需付费成为VIP用户
4. 非VIP用户只能使用普通功能（正负向提示词、高性能模式、时长5秒、视频比例、运镜等），每天登录可灵送66灵感值（有效期一天），大约可生成6个视频
5. 会员充值限时一周半价，黄金会员33元/单月、396元/一年，铂金会员133元/单月、1596元/一年，钻石会员333元/单月、3996元/一年

图生视频效

测试者：AJ+三思



上面有个美女集合长视频

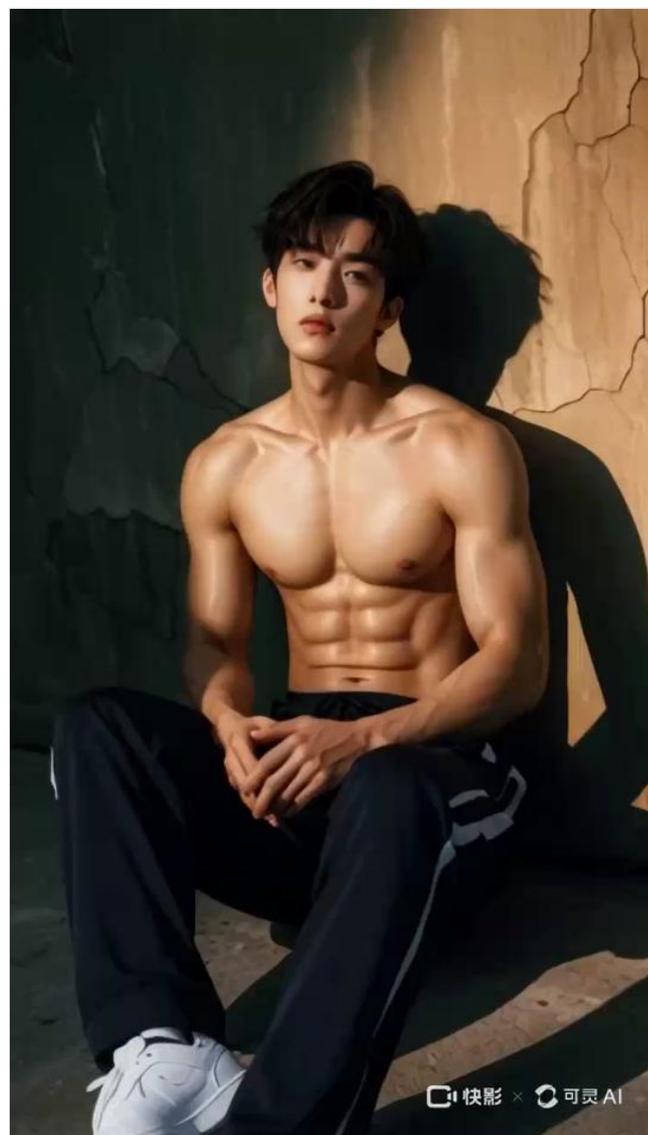


测试者：guahunyo（一定要看）



测试者：麦橘





测试者：葡了个萄



测试者：大佬的鼠DonRat



测试者：八月DMNK



测试者： 纸墨

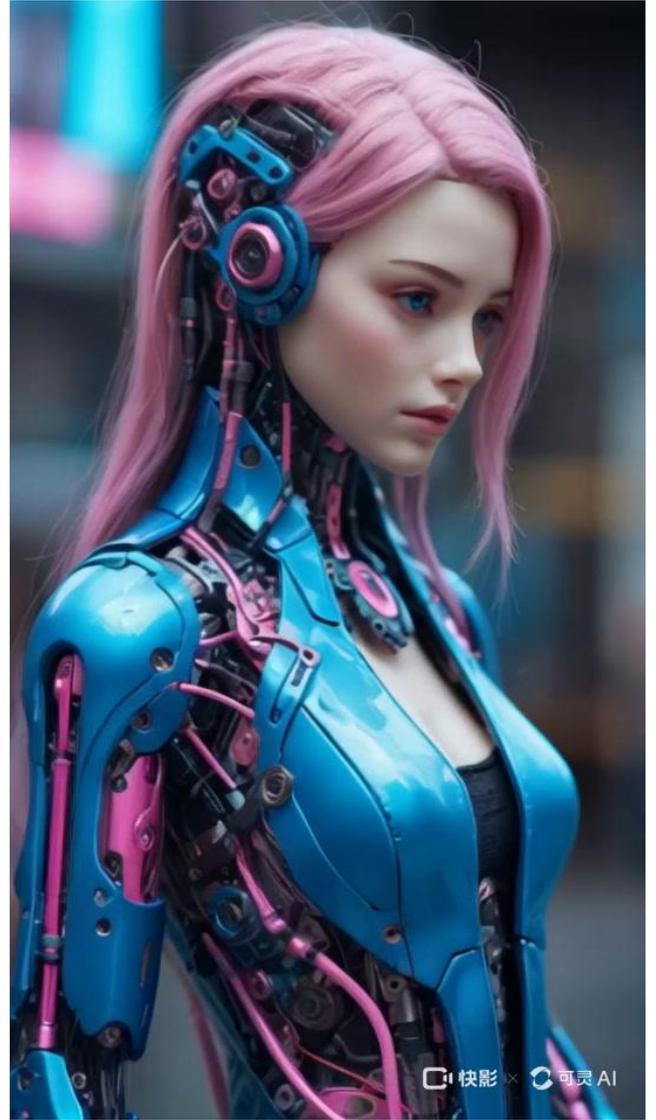


测试者： 汤团猪



测试者：德里克文





测试者：AlainYe



测试者：OceanYoung





测试者：洪九九





测试者：吵爷



测试者：猫叔



 WeChat_20240621183535.mp4





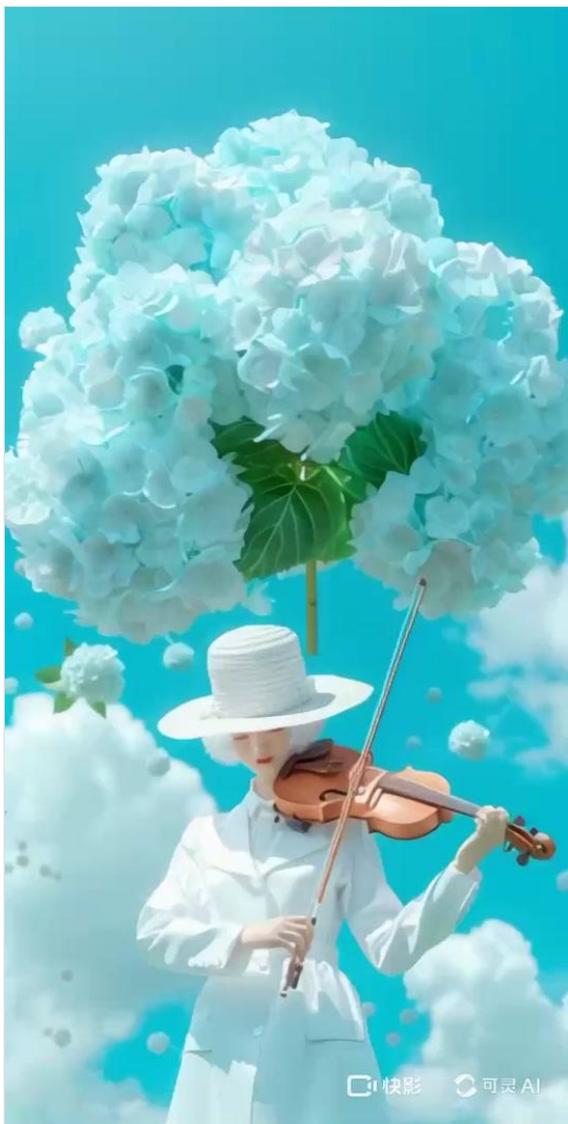




测试者：超楠



测试者：咖菲猫咪







测试者：张余



测试者：梁明晶





测试者：上官文卿







快影 × 可灵 AI

AI 生成内容



测试者：杉杉杉木



测试者：阿鸿师兄 3.6.0



测试者：汤小白





测试者：爱育





大鹏一日同风起



测试者：凌云



高性能_生成视频 (1).mp4
4.37MB



测试者：Silver



测试者：

测试者：

测试者：

文生视频效果

文生视频大模型

可灵大模型 (Kling) 是由快手大模型团队自主研发的视频生成大模型，具备强大的视频生成能力
 让用户可以轻松高效地完成艺术视频创作

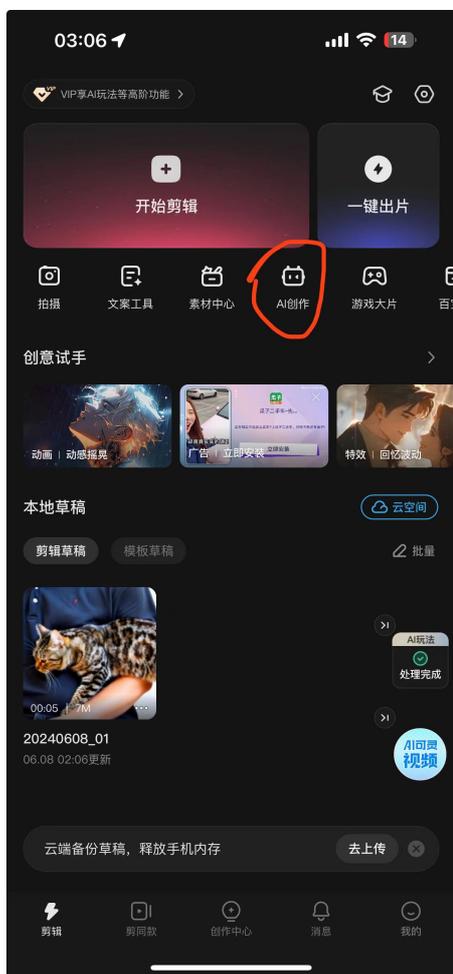
大幅度的合理运动

可灵采用3D时空联合注意力机制，能够更好地建模复杂时空运动，生成较大幅度运动的视频内容，同时能够符合运动规律。



长达2分钟的视频生成

得益于高效的训练基础设施、极致的推理优化和可扩展的基础架构，可灵大模型能够生成长达2分钟的视频，且帧率达到30fps。



🌟【**归藏**】提升质量关键词：获奖作品，丰富细节等与文生图类似。同时加上一些表现情绪和氛围的词也会变好，另外加上明亮的颜色也会让视频变清晰

测试者：三思

我不知道该测试哪些页面，所以问了下GPT-4o

测试内容

在应用层面测试AI视频模型时，作为一个测试者，您可以专注于以下画面内容，以全面评估模型的性能和效果：

类别	具体内容	说明
场景识别	室内场景、室外场景	测试模型在不同环境中的表现
物体识别	静态物体、动态物体	评估模型对不同类型物体的识别能力
人物识别	面部识别、姿态识别	测试模型对不同人物特征和姿态的识别能力
动作识别	基本动作、复杂动作	评估模型对不同复杂度动作的识别能力
事件检测	简单事件、复杂事件	测试模型对不同事件的检测能力
环境变化	光线变化、天气变化	评估模型在不同环境变化下的表现
多对象场景	单一对象场景、多对象场景	测试模型在复杂场景中的表现
特殊情况	遮挡情况、非标准视角	评估模型在特殊情况下的识别能力
细节检测	小物体、微表情	测试模型对细节的检测能力

测试结果：

1. 场景识别

- 室内场景:** 测试模型在不同室内环境中的表现，例如办公室、厨房、卧室等。
- 一个穿着西装的人正在办公室里，坐在电脑前工作。背景有办公桌、文件柜和窗户。



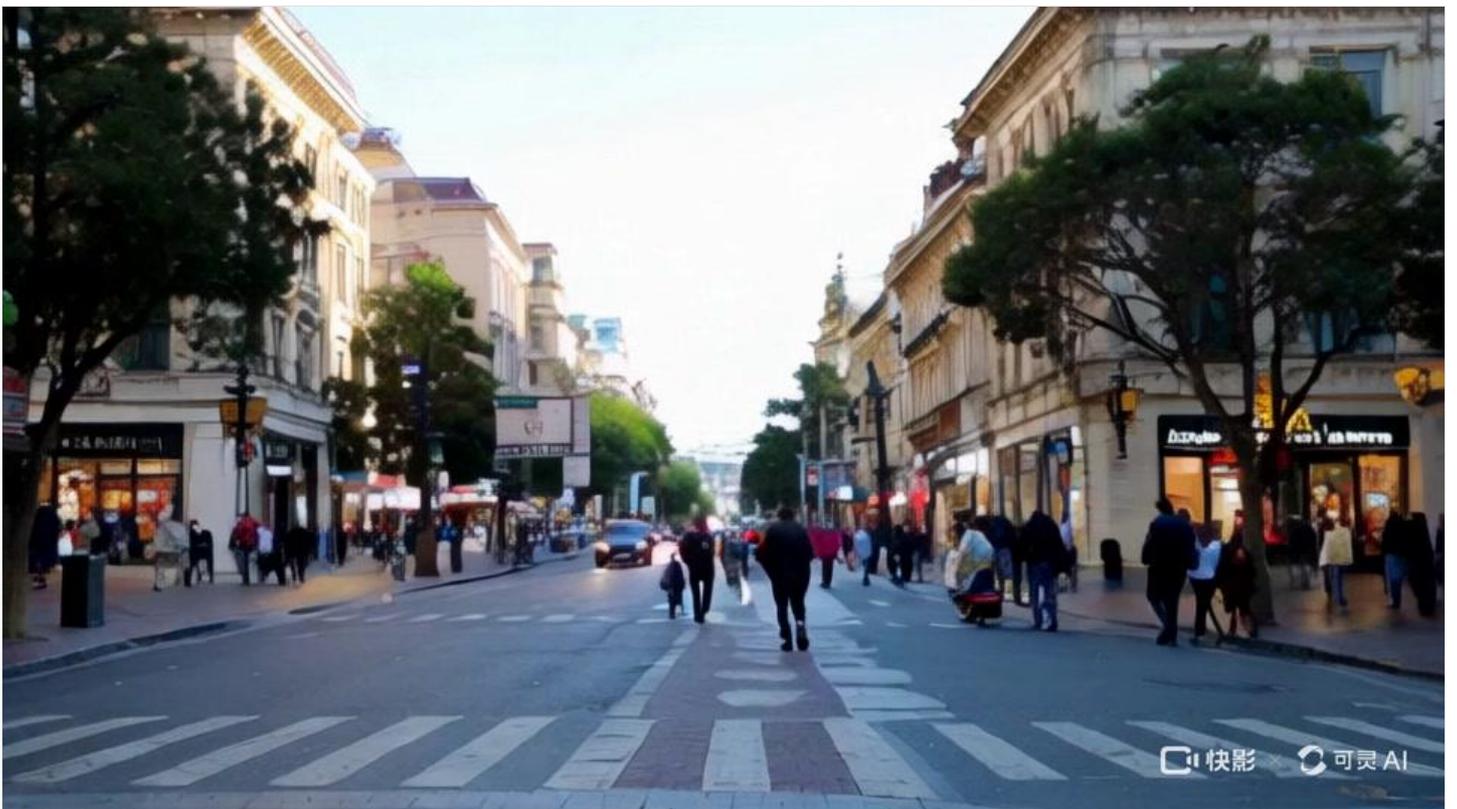
- 一个人在厨房里烹饪，背景有炉灶、冰箱和厨具。（人没出现）



- **室外场景：**测试模型在不同室外环境中的表现，例如街道、公园、商场等。
- <https://lumalabs.ai/dream-machine>



- 一群人在街道上行走，背景有车辆、行人和商店。



2. 物体识别

- **静态物体:** 评估模型对静止物体的识别能力，如家具、车辆、建筑物等。



- **动态物体:** 评估模型对移动物体的识别能力，如行人、跑步者、骑自行车的人等。



3. 人物识别

- **面部识别:** 测试模型对不同面部特征的识别能力，包括不同年龄、性别、种族等。
- 一组不同年龄段的人，包括儿童、青少年、成年人和老人。



不同种族的人物面部图像。（好像是一样的）



- **姿态识别:** 评估模型对不同人物姿态的识别能力，例如站立、坐下、躺卧等。
- 一个人站立在房间中。



一个人躺在床上（啊？2个人）



4. 动作识别

- **基本动作:** 评估模型对基本动作的识别能力，如走路、跑步、跳跃等。



一个人在草地上跳跃。



- **复杂动作:** 评估模型对复杂动作的识别能力，如舞蹈、体育活动、手势等。



一群人在舞蹈室里跳舞。



一群人在篮球场上打篮球。







5. 事件检测

- **简单事件:** 测试模型对简单事件的检测能力，如开门、关灯、拿起物品等。（其实试了2次书没有拿起来，只在摸了下）
- 一个人从桌子上拿起一本书。



一个人在房间里开门。



- **复杂事件:** 测试模型对复杂事件的检测能力，如交通事故、打斗、拥挤等。
- 两个人在公园里发生打斗。



6. 环境变化

- **光线变化:** 评估模型在不同光照条件下的表现，包括白天、夜晚、阴天、阳光直射等。
- 一个房间在白天的自然光照下。



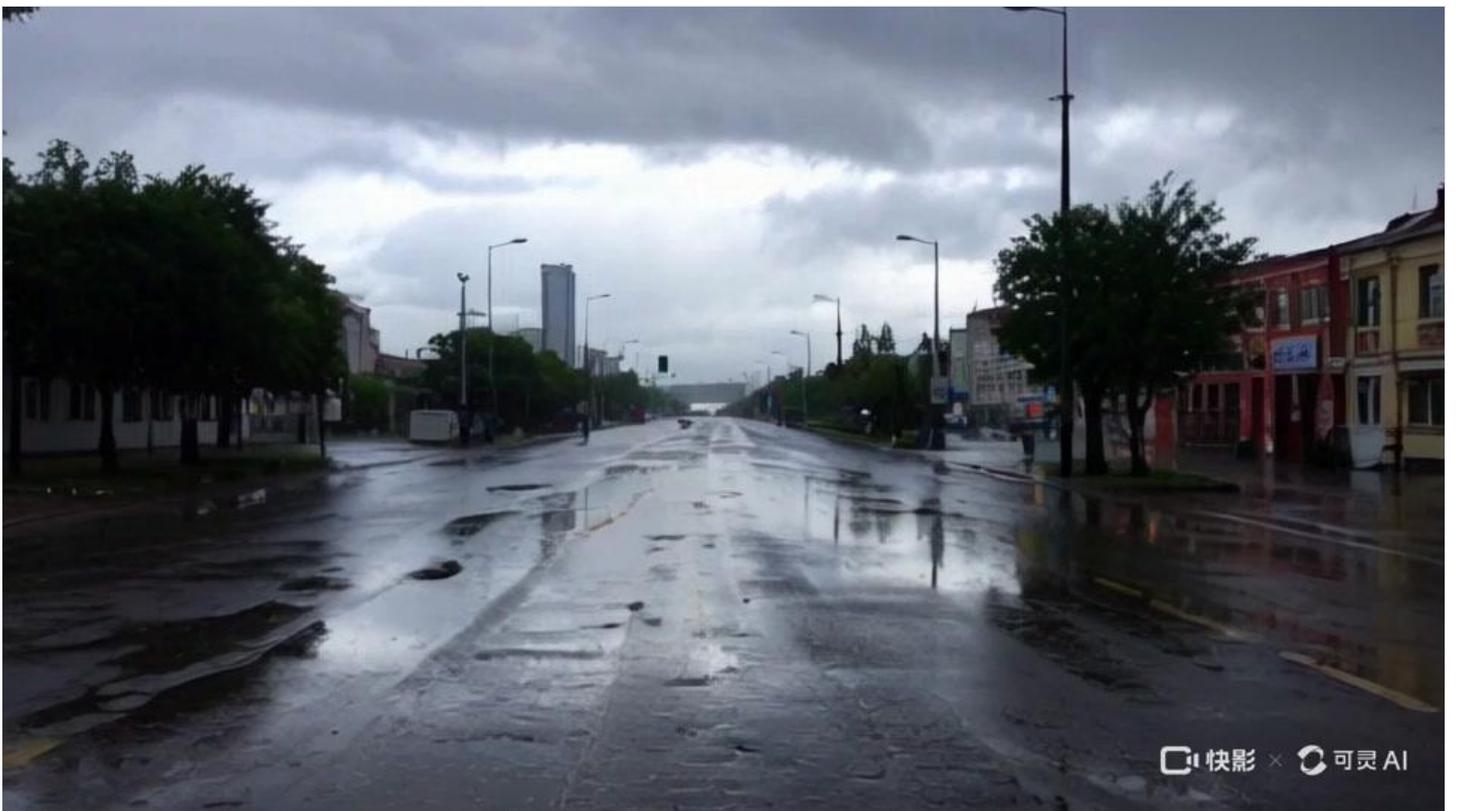
一个房间在夜晚灯光照明下。



- **天气变化:** 测试模型在不同天气条件下的表现，如晴天、雨天、雪天、雾天等。
- 公园在下雪的情况下。



街道在下雨的情况下。





7. 多对象场景

- **单一对象场景:** 评估模型在单一对象场景中的表现。
- 一个房间里只有一个人在看书。

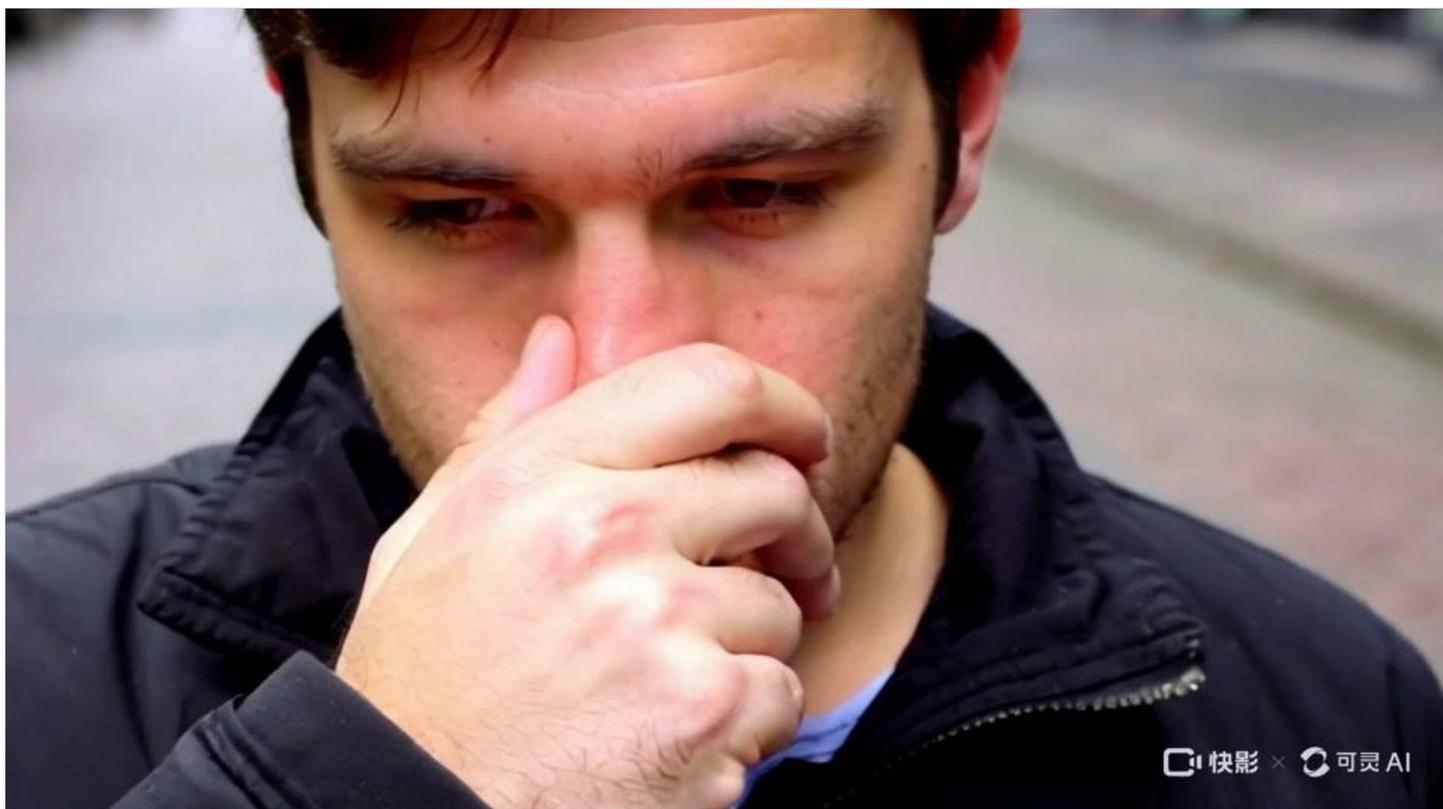


- **多对象场景:** 测试模型在包含多个对象的复杂场景中的表现，例如多人聚会、车流、商场购物等。
- 客厅里有几个人在聚会。



8. 特殊情况

- **遮挡情况:** 评估模型在部分物体或人物被遮挡时的识别能力。
- 一个行人的面部被手遮挡。



- **非标准视角:** 测试模型在非标准视角下的表现，例如斜视、俯视、仰视等。



9. 细节检测

- **小物体:** 测试模型对小物体的识别能力，例如钥匙、手机、书本等。



-
- **微表情:** 评估模型对面部微表情的识别能力，如皱眉、微笑、惊讶等。



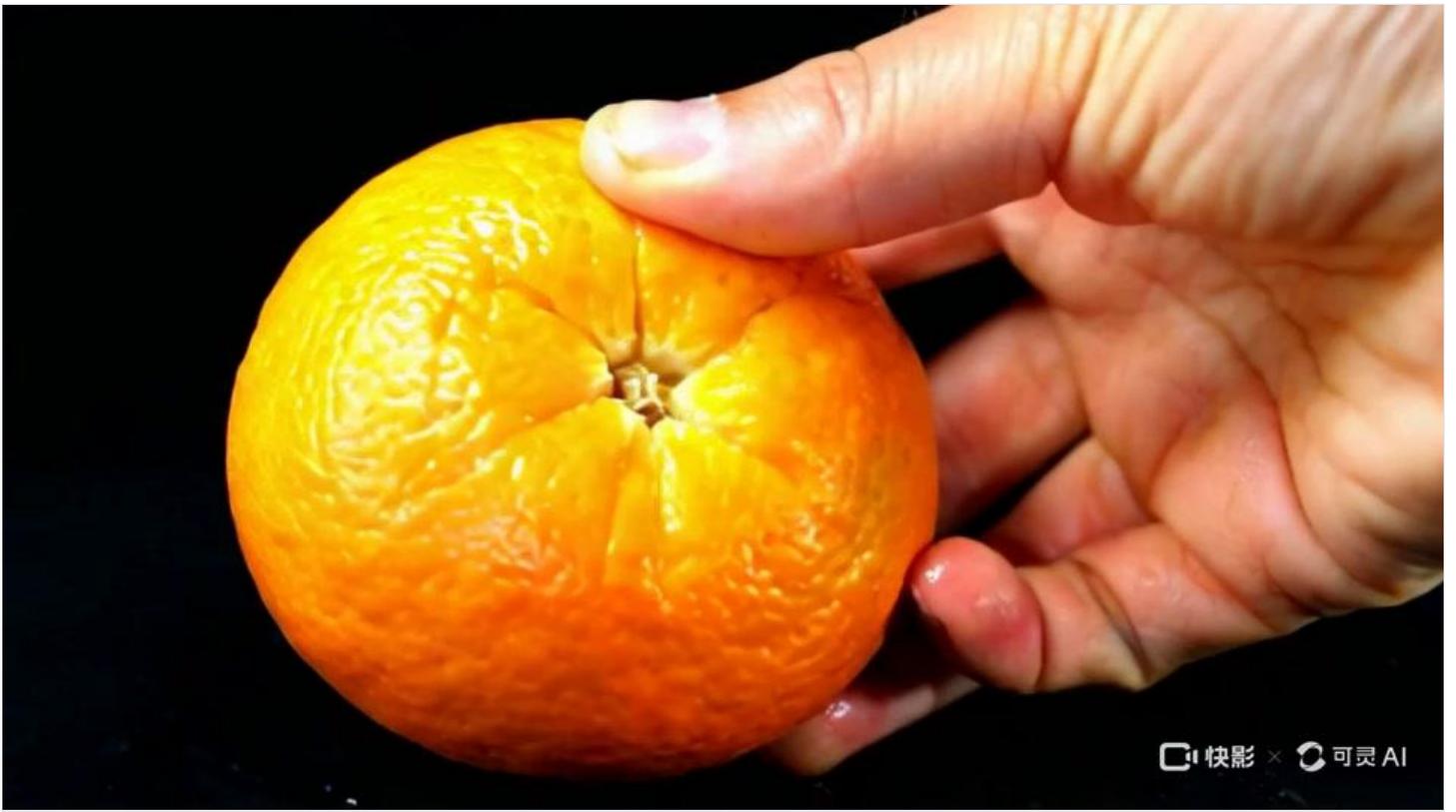


失败的案例：

一直兔子穿着西服，拿着公文包（？）



手剥橘子这种细致的动作还是会失败，试了几次



玫瑰花逐渐枯萎 (试了几次都不行)



黄种人和白种人和黑种人合影 (???)



穿着运动服的鹿头人，带着黑色耳机。拿着一个巨大的橘子（不稳定，动物拟人穿衣服很难）







测试者：





测试者：尾鳍Vicky



测试者：大雷





测试者：黄社长









测试者：十三







快影 × 可灵 AI



快影 × 可灵 AI



测试者：袁博





测试者：猫叔







测试者：温维斯Wenvis

Prompt	Video
床前明月光，疑似地上霜	 A bed in a snowy landscape under a full moon. The bed is covered with a white sheet and a pillow. The background is a dark blue sky with a full moon and some stars. In the bottom right corner, there is a small logo for '快影' and '可灵 AI'.
床边的闹钟响了，北极熊伸手把闹钟关了	



猴子在金碧辉煌的宫殿打麻将



全身镜头，一个熊猫在浴室洗澡，隔着玻璃，花洒，热气腾腾

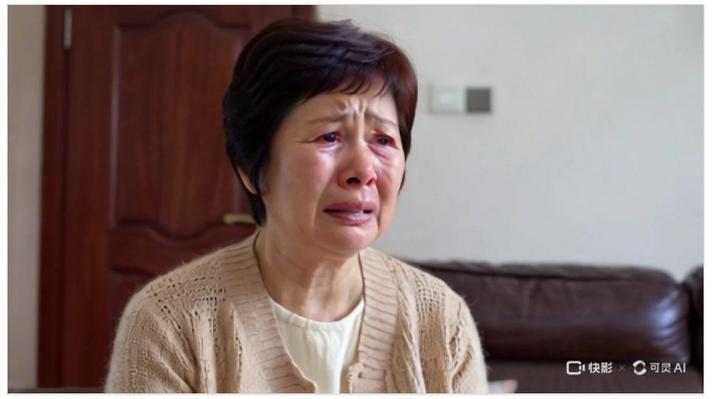


测试者：杉杉

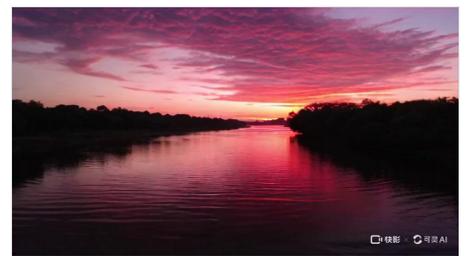
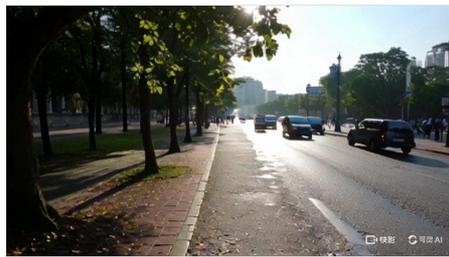
最开始想做睁眼到起床的镜头，需要直观的描述画面内的元素及动态。



需要明确描述中国人、亚洲人等才可以稳定的出中国面孔。



自然场景的光线很变化很流畅也比较自然。



场景可以明确国家地点，有些可能涉及屏蔽词。



风格化运动比较难实现，更类似于在真实世界用去拍摄艺术作品。（浅试了下千禧风格，Y2k）



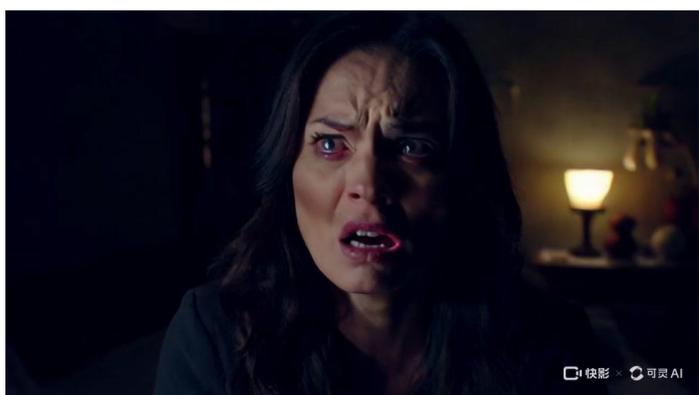
诺兰，城市的街道弯曲变形，建筑颠倒错落，旋转的全景镜头，呈现出奇幻而混乱的空间。



真实及虚拟的动物尝试。



剧情向镜头尝试，目前5s虽稳定，但是节奏过于慢了点。

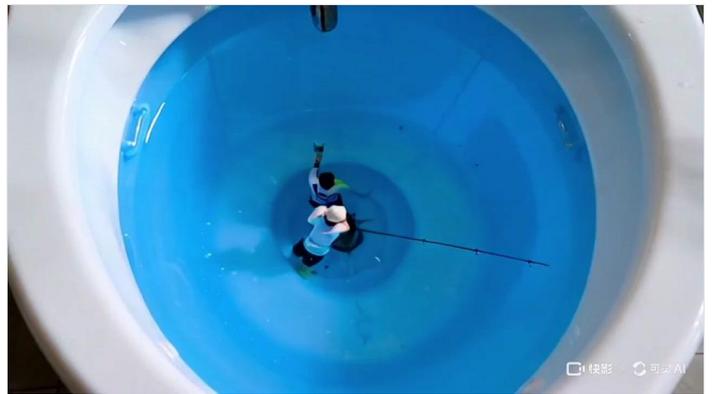




其他



测试者：电子酒







测试者：豆酱



我的世界风格，还OK



云组成的龙，大失败



微缩世界全景，OK



微缩世界特写，OK



写实烤火，OK，对于写实的表现最好



未来主义风格，大失败



倒酒, OK



特写液体, OK



银河广角, OK



眼部特写, OK

测试者: 德里克文



大象近景, 写实



大象中景, 动画风



火鸡打篮球



测试SORA, 花瓣老虎飞天



科技未来都市，发生爆炸



科技未来都市，天空中飞机发生爆炸



一群兔子在跳兔子舞，背景是赛博朋克都市



镜头推进，东方韵味，千里江山图



一个美丽性感的中国女人在泳池里吃苹果



一个美丽性感的中国女人在泳池里吃鸡腿



一个美丽性感的中国女人在泳池里吃橘子



一个小朋友在餐盘吃苹果派

测试者：卡尔的AI沃茨

17个镜头剪辑 | 与Sora平行对比

测试者：波风若川

不符合现实逻辑的画面测试



蓝鲸在天空中飞，晚霞



金鱼在天空中飞



天空的彩虹上，独角兽在奔跑



海底，水中的彩虹，珊瑚



机器人在海水里弹吉他，珊瑚礁，鱼



猫猫在海里潜水，抓鱼



海盗船在天空中快速的飞，云



大象在云上散步



一本打开的书在城市的空中



飞翔的书



火车在太空银河中飞行



无尽的旋转楼梯

测试者：今夕何夕







测试者：花花hana

室外场景识别



(少女背靠一颗大树，四周都是广袤无垠的草原风光，少女手拿一本书和铅笔，旁边放着一顶遮阳帽-plus: 少女是微笑的脸庞)

人物面部识别





(女生从生气到释然的表情变化-plus: 女生从生气紧皱眉头, 到开心微笑、舒展眉头的表情变化, ps: 都有点奇怪lmao...)

测试者：噍哩啪啦猴

单一对象多场景测试



一只白色的猴子戴着黑色大眼镜，骑着单车送快递



一只白色的猴子戴着黑色大眼镜，骑着摩托疾驰



一只白色的猴子戴着黑色大眼镜，骑着单车送快递



一只白色的猴子戴着黑色大眼镜，在咖啡厅喝冰咖啡，周围有很多动物也在喝咖啡



一只白色的猴子戴着黑色大眼镜，开着一辆跑车疾驰



一只白色的猴子戴着黑色大眼镜，在便利店吃泡面

测试者：AlainYe



镜头缓慢推进，一只熊猫穿着高科技战甲，坐在一辆造型独特的摩托车上，营造出一种冒险即将开始的氛围。



熊猫战甲上的灯光在昏暗的环境中闪烁。特写穿着机甲熊猫，面部表情



一个摩托车手，车头镜头，驾驶快速通过赛道，身边的风景快速后退



穿着机甲的熊猫战士开着摩托在城市中高速穿梭，周围是各种未来科技的展示。



一片废墟，熊猫战士开着摩托在废墟中穿行，偶尔有爆炸声和火光。



切换到熊猫的赛博战甲，展示其在战斗中受损但依然坚挺。



画面转为黄色调，展示一望无际的沙漠，熊猫战士的摩托在沙丘上留下深深的车辙。



一只熊猫战士停在一座高楼的顶端，特写眺望远方。寂寞

测试者：Austin 周安鑫

1、场景识别：室内&室外



客厅的沙发，液晶电视，漂亮的波斯猫，超现实主义。碧绿的草坪。



蜿蜒的河流，茂密的森林，灿烂的花海，

2、物体识别：静态&动态



水晶球，展览在舞台上，超现实主义。飞溅的水珠。



跳动的火焰，翱翔的飞鸟，闪烁的流星，飞

3、人物识别：面部&姿态



一个人笑容灿烂，一个人惊恐愤怒。
美。



一位芭蕾舞女孩在冰面上跳舞，动作华丽优美。

4、动作识别：基本&复杂



短跑运动员在快速的奔跑。
热情呐喊。



赛车手驾驶着跑车在激烈的漂移，场边的观众
热情呐喊。

5、事件检测：简单&复杂



一位帅哥在喝水。
烈。



外星文明和地球，在进行星际大战，战况激
烈。

6、环境变化：光线&天气



奇幻的极光在舞动变幻，城市写字楼的灯光忽明忽暗。



闪电交加，乌云密布，台风，大雨。

7、多对象场景：单一&多对象



1个飞行员驾驶着战斗机翱翔太空。



海军指挥着航母战斗群展开了激烈战斗。

8、特殊场景：遮挡&非标准



城市大雾遮挡了摩天大楼。

一边是神秘的海底世界，一边是梦幻的糖果屋。



9、细节检测：小物体&微表情



崭新的苹果电脑，屏幕上播放着电影。



瞳孔收缩，脸颊泛红，嘴唇微微颤抖。

测试者：mcalide 孙风波

序号	分组	提示词	视频
1	通用提示词	端午节划龙舟，吃粽子，8k，大师作品，最佳画作，丰富细节	
2		一只松鼠庆祝生日，8k，长焦镜头，小狗，丰富细节，超现实主义	
3		长焦镜头，超现实主义，8k，高清画质，阳光下，托腮，五官精致，丰富细节	

测试者：tang



提示词：一只雨林大猴子吃了一堆香蕉后，变成了歌斯拉，然后他又吃了一辆卡车。

效果：吃香蕉的动作有了，但是是吃香蕉后猴子才变成歌斯拉的，而且吃卡车的动作未实现。



提示词：一只小白兔，正在啃胡萝卜，一口就把一个胡萝卜吃掉了。

效果：小白兔啃胡萝卜的动作没问题，但一口就把一个胡萝卜吃掉了的动作未实现。



提示词：天空乌云密布，电闪雷鸣，接着就下起了狂风暴雨，突然从乌云中冲出一架客机

效果：电闪雷鸣有了，但狂风暴雨没有看到，而且飞机是直接就出现的，不是从乌云冲出的



提示词：美丽的汉服少女，年轻女子，光透过树叶形成斑驳的影子，柔光落在她的脸上。

效果：这个效果还不错。

测试者：黄桃子



公园里的欢乐时光：孩子们在秋千上欢笑的场景生动有趣，注意看地面上的光影变化处理得恰到好处。





老人的微笑瞬间：在捕捉老人神情和皮肤细节方面，“可灵”做得相当出色，岁月感也得到了很好的体现。

草地上的学车儿童：在处理稍长一些的动作时，“可灵”遇到了一些挑战，但整体而言仍具备一定的表现力。

（我输入的指令是：一个孩子在草地上学习骑自行车，摔倒后又勇敢地站起来。可能是描述动作超过了5S，反复尝试未能实现。）



舞台上的舞者：在多对象场景中，“可灵”能够保持较好的形体轮廓，但在处理远距离和密集人群时仍有局限，效果呈现上还是会有粘连感。



通过真实场景和卡通场景对比来看，“可灵”训练的重点还是在于偏真实场景的视频，可能是基于快手自己平台的长期积累，尤其是在吃东西上的表现独领风骚，但在动画方面不是强项。

测试者：Aaron W

河里的猫：一只猫站在粽子上，月光洒下来，猫的毛色黑得发亮，粽子飘在河里，倒影出此时此刻从第一个测试来看，提示词遵从能力只有60%，不过效果确实很惊艳。



简单输入一个“男人吃面”，感觉和官方的示例是同一个男人，不过这次拿的是勺子。



再试了一个“女人吃饺子”，感觉还是和其他人测的视频差不多



把Sora的提示词翻译成中文测试：一位时髦女子走在东京街头，两旁是闪着温暖光芒的霓虹灯和充满活力的城市标志。她穿着黑色皮夹克、红色长裙和黑色靴子，拿着黑色钱包。她戴着墨镜和红色口红。她自信而随意地走着。这条街很潮湿，反射着色彩斑斓的灯光，营造出一种镜像效果。许多行人走在一旁



把Sora的提示词翻译成中文测试：一个玻璃球的近景，里面有一个禅园，球里有一个小矮人正在挖禅园，在沙子里创造图案。



测试者：桔了个仔

和sora对比：

无人机拍摄的海浪冲击大苏尔加雷角海滩崎岖悬崖的景象。蔚蓝的海水激起白色的波浪，夕阳的金色光芒照亮了岩石海岸。远处有一座小岛，岛上有一座灯塔，悬崖边长满了绿色的灌木丛。从公路到海滩的陡峭落差是一项戏剧性的壮举，悬崖边缘伸出海面。这一景观捕捉到了海岸的原始之美和太平洋海岸公路的崎岖景观。



玻璃球的特写视图，里面有一个禅宗花园。球体中有一个小矮人正在耙禅宗花园并在沙子上创造图案



一位时尚女性走在充满温暖霓虹灯和动画城市标牌的东京街道上。她穿着黑色皮夹克、红色长裙和黑色靴子，拎着黑色钱包。她戴着太阳镜，涂着红色口红。她走路自信又随意。街道潮湿且反光，在彩色灯光的照射下形成镜面效果。许多行人走来走去。



3D 动画展示了一个小而圆、毛茸茸的生物，长着富有表现力的大眼睛，探索充满活力的魔法森林。这种生物是兔子和松鼠的异想天开的混合体，拥有柔软的蓝色皮毛和浓密的条纹尾巴。它沿着波光粼粼的小溪跳跃，眼睛睁得大大的，充满惊奇。森林里充满了神奇的元素：会发光和变色的花朵，叶子呈紫色和银色的树木，还有像萤火虫一样的小漂浮灯。该生物停下来与一群围绕着蘑菇环跳舞的小仙女进行有趣的互动。



其他prompt测试

一只穿着背带裤的鸡，手里拿着篮球在拍，做了转身的动作



长焦镜头，一个人站在山上，布达拉宫从他背后缓缓升起



在2077年的深圳，人们乘坐飞行器去上班，赛博朋克风格，氛围光照，最佳画质



测试者：来来

测试AI视频生成模型不同生成能力的提示词，这些提示词涵盖了不同的测试类型，帮助全面评估AI视频生成模型的各种能力。

测试类型	提示词 (Prompt)
生成质量	"A high-resolution 4K video of a sunset over the ocean with smooth transitions." "A detailed animation of a cityscape at night, showcasing neon lights and moving traffic."
生成速度	"A short 10-second clip of a running dog in a park." "A 1-minute video of a timelapse showing the changing seasons in a forest."
一致性	"A person walking from left to right across the screen with consistent motion and background." "A butterfly flying around a garden, with smooth transitions between frames."
多样性	"A fantasy world with dragons flying in the sky and wizards casting spells." "A futuristic city with flying cars and robots interacting with humans."
可控性	"Generate a video of a beach scene with a setting sun and seagulls flying, with the option to change the weather from sunny to cloudy." "Create a video of a person dancing, with options to change the dance style from hip-hop to ballet."
音频生成	"A video of a thunderstorm with synchronized sound effects of thunder and rain." "An animation of a music concert with coordinated audio and video elements."
界面友好度	"Test the user interface by generating a video of a mountain landscape with easy-to-use controls for adjusting the time of day and weather conditions."

可定制性	"Create a video of a city street with customizable options for the number of people, vehicles, and time of day."
计算资源	"Generate a 30-second video of a busy market scene and measure the CPU, GPU, and memory usage."
	"Create an animated video of a fantasy forest with detailed characters and backgrounds, then assess the resource consumption."
运行成本	"Produce a 2-minute educational video on the water cycle, evaluating the cost efficiency compared to longer videos."
错误处理	"Generate a video with an incomplete prompt to test how the model handles missing information."
	"Input a deliberately incorrect or nonsensical prompt and observe the model's response and error handling."
稳定性	"Run the model continuously to generate a series of videos over a 24-hour period to check for any stability issues."
新功能	"Test a unique feature like real-time video generation based on live inputs, such as generating a video based on a live webcam feed."
行业适应性	"Generate a marketing video for a new product launch, showcasing dynamic product features and text overlays."
	"Create an educational video for children about space exploration, including animated characters and engaging visuals."
比较分析	"Generate a video using the same prompt across different models for side-by-side performance comparison."
用户反馈	"Collect feedback from users after generating personalized birthday greeting videos to improve the quality of messages and animations."

测试者：张余



测试者: Eddy

