

可灵 AI 使用指南 - 视频生成

关于可灵

可灵（Kling）是由快手大模型团队自研打造的视频生成大模型，现已支持文生视频、图生视频、视频续写、运镜控制、首尾帧等多个能力，让用户轻松高效地完成艺术视频创作。

一、基础功能

文生视频

输入一段文字，可灵大模型根据文本表达生成5s或10s视频，将文字转变为视频画面。现已支持"标准"与"高品质"两个生成模式，标准模式生成速度更快，高品质模式画面质量更佳。同时支持16:9、9:16与1:1三种画幅比例，更多元满足大家的视频创作需求。

提示词公式

提示词 = 主体(主体描述) + 运动 + 场景(场景描述) + (镜头语言 + 光影 + 氛围)
注：括号里的内容可选填

各要素解释

- 主体**：视频中的主要表现对象，是画面主题的重要体现者。如人、动物、植物，以及物体等
- 主体描述**：对主体外貌细节和肢体姿态等的描述，可通过多个短句进行列举。如运动表现、发型发色、服饰穿搭、五官形态、肢体姿态等
- 主体运动**：对主体运动状态的描述，包括静止和运动等，运动状态不宜过于复杂，符合5s视频内可以展现的画面即可
- 场景**：场景是主体所处的环境，包括前景、背景等
- 场景描述**：对主体所处环境的细节描述，可通过多个短句进行列举，但不宜过多，符合5s视频内可以展现的画面即可。如室内场景、室外场景、自然场景等
- 镜头语言**：通过镜头的各种应用以及镜头之间的衔接和切换来传达故事或信息，并创造出特定的视觉效果和情感氛围。如超大远景拍摄，背景虚化、特写、长焦镜头拍摄、地面拍摄、顶部拍摄、航拍、景深等
- 光影**：如氛围光照、晨光、夕阳、光影、丁达尔效应、灯光等
- 氛围**：对预期视频画面的氛围描述。如热闹的场景、电影级调色、温馨美好等

示例提示词对比

基础版	增加细节描述	增加镜头语言和光影氛围
一只大熊猫在咖啡厅看书	一只大熊猫戴着黑框眼镜在咖啡厅看书，书本放在桌子上，桌子上还有一杯咖啡，冒着热气，旁边是咖啡厅的窗户	镜头中景拍摄，背景虚化，氛围光照，一只大熊猫戴着黑框眼镜在咖啡厅看书，书本放在桌子上，桌子上还有一杯咖啡，冒着热气，旁边是咖啡厅的窗户，电影级调色

使用小技巧

- 尽量使用简单词语和句子结构，避免使用过于复杂的语言
- 画面内容尽可能简单，可以在5s到10s内完成
- 用"东方意境、中国、亚洲"等词语更容易生成中国风和中国人
- 当前视频大模型对数字还不敏感，比如"10个小狗在海滩上"，数量很难保持一致
- 分屏场景，可以使用prompt："4个机位，春夏秋冬"
- 现阶段较难生成复杂的物理运动，比如球类的弹跳、高空抛物等

图生视频

输入一张图片，可灵大模型根据图片理解生成5s或10s视频，将图片转变为视频画面；也可输入一张图片加文本描述。支持"标准"与"高品质"两个生成模式，以及16:9、9:16与1:1三种画幅比例。

提示词公式

提示词 = 主体 + 运动，背景 + 运动

各要素解释

- **主体**：画面中的人物、动物、物体等主体
- **运动**：指目标主体希望实现的运动轨迹
- **背景**：画面中的背景

示例提示词对比

简单版	主体运动描述	增加背景描述
戴墨镜	蒙娜丽莎用手戴上墨镜	蒙娜丽莎用手戴上墨镜，背景出现一道光

使用小技巧

- 尽量使用简单词语和句子结构，避免使用过于复杂的语言
- 运动符合物理规律，尽量用图片中可能发生的运动描述
- 描述与图片相差较大，可能会引起镜头切换
- 现阶段较难生成复杂的物理运动，比如球类的弹跳、高空抛物等

视频延长

对AI生成后的视频可续写4~5秒，支持多次续写（最长3分钟），可通过微调提示词进行视频续写创作。

功能模式

- **自动延长**：无需输入Prompt，模型根据对视频本身的理解进行视频续写
- **自定义创意延长**：用户可以通过文本控制延长后的视频

提示词公式（自定义创意延长）

提示词 = 主体 + 运动

使用小技巧

- 视频"自定义创意延长"里Prompt需要与原视频主体保持一致，不相关的文本可能会引起镜头切换
- 延长具有一定概率，可能需要多次延长才能生成符合预期的视频

二、进阶功能

标准模式与高品质模式

模式对比示例

创意输入	标准模式	高品质模式	评价
Prompt：一只大熊猫在湖边弹吉他	高性能 16x9...1).mp4(6.35M)	高表现 16x9...1).mp4(6.29M)	标准模式：🌟🌟🌟 🌟 大熊猫动态自然，憨态可掬，视频画面色调柔和；高品质模式：🌟🌟🌟🌟🌟 吉他细节更丰富，场景想象力提高，运镜更稳定。
Prompt：一艘中世纪的帆船在海上航行，雾蒙蒙的夜晚，明亮的月光	demo_14-a...28.mp4(252.46K)	高表现 16x9...1).mp4(6.01M)	标准模式：🌟🌟🌟 画面动态自然，细节较不明显，视频画面色调柔和；高品质模式：🌟🌟🌟🌟🌟 画面细节丰富，船帆符合物理动态，有电影质感，运镜更稳定。

标准模式特点

- 视频生成速度快，推理成本更低
- 擅长生成人像、动物、以及动态幅度较大的场景
- 生成的动物更亲切，画面调色柔和

高品质模式特点

- 视频生成细节更丰富，推理成本更高
- 擅长生成人像、动物、建筑、风景类等视频
- 细节更丰富，构图与色调氛围更高级

运镜控制

支持10种运镜方式：

- 6个基本运镜：水平运镜、垂直运镜、推进/拉远、垂直摇镜、旋转摇镜、水平摇镜
- 4个大师运镜：左旋推进、右旋推进、推进上移、下移拉远

首尾帧能力

通过上传两张图片作为首帧和尾帧来生成视频，可实现对视频的更精细控制。

使用小技巧

- 尽量选择两张相同主题且近似的图，这样模型容易在5s内进行流畅衔接
- 如果两张图片相差较大，可能会触发镜头切换
- 很多创作者会通过图像生成来进行相似图片的选择，随后利用首尾帧能力进行视频生成

运动笔刷功能

上传图片后，可通过"自动选区"或"涂抹"对特定区域或主体进行选中，添加运动轨迹。同时支持"静态笔刷"功能，用于固定不需要运动的区域。

使用建议

关于图像选区：

- 单个动态笔刷只选中类别一致的单个物体【推荐】
- 单个动态笔刷只画一个相互联通的区域【推荐】
- 静态笔刷可以选中多个彼此不联通的区域，但建议每个独立选区内是同一类

关于运动轨迹：

- 轨迹曲线的方向和长度都会起作用
- 选中物体的中间运动过程会严格按照绘制的轨迹移动

对口型功能

支持在生成人物视频后，通过上传配音/歌唱文件或使用文本朗读功能，让视频人物口型和音频同步。

使用说明

1. 生成包含完整人物面容的视频
2. 使用文本朗读生成配音或上传本地配音/歌唱文件
3. 点击对口型按钮等待生成

注意事项

- 为付费功能，5s视频需5灵感值，10s视频需10灵感值
- 支持人物类角色（真实/3D/2D），动物类角色暂不支持
- 如果音频超过画面长度，提供裁剪功能