# 多模态

# 今日课程内容

1、视觉和语言的打通

2、视觉识别 与 视觉推理

3、视觉生成

# 一、视觉和语言的打通

# 视觉和语言的打通

如何打通？打通之后有什么好处？

# 视觉和语言的打通

如何打通？打通之后有什么好处？

如何打通：一个模型，能同时看懂语言和视觉，

进阶能力：能输出文字，也能输出图片、视频

好处：视觉转译、融合推理、视觉编辑

# 二、视觉识别 与 视觉推理

# 视觉识别 与 视觉推理

传统视觉识别模型 vs 多模态模型

# 视觉识别 与 视觉推理

传统视觉识别模型 vs 多模态模型

传统视觉识别模型：Yolo、UNet

Yolo：目标物体的识别

UNet：具体区域的分割

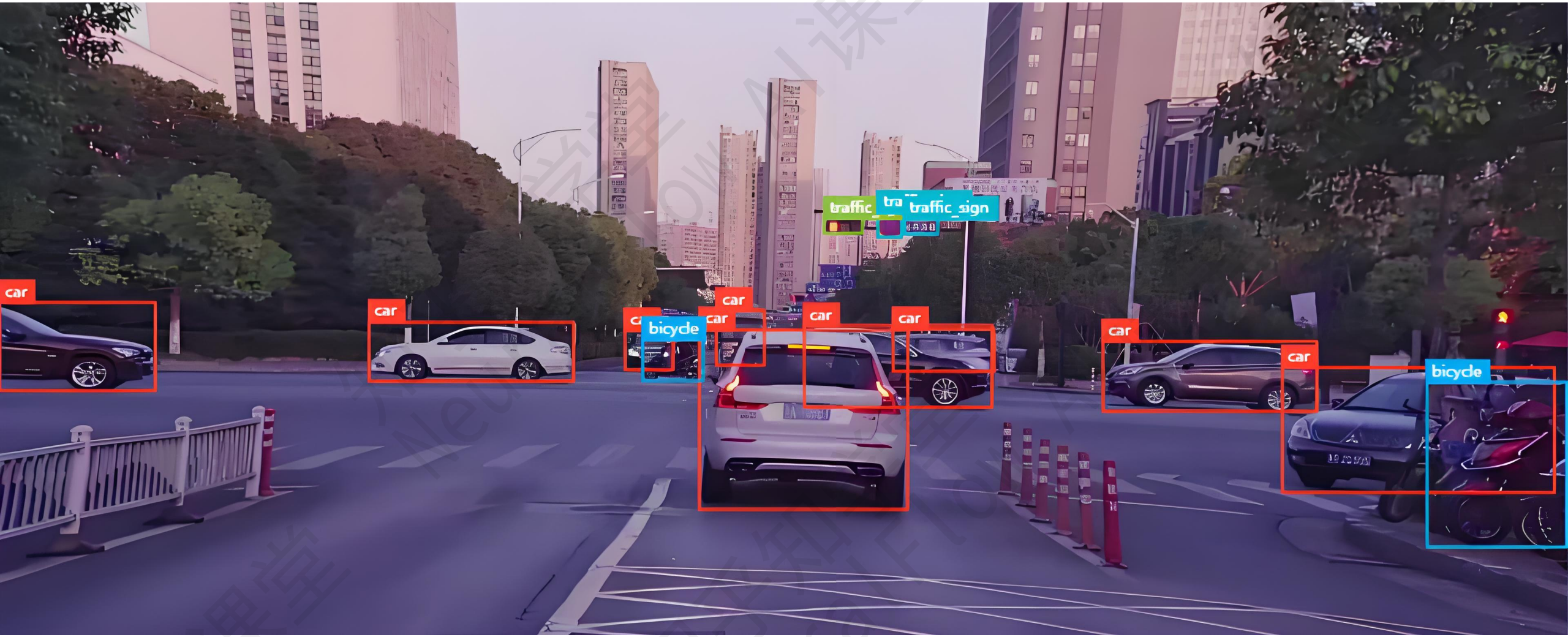
# 视觉识别 与 视觉推理

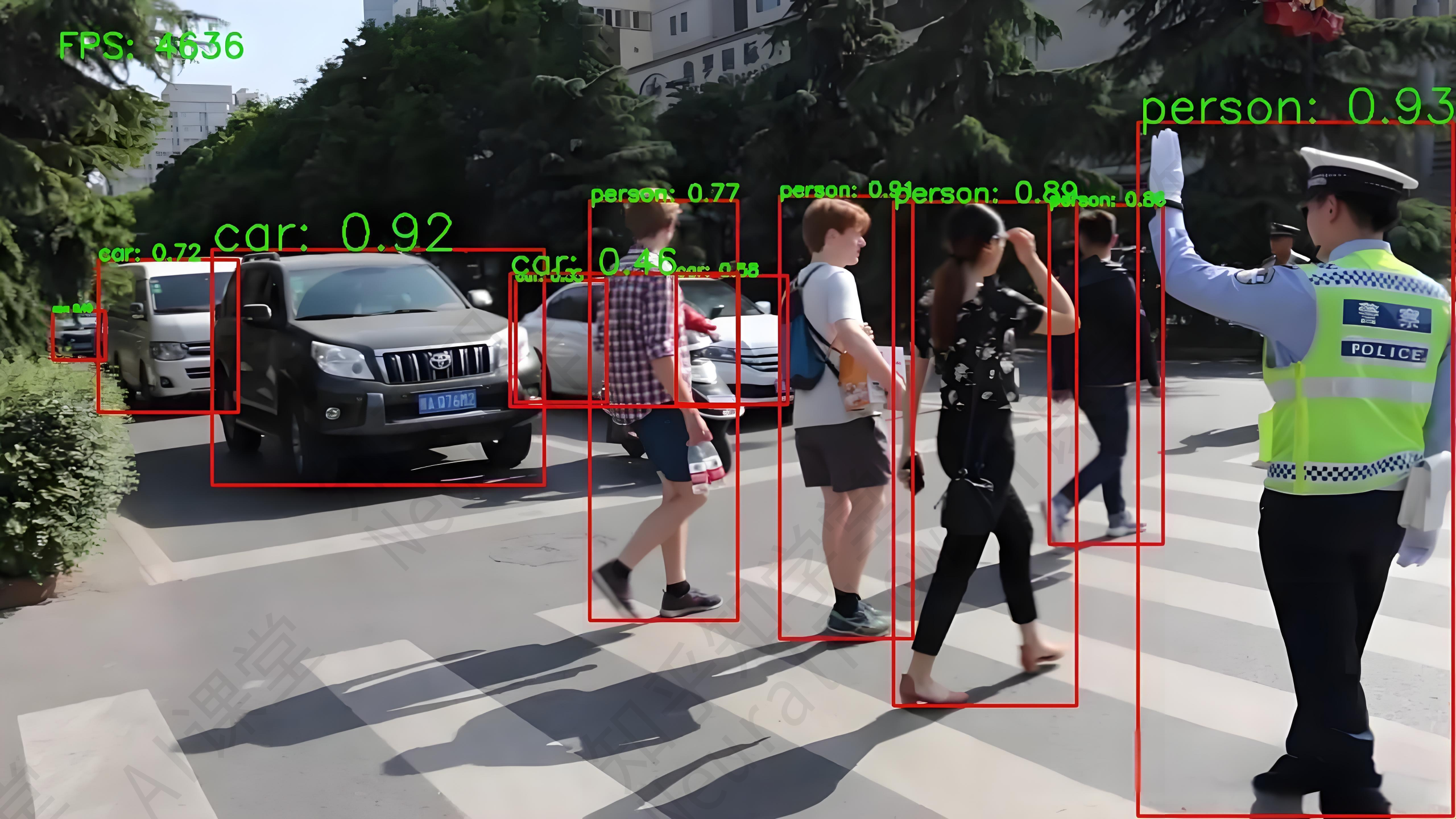
传统视觉识别模型 vs 多模态模型

传统视觉识别模型：Yolo、UNet

Yolo：目标物体的识别

UNet：具体区域的分割


car
car
car
bicycle
car
car
car
car
car
bicycle
traffic
traffic_sign


FPS: 4636
car: 0.92
car: 0.72
car: 0.46
person: 0.77
person: 0.91
person: 0.89
person: 0.93
POLICE

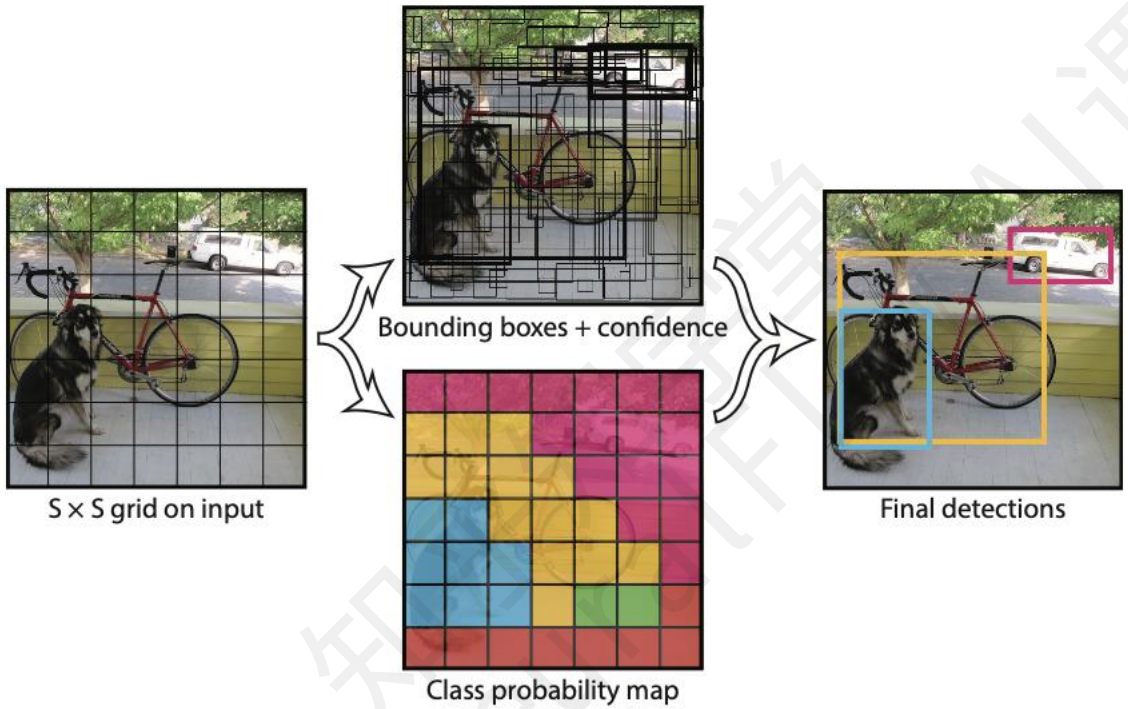



**Figure 2: The Model.** Our system models detection as a regression problem. It divides the image into an $S \times S$ grid and for each grid cell predicts $B$ bounding boxes, confidence for those boxes, and $C$ class probabilities. These predictions are encoded as an $S \times S \times (B * 5 + C)$ tensor.

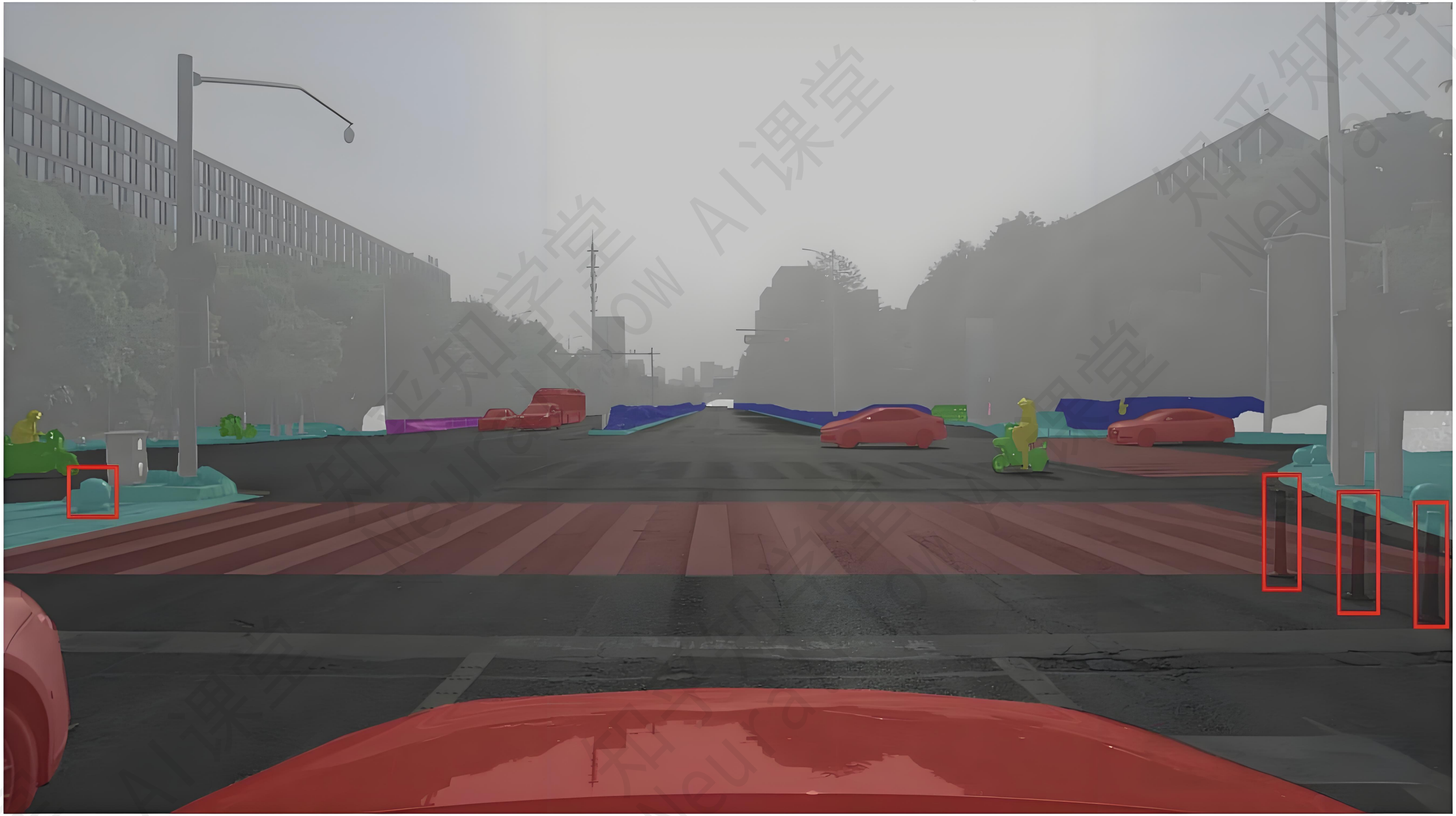
# 视觉识别 与 视觉推理

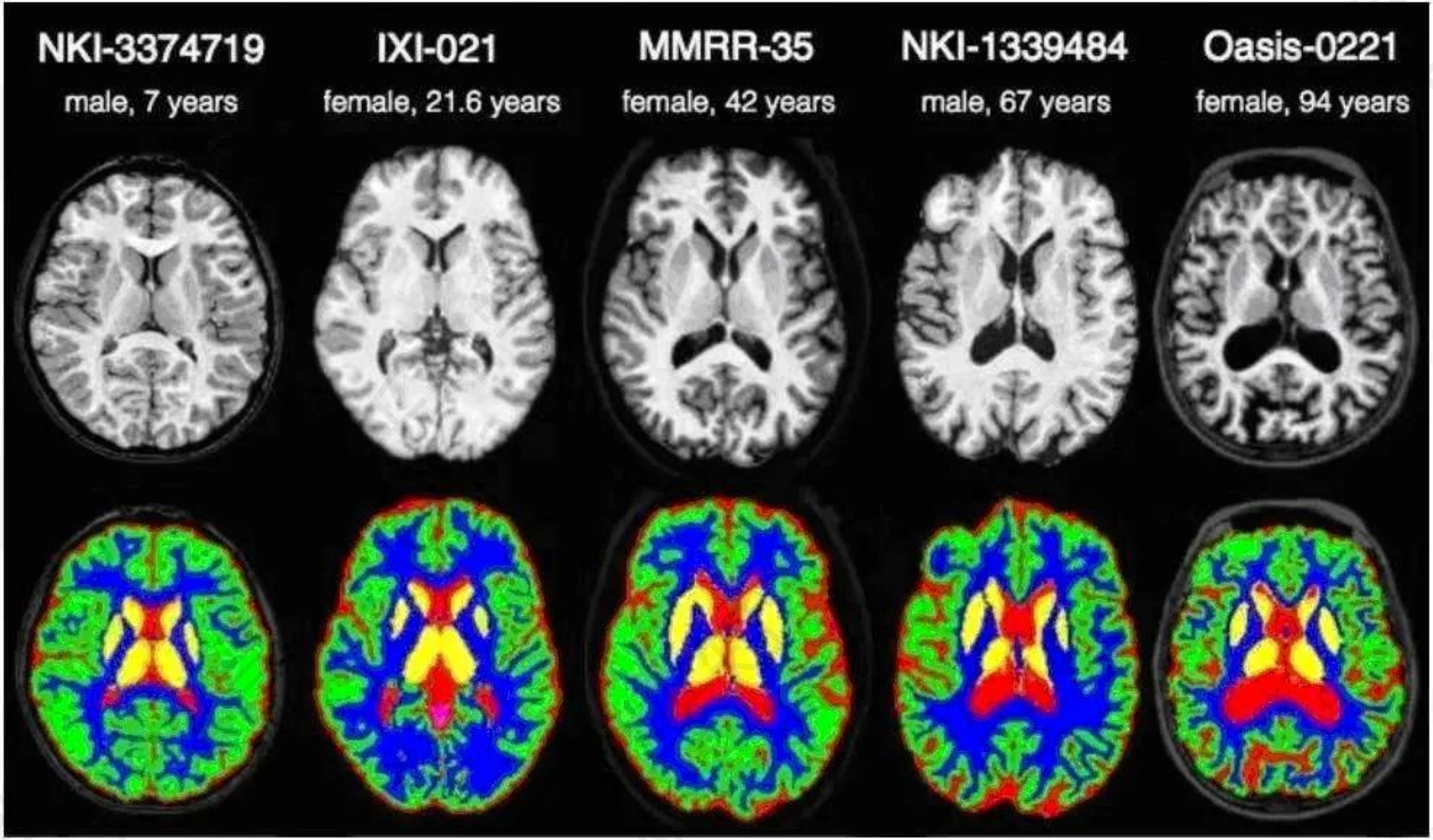
传统视觉识别模型 vs 多模态模型

传统视觉识别模型：Yolo、UNet

Yolo：目标物体的识别

UNet：具体区域的分割




NKI-3374719
male, 7 years
IXI-021
female, 21.6 years
MMRR-35
female, 42 years
NKI-1339484
male, 67 years
Oasis-0221
female, 94 years

# 视觉识别 与 视觉推理

传统视觉识别模型 vs 多模态模型

传统视觉识别模型：Yolo、UNet

Yolo：目标物体的识别

UNet：具体区域的分割

优势：模型小、部署和使用成本低、识别精度高

劣势：需要单独标注数据、训练模型

# 视觉识别 与 视觉推理

传统视觉识别模型 vs 多模态模型

传统视觉识别模型：Yolo、UNet

Yolo：目标物体的识别

UNet：具体区域的分割

优势：模型小、部署和使用成本低、识别精度高

劣势：需要单独标注数据、训练模型

多模态模型

Gemini、GPT

Qwen VL、豆包 Seed

优势：无需标注、无需训练、直接使用、有推理能力

劣势：部署和使用成本较高，精度中等

# 三、视觉生成

# 视觉生成

模型能力不足时的综合方案

# 视觉生成

模型能力不足时的综合方案

模型能力不足：只是写提示词给模型，生成的视频无法满足需求

举例

1、海报生成

2、漫剧视频

3、电商视频

# 用 AI 做海报


来盒马 吃海鲜火锅
盒马火锅
吃好蟹 上盒马
百度AI生成


12:53
泸溪河桃酥(深圳One Avenue·...
盒马鲜生
消毒纸巾
搜索
年货大街
火锅到家
厨卫百货
天天会员价
美妆个护
入群抽奖
低价拼团
今日疯抢
天天省钱
超值好价越买越省
¥9.9 ¥12.9
¥7.98 ¥9.9
尝新盒子
限定米酒礼
¥88
健康生活
澳谷高蛋白
¥79.9
冬日滋补食谱
好肉炖出满屋香
立即抢购
智利车厘子 3J
2.5kg 礼盒
3J超大果 轻咬甜蜜脆爽
月回头客8000+人
¥159/箱
588
最高588元
车厘子热卖榜 TOP1
¥0
去下单
首页
分类
会员中心
购物车
我的

# 用 AI 做海报

使用

1、豆包、即梦 等App

2、写提示词

3、反复抽卡，修改提示词

# 用 AI 做海报

使用

1、豆包、即梦 等App

2、写提示词

3、反复抽卡，修改提示词


# 用 AI 做海报

搭建

1、新的 AI 应用

2、新的工作流程

3、新的人机协作

# 视觉生成

模型能力不足时的综合方案

模型能力不足：只是写提示词给模型，生成的视频无法满足需求

举例

1、海报生成

2、漫剧视频

3、电商视频

# 电商视频生成

1

## 中长尾商品素材匮乏

针对长尾产品素材天然的缺失，有的可能只有模特图甚至是白底图，没有产品视频，素材严重匮乏

2

## 爆款商品延展性差

在某单个电商平台爆火的商品，不能很好的根据不同渠道特点进行一站式的素材定制和管理

3

## 内容需求量大制作成本高

商品内容供不应求，无法批量低成本的快速产出高质量内容以应对快节奏的市场变化

4

## 生产流程长协同效率低

视觉及内容的生产要历经几个月周期跨越多个部门，低效耗时，且实时进度跟进难

5

## 素材管理智能程度低

内容存在于各业务团队，点状零散，无法高效支撑触点与消费者的链接导致内容使用成本高，复用率低，分发效率低

6

## 优质内容的筛选和复用难

难以及时发现优质内容
内容的评判标准变更优选难度变大
优质内容的效果延续和案例传递难

7

## 效果反馈不及时

电商数据依赖人为操作各营销触点的效果反馈依赖事后收集，数据更新滞后。数据难沉淀整合，无法指导市场策略

8

## 精细化内容营销难

不具备完整的链路，针对素材类型、形式、渠道等差异、内部并无差异化的营销策略，千篇一律，增加内耗。

# FancyTech 赋能营销全流程、掌控营销全局

FancyTech

**触达渠道**

公域 私域 电商 广告 本地

**业务场景层**

需求发起 → 商品录入分析 → 脚本生成 → 外部素材调用 / 原生内容生成 → 混剪生成 → 渠道发布 → 数据回流

持续分析数据

**应用功能层**

| | | | |
|---|---|---|---|
| 持续数据采集 | 商品信息 | 营销素材 | 投放数据 |
| 持续数据整理 | 内容理解 | 智能标注 | 智能抠图 |
| | 文本分割 | 视频切片 | ..... |
| 持续数据沉淀 | 文本 | 图片 | 视频 |
| | 模板 | 动效 | ..... |

超 2 年 持续积累
分钟级 持续更新
500w+商品信息
亿级 图、文、视频素材

规模最大、信息维度最完善、迭代速度最快的营销数据&素材库

高效素材收集

**底层功能层**

| 内容理解模型 | 脚本生成模型 | 文案生成模型 | 图片生成模型 | 视频生成模型 | 工具调用模型 | 视频打分模型 |
|---|---|---|---|---|---|---|
| 自研 LLM | 商品理解 | 自研LLM | 场景重绘 | UNET&DiT | API 调用 | 机器学习 |
| 多模态理解 | 特征提取 | Lora SFT+DPO | 材质增强 | 多模态输入 | 外部工具链接 | 投放数据训练 |
| 百万级对齐数据 | LLM+NER+OCR | 百亿参数量级 | 千万级训练数据 | 极高还原度 | 自动参数构造 | 持续强化 |

效率最高、效果最好、场景适配最全的商业视频生成模型库

底层训练数据

# 营销内容

电商及零售行业

主体思路是视频片段组合

# 营销内容

电商及零售行业

主体思路是视频片段组合

1、视频由多条视频片段拼接而成

# 营销内容

电商及零售行业

主体思路是视频片段组合

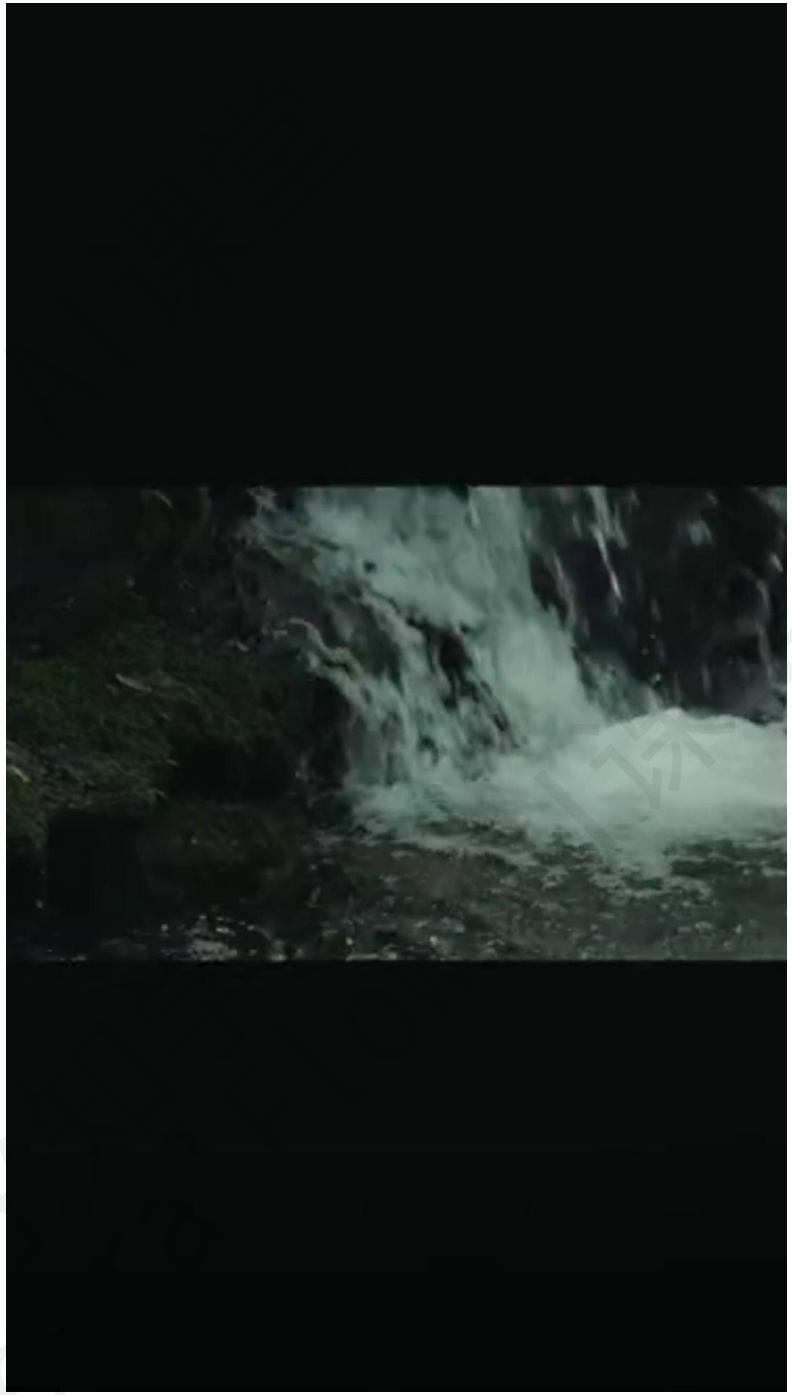
1、视频由多条视频片段拼接而成

# 营销内容

电商及零售行业

主体思路是视频片段组合

1、视频由多条视频片段拼接而成



# 营销内容

电商及零售行业

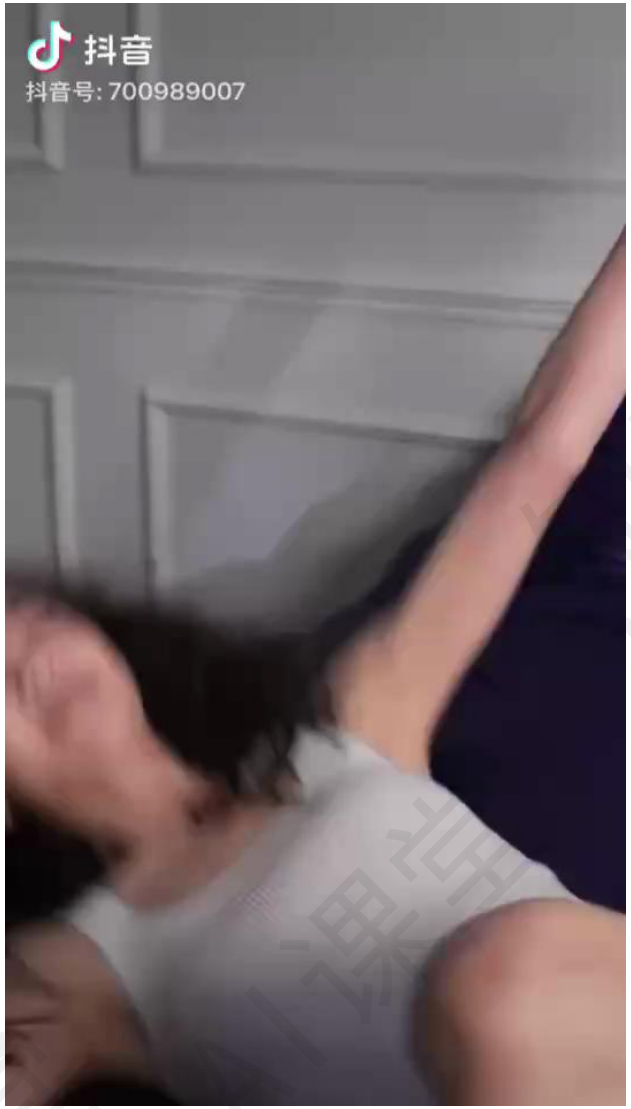
主体思路是视频片段组合

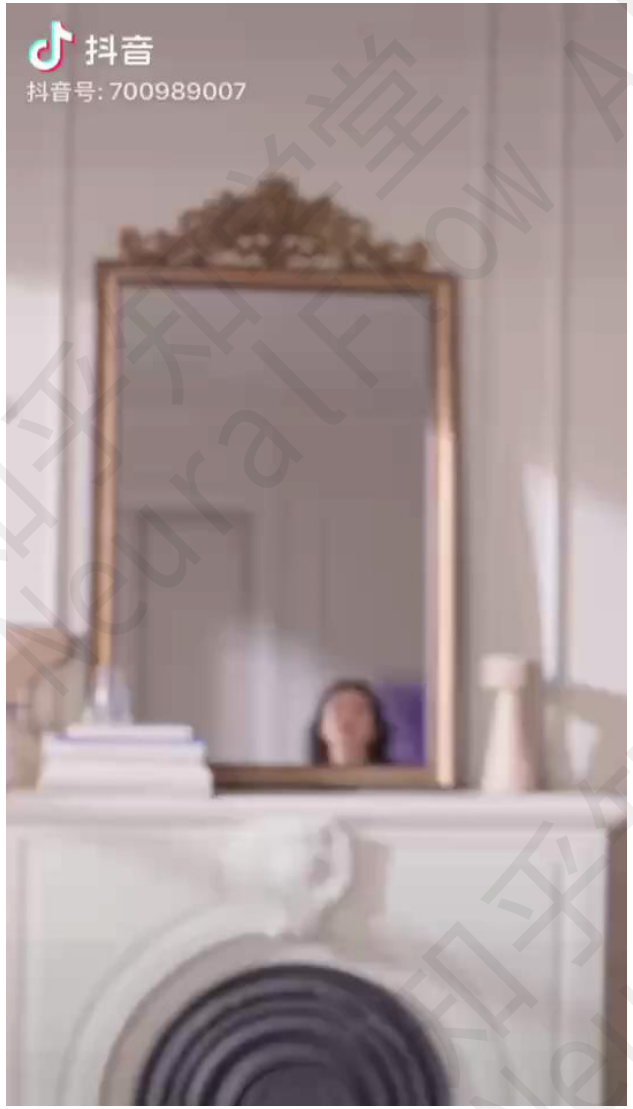
1、视频由多条视频片段拼接而成

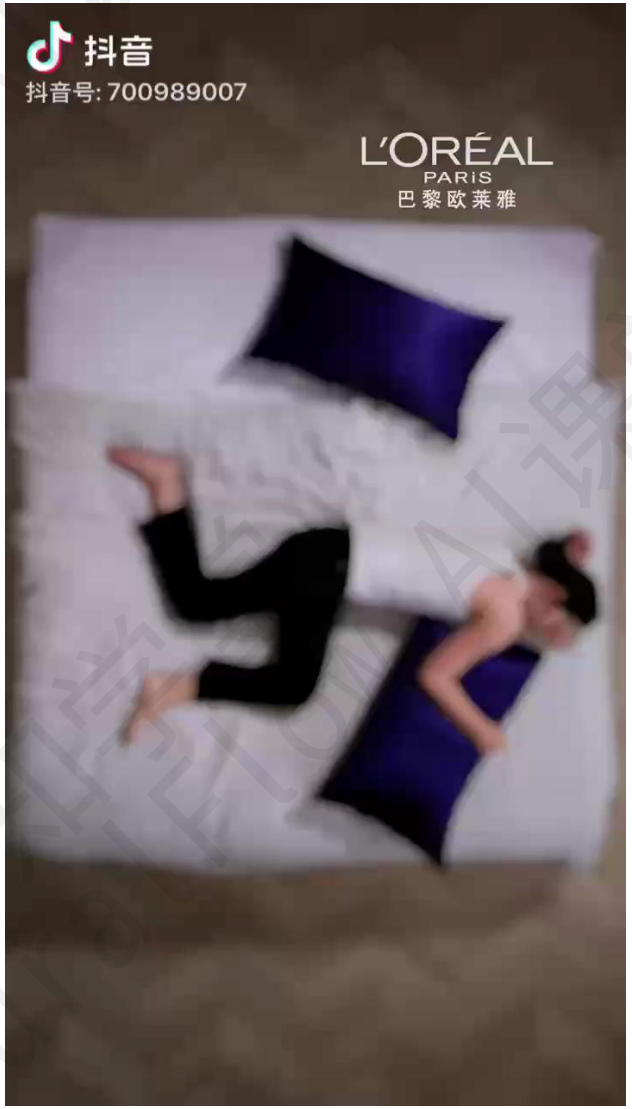
2、如何得到视频片段：品牌视频切片、产品展示切片、模特展示切片、直播切片 等

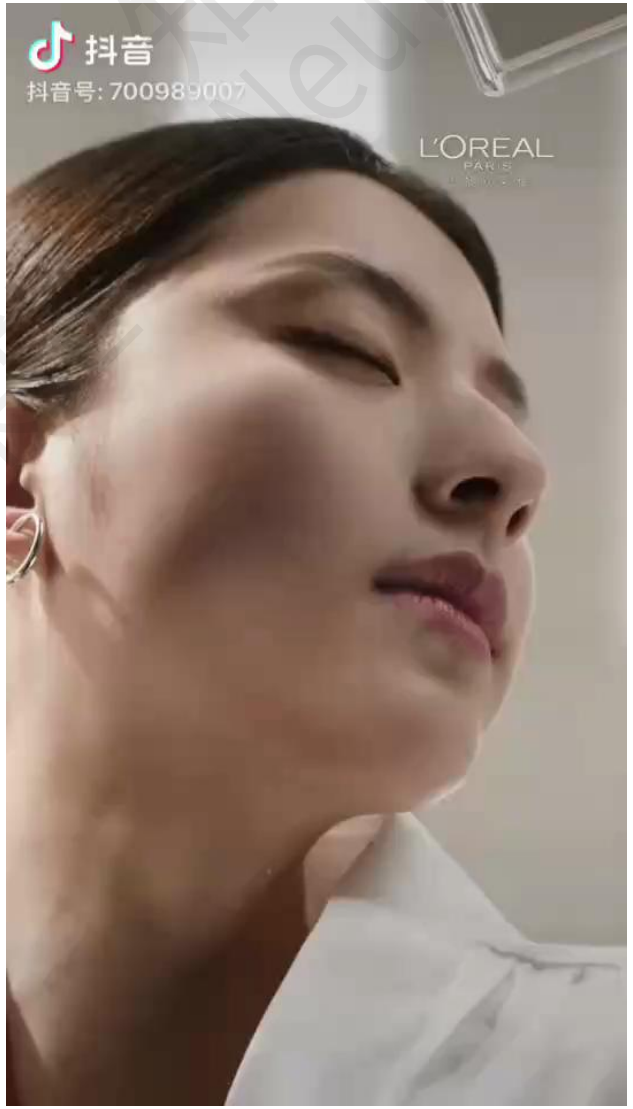
# 探索有效的切片方式

品牌视频切片、产品展示切片、模特展示切片、直播切片



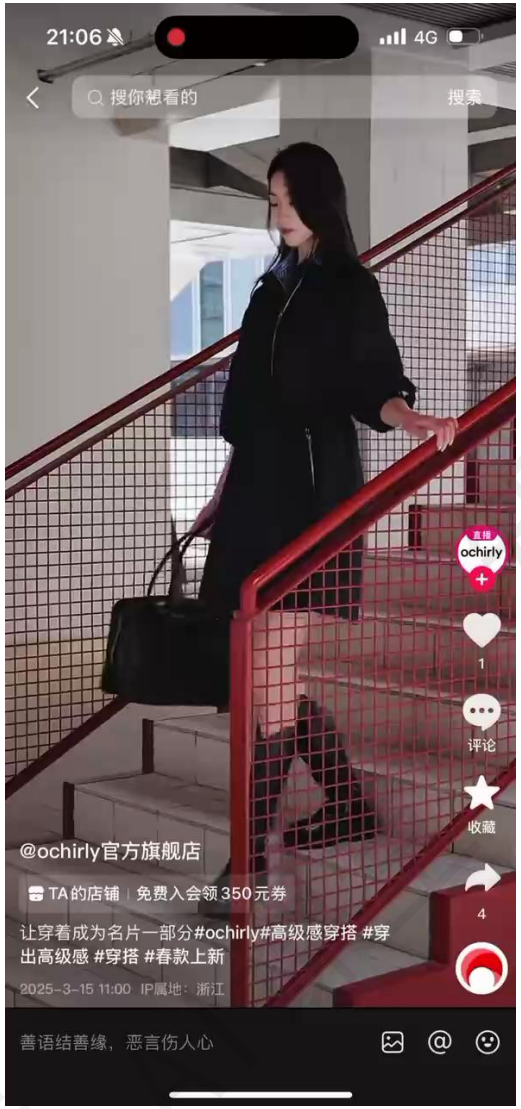
# 探索有效的切片方式

品牌视频切片、产品展示切片、模特展示切片、直播切片

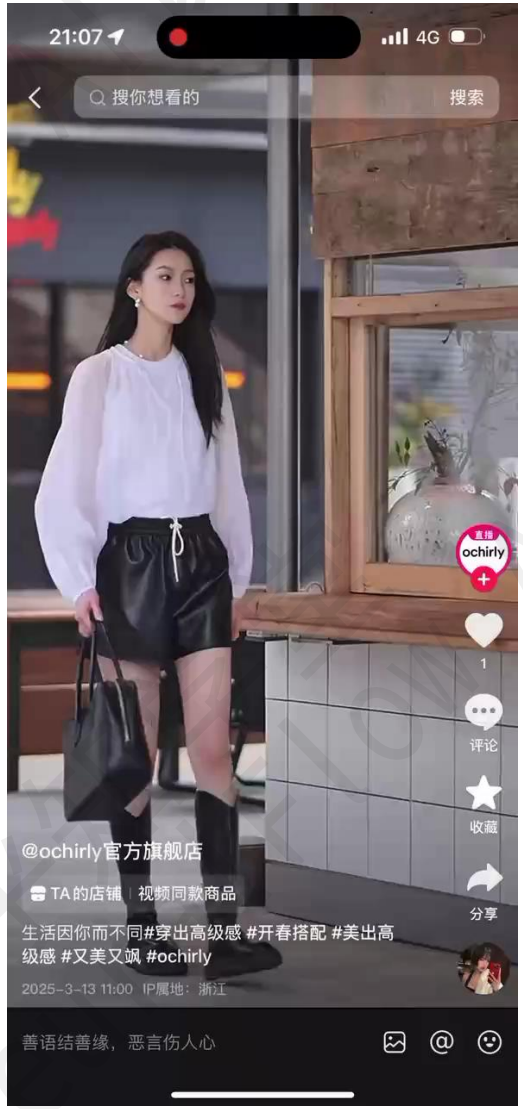


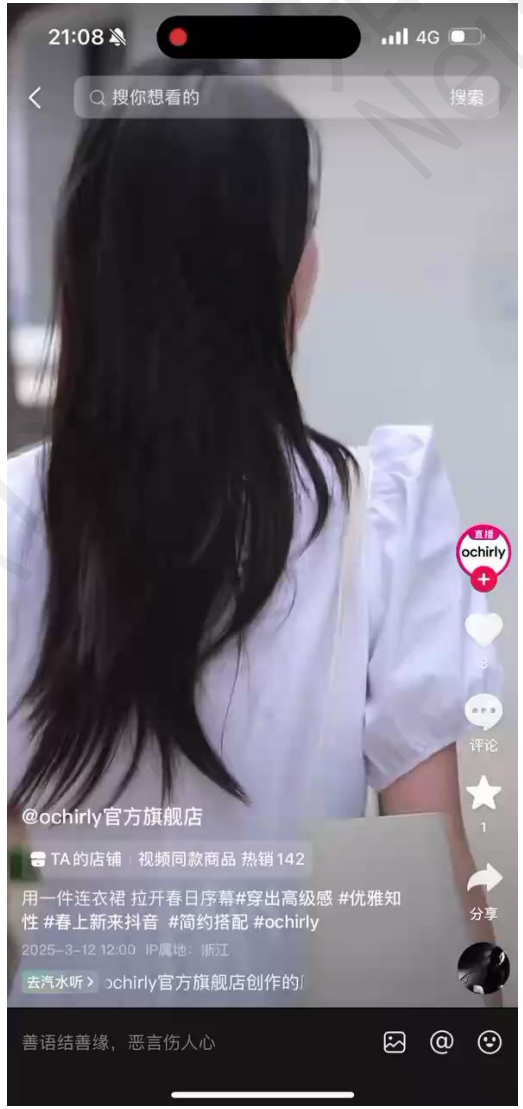

# 探索有效的切片方式

品牌视频切片、产品展示切片、模特展示切片、直播切片



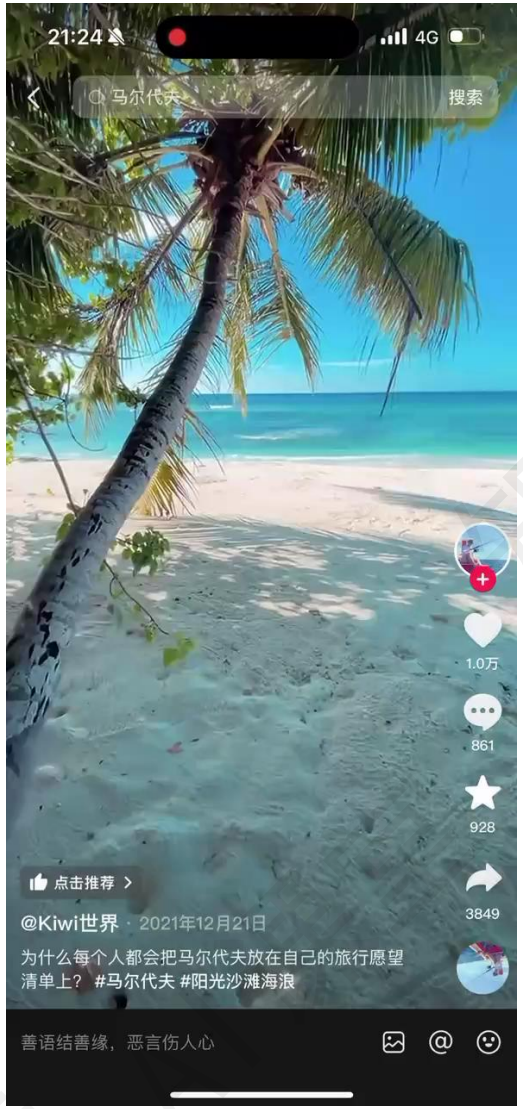




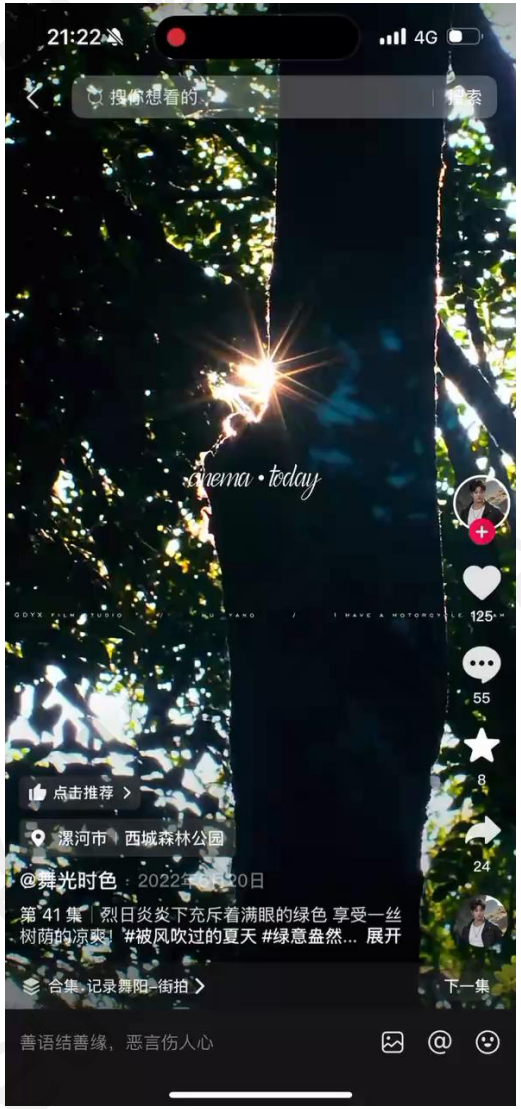



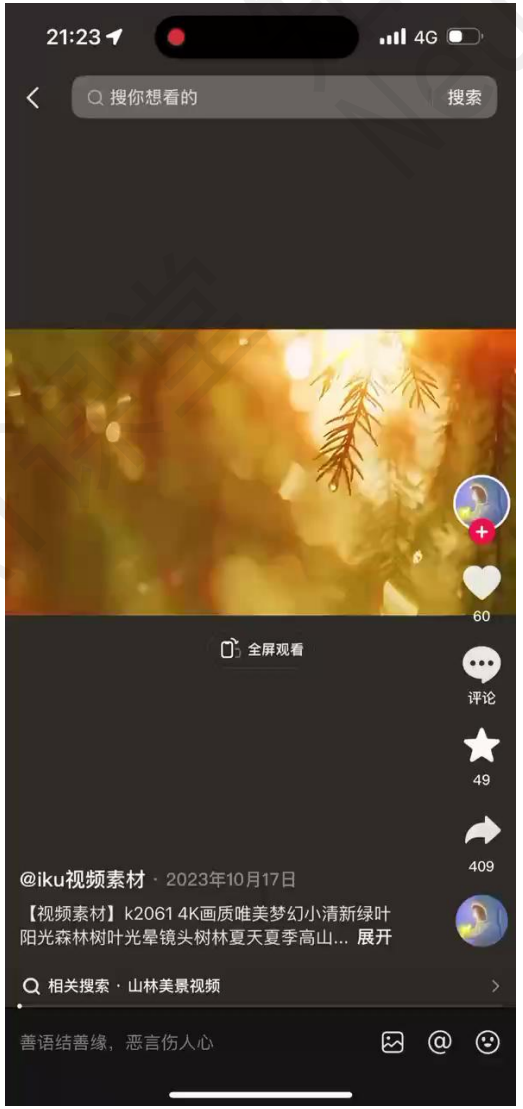
# 探索有效的切片方式

品牌视频切片、产品展示切片、模特展示切片、直播切片







# 探索有效的切片方式

品牌视频切片、产品展示切片、模特展示切片、直播切片

1、AI+人工切片，1-10秒视频

2、通过代码或工具，将视频中的音频分离

3、从音频中提取文字

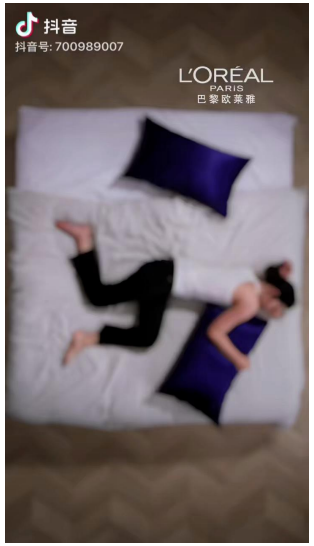
4、通过多模态模型对画面进行文字描述

5、人工补充修改文字描述

6、整理成结构化信息

# 探索有效的切片方式

品牌视频切片、产品展示切片、模特展示切片、直播切片




视频内容：痛点描述

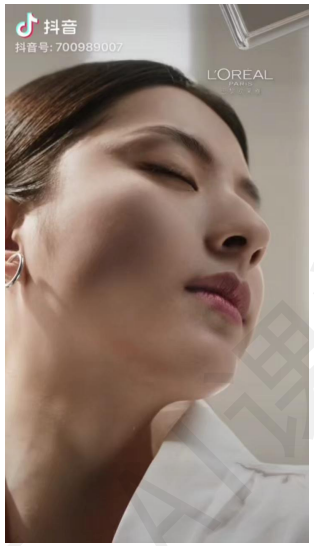
视频形式：真人演绎、特效动画

视频时长：6秒

相关产品：欧莱雅紫熨斗全脸淡纹眼霜

音频文字：这是侧睡时的你，就像4千克的哑铃在脸上挤压，压垮肌肤弹簧

视觉描述：视频开始是一个女性侧躺在床上睡觉，然后跳出一个白色4kg大哑铃，将侧睡比喻为皮肤承受了4kg哑铃的重量，最后出现动画特效，表示肌肤的弹簧会被压垮




视频内容：痛点描述

视频形式：真人演绎、特效动画

视频时长：10秒

相关产品：欧莱雅紫熨斗全脸淡纹眼霜

音频文字：肌肤压力警告，侧睡纹生成中。侧睡压力4公斤。

视觉描述：视频开头有一个女人的脸贴在了玻璃上，表现侧睡是肌肤的状态，然后出现电脑、咖啡杯、文件夹、绿植等物品，表示这些东西加起来的重量，才能到4公斤。

# 探索有效的切片方式

品牌视频切片、产品展示切片、模特展示切片、直播切片


视频内容：痛点描述

视频形式：直播切片

视频时长：10秒

相关产品：欧莱雅紫熨斗全脸淡纹眼霜

音频文字：我们其实很多宝宝会侧着睡觉对吧，侧着睡觉的话就相当于在8个小时的时间，脸上压了5公斤重左右的哑铃，所以说我们就把这只眼霜用起来。

视觉描述：男性主播，形象清秀穿白色衣服，在直播台前介绍产品


视频内容：痛点描述

视频形式：直播切片

视频时长：5.2秒

相关产品：欧莱雅紫熨斗全脸淡纹眼霜

音频文字：就会比同龄人老个5-6岁，对啊我们不能这样，我们要比同龄人年轻个5-6岁。

视觉描述：男性主播，形象清秀穿白色衣服，在直播台前介绍产品

# 探索有效的切片方式

品牌视频切片、产品展示切片、模特展示切片、直播切片


视频内容：品牌介绍

视频形式：直播切片

视频时长：8秒

相关产品：欧莱雅紫熨斗全脸淡纹眼霜

音频文字：欧莱雅做波心做了20多年，欧莱雅做波心没有平替的。抗皱的、淡纹的、提拉紧致补水保湿修护肌肤的。

视觉描述：男性主播，形象清秀穿白色衣服，在直播台前介绍产品


视频内容：产品介绍

视频形式：直播切片

视频时长：6.2秒

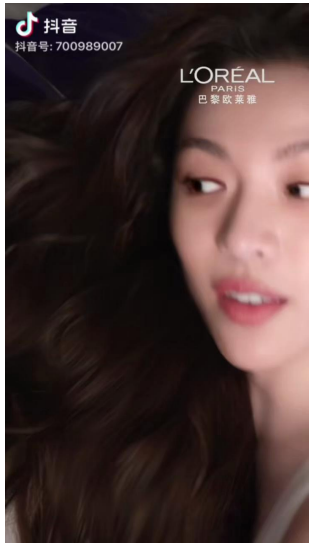
相关产品：欧莱雅紫熨斗全脸淡纹眼霜

音频文字：全新升级的第三代紫熨斗眼霜，它是一觉淡褪侧睡纹，一支全脸淡纹的眼霜。

视觉描述：男性主播，形象清秀穿白色衣服，在直播台前介绍产品

# 探索有效的切片方式

品牌视频切片、产品展示切片、模特展示切片、直播切片




视频内容：痛点描述

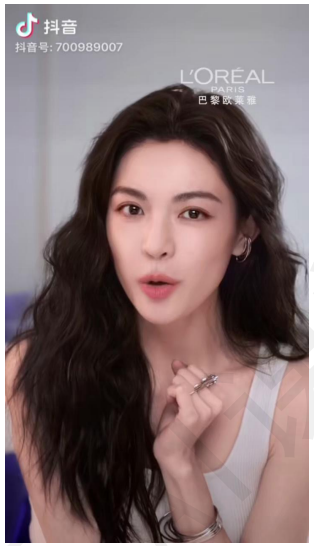
视频形式：真人演绎、特效动画

视频时长：3.4秒

相关产品：欧莱雅紫熨斗全脸淡纹眼霜

音频文字：知道么，睡觉也是会压出细纹的哦

视觉描述：品牌代言人钟楚曦，长发女性，侧躺在沙发，提出睡觉也会压出细纹的观点




视频内容：痛点描述

视频形式：真人演绎、特效动画

视频时长：6.5秒

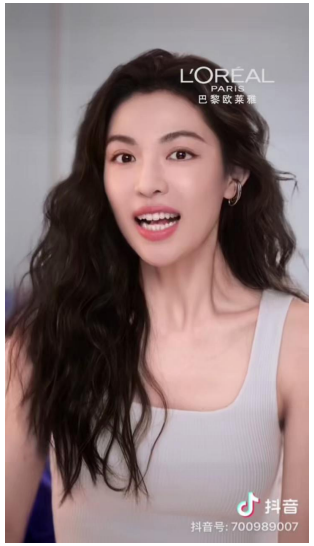
相关产品：欧莱雅紫熨斗全脸淡纹眼霜

音频文字：我们的肌肤是有许许多多的小弹簧的，一支像这样子挤压着它们的话，肌肤会失去弹性

视觉描述：品牌代言人钟楚曦，传白色吊带背心，卷发，面对镜头进行表述，穿插动画特效

# 探索有效的切片方式

品牌视频切片、产品展示切片、模特展示切片、直播切片




视频内容：代言人展示产品

视频形式：真人演绎、特效动画

视频时长：3.5秒

相关产品：欧莱雅紫熨斗全脸淡纹眼霜

音频文字：然后这一支呢，现在又、又、又升级啦

视觉描述：品牌代言人钟楚曦，传白色吊带背心，卷发，面对镜头进行表述，穿插动画特效


视频内容：产品特性展示

视频形式：真人演绎、特效动画

视频时长：5.7秒

相关产品：欧莱雅紫熨斗全脸淡纹眼霜

音频文字：多添加了一个功效放大器，能够把肌肤里的小弹簧啊，通通都撑起来

视觉描述：品牌代言人钟楚曦，传白色吊带背心，卷发，面对镜头进行表述，穿插动画特效

# 营销内容

电商及零售行业

主体思路是视频片段组合

1、视频由多条视频片段拼接而成

2、如何得到视频片段：品牌视频切片、产品展示切片、模特展示切片、直播切片 等

# 营销内容

电商及零售行业

主体思路是视频片段组合

1、视频由多条视频片段拼接而成

2、如何得到视频片段：品牌视频切片、产品展示切片、模特展示切片、直播切片 等

3、如果每条视频片段都有足够丰富的文字描述，则 LLM 应该有能力去组合视频片段

# 没有用模特拍过视频的商品怎么办？

大量商家95%以上的商品没有拍过模特展示视频

# 没有用模特拍过视频的商品怎么办？

大量商家95%以上的商品没有拍过模特展示视频

1、文生视频，可控性普遍还不够

# 没有用模特拍过视频的商品怎么办？

大量商家95%以上的商品没有拍过模特展示视频

1、文生视频，可控性普遍还不够

2、图生视频，生成几秒钟的展示视频，效果是可用的

# 没有用模特拍过视频的商品怎么办？

大量商家95%以上的商品没有拍过模特展示视频

1、文生视频，可控性普遍还不够

2、图生视频，生成几秒钟的展示视频，效果是可用的

# 没有用模特拍过视频的商品怎么办？

大量商家95%以上的商品没有拍过模特展示视频

1、文生视频，可控性普遍还不够

2、图生视频，生成几秒钟的展示视频，效果是可用的

# 没有用模特拍过视频的商品怎么办？

大量商家95%以上的商品没有拍过模特展示视频

1、文生视频，可控性普遍还不够

2、图生视频，生成几秒钟的展示视频，效果是可用的

3、你的商品如何上身到模特？

# 没有用模特拍过视频的商品怎么办?

大量商家95%以上的商品没有拍过模特展示视频

1、文生视频，可控性普遍还不够

2、图生视频，生成几秒钟的展示视频，效果是可用的

3、你的商品如何上身到模特?

4、Flux 是很强的开源生图模型，而且有很多基于Flux 或 Stable Diffusion 的变体模型

# 没有用模特拍过视频的商品怎么办？

大量商家95%以上的商品没有拍过模特展示视频

基于以上前提，我们再来梳理一下从电商商家出发，这样的 AI 产品都需要哪些步骤

# 没有用模特拍过视频的商品怎么办？

大量商家95%以上的商品没有拍过模特展示视频

基于以上前提，我们再来梳理一下从电商商家出发，这样的 AI 产品都需要哪些步骤

1、商家都有什么？商家都缺什么？

# 没有用模特拍过视频的商品怎么办？

大量商家95%以上的商品没有拍过模特展示视频

基于以上前提，我们再来梳理一下从电商商家出发，这样的 AI 产品都需要哪些步骤

1、商家都有什么？商家都缺什么？

# 没有用模特拍过视频的商品怎么办？

大量商家95%以上的商品没有拍过模特展示视频

基于以上前提，我们再来梳理一下从电商商家出发，这样的 AI 产品都需要哪些步骤

1、商家都有什么？商家都缺什么？

商家有很多商品图片，商家缺版权模特

# 没有用模特拍过视频的商品怎么办？

大量商家95%以上的商品没有拍过模特展示视频

基于以上前提，我们再来梳理一下从电商商家出发，这样的 AI 产品都需要哪些步骤

1、商家都有什么？商家都缺什么？

商家有很多商品图片，商家缺版权模特

解决方案：用 Flux 模型帮助商家生成足够多的模特，且没有版权问题

# 没有用模特拍过视频的商品怎么办？

大量商家95%以上的商品没有拍过模特展示视频

基于以上前提，我们再来梳理一下从电商商家出发，这样的 AI 产品都需要哪些步骤

1、商家都有什么？商家都缺什么？

商家有很多商品图片，商家缺版权模特

解决方案：用 Flux 模型帮助商家生成足够多的模特，且没有版权问题





# 没有用模特拍过视频的商品怎么办？

大量商家95%以上的商品没有拍过模特展示视频

基于以上前提，我们再来梳理一下从电商商家出发，这样的 AI 产品都需要哪些步骤

1、商家都有什么？商家都缺什么？

商家有很多商品图片，商家缺版权模特

解决方案：用 Flux 模型帮助商家生成足够多的模特，且没有版权问题

2、需要一个模型，将服饰传到模特身上

# 没有用模特拍过视频的商品怎么办？

大量商家95%以上的商品没有拍过模特展示视频

基于以上前提，我们再来梳理一下从电商商家出发，这样的 AI 产品都需要哪些步骤

1、商家都有什么？商家都缺什么？

商家有很多商品图片，商家缺版权模特

解决方案：用 Flux 模型帮助商家生成足够多的模特，且没有版权问题

2、需要一个模型，将服饰传到模特身上

解决方案：调研大量开源模型、闭源模型API，找到合适的换装模型，这里我们使用CatVTON


知乎知学堂
NeuralFlow AI课堂


知乎知学堂 NeuralFlow AI课堂

# 没有用模特拍过视频的商品怎么办？

大量商家95%以上的商品没有拍过模特展示视频

基于以上前提，我们再来梳理一下从电商商家出发，这样的 AI 产品都需要哪些步骤

1、商家都有什么？商家都缺什么？

商家有很多商品图片，商家缺版权模特

解决方案：用 Flux 模型帮助商家生成足够多的模特，且没有版权问题

2、需要一个模型，将服饰传到模特身上

解决方案：调研大量开源模型、闭源模型API，找到合适的换装模型，这里我们使用CatVTON

如果测试的足够多，就能找到这类模型的特点

# 没有用模特拍过视频的商品怎么办？

商品上身模特：服装、帽子、鞋子、包、配饰、首饰、手表等；

梳理和标注所有的商品图，模特图

# 没有用模特拍过视频的商品怎么办？

商品上身模特：服装、帽子、鞋子、包、配饰、首饰、手表等

梳理和标注所有的商品图，模特图

1、上衣内容类别：连衣裙、T恤、衬衫、外套、针织衫、风衣、西服、卫衣、马夹、大衣、皮衣、皮草、毛衣、羽绒服等等

2、上衣款式：贴身款、修身款、合身款、宽松款、超宽松款

3、上衣长度：超短款、短款、常规款、中长款、长款、超长款

4、上衣袖子：长袖、半袖、短袖、无袖

5、下衣内容类别：连衣裙、半身裙、休闲裤、牛仔裤、短裤、直筒裤、工装裤、西裤、运动裤等等

6、下衣款式：紧身款、修身款、合身款、宽松款、超宽松款

7、下衣长度：拖地、长、7分、膝盖、短、超短

# 没有用模特拍过视频的商品怎么办？

商品上身模特：服装、帽子、鞋子、包、配饰、首饰、手表等

整理一套适用于该模型的规则，模特与商品的可用关系

1、长配长、短配短

2、外套配外套、裙子配裙子

3、修身配修身、肥大配肥大

4、等等其他有效规则

# 没有用模特拍过视频的商品怎么办？

大量商家95%以上的商品没有拍过模特展示视频

基于以上前提，我们再来梳理一下从电商商家出发，这样的 AI 产品都需要哪些步骤

1、商家都有什么？商家都缺什么？

商家有很多商品图片，商家缺版权模特

解决方案：用 Flux 模型帮助商家生成足够多的模特，且没有版权问题

2、需要一个模型，将服饰传到模特身上

解决方案：调研大量开源模型、闭源模型API，找到合适的换装模型，这里我们使用CatVTON

如果测试的足够多，就能找到这类模型的特点

3、给图片增加合适的场景

## 替换背景

上传图片

输出图片

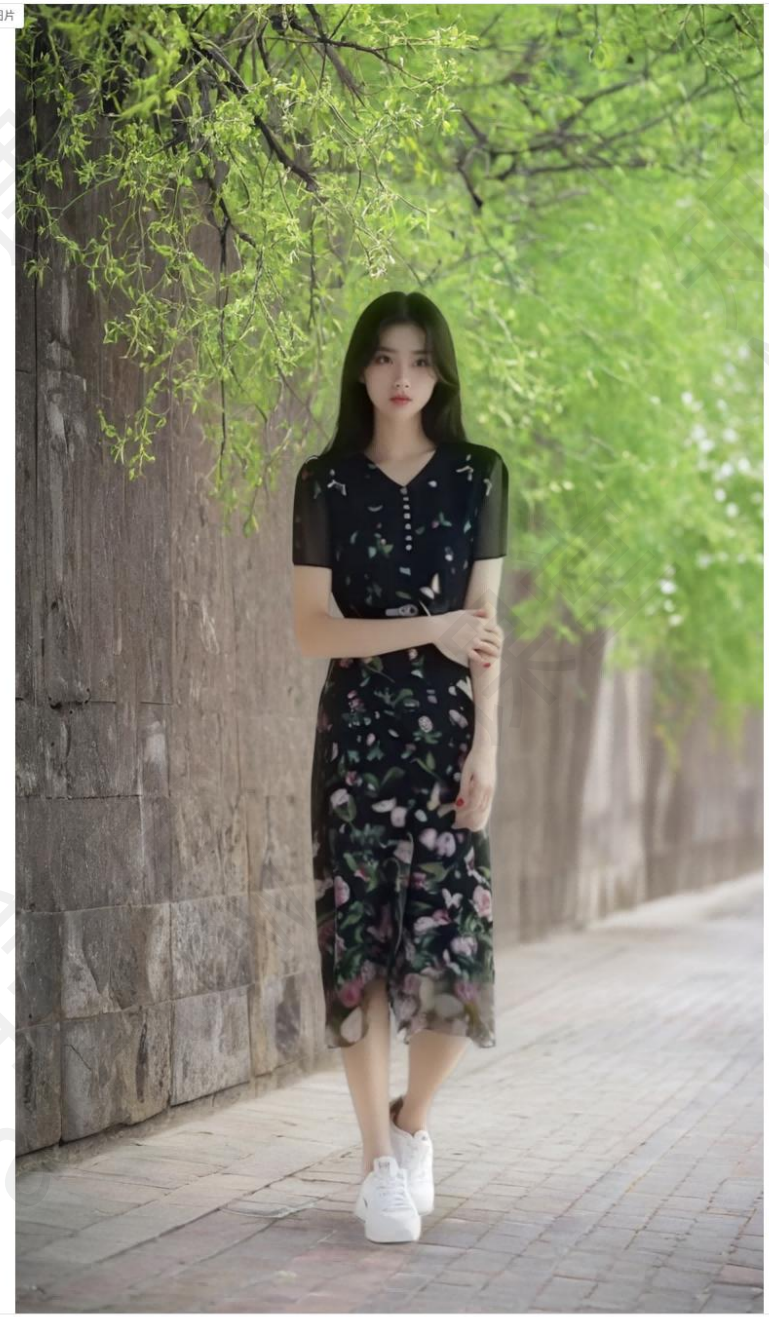
输入新背景提示词

The Bund in Shanghai, with the sunset by the river.

## 替换背景

上传图片

输出图片



输入新背景提示词

In Beijing, outside the walls of the Forbidden City during spring, with green trees and blooming flowers.

## 替换背景

上传图片

输出图片

输入新背景提示词

At the seafood market, there are many stalls selling fish, shrimp, and shellfish.

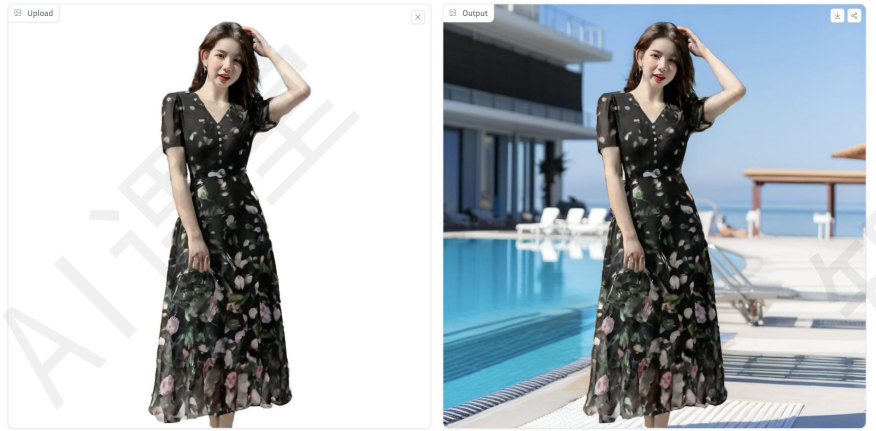
# 营销内容生成

给商品图、模特图换场景

不同类型的商品和模特适合配不同的场景

1、正式商务装、休闲运动装、日常休闲装、晚礼服、度假装、复古风、街头潮流装、夏季轻装、婚纱礼服、秋冬款

2、办公室环境、城市商务街区、健身房/运动场、公园/户外、咖啡店/餐厅、街头/城市街区、宴会厅/高端酒店、豪华酒吧/夜店、海滩/度假胜地、热带植物园、古老咖啡馆/复古街区、博物馆/艺术馆、工业区/仓库、雪地/山区





# 营销内容生成

给商品图、模特图换场景

| 女装类型 | 推荐拍摄场景 | 示例 |
| --- | --- | --- |
| 正式商务装 | 办公室环境、城市商务街区 | 西装裙在会议室内展示自信，或在都市街头走动 |
| 休闲运动装 | 健身房/运动场、公园/户外 | 穿运动装在跑步机上、户外晨跑或做瑜伽练习 |
| 日常休闲装 | 咖啡店/餐厅、街头/城市街区 | 在咖啡厅喝咖啡，或穿牛仔裤走在城市街头 |
| 晚礼服/派对装 | 宴会厅/高端酒店、豪华酒吧/夜店 | 身着晚礼服在宴会厅跳舞，或在酒店大堂走动 |
| 度假装 | 海滩/度假胜地、热带植物园 | 穿沙滩裙在海滩晒太阳，或在泳池旁享受阳光 |
| 复古风 | 古老咖啡馆/复古街区、博物馆/艺术馆 | 复古裙子在街头漫步，或在博物馆展示经典时尚 |
| 街头潮流装 | 都市街头/街头艺术墙、工业区/仓库 | 在涂鸦墙前拍摄，或在废弃仓库中展示街头风格 |
| 夏季轻装 | 花园/露台、海滨度假村 | 穿吊带裙在花园中漫步，或在度假村享受阳光 |
| 婚纱/新娘礼服 | 教堂/婚礼现场、花园/庄园 | 在教堂走道上穿婚纱，或在花园内与爱人共度甜蜜时光 |
| 秋冬款 | 雪地/山区、温馨的家庭环境 | 穿羽绒服在雪地中漫步，或在室内火炉旁享受温暖 |

# 没有用模特拍过视频的商品怎么办？

大量商家95%以上的商品没有拍过模特展示视频

基于以上前提，我们再来梳理一下从电商商家出发，这样的 AI 产品都需要哪些步骤

1、商家都有什么？商家都缺什么？

商家有很多商品图片，商家缺版权模特

解决方案：用 Flux 模型帮助商家生成足够多的模特，且没有版权问题

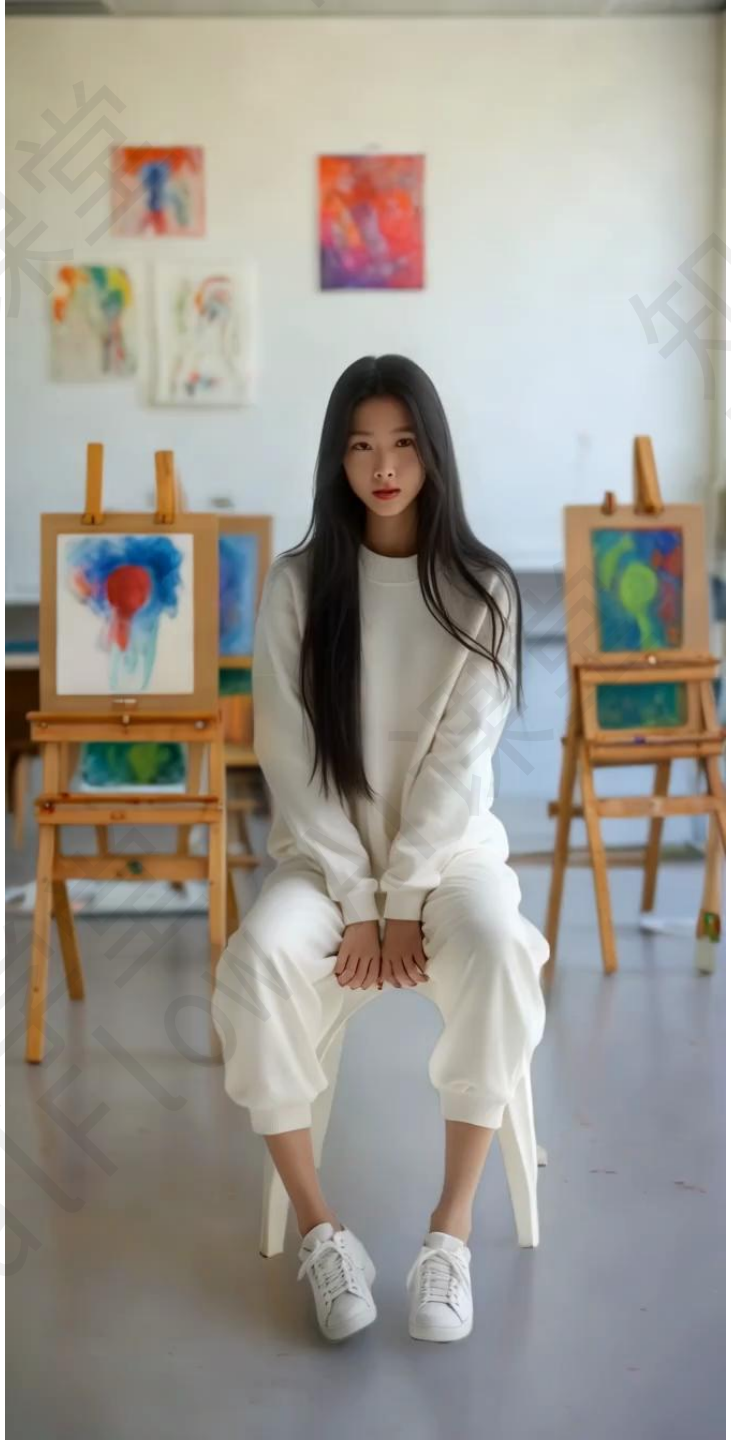
2、需要一个模型，将服饰传到模特身上

解决方案：调研大量开源模型、闭源模型API，找到合适的换装模型，这里我们使用CatVTON

如果测试的足够多，就能找到这类模型的特点

3、给图片增加合适的场景

4、让商品图、模特图动起来（海螺）

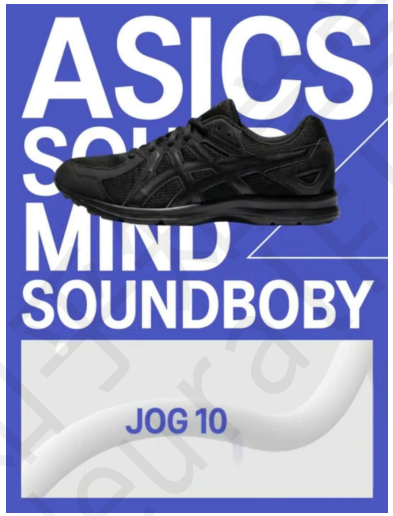


# 营销内容生成

生成视频片段：让商品图、模特图动起来

套模版




# 没有用模特拍过视频的商品怎么办？

大量商家95%以上的商品没有拍过模特展示视频

基于以上前提，我们再来梳理一下从电商商家出发，这样的 AI 产品都需要哪些步骤

1、商家都有什么？商家都缺什么？

商家有很多商品图片，商家缺版权模特

解决方案：用 Flux 模型帮助商家生成足够多的模特，且没有版权问题

2、需要一个模型，将服饰传到模特身上

解决方案：调研大量开源模型、闭源模型API，找到合适的换装模型，这里我们使用CatVTON

如果测试的足够多，就能找到这类模型的特点

3、给图片增加合适的场景

4、让商品图、模特图动起来（海螺）

5、给文案配音，并且配上音乐

# 营销内容生成

混剪的最终效果

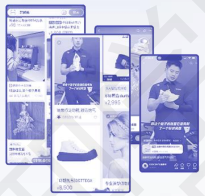

# 千人千面内容制作内容精准触达

FancyTech

更好的生成效果带来更多用户和数据

镜头 1: .....
镜头 2: .....
镜头 3: .....

脚本设计



外部素材库匹配

Ptompt Output

自研模型内容生成

完全自动化

无需客户投入精力

投放数据回流

更多数据持续优化更匹配商品特点、热点趋势、用户偏好的技术方案

每日数据回流

实现T+1反馈、T+1优化生产

更高的播放量

视频内容播放量提升30%

更好的点击率

视频内容平均点击率提升50%