

视频前处理算法探索与实践

腾讯云高级工程师 李峰



视频前处理场景探索

虚拟背景解决方案的探索与实践

美颜解决方案的探索与实践



视频前处理场景探索

关注「腾讯产业互联网」公众号

回复「加群」进入技术交流群



云视频会议

依靠云计算、虚拟化部署、新一代音视频算法等技术，基于通用网络提供高质量音视频服务的一种在线会议解决方案。腾讯会议是一款云视频会议产品。

视频前处理

视频是连续的，在转播时需要经过编码和解码的流程，所以视频处理需要分为前处理和后处理。所谓前处理是指编码前的视频处理，比如虚拟背景；所谓后处理是指解码后的视频处理，比如视频超分。

云视频会议场景下的前处理探索

虚拟背景：可以很好的保护用户隐私，创造一个公平的环境；

美颜：鼓励大家参与到视频通话的场景中来；

视频降噪：可以降低摄像头噪声、消除灯光造成的闪烁等问题，进而提升视频画面质量；

暗场景增强：可以提升暗光场景下的视频体验。

非常感谢腾讯工程师为腾讯会议开发的背景虚化功能，为许多孩子取消的歧视，让家庭背景，经济条件不再是孩子的负担，非常感谢！

虚拟背景解决方案——目标

虚拟背景功能允许用户在使用腾讯会议期间上传自定义的图片或者视频作为视频场景下的虚拟背景，或者将视频背景模糊掉，满足用户保护隐私和个性化视频的需求。

关注「腾讯产业互联网」公众号

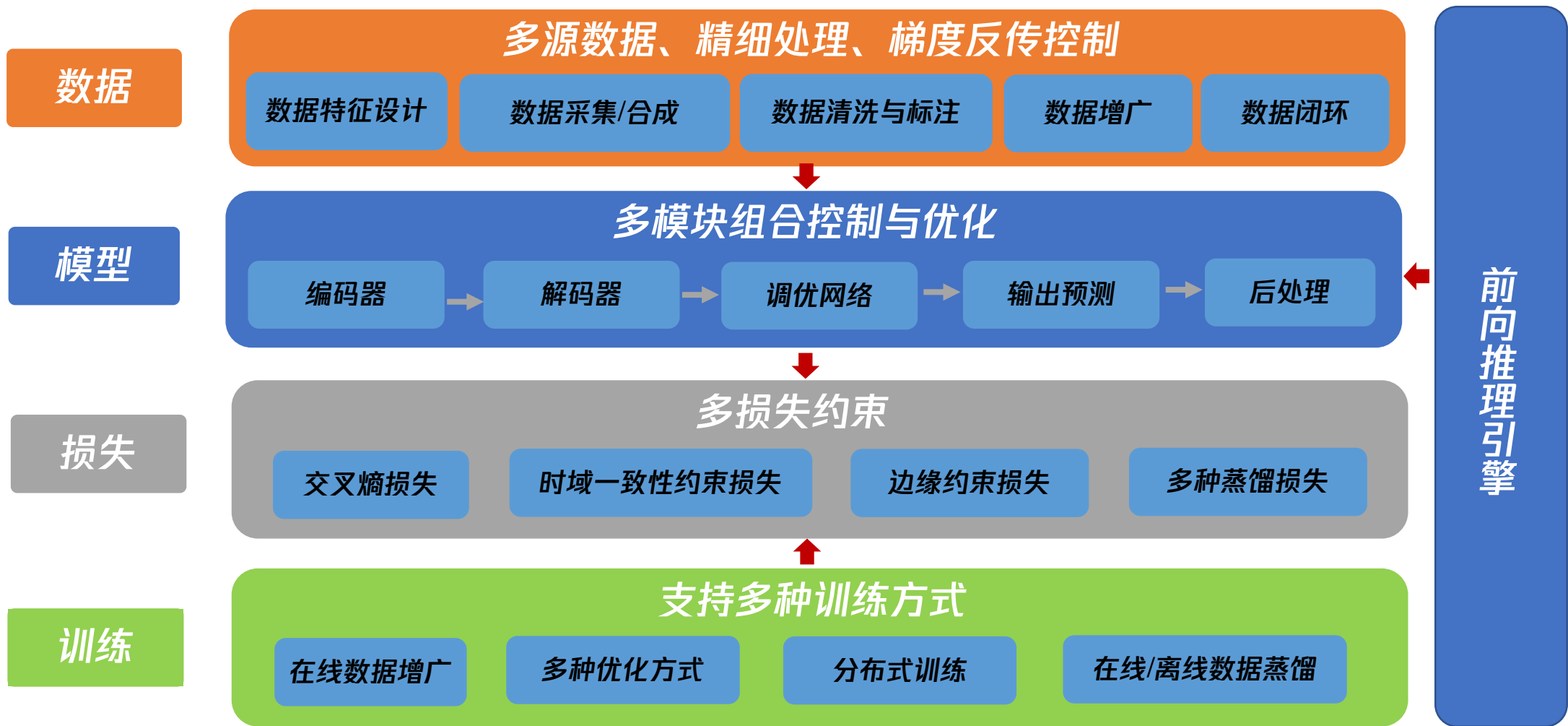
回复「加群」进入技术交流群



虚拟背景解决方案——框架

关注「腾讯产业互联网」公众号

回复「加群」进入技术交流群



虚拟背景解决方案——数据

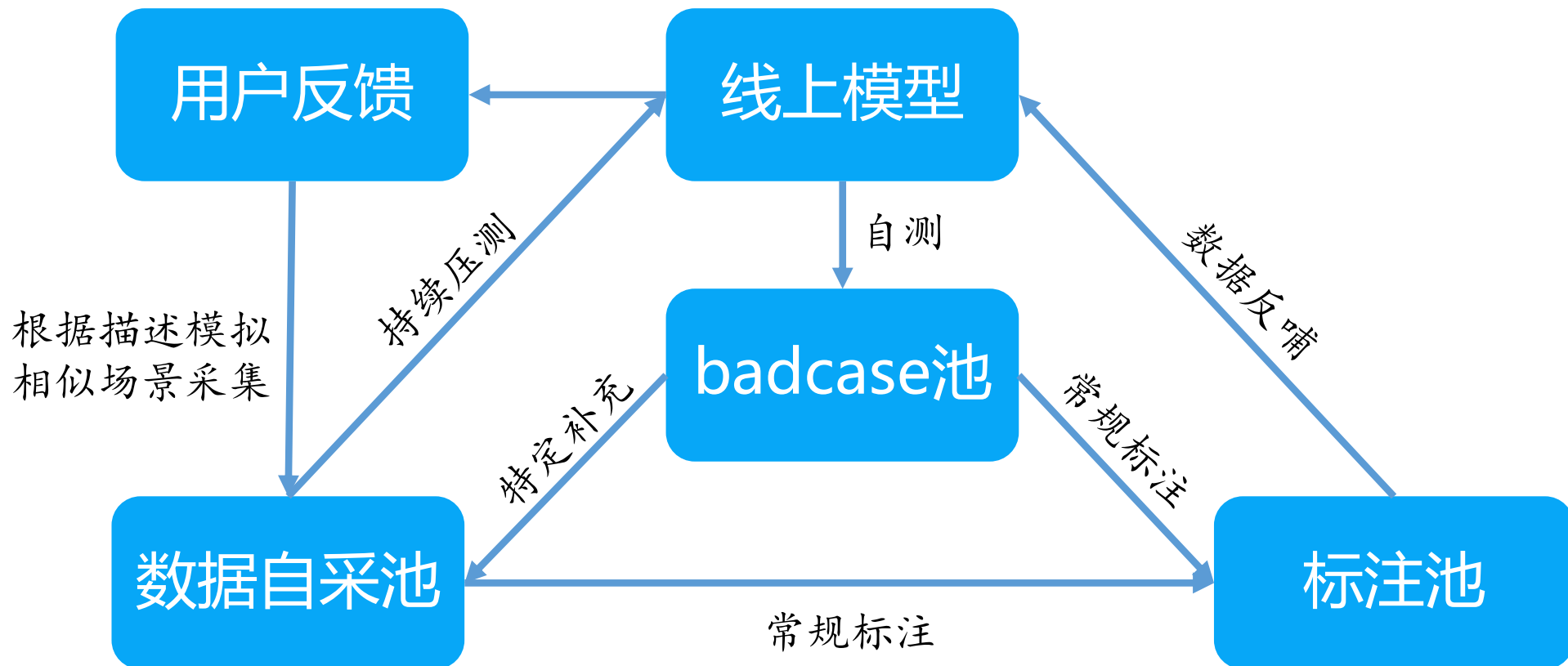
关注「腾讯产业互联网」公众号

回复「加群」进入技术交流群



数据集建设

- 数据特征设计
- 数据采集
- 数据合成
- 数据清洗与标注
- 多源数据集融合
- 数据增广



虚拟背景解决方案——模型

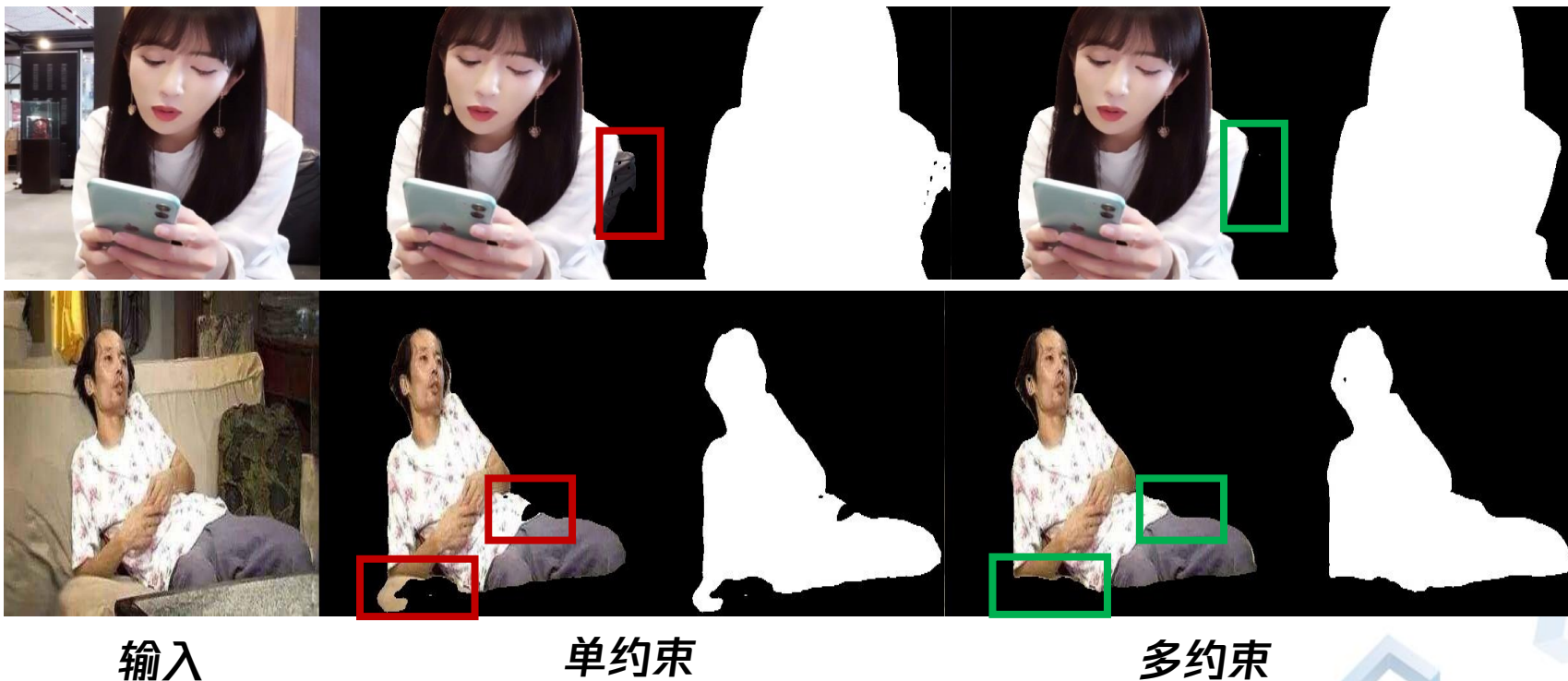
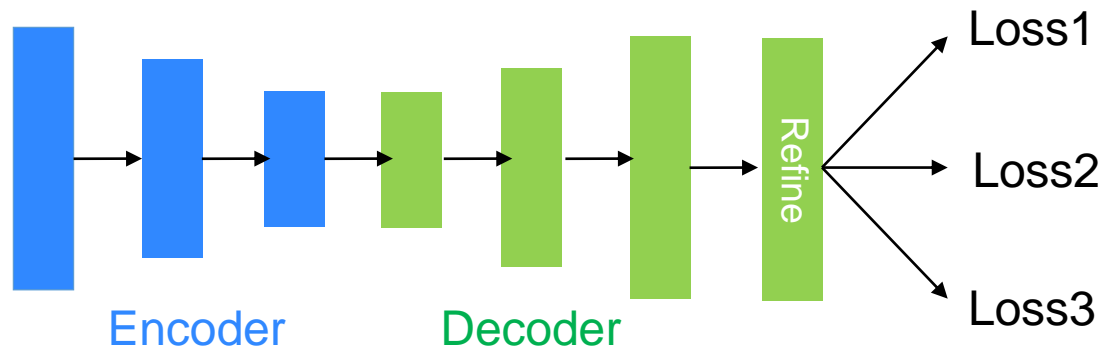
关注「腾讯产业互联网」公众号

回复「加群」进入技术交流群



模型

- 网络输入
- 编码器
- 解码器
- 调优网络
- 多任务约束
- 后处理



虚拟背景解决方案——蒸馏 [1]

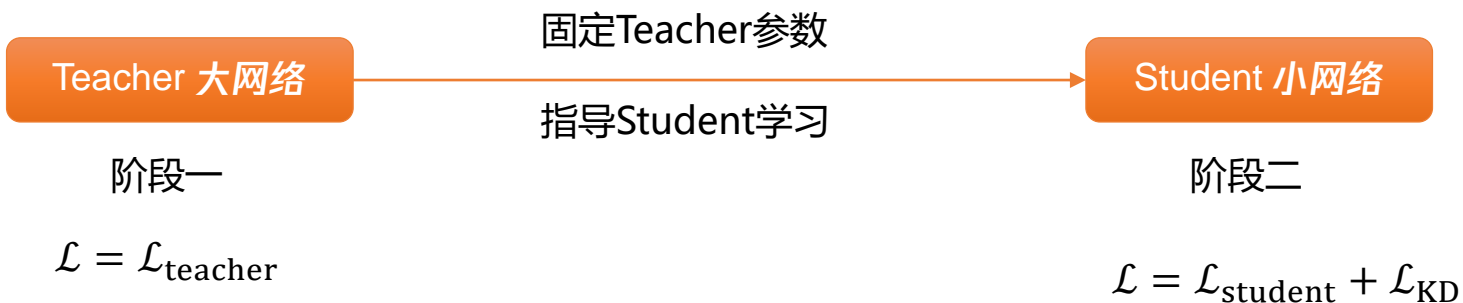
关注「腾讯产业互联网」公众号

回复「加群」进入技术交流群

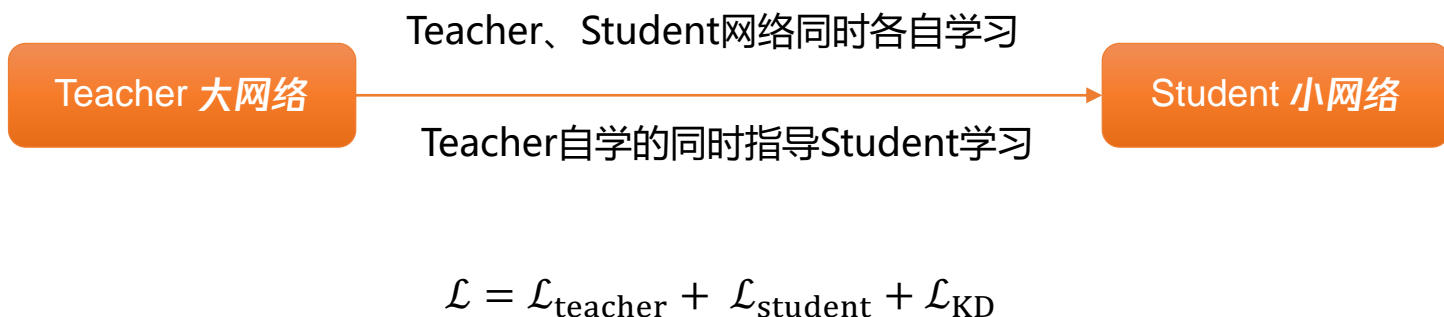


为什么要蒸馏

离线蒸馏



在线蒸馏



虚拟背景解决方案——蒸馏 [2]

关注「腾讯产业互联网」公众号

回复「加群」进入技术交流群



Pixelwise蒸馏

$$\mathcal{L}_{pixelwise} = \frac{1}{N} \sum_{i=1}^N D_{KL}(p_i^t \parallel p_i^s)$$

Teacher 网络输出层(class probability)

作用在输出层(class probability)

Student 网络输出层(class probability)

Affinity蒸馏

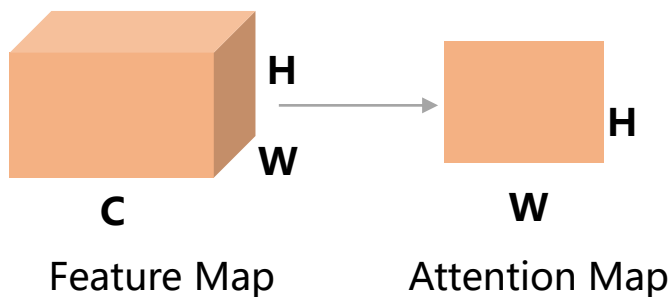
$$\mathcal{L}_{pairwise} = \frac{1}{N^2} \sum_{i \in N} \sum_{j \in N} (f_{ij}^t - f_{ij}^s)^2$$

Teacher 网络特征层(class probability)

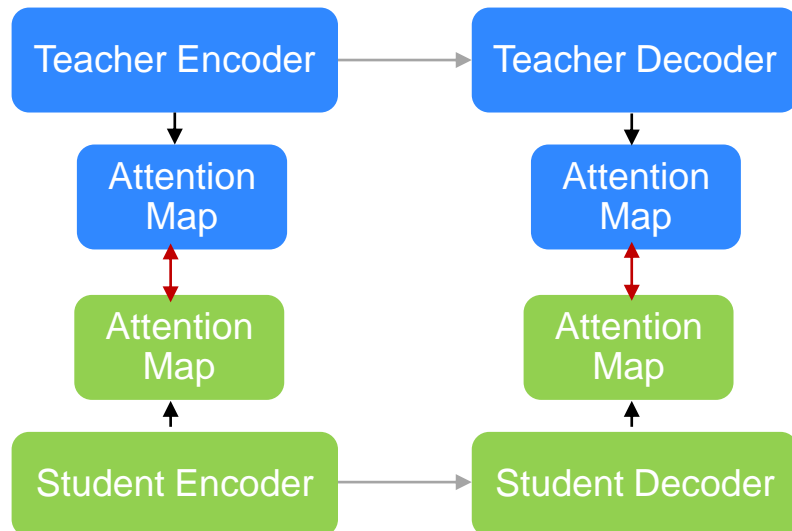
Affinity矩阵耗计算资源

Student 网络特征层(class probability)

Attention蒸馏



$$F: \mathcal{R}^{CHW} \rightarrow \mathcal{R}^{HW}$$



高效作用在任意位置

[1] Hinton G, Vinyals O, Dean J. Distilling the knowledge in a neural network[J]. arXiv preprint arXiv:1503.02531, 2015.

[2] Huang Z, Wang N. Like what you like: Knowledge distill via neuron selectivity transfer[J]. arXiv preprint arXiv:1707.01219, 2017.

[3] Zagoruyko S, Komodakis N. Paying more attention to attention: Improving the performance of convolutional neural networks via attention transfer[J]. arXiv preprint arXiv:1612.03928, 2016.

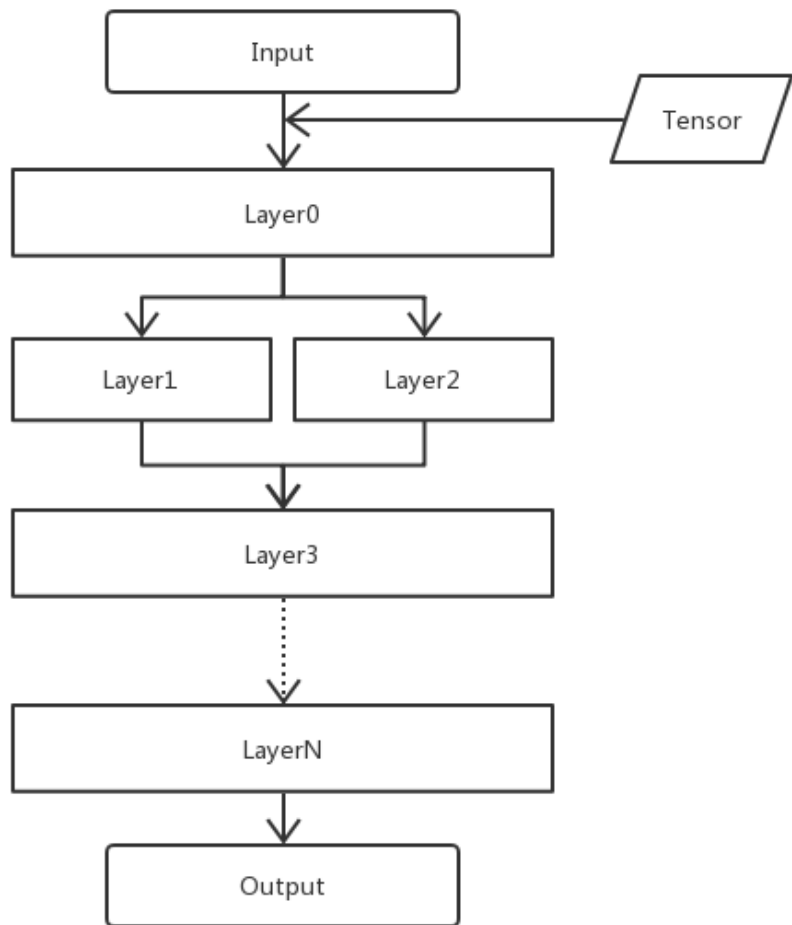
虚拟背景解决方案——前向推理加速 [1]

关注「腾讯产业互联网」公众号

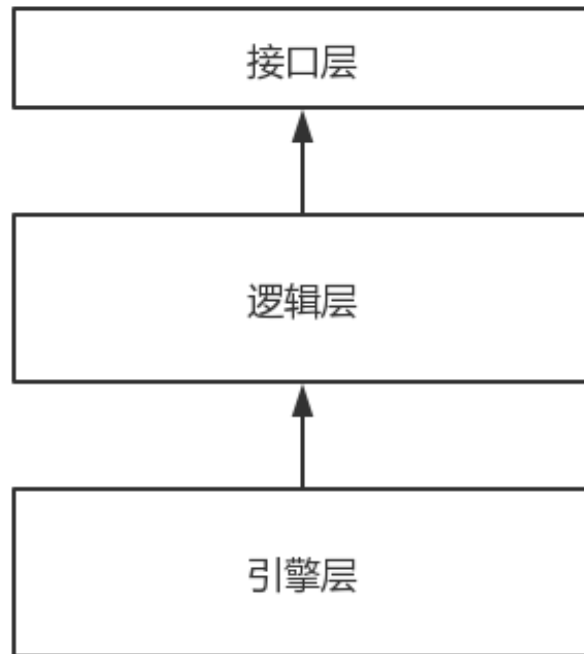
回复「加群」进入技术交流群



前向推理

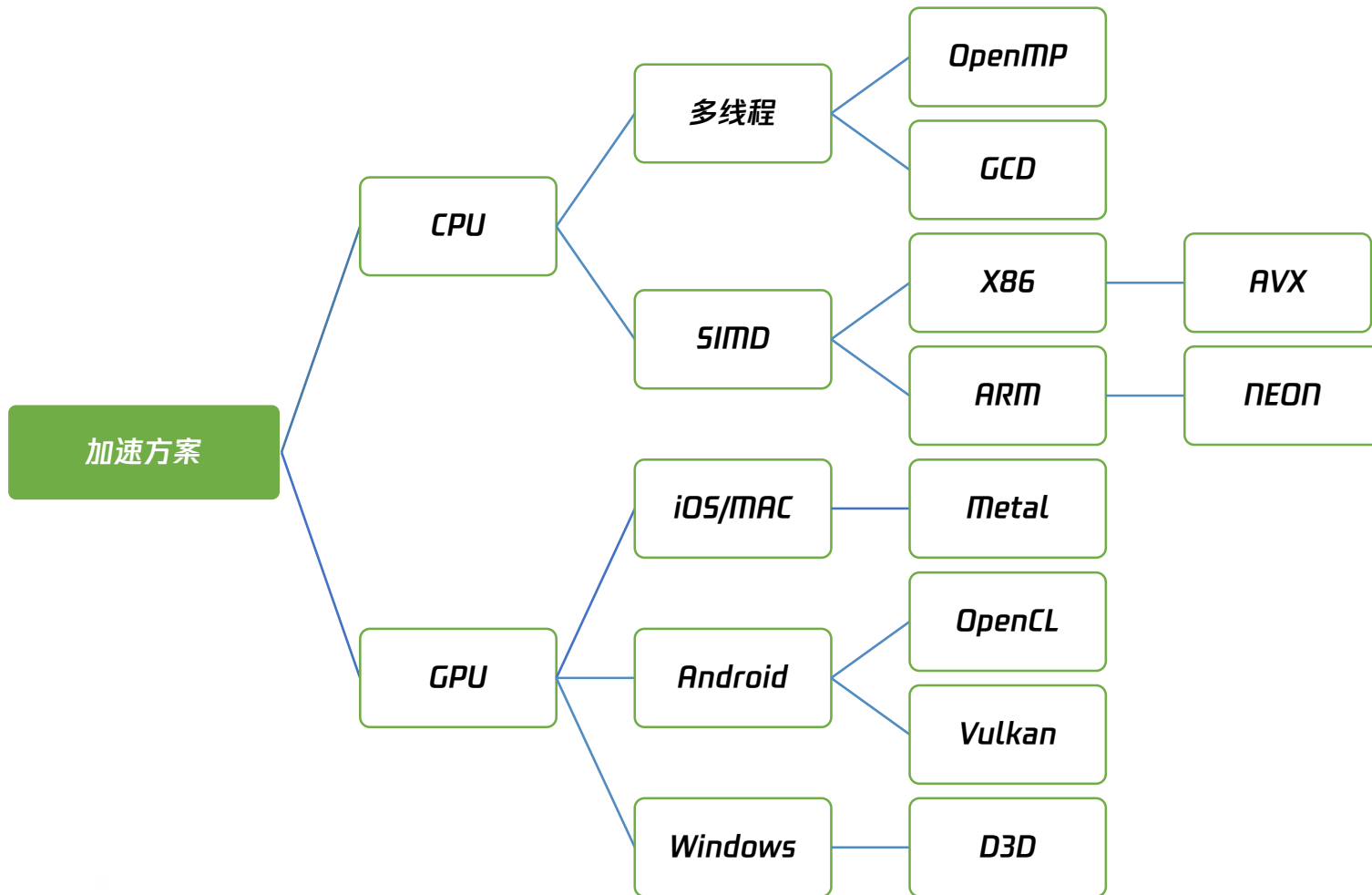


推理框架



虚拟背景解决方案——前向推理加速 [2]

异构计算加速



关注「腾讯产业互联网」公众号

回复「加群」进入技术交流群

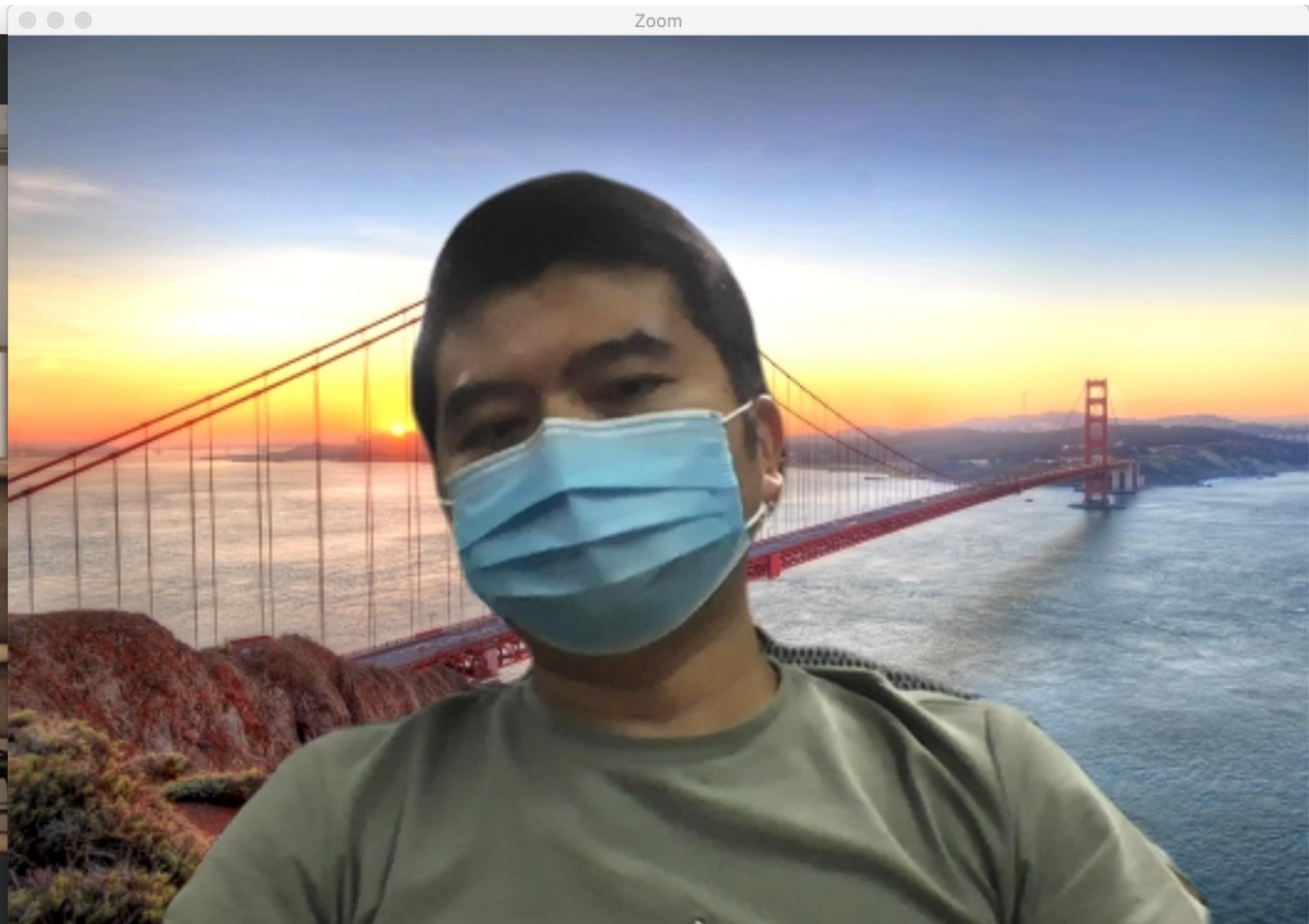


虚拟背景解决方案——直观体验

与友商的直观体验比较

关注「腾讯产业互联网」公众号

回复「加群」进入技术交流群



美颜解决方案的探索与实践

美颜

视频美颜功能可以提升用户视频体验，鼓励用户积极参与到视频通话的场景中来。通常来说，美颜包括磨皮、美型和美妆三个部分。

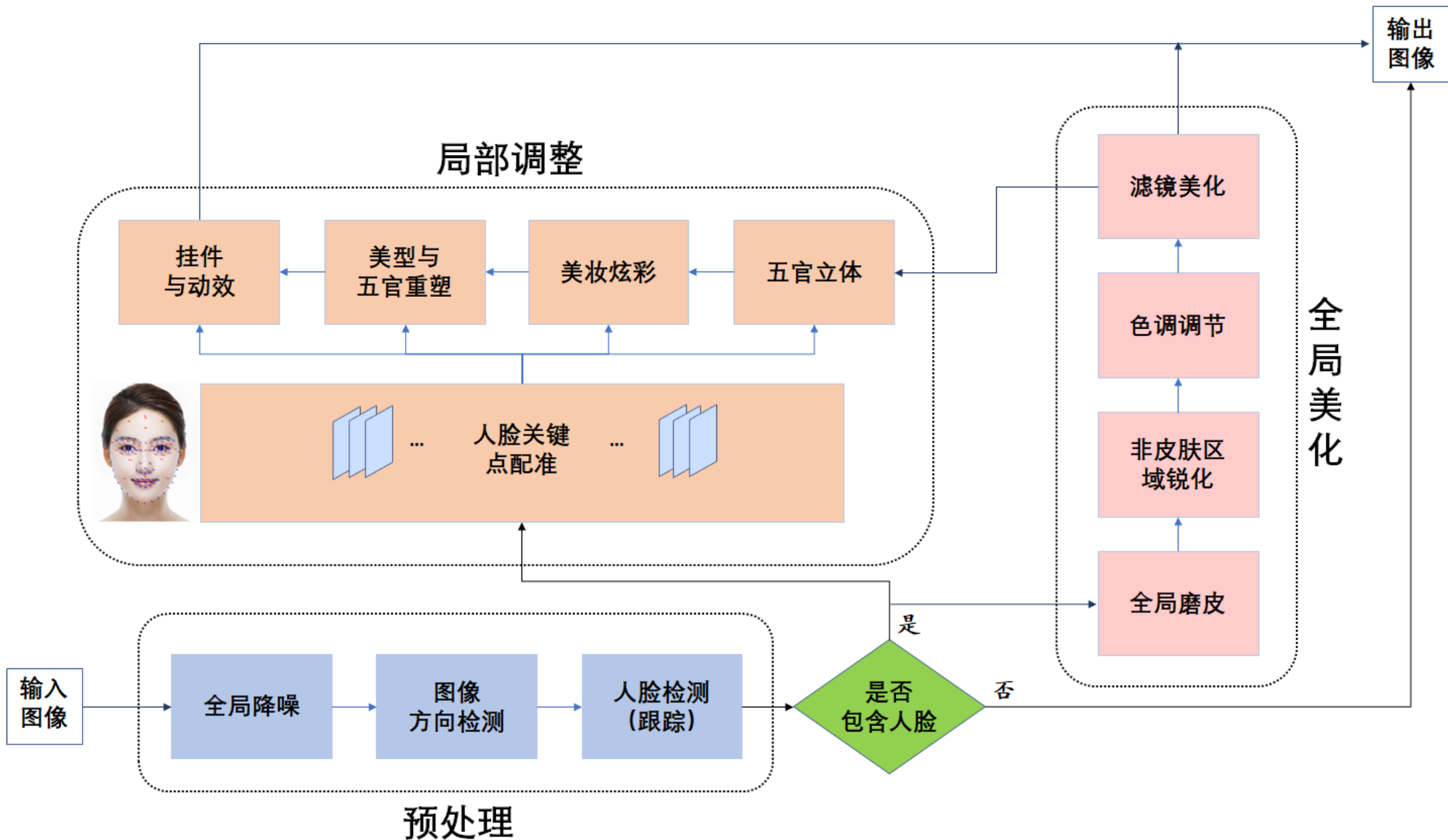
关注「腾讯产业互联网」公众号

回复「加群」进入技术交流群



美颜解决方案——框架

关注「腾讯产业互联网」公众号
回复「加群」进入技术交流群



美颜解决方案的探索与实践

问题

- 如何在性能和效果之间取得平衡?
- 强噪声的摄像头数据下如何有效保证美颜体验?
- 为什么人脸检测与人脸配准要分开设计?
- 美颜策略如何适配巨大的分辨率跨度?
- 应用场景中如何精简美颜处理复杂度?

关注「腾讯产业互联网」公众号

回复「加群」进入技术交流群



关注「腾讯产业互联网」公众号
回复「加群」进入技术交流群



Thanks!