

Ming Guang, Yong, Product Manager, Research
Perception, Google

MediaPipe: Building real time cross platform (mobile, web, edge, desktop) video audio
ML pipelines

Abstract

Applying machine learning models like action recognition, pose estimation, face detection, hand tracking on video/audio (multimodal) data is becoming more common in modern applications and devices (eg Tiktok, Shazam, Google Home Hub). However, creating these multimodal ML applications are challenging as developers need to deal with real time synchronization of time series data during model inference and doing it cross platform on mobile and edge devices.

Google open sourced in Jun 2019, MediaPipe (<https://mediapipe.dev>), a cross platform applied machine learning pipeline framework that simplifies the development process for mobile and edge devices like Raspberry Pi, Coral (EdgeTPU), Nvidia Jetson/Nano and Qualcomm 8155. My talk will be introducing the open source MediaPipe framework, walking through mobile and edge (EdgeTPU coral) demos like hand tracking, face detection and gaze estimation

Related links

<https://mediapipe.page.link/handgoogleaiblog>

<https://arxiv.org/pdf/1906.08172.pdf>

Software & Systems

Contact info:

Ming Guang, Yong

Product Manager, Research Perception

Google

1600 Amphitheatre Parkway, Mountain View CA 94043

+1-650-447-4116

mgjong@google.com