用TensorFlow开发Android手机本地深度学习应用 头条

原文

ctolib 发布于5天前 阅读53次

社区推荐：掘金是一个面向开发者的技术社区，内容涵盖了前后端，Android 和 iOS 开发，每天更新深度文章，最新技术资讯，优质开源库推送。无论是入门还是进阶，掘金都可以帮助你成为更优秀的开发者。

ctolib
@CEEO.码库

头条目录 只显示目录

Chia-Chun (JJ) Fu holds a PhD in Chemical Engineering from UC Santa Barbara. At Insight, she deployed a WaveNet model on Android using TensorFlow, and in the process rewrote into Java a Python module that extracts features from audio.

Want to learn applied Artificial Intelligence from top professionals in Silicon Valley or New York? Learn more about the Artificial Intelligence program.

There are many situations when running deep learning inferences on local devices is preferable

for both individuals and companies: imagine traveling with no reliable internet connection available or dealing with privacy concerns and latency issues on transferring data to cloud-based services. *Edge computing* provides solutions to these problems by processin

g and analyzing data at the edge of network.

Take the “Ok Google” feature as an example—by training “Ok Google” with a user’s voice, that user’s mobile phone will be activated when capturing the keywords. This kind of small-footprint k

eyword-spotting (KWS) inference usually happens on-device so you don’t have to worry that the service providers are listening to you all the time. The cloud-based services will only be initiated after you make the command s. Similar concepts can be extended to application

s on smart home appliances or other IoT devices, where we need hand-free voice control without internet.

What's more, edge computing not only brings AI to the IoT world, but provides many other possibilitie s and benefits. For example, we can preproces

s images or voice data into a compressed representation on-device and then send it to the cloud. This approach resolves both privacy and latency issues.

During my time at Insight, I deployed a pretrained WaveNet model on Android using TensorFlow. My

goal was to explore the engineering challenge of bringing deep learning models onto devices and making things work! In this post, I'll quickly walk you through the process of building a general speech-to-text recognition application on Android with

TensorFlow. I hope after this post you'll be able to build your own DL-powered applications next time!

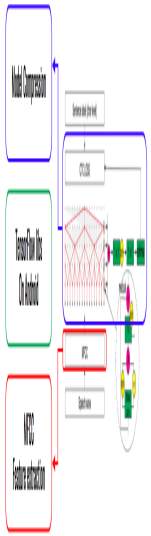

Figure 1. An overview of the process. Let's look into these