

开源 | 基于Metal的机器学习框架Bender：可在iOS上运行TensorFlow模型

2017年06月07日 12:15:04 机器之心


0| | | |

选自GitHub

机器之心编译

参与：吴攀

在正在举行的 WWDC 2017 上，苹果宣布发布了 Metal 2，详情可参阅机器之心的报道《苹果开发者大会 WWDC 2017：首次全面展示苹果的人工智能实力》。而差不多在 WWDC 2017 大会开幕时，XMART LABS 在 GitHub 上开源了 Bender——一个基于 Metal 的现代机器学习框架，而值得一提的是，Bender 也是著名动画《飞出个未来》里面的一个金属（metal）机器人主角。机器之心在这篇文章中对该项目进行了编译介绍。

 机器之心

专业的人工智能媒体与产业服务平台。

热文排行

- 日榜周榜月榜
- 1

养老保险交15年和25年，领钱的时候差...
- 2

2017年中国和日本工业实力比较，结果...
- 3

中国这一招灭了美国威风？原油定价权我...
- 4

银行拼了！理财产品收益飙到7%，背后...
- 5

安迪做空包氏集团，金融圈大牙都要笑掉...
- 6

贾跃亭的中国好同学，今天公司一分钟蒸...
- 7

洗牌下的中国城市战争将爆发，你在的城...
- 8

最神秘高考教材老板：疑似牛散身家过亿...
- 9

靠“高考”养活的小镇，租金连年高涨直逼...
- 10

一块饼干60元，杜海涛代言的减脂产品摊...



项目地址：<https://github.com/xmartlabs/Bender>

Bender 是 MetalPerformanceShaders 之上的一个抽象层（abstraction layer），可用于操作神经网络。在 Bender 出现后，苹果机器学习环境中有了自己的框架。

本项目的文档在 Documentation 文件中：

API 包含了上手使用所需的最重要的信息；

Supported Layers 解释了其支持哪些层以及它们是如何映射到 TensorFlow 指令的；

Importing 解释了如何将 TensorFlow 等其它框架的模型导入进来。你可以找到如何为自定义实现增强这个功能的信息。

介绍

Bender 是 MetalPerformanceShaders 之上的一个抽象层（abstraction layer），可用于操作神经网络。在人工智能领域，人们对在移动设备上执行神经网络的兴趣越来越大，即便其训练过程是预先在其它地方完成的。我们希望人们能更轻松地在 iOS 上运行预训练的网络。



Bender 可以让你轻松使用卷积、池化、全连接和一些规范化等最常见的层，从而轻松地定义和运行神经网络。而且这些层接收参数的方式也很灵活。我们希望能加载在 TensorFlow 或 Caffe2 等框架上训练好的模型。目前 Bender 有一个用于 TensorFlow 的适配器（adapter），其可以加载带有变量的图（graph），并将其「翻译」成 Bender 的层（layer）。这个功能仅支持一部分 TensorFlow 指令，但我们计划增强它，使其覆盖更大的范围。

Bender 已经可用了，但也仍处在积极的开发阶段，我们也希望看到社区的需求能推动它的发展。注意，它的 API 还不是很稳定。

我们为什么需要 Bender？

在 Xmartlabs，我们曾经想要启动一个机器学习项目，然后我们调查了可以在 iOS 中使用的框架。我们发现 MetalPerformanceShaders 很有用，但对用户并不太友好，我们发现我们重复了大量代码和信息。这就是我们为什么要开始构建一个能够处理这类事物的框架。

我们还发现我们需要将使用 TensorFlow 训练的模型翻译成 iOS 上可用，为此我们需要编写脚本，使之可以将权重转换成 MPSCNN 格式，并将 TensorFlow 的层中不同类型的参数映射成 MPSCNN 核所使用的参数。TensorFlow 可以为 iOS 进行编译，但目前它并不支持在 GPU 上运行，而这却正是我们需要的。我们也并不想将 TensorFlow 的静态库包含在我们的项目中。所以我们开发了一个适配器，可以将 TF graph 解析并翻译成我们的 Bender layer。

使用方法

你可以使用我们自定义的算子在 Bender 中定义你自己的网络，或者你可以加载从 TensorFlow 导出的模型。你可以通过如下方式定义一个网络和加载一个模型：

```
import Bender// Define a Network and how it will load its weights / parameters

let randomLoader = RandomParameterLoader(maxSize: 7*7*64*1024)network = Network(d

// Convert a graph from TensorFlow

let url = Bundle.main.url(forResource: "myGraph", withExtension: "pb")!

let converter = TFConverter.default()network.nodes = converter.convertGraph(file: url, type:

// Initialize the network

network.initialize()

// Run the network

let inputImage = ...network.run(inputImage: image, queue: commandQueue) { output in ...

}
```

你可以在 Importing 文档了解更多信息。

如果你想自己定义你的网络，你可以如下操作：

```
let network = ...network.start - Convolution(convSize: ConvSize(outputChannels: 16, kernelSize: 3, stride: 2), neuronType: .l
```

```
- InstanceNorm()
```

```
- Convolution(convSize: ConvSize(outputChannels: 32, kernelSize: 3, stride: 2), neuronType: .l
```

```
- InstanceNorm()
```

```
- FullyConnected(neurons: 128)
```

```
- Neuron(type: .tanh)
```

```
- FullyConnected(neurons: 10)
```

```
- Softmax()...
```

了解更多请查看 API 文档。

环境要求

iOS 10.0+

Xcode 8.3+

参与

你可以在 GitHub 上参与和贡献；

如果你希望增加新功能，请开启一个问题；

如果你发现了 bug 或需要帮助，请先查阅过去的问题、FAQ 和 StackOverflow 上的流程（标签：Bender），然后再提交问题。

在贡献之前，请先查阅 CONTRIBUTING 文件了解更多信息。如果你在你的 App 中使用了 Bender，请告知我们。

样例

按以下三个步骤，可以运行一个样例项目：克隆 Bender 库、打开 Bender 工作空间，运行 Example 项目。

安装

CocoaPods

安装 Bender，在你的 Podfile 中添加以下代码即可：

```
pod 'MetalBender', :git = 'https://github.com/xmartlabs/Bender.git'
```

记住 Bender 是为 iOS 10 编译的。所以你必须要在你的 Podfile 中加入 platform :ios, '10.0'

作者

Xmartlabs SRL：https://github.com/xmartlabs

@xmartlabs：https://twitter.com/xmartlabs

更改日志

参阅 CHANGELOG.md 文件。






更多有关GMIS 2017大会的内容，请点击「阅读原文」查看机器之心官网↓↓↓



作者历史文章


深度 | 详解苹果Core ML：如何为iOS创建机器学习应用？



选自developer.apple机器之心编译参与：吴攀在昨天开幕的WWDC2017开发者大会上，苹果宣布了一系列新的面向开发者的机器学习API，包括面部识别的[详细]

2017年 06月07日 12:15


关系推理水平超越人类：DeepMind展示全新神经网络推理预测技术



选自DeepMind作者：AdamSantoro等机器之心编译参与：机器之心编辑部想象一下在阿加莎·克里斯蒂（《东方快车谋杀案》作者）的侦探小说里收集所有证据找[详细]

2017年 06月07日 12:15


学界 | AAI-17论文提出深度学习C语言修改器DeepFix：用人工智能加



选自aai.org机器之心编译参与：吴攀如果你是一个程序员，你一定体验过编译时报错的痛苦，有时候一个小错误可能就需要你很多时间检查许多代码才能找到。近日，印度[详细]

2017年 06月07日 12:15


报名 | 图普科技「让世界充满 AI！」，与您相约浙大



这周6月8日，图普科技与您相约美丽的浙江大学2014年末，诺兰导演的一部《星际穿越》让无数影迷叹为观止，而全程人气最高的角色——智能冰箱更是让人耳目一新。“等下[详细]

2017年 06月07日 12:15


专栏 | 大漠孤烟，长河落日：面向景深结构的风景照生成技术



机器之心专栏上海交通大学电子工程系作者：杨蕊简介2014年以来，生成对抗网络（GenerativeAdversarialNetworks）已经在各式各样的图像生[详细]

2017年 06月06日 06:15


苹果开发者大会WWDC 2017：首次全面展示苹果的人工智能实力



机器之心报道参与：李亚州、XavierMassa当地时间6月5日，苹果开发者年度盛会WWDC2017在美国加州举行。在这个舞台上，我们看到了苹果软件、硬件有哪些[详细]

2017年 06月06日 06:15


实验研究工作流程详解：如何把你的机器学习想法变成现实



选自dustintran作者：DUSTINTRAN 机器之心编译参与：李泽南、Smith从研究思想的提出到实验的具体实现是工程中的基础环节。但是这一过程常常被一[详细]

2017年 06月05日 13:15


开源 | 浏览器上最快的DNN执行框架WebDNN：从基本特性到性能测评



选自Github机器之心编译参与：蒋思源、晏奇WebDNN是网页浏览器中最快的DNN执行框架，而本文首先简单介绍了WebDNN特征与其框架结构，即表明了为什么W[详细]

2017年 06月05日 13:15

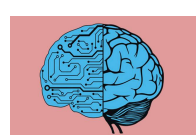
学界 | CMU新研究试图统一深度生成模型：搭建GAN和VAE之间的桥梁



选自arXiv机器之心编译参与：吴攀不同的深度生成模型之间存在怎样的共性？近日，来自CMU和Petuum的四位研究者ZhitingHu、ZichaoYang、R[详细]

2017年 06月05日 13:15

资源 | 《人工智能与游戏》发行初版：从三个方面概述游戏人工智能（附下



选自gameaibook机器之心编译参与：黄小天近日，由马耳他大学副教授、情感计算专家GeorgiosN.Yannakakis和纽约大学副教授、人工智能与游戏研[详细]

2017年 06月05日 13:15

