

新闻稿

Khronos 发布训练好的神经网络优化部署标准 NNEF 1.0

Khronos 会员标准采用不断增长，开源工具生态系统不断扩展

SIGGRAPH, 温哥华– 2018 年 8 月 13 日 – 6:00 AM PT – [The Khronos™ Group](#), 一个由领先硬件和软件公司组成的、致力于创建高级加速标准的开放组织，今天宣布公开发布 [NNEF™ 1.0](#)（神经网络交换格式）规范。经过搜集来自业界对预测版规范的反馈建议，Khronos 发布了 NNEF 1.0，为硬件制造商带了一个稳定、灵活可扩展的开放标准，以部署优化加速神经网络推论到各种高端设备上。随同该标准的发布，现在工具生态系统也在 [GitHub](#) 上发布，包括从 Tensorflow 和 Caffe 的 NNEF 解析器和转换器。受欢迎的推论环境导入器，包括安卓的神经网络 API（NNAPI）和 Khronos 的 OpenVX™，都已完成开发。

NNEF 的创建是为了通过促进训练框架和推论引擎之间神经网络的交换，提高开发者混合和匹配推论和训练解决方案的自由度，最终减少业界分裂。这个开放、免版权标准，是 Khronos NNEF 工作组的业界领先会员单位的共同合作成果，为硬件和软件公司在确保响应 machine learning 产业迅速发展的灵活性的同时，可以拥有坚实基础，完成产品部署发展。

NNEF 可用于各种使用案例和网络类型。其基础格式是针对人类可读和简化编写和优化设计的，包括操作器精准度和复杂操作的灵活表示方法，以完成到各种加速器架构的最佳匹配。NNEF 扩展实现了大型网络的人类编辑和具体问题的处理，例如为经训练的网络值定制数据格式。

“Khronos 意识到了公司在部署训练好的神经网络到高端设备过程中不断发展的格式僵局。我们开始建立首个开放标准解决方案，为工程师们带来各种推理引擎的优化和部署训练好的网络。核心 NNEF 1.0 将实现当今的高端解决方案并灵活地融入到其扩展机制中”，Khronos NNEF 工作组主席 Peter McGuinness 讲到。“2017 年 12 月，我们发布了开发者预览版 NNEF 规范并公开征集业界反馈。社群积极响应，表明了对这个标准的强烈需求并帮助我们完成了 NNEF 1.0 规范”。

一个完整的工作流，从优化到部署，都可以使用 NNEF 作为一个标准化传输格式。自标准发布之时，即将由两个开源 Tensorflow 转换器支持，既可用于基于 protobuf 和 python 代码的网络描述，也可用于 Caffe 的转换器。[Au-Zone Technologies](#) 正在开发一个 Caffe 2 开源导入/导出器，

即将在 2018 年第三季度推出。各种来自会员公司的工具，例如 Almotive 和 AMD，都在开发过程中，包括台湾国立清华大学的安卓 NNAPI 导入器。

NNEF 工作组致力于实现并鼓励在高速发展的训练框架之间的可靠的交互转换，包括 Torch、Chainer、Theano、PyTorch 和带有开源导入器和导出器的 MXNet。另外还有其他开源工具，包括 NNEF 语法分析器和验证器，现已在市场推出，简化了定制、移动和嵌入式高端推论环境的导入器的开发，例如 Apple 的 Core ML 和 Khronos 的用于视觉和推论 runtime 加速的 OpenVX 标准。

[NNEF 1.0 规范 and 文件](#)可在 Khronos 网站上免费获取，NNEF 开源工具和项目可以在 [Khronos NNEF 工具库](#) 获取。[Khronos 文件格式采用者项目](#) 现已推出，完全免费，实现 NNEF 用户拥有一个正式许可可以使用 NNEF 商标并受到 Khronos 知识产权框架保护的愿望。

业界支持

“最近几个月，我们的一些汽车和其他领域的合作伙伴就他们对全新的神经网络加速硬件平台所带来的挑战和工作表示担心。我们多次公开表示 NNEF 绝对是这些挑战的解决方案。作为 NNEF 工作组发起成员之一，我们很高兴看到这个标准的完成，并已经有很多主要厂商采用。作为首个神经网络硬件 IP 供应商之一，带来了完全支持 NNEF 的 SDK，我们坚信 NNEF 将在全球神经网络生态系统中不断受到广泛采用，成为神经网络交换的值得信赖的国际标准，” Almotive 公司硬件工程部负责人 Marton Feher。

“Imagination 恭祝 NNEF 的发布。多年来 PowerVR 一直是 Khronos 标准的采用者，并在我们的 PowerVR GPU 和我们的 PowerVR NNA 硬件上支持神经网络，交互操作一直都是我们客户关心的问题。NNEF 标准的推出将实现更好的神经网络在不同业界的交互转换，” Imagination 公司市场部副总裁 David Harold。

关于 The Khronos Group

The Khronos Group 是一个开源组织，致力于创建开放标准，以实现并行计算、图形、视觉和神经网络在各种平台和设备上的编写和加速。Khronos 标准包括 Vulkan®, OpenGL®, OpenGL® ES, OpenGL® SC, WebGL™, SPIR-V™, OpenCL™, SYCL™, OpenVX™, NNEF™, COLLADA™, OpenXR™ 和 glTF™。Khronos 会员可以为 Khronos 规范的开发做出贡献，在标准公开部署之前的不同阶段拥有投票权，并通过提前获取规范草本和一致性测试加速其高端加速平台和应用的推出。更多关于 The Khronos Group 的信息，请浏览：[Khronos.org](https://www.khronos.org)。

###

Khronos, EGL, glTF, NNEF, OpenVG, OpenVX, OpenXR, SPIR, SPIR-V, SYCL, Vulkan 和 WebGL 是 The Khronos Group Inc. 的商标或注册商标。OpenCL 是 Apple Inc. 的商标，OpenGL 是注册商标，OpenGL ES 和 OpenGL SC 标识是 Hewlett Packard Enterprise 的商标，由

Khronos 在许可下使用。全部其他产品名称、商标和/或公司名称仅用于识别并归属于各自所有者。

媒体联系人:

Alex Crabb, Caster Communications Inc.

alex@castercomm.com

401-792-7080