



使用 Omniverse Audio2Face 和 Riva 构建数字人

NVIDIA 深度学习培训中心 (DLI)



课程概述

数字人技术是以人工智能为基础的虚拟智能机器人,它结合了最新的 3D 建模技术和交互式 AI 语音技术的优势,具备出色的亲和性,符合人类自然交流习惯,在元宇宙的大潮中备受瞩目,非常有潜力成为下一代人机交流主要界面。目前,数字人技术已经在泛娱乐、金融、教育、政务、汽车、医疗、零售等领域取得了相当不错的发展势头,成为业界的技术发展热点。凭借其灵活的可定制性和高度智能化的特点,数字人技术可以为企业 提供高质量、高效率、个性化且低成本的交互服务,从而有效地优化企业运营流程、提升客户体验,增强品牌形象,创造更加美好的商业未来。

本课程主要面向希望学习使用 NVIDIA Omniverse 来构建虚拟数字人的开发人员。从端到端应用开发的角度,课程重点讲解如何使用语音驱动自定义的数字人的面部动作和表情,以及用微服务方式控制数字人的身体动作。课程同时介绍如何通过 NVIDIA Riva 为数字人构建一套交互式语音系统后端,包含自动语音识别 (ASR)、NLP 问答,以及语音合成 (TTS) 等模块。课程最后学习如何将前期开发的各个系统有机串联起来,以构建一个完整的数字人交互系统。

学习目标

参加本次培训,您将学到:

- 如何在 Audio2Face 中通过角色迁移驱动任意虚拟人物头像的面部动作。
- 如何通过 Extensions 以及微服务 (Microservices) 的方式实现虚拟人物肢体动作的控制。
- 如何使用 Riva 快速构建自动语音识别 (ASR) 系统、基于自然语言的问答系统,以及语音合成 (TTS) 的发声服务系统。
- 如何有效连接虚拟人和语音系统,从而构建一个基础的交互式虚拟数字人应用。

课程概要

课程时长	8 小时 (课后 6 个月内可以继续访问课件,其中 GPU 实验资源用量有限额)
课程模式	讲师授课,每位学员使用云端完全配置的 GPU 加速工作站实验练习
预备知识	<ul style="list-style-type: none">➢ 基础 Python 编程经验➢ 基本了解 深度学习神经网络➢ 具备 Omniverse 基础知识,对于学习本课程会很有帮助<ul style="list-style-type: none">• Omniverse Audio2face 中文官网介绍• DLI 中文免费课程《在 Omniverse 中定制化开发和发布 Extensions》• Omniverse Extensions 视频介绍和文档• Omniverse 微服务视频介绍和文档
使用的工具、库和框架	NVIDIA Omniverse Audio2Face, NVIDIA Riva, PyTorch
学员评测方式	<ul style="list-style-type: none">➢ 多项选择题:评估学员关于数字人架构的理解程度➢ 编程技能评测:评估学员构建交互式数字人应用的能力
培训证书	成功完成本课程和测试后,学员将获得 NVIDIA DLI 培训证书,证明在相关领域的技能,为职业发展提供证明
课程语言	中文

学习此课程的硬件要求	您需要一台能够上网的笔记本电脑或台式机,且能够运行最新版 Chrome 或 Firefox 浏览器;我们为您提供在云端完全配置的 GPU 加速工作站的专用访问权限
课程价格	微信添加 DLI 小助手(微信号 NVIDIALearn),沟通培训需求

课程大纲

议题	说明
课程介绍 (15 分钟)	<ul style="list-style-type: none"> > 讲师介绍 > 登录 DLI 学习平台
基于Audio2Face 开发数字人 (240 分钟)	<ul style="list-style-type: none"> > 了解数字人、开发流程和工具 <ul style="list-style-type: none"> • 基于 Audio2Face的 3D 数字人开发流程 • 扩展篇:2D Live Portrait 及 ACE 微服务 > 面部动画 <ul style="list-style-type: none"> • Audio2Face及Audio2Emotion 算法原理介绍 • Audio2Face 中数字资产的设置与驱动 <ul style="list-style-type: none"> - 基于 Reallusion CC4 的数字资产准备 - 角色设置及角色迁移 - Audio2Face 面部驱动 - 与身体联动 - Audio2Emotion 面部表情控制 • 实验 <ul style="list-style-type: none"> - Audio2Face 初步体验 - 角色迁移 - Audio2Face 流播放器 > 身体动画 <ul style="list-style-type: none"> • 身体动画生成 • 插件(extension)的编写 • 身体动画的控制及构建服务 • 面部动画与身体动画的结合 • 实验:身体动画的控制
休息 (60 分钟)	
构建对话式语音系统 (120 分钟)	<ul style="list-style-type: none"> > 自动语音识别 (ASR) <ul style="list-style-type: none"> • 语音识别基本原理 • 实验:使用 NVIDIA Riva 搭建语音识别系统 > FAQ 问答系统(NLP-FAQ) <ul style="list-style-type: none"> • 问答系统的类型与应用 • 基于向量检索的 FAQ 问答系统技术方法 • 系统组成模块 • 总体运行流程 • 实验 <ul style="list-style-type: none"> - 文件结构和数据 - 构建离线向量索引 - 运行在线检索 > 语音合成(TTS) <ul style="list-style-type: none"> • 语音合成基本原理 • 实验:调用语音生成服务从文字生成语音

休息 (15 分钟)	
构建数字人全流程 (60 分钟)	<ul style="list-style-type: none">> 数字人全流程方案简介> 数字人的技术框架> 实验:根据提示信息构建完整的数字人流水线
总结和评估 (15 分钟)	<ul style="list-style-type: none">> 回顾所学要点并解答问题> 完成评估并获得证书> 填写培训调查表
下一步	<ul style="list-style-type: none">> 构建基于 Transformer 的自然语言处理应用> 数据并行 — 用多 GPU 训练神经网络> 模型并行 — 构建和部署大型神经网络

为何选择 NVIDIA 深度学习培训中心 (DLI) 的实战培训?

- > 学习 NVIDIA 与技术专家和行业领导者合作开发的课程, 获取全球同步、技术领先和现实可用的专业开发技能和经验。
- > 学习使用行业通用、标准的软件、工具和框架进行端到端的应用开发, 能够在广泛的行业中构建基于深度学习、加速计算、图形与仿真和数据科学的应用。
- > 系统化地学习理论, 并使用云端完全配置的 GPU 实验环境同步边练, 高效提升实战开发能力。
- > 获得 NVIDIA 全球开发者培训证书, 加持专业培训认证, 助力职业发展。

准备好开始学习了吗?

更多课程, 请访问 nvidia.cn/dli。

如有疑问, 请通过微信联系 DLI 小助手 (微信号 NVIDIALearn)。