

AI for Safety and Security

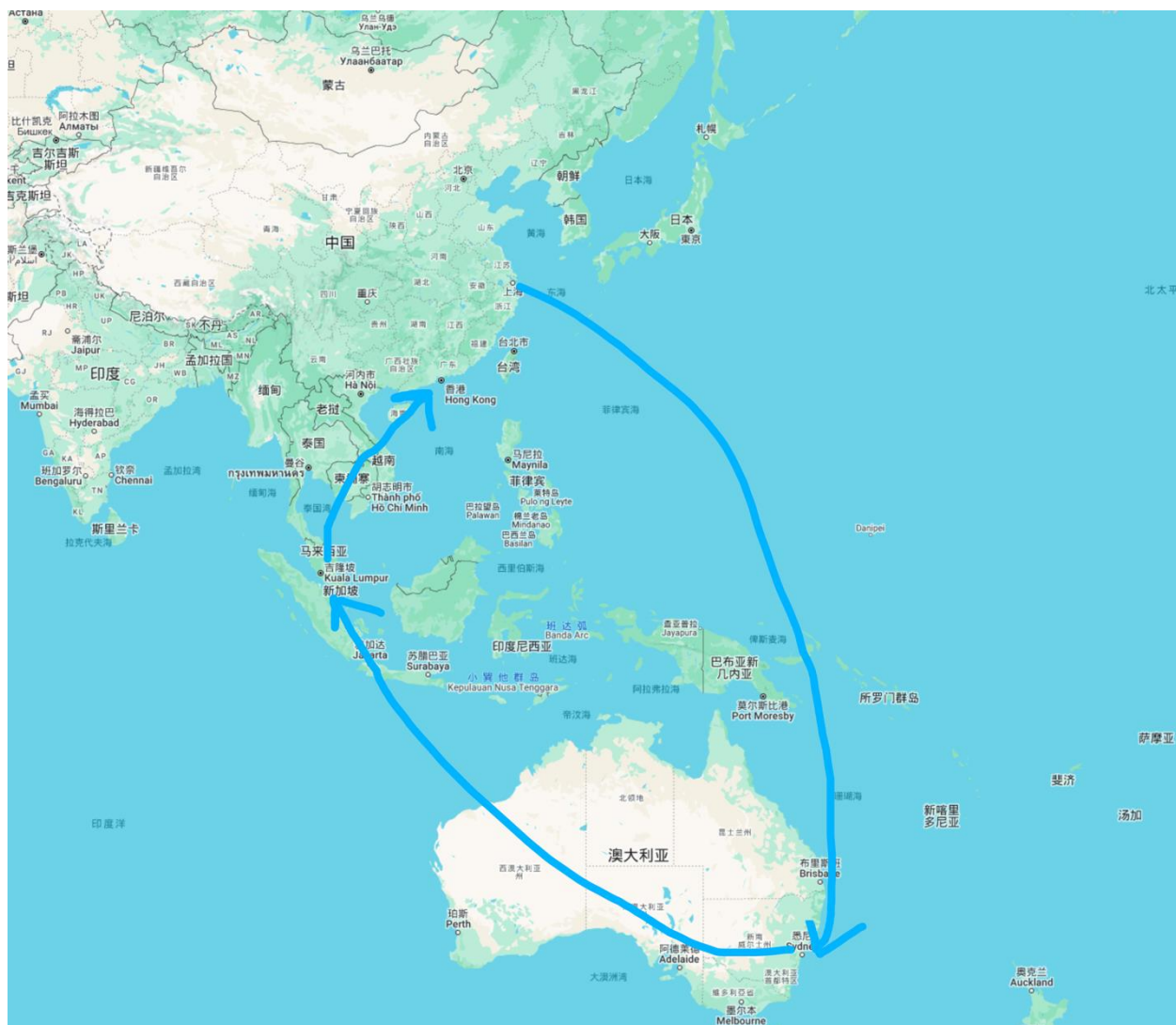
Instructor:
Zhedong Zheng

Autumn 2025
University of Macau

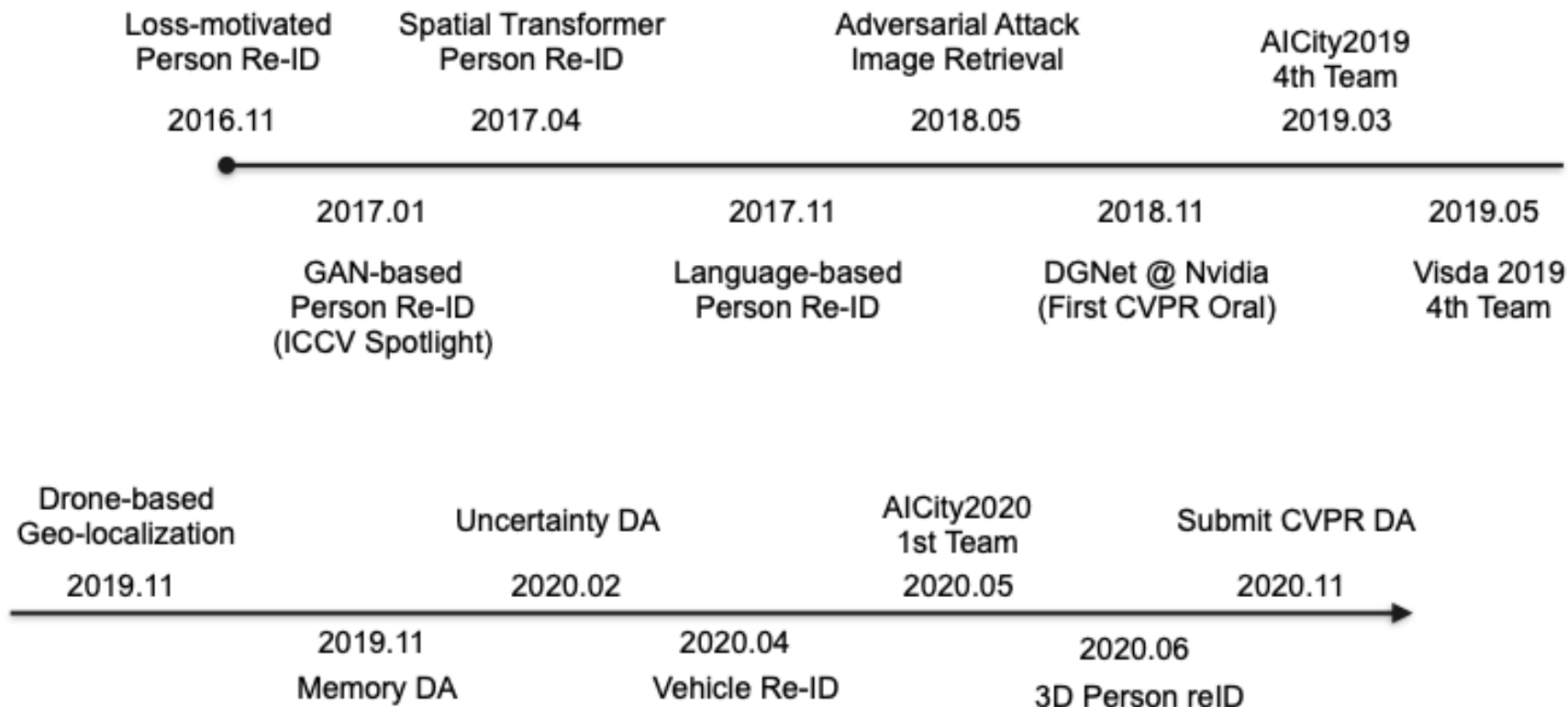


Hikaru's go

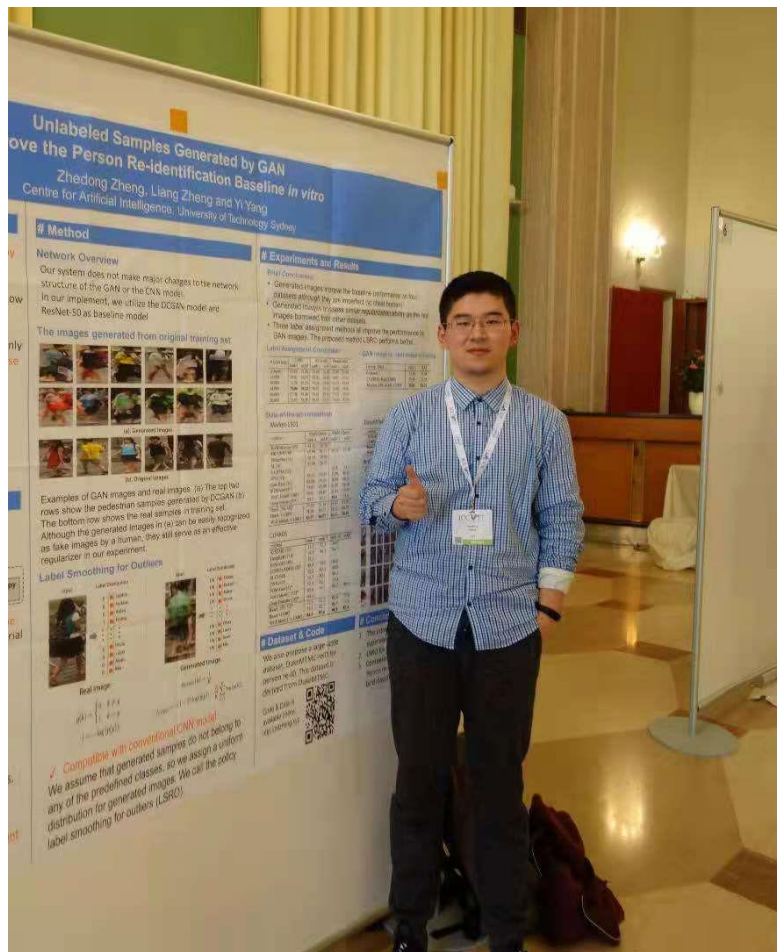
About Me



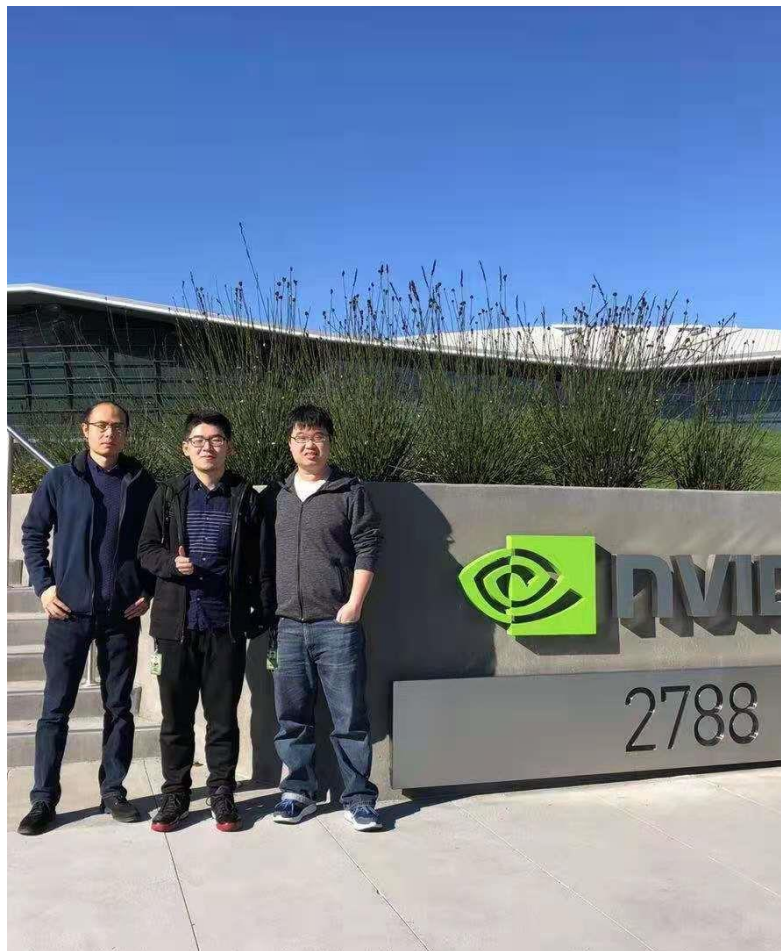
About Me



About Me



About Me



Today

- What is this class about?
- What are you getting into?
- What are you getting out of it?
- What are the expectations?

About

- 人工智能基礎概念; Fundamental Concepts of Artificial Intelligence
- 人工智能的生成與發展; The Emergence and Development of AI
- 人工智能基礎理論與技術; Foundational Theories and Technologies of AI
- 人工智能於警務應用的原理(包括犯罪地圖生成、犯罪熱點預測、人臉辨識追蹤、自然語言分析等); Principles of AI Applications in Policing
- 人工智能未來發展及與社會的關係。Future Development of Artificial Intelligence and Its Relationship with Society

Tentative Schedule



Day 1 : 认知基础 + 技术原理 + 实验室初探



上午 09:00 – 13:00

| 时间 | 内容模块 | 形式/备注 |
|---------------|--|--|
| 09:00 – 09:10 | 开场 & 课程导览 | 明确“警务+AI+未来”主线，介绍课程安排 |
| 09:10 – 09:55 | 认知心理学作为智能的基础 | 人类感知、记忆、决策 → 对比机器如何“模仿” |
| 09:55 – 10:25 | AI换脸结果展示 + 伦理冲击 | 展示伪造视频 → 引导思考“眼见是否为实？” → 警务鉴伪需求 |
| 10:25 – 10:40 |  下课休息 | |
| 10:40 – 11:25 | 大模型发展简史 | 大模型时代，突出“数据驱动”转折点  为 Day 2 Data-Centric埋线 |
| 11:25 – 12:00 | 机器学习基础概念（监督/无监督/强化） | 用警务案例解释：罪案分类（监督）、热点聚类（无监督）、巡逻路径优化（强化） |
| 12:00 – 12:10 |  下课休息 | |
| 12:10 – 13:00 | 动手实验1: Prompt工程 + NLP警务实战 |  自然语言分析融入 任务1: 用Prompt让AI总结报案笔录 任务2: 分析社交媒体舆情情绪 任务3: 模拟审讯问答生成（如：“嫌疑人可能隐瞒什么？”） |

Tentative Schedule



Day 1 : 认知基础 + 技术原理 + 实验室初探



下午 14:30 – 17:30

| 时间 | 内容模块 | 形式/备注 |
|---------------|---|---|
| 14:30 – 15:15 | 神经网络与扩散模型可视化入门 | 动画演示：从像素→特征→生成图像，强调“为什么Diffusion适合生成模拟画像” |
| 15:15 – 16:00 | 动手实验2: Diffusion生成 + 3D可视化 | 生成“年龄变化嫌疑人”、“模糊监控增强图”、“三维犯罪现场重建草图” |
| 16:00 – 16:30 |  实验室参观 | 重点看:澳大NLP语音转写设备（2楼） |
| 16:30 – 17:15 | 无人机在AI系统中的角色 | 数据采集→多天气鲁棒性→语言引导→目标识别 |
| 17:15 – 17:30 |  Day 1 小结 & 明日预告 | 强调“数据是AI的燃料，模型是引擎”→引出Day 2技术深化 |

Tentative Schedule



Day 2: 身份追踪技术 + 数据与模型革新 + 未来思辨

上午 09:00 – 13:00

| 时间 | 内容模块 | 形式/备注 |
|---------------|--|---|
| 09:00 – 09:10 | 回顾 + 学员“最想用AI解决的警务痛点”分享 | 快速收集需求，为下午讨论铺垫 |
| 09:10 – 10:10 | AI于警务应用：人脸 + Person ReID 跨镜追踪技术 | <ul style="list-style-type: none">• 人脸识别（单图比对）• ReID（跨摄像头、换装、低清场景）• 实战案例：地铁逃犯追踪、商场盗窃串并案• 局限与误判风险（光照、遮挡、相似体型） |
| 10:10 – 10:40 |  下课休息 | |
| 10:40 – 11:10 | 前沿架构入门：MoE（Mixture of Experts） | <p>比喻：“一个案件，多个专家会诊”</p> <ul style="list-style-type: none">• 为什么MoE适合大模型？→ 效率高、专精模块• 警务潜在应用：不同罪案类型调用不同“专家子模型” |
| 11:10 – 11:40 | 方法论革新：Data-Centric Learning | <p>从“调模型”转向“调数据”</p> <ul style="list-style-type: none">• 案例：标注质量提升 → 模型准确率飙升• 警务启示：清洗报案记录、标准化笔录格式比换模型更重要！ |
| 11:40 - 12:00 |  下课休息 | |
| 12:00 – 13:00 | 动手实验3：设计一个“数据驱动”的AI警务流程 | 选一个场景（如“走失老人搜寻”），设计数据采集→标注→模型选择→部署流程，强调Data-Centric思维 |

Tentative Schedule



Day 2: 身份追踪技术 + 数据与模型革新 + 未来思辨



下午 14:30 – 17:30

| 时间 | 内容模块 | 形式/备注 |
|---------------|--|--|
| 14:30 – 15:15 | 因果理论：从相关性陷阱到决策依据 | 您的45分钟内容 • 警务经典误判案例（如：警车多→犯罪率高？） • 因果图入门 + 干预 vs 观测 • AI如何辅助因果推断？（简要提Do-Calculus） |
| 15:15 – 16:00 | 人工智能的未来：AGI、社会冲击、警务新角色 | • 自动驾驶责任归属 • AI法官的公平性质疑 • 警务人员：从“执行者”转向“AI监督者+伦理把关人” |
| 16:00 – 16:45 | 圆桌论坛：效率、隐私、人性 — AI警务的边界 | 辩题示例： “是否应允许AI预测某人未来犯罪概率？” “ReID追踪是否应覆盖全城摄像头？” |
| 16:45 – 17:15 |  课程总复习 + 知识地图 | 一张图串讲：认知→历史→技术→应用→数据→未来→伦理 |

Goals

1. Know AI
2. Think of AI
3. Apply AI

Raise the Interest for AI.

Understand some AI papers.

Finally...



Please raise your hand. I will illustrate again.