



“科”“技”並重下人工智能教育的探索實踐

Practices and Innovations in AI Education under the Dual Emphasis on
Science and Technology

方勁 | MR. Fang

中山市小欖中學(中山市外國語學校)“小平科技創新實驗室”

Xiaoping Technology Innovation Lab

Zhongshan Xiaolan Senior High School (Zhongshan Foreign Language School)

2025年7月3日



個人介紹

Personal Introduction



中山市小欖中學
中山市外国语学校
ZHONGSHAN XIAOLAN SENIOR HIGH SCHOOL
ZHONGSHAN FOREIGN LANGUAGE SCHOOL

計算機科學與技術碩士研究生

中山市小欖中學信息技術學科組長

廣東省高中信息技術學科教研基地成員

中山市“劉建華名師工作室”成員

長期致力於高中階段人工智能教育，指導學生在科技創新領域獲各級獎項200餘項，其中國家級獎項8項，省級30餘人次



個人榮譽

- 全國中小學實驗教學能手
- 廣東省十佳科技輔導員
- 廣東省科技勞動教育實踐活動優秀指導教師
- 中山市優秀教師
- 中山市學生科技創新大賽優秀指導教師

主要科研成果

- 發表計算機類SCI一篇，國際會議論文一篇，發表教育類論文國家級論文3篇，省級論文5篇
- 2項發明專利，1項軟體著作權
- 主持（參與）國家級課題1項，省級課題2項，市級課題4項
- 編寫出版教材《人工智能教育》並通過審定

目錄

1

政策引領與發展
趨勢

2

雙軌課堂建設

3

實驗室課程體系
的探索和建設

4

小平科技創新實
驗室運行機制

5

實踐與成果展示

CONTENTS



一、國際人工智能教育政策

● 美國

- 2023年，美國白宮發布最新版《國家人工智能研發戰略計劃》，將AI素養和K-12人才培養納入核心目標^[1]。
- 2019年，AI4K12首次發布了《K-12人工智能五大核心理念》框架。2020年以來，項目持續開發教師指南和配套教學資源，推動AI素養教育在美國多個州和學區的K-12課程標準與教師培訓中廣泛落地，成為美國基礎教育領域具有重要影響力的AI教育標準之一，由NSF與CSTA等專業團體協同推動^[2]。

[1] National Science and Technology Council, "The National Artificial Intelligence Research and Development Strategic Plan: 2023 Update," 2023.

[2] D. Touretzky, C. Gardner-McCune, F. Martin, and D. Seehorn, "Envisioning AI for K-12: What Should Every Child Know about AI?" , AAAI, vol. 33, no. 01, pp. 9795-9799, Jul. 2019.

[3] European Commission, "Digital Education Action Plan (2021–2027)," Oct. 2021.

● 歐盟

- 2021年歐盟啟動《數字教育行動計劃2021-2027》，將AI與數據素養正式納入K-12教育核心，通過倫理指南、教師培訓與AI教育內容開發推動AI課程落地^{[3][4]}。
- 同期，Intel等企業與歐盟多國合作開展AI4Youth項目，在多國通過實踐課程與競賽驗證AI課程體系，強調跨學科融合和社會創新，並被納入UNESCO UNEVOC案例庫^{[5][6][7]}。

[4] European Commission – Directorate-General for Communications Networks, Content & Technology, "Ethical guidelines on the use of AI and data in teaching and learning," Dec. 16, 2022.

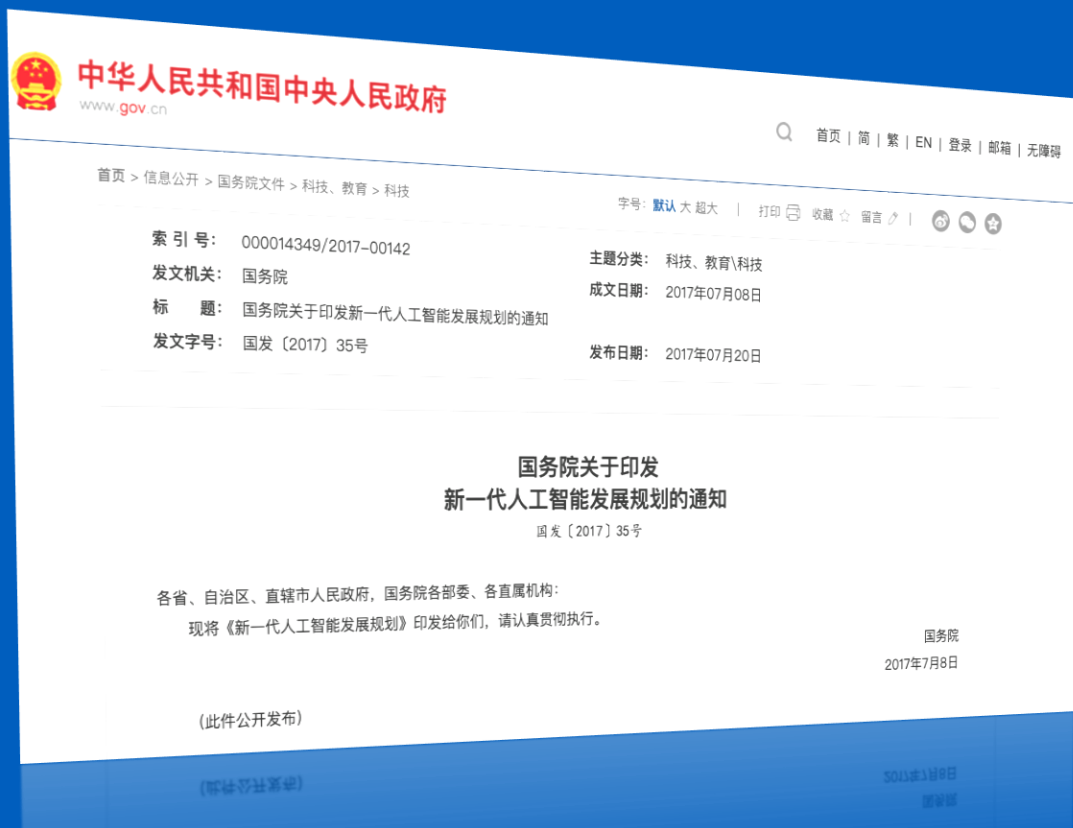
[5] Intel, "AI4Youth program" (EU pilot countries), 2020–2023.

[6] PSNC & Intel, "Final gala of the 'AI4Youth Artificial Intelligence' competition," Nov. 3, 2022.

[7] PSNC, "AI4Youth included in UNESCO-UNEVOC list," May 7, 2024.



二、國內人工智能教育政策



國務院印發《新一代人工智能發展規劃（2017）

明確提出“推動人工智能教育向基礎教育階段延伸”，支持

編寫AI教材、建設AI實驗室、培養師資。



二、國內人工智能教育政策

普通高中 信息技术课程标准

(2017年版 2020年修订)

中华人民共和国教育部制定

人民教育出版社
·北京·

义务教育 信息技术课程标准

(2022年版)

中华人民共和国教育部制定

北京师范大学出版集团
北京师范大学出版社

普通高中信息技術課程標準 (2020 修訂)

- 選擇性必修4《人工智能初步》
- 理解 AI 基本概念與方法，應用 AI 解決問題
- 開展項目式實踐與選修課程

義務教育信息科技課程標準 (2022 版)

- 將人工智能設為六大主線之一
- 各學段縱貫人工智能知識、技能與倫理
- 強調人工智能社會影響、跨學科與創新實踐



二、國內人工智能教育政策

教育部办公厅

教基厅函〔2024〕32号

教育部办公厅关于加强中小学 人工智能教育的通知

各省、自治区、直辖市教育厅（教委），新疆生产建设兵团教育局，部属各高等学校、部省合建各高等学校：

为深入贯彻党的二十大和二十届二中、三中全会精神，认真落实习近平总书记在全国教育大会上的重要讲话精神，为促进新质生产力发展储备人才，探索中小学人工智能教育的实施路径，培育具有创新潜质的青少年群体，结合当前实际，现就有关事项通知如下。

一、总体要求

（一）坚持立德树人，把准方向。全面贯彻党的教育方针，落实立德树人根本任务，紧扣新时代新征程教育使命，满足面向未来的创新型人才培养需求，确保党的事业和社会主义现代化强国建设后继有人。

（二）坚持以人为本，全面发展。遵循教育规律和人才成长规律，以人工智能引领构建以人为本的创新教育生态，为促进学

教育部部署加強中小學人工智能教育

- 系統化課程體系 & 常態化教學評價（小學—感知；初中—理解與應用；高中—項目創作）
- 設立人工教育基地，高校/企業資源向中小學開放
- 增強基礎設施 & 師資培訓，專家參與提升教師AI能力
- 採用項目式教學方式，高中學生聚焦AI應用創新
- 目標到2030年實現中小學AI教育基本普及



二、國內人工智能教育政策

广东省教育厅
DEPARTMENT OF EDUCATION OF GUANGDONG PROVINCE

首页 教育资讯 政务公开 政务服务 互动交流 专题专栏 请输入您想查询的内容

首页 > 政务公开 > 教育政策法规 > 地方教育政策法规

广东省教育厅关于印发《广东省中小学教师人工智能素养框架（试行）》《广东省中小学学生人工智能素养框架（试行）》《广东省中小学人工智能课程指导纲要（试行）》以及《国家中小学智慧教育平台与人工智能融合应用指南（试行）》的通知

时间：2025-04-10 11:12:22 资料来源：本网

【打印】 【小 中 大】 分享到：

粤教基函〔2025〕12号

各地级以上市教育局，横琴粤澳深度合作区民生事务局，省属普通中小学校：

现将《广东省中小学教师人工智能素养框架（试行）》《广东省中小学学生人工智能素养框架（试行）》《广东省中小学人工智能课程指导纲要（试行）》以及《国家中小学智慧教育平台与人工智能融合应用指南（试行）》印发给你们，请认真组织实施。

此致
广东省教育厅
2025年4月10日

粤教基函〔2025〕12号

廣東省中小學AI教育211落地架構

- 《廣東省中小學教師人工智能素養**框架**(試行)》
- 《廣東省中小學學生人工智能素養**框架**(試行)》
- 《廣東省中小學人工智能課程**指導綱要**(試行)》
- 《**國家**中小學智慧教育平臺與人工智能融合應用**指南**(試行)》

推出廣東省中小學“AI教育211”落地架構，打出一套人工智能教育課程設置、資源開發、師資建設、素養提升、育人賦能的組合拳。



三、國內外人工智能教育趨勢

維度

發展趨勢與核心內容

政策導向

各國/地區**高度重視AI教育**，制定專門政策與發展戰略，將**人工智能素養納入基礎教育核心目標**，推動課程標準更新，明確分階段培養目標，重視公平和倫理規範。

課程體系

構建遞進式課程體系，人工智能內容從小學啟蒙、初中普及到高中項目實踐，內容涵蓋AI基本概念、編程、機器學習、智能應用及倫理思考，強調跨學科融合與創新能力培養。

師資建設

系統加強教師培訓與專業發展，推動AI素養納入師範教育，開發在線線下教師培訓項目，高校、企業、科研機構深度參與教師培養，提升教師AI理論與實踐指導能力。

教學資源

推動數字化、開放共享的AI教育資源平臺建設，開發適合中小學生的教材、軟件、實驗工具和項目案例，鼓勵產教融合與校企合作，促進優質資源普及與區域均衡。

教學方式

廣泛采用項目制、探究式、跨學科融合教學模式，注重動手實踐、問題解決與創新能力提升，組織AI創客活動、競賽與展示，強調學生主動探索與團隊合作。

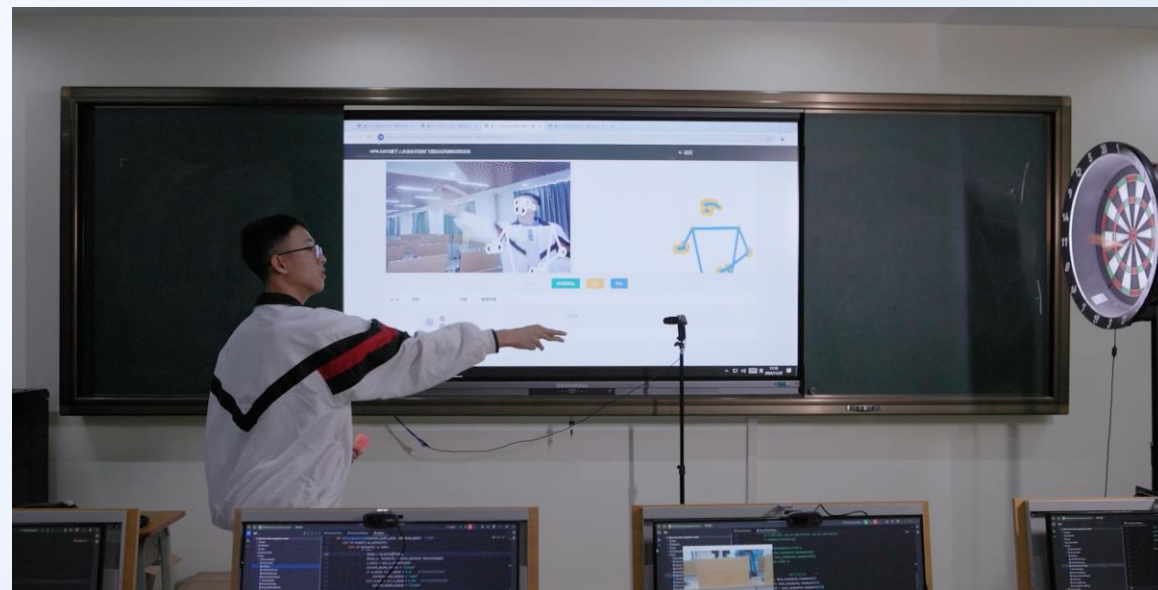
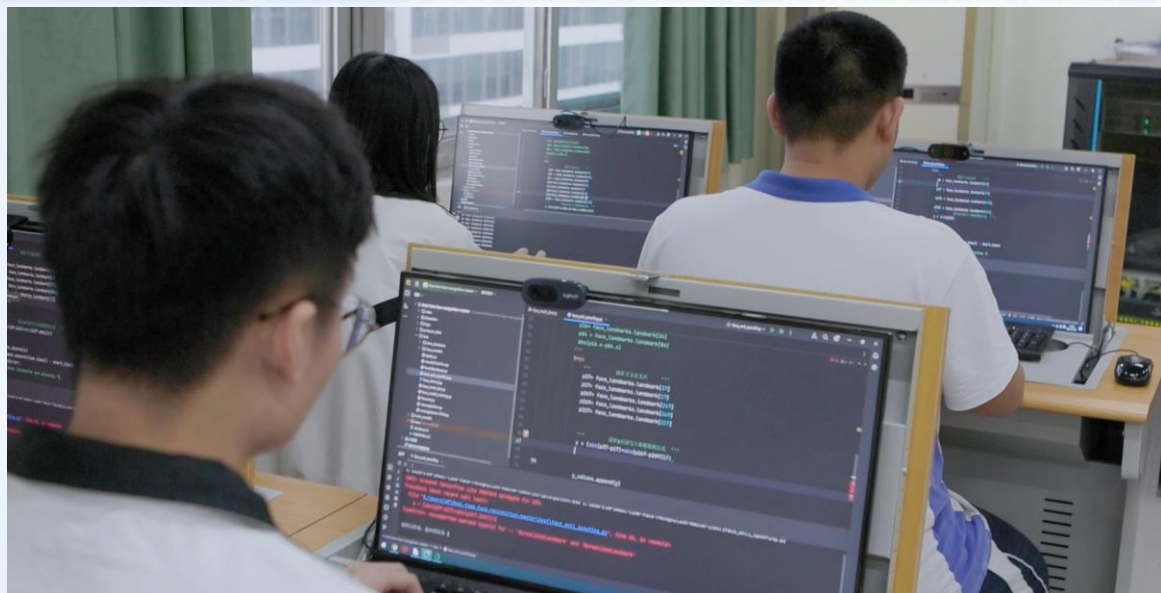
評價體系

構建多元、能力導向的AI教育評價體系，過程性評價與成果展示相結合，關注學生的AI知識、實踐能力、創新意識與倫理素養，逐步納入綜合素質評價和畢業要求。



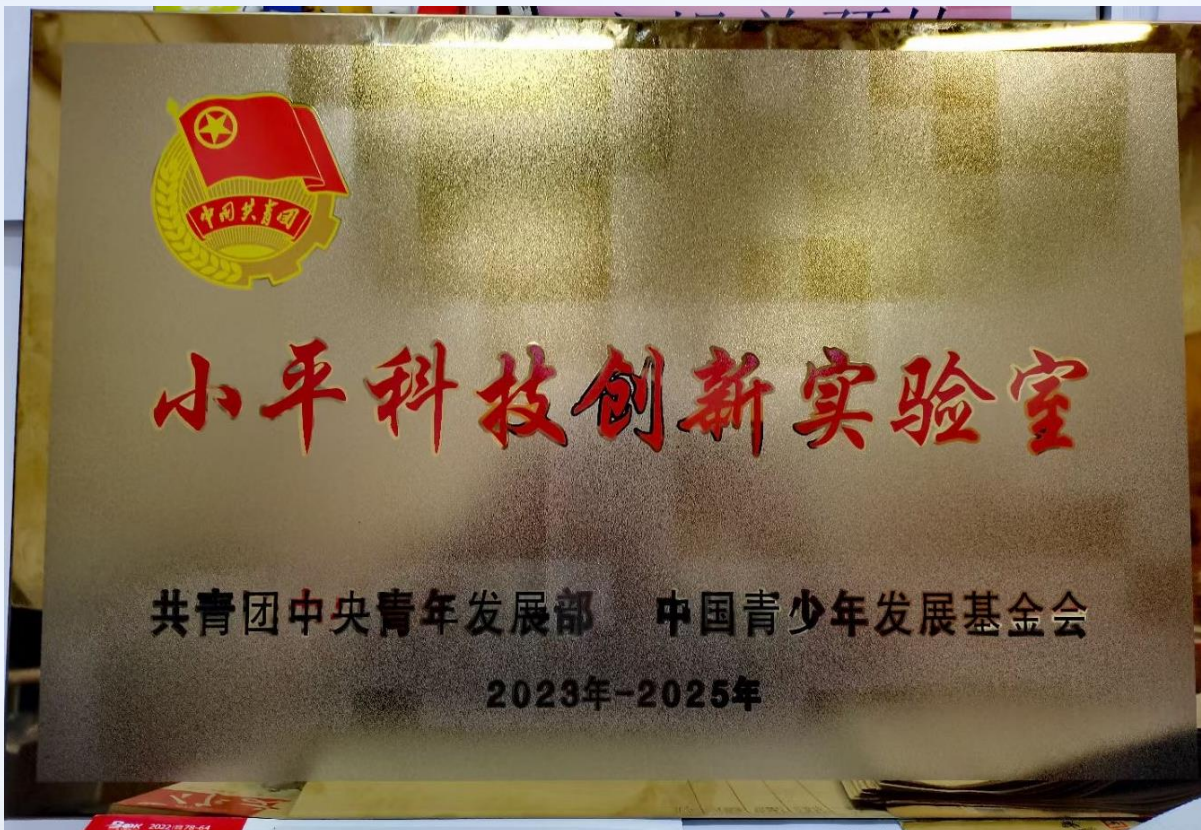
四、我校AI教育的發展

從2018年開始從零創建了學校的科創社團，逐步從興趣社團向 **人工智能** 科研實驗室發展

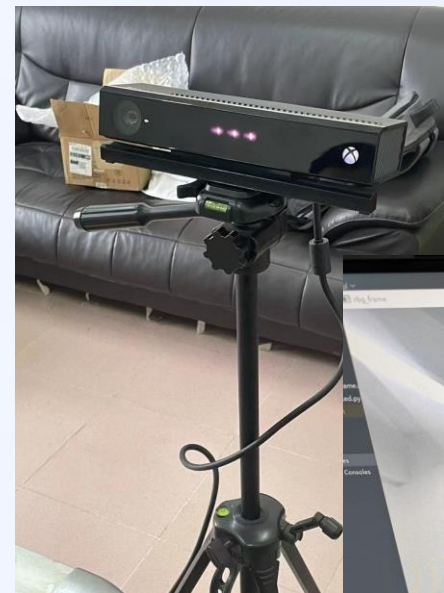




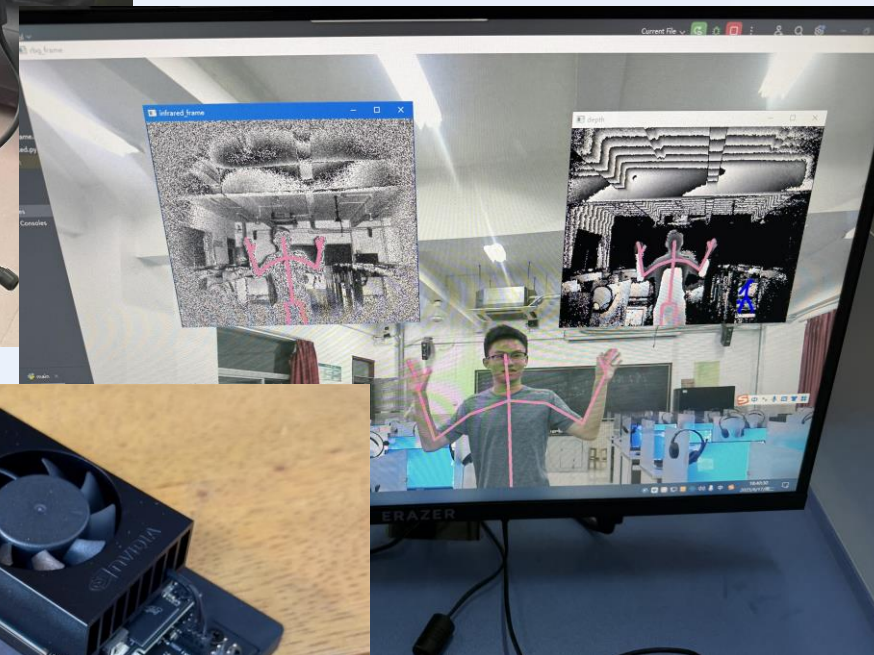
四、我校AI教育的發展



2023年被團中央授予“小平科技創新實驗室”



Kinect2.0



Jetson Nano

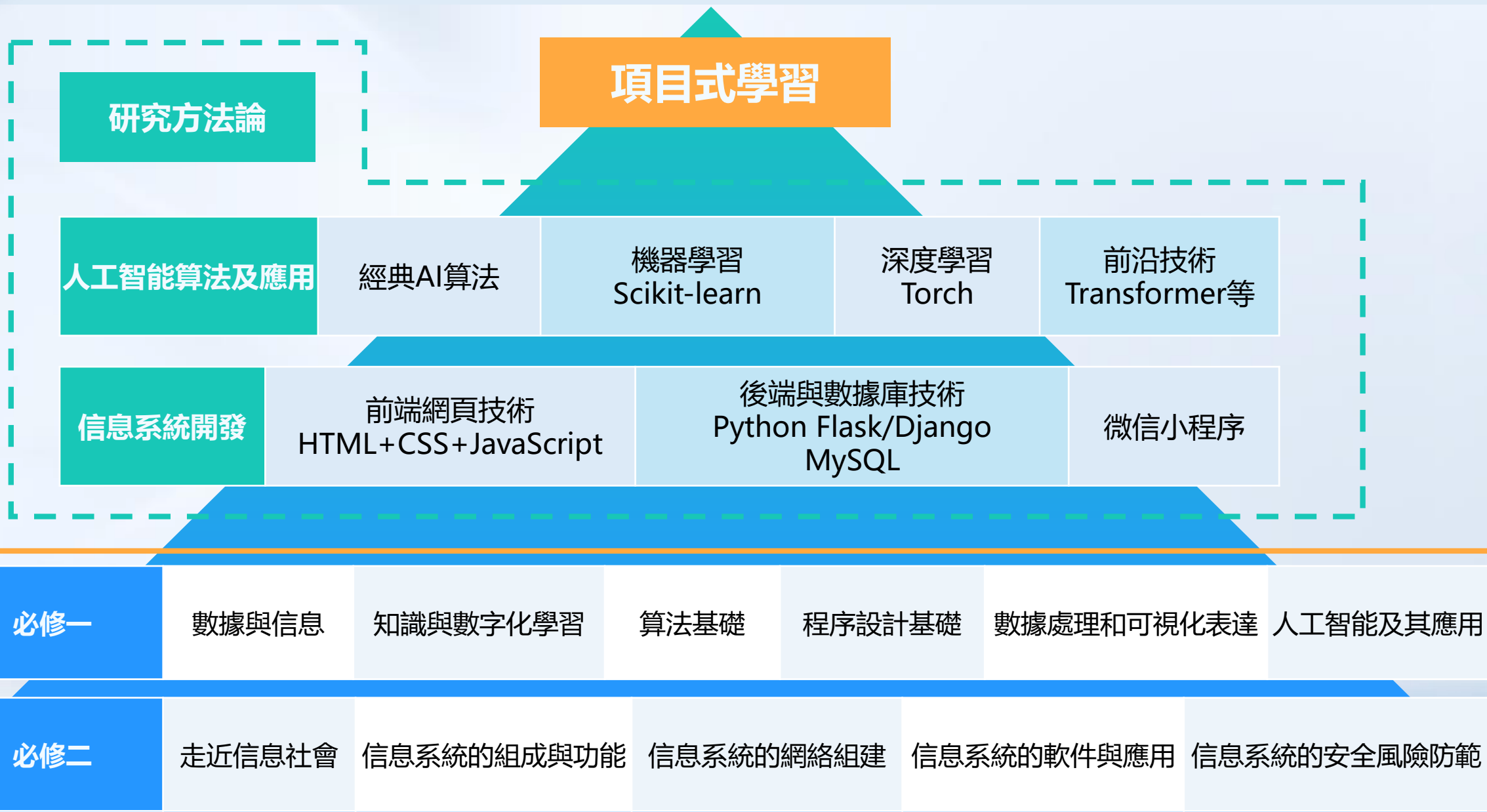
配備更多前沿硬設備



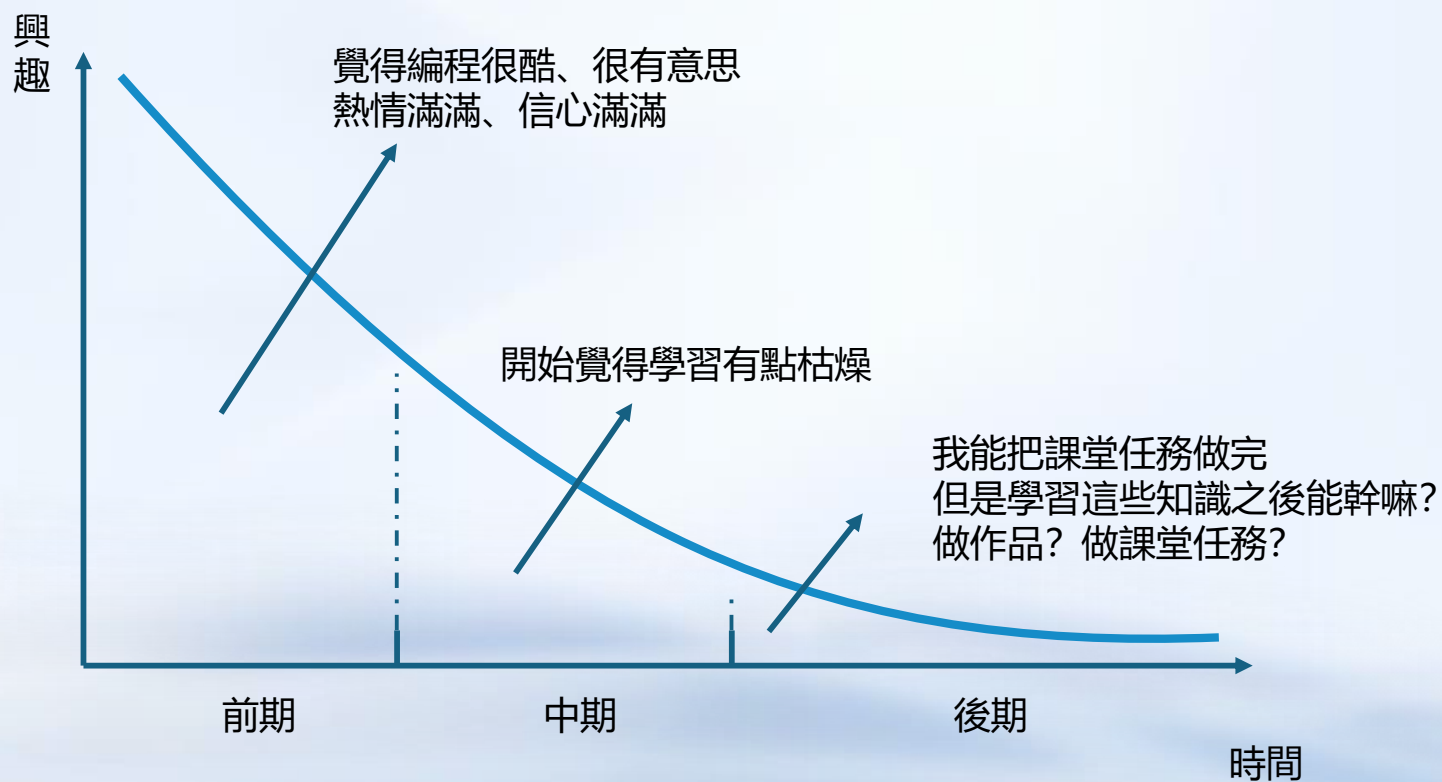
雙軌課程，根深葉茂。

以國家課程奠定基礎，以校本課程促進創新，協同推進信息科技與人工智能教育高質量發展。





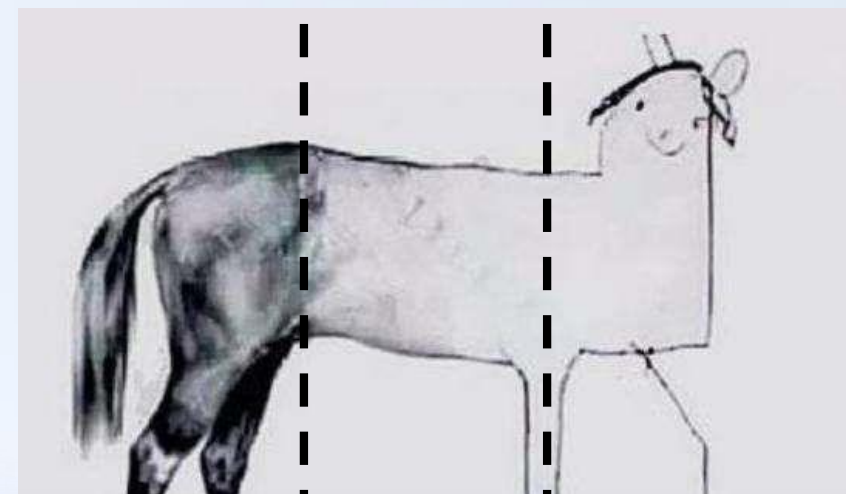
在一線教學過程中，我們不難發現一種現象：



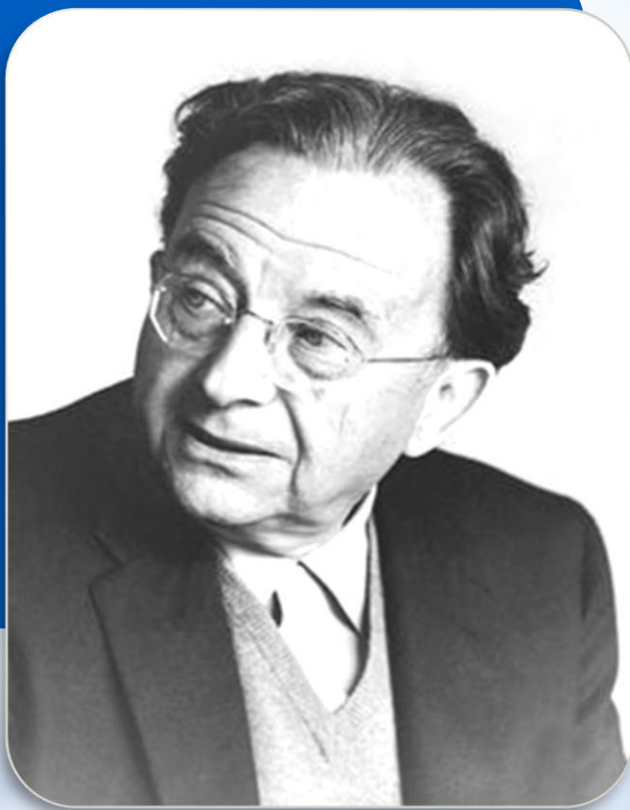
初學
課堂小任務

中期
課堂任務

制作項目/作品



如何扭轉學生學習興趣不足的變化過程？



David Paul Ausubel
(October 25, 1918 – July 9, 2008)

Ausubel 在《Educational Psychology: A Cognitive View》中指出，學校情境中的成就動機至少包括三方面的內驅力^[8]，即：



認知內驅力 (Cognitive Drive)

——對理解、知識掌握的內在渴求



自我內驅力 (Self-Improvement Drive)

——通過學習實現自我提升和勝任感



附屬內驅力 (Affiliative Drive)

——通過學習獲得他人認同、歸屬感與社會地位

[8] D. P. Ausubel, Educational Psychology: A Cognitive View, New York: Holt, Rinehart & Winston, 1968.

激發學生編程學習的認知內驅力的策略



1. 注重實時成效反饋，增強學習獲得感

通過及時呈現編程成果和可視化效果，讓學生直觀體驗知識應用價值，激發求知欲與學習興趣。

2. 設計貼近真實情境的任務與案例

將編程任務與學生日常生活、社會實際問題緊密結合，提升學習的現實意義，增強自主探索的動力。

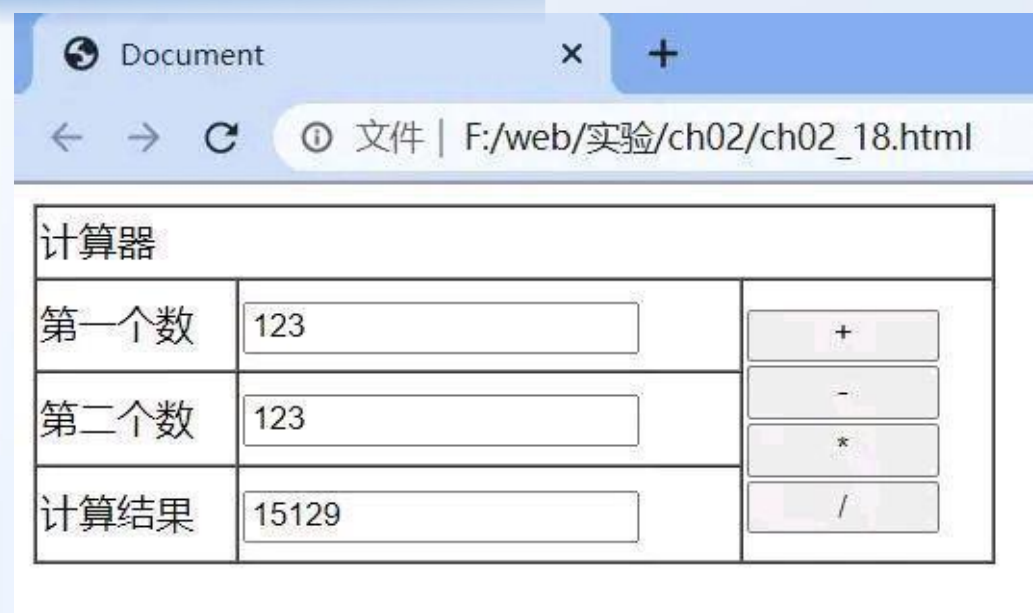
3. 加強獨立解決問題能力的培養

系統訓練學生自主分析、獨立解決問題的能力，鼓勵探索不同算法和編程思路。引導學生有效利用生成式AI作為學習輔助工具，提升自主攻關複雜問題的能力，逐步形成持續內驅的學習習慣。



使用C語言和網頁編程制作的計算器

```
C:\Users\Administrator.XB-201707021900\Desktop\c\简易计算器\项目1.exe
简易计算器
请选择加减乘除算法 + - * / :+
请输入第一个数:20
请输入相加数:50
结果 = 70
请选择加减乘除算法 + - * / :-
请输入第一个数:70
请输入被减数:10
结果 = 60
请选择加减乘除算法 + - * / :*
请输入第一个数:900
请输入相乘数:2
结果 = 1800
请选择加减乘除算法 + - * / :/
请输入第一个数:1000
请输入被除数:10
结果 = 100
请选择加减乘除算法 + - * / :.
```



教學實踐中發現，使用網頁編程進行入門能讓學生更直觀地體驗知識應用



使用網頁呈現 AI預測車流量→紅綠燈配時建模

本项目通过部署到云服务器上，所以要运行网站版需要[联网](#)

数据集下载

下面我们已经做好的训练集。若想了解我们训练集的数据是怎么形成的，可以查看我们的[说明文档](#)。

数据集下载: [点击下载](#)

数据集上传

[上传东侧数据集](#) [上传西侧数据集](#) [上传南侧数据集](#) [上传北侧数据集](#)

配时策略生成

[配时策略生成](#)

当前实时4个路口的红绿灯的情况

[点击运行](#)

当前时间:2025-06-18星期三 09:05:01

北
07

西
07

东
07

南
07

The interface displays a traffic light status diagram with four directions: North (北), South (南), West (西), and East (东). Each direction has a 07-second cycle. The North and South directions show a red light (red circle) and two black circles, indicating they are currently red. The West and East directions show two black circles and a green light (green circle), indicating they are currently green.

AI項目需要可視化的方式呈現



以信息系統開發為入門的人工智能實驗室課程



引言

第一章：人工智能导论

第一节：什么是人工智能？强AI vs 弱AI

第二节：AI发展简史：从符号主义到深度学习、大模型

第三节：AI的主要分支与应用领域概览

第二章：问题求解与启发式搜索

第一节：问题建模与状态空间表示

第二节：常见搜索算法

第三节：案例：八数码问题

第四节：案例：迷宫寻路问题 (Maze Solving)

第五节：进阶专题——图染色问题与排监考建模

第三章：博弈搜索与智能决策

第一节：博弈树构建与分析

第二节：极大极小算法 (Minimax)

第三节： α - β 剪枝优化

第四节：实战案例-井字棋AI对弈程序

第四章：智能优化算法与组合问题求解 (多解空间全局优化类)

第一节：组合优化问题与建模基础

第二节：贪心算法 —— 局部策略的快速求解方法

第三节：爬山算法 —— 直冲最优的局部搜索法

第四节：模拟退火算法 —— 概率扰动跳出局部最优

第五节：遗传算法 —— 模拟进化寻找最优解

第六节：蚁群算法 —— 模拟自然界的群体智慧

第七节：小结与算法选择策略

前置学习：Python数据分析与可视化基础

一、课程导引：为什么一定要学数据分析工具？

- AI课程不只是写算法，更关键的是：看懂、整理、分析数据。
- 机器学习和深度学习实验几乎全程需要 pandas/numpy 做数据处理，matplotlib/seaborn 做可视化展示。
- 实际数据几乎全是Excel/CSV格式，只有会处理多表/多sheet/多类型，才能真正分析清楚。
- 学好这一节，后续所有 AI、神经网络、PyTorch 项目都能上手快、调试顺。
- 看得懂数据 ≠ 只会print前几行，而是能高效筛查异常、统计规律、汇总多张表。
- 所有AI实验、论文复现、课题实践，第一步永远是“数据清洗”。

二、Python数据科学的四大法宝简介

1. numpy (np) —— 数字世界的基础

- 用于高效处理大量数字数据 (如向量、矩阵、图片像素)。
- 后续AI算法的“底层引擎”。
- list 慢、功能少，np.array 快100倍，支持数学表达式直接批量操作，适合大数据。
- numpy内置随机数、线性代数、数组变换、广播机制，非常适合图像/信号/数据增强等AI任务。

2. pandas (pd) —— 操作表格像Excel一样强大

- 用于读取/写入CSV/Excel、数据筛选、批量处理。
- 做数据预处理、数据分析必备。
- DataFrame 是“电子表格”对象，最像Excel的结构。
- 支持多Sheet读取、批量合并、多级索引、缺失值智能处理、透视表生成，极大。
- 可以轻松解决“Excel操作不便的100个问题”，比如数据去重、自动化筛选、批量替换、复杂分组统计等。

3. matplotlib (plt) —— 基础画图神器

- 画折线图、直方图、散点图等。
- AI实验结果常用。

4. seaborn (sns) —— 让统计图表又美又快

- 在matplotlib基础上美化统计类可视化，适合“看分布/对比/分类”场景。
- seaborn支持“一行代码”出箱线图、热力图、带回归分析的散点图，极适合快速数据探索与结果汇报。

Numpy

Pandas

Matplotlib

Seaborn

經典智能算法

數據分析與可視化前置學習



第五章：机器学习入门

前置学习：Python数据分析与可视化基础

第一节：走进机器学习的真实世界

第二节：三大任务类型——分类、回归与聚类

第三节：分类任务详解——让AI学会“分门别类”

一、分类任务定义与典型应用

二、分类数据集结构及特征分析

三、特征向量、特征空间与分类器决策流程

四、K近邻算法（KNN）——淘宝购物推荐背后的秘密

五：朴素贝叶斯（Naive Bayes）——用概率分类的生活智慧

六、决策树（Decision Tree）——用“树结构”解释你的每一步决策

七、逻辑回归（Logistic Regression）——二分类问题的概率解法

八、支持向量机（SVM）——最大间隔与核方法的分类利器

第四节：回归任务详解——让AI学会“数值预测”

一、回归任务的定义与生活案例

二、回归数据集结构及特征分析

三、回归算法基础原理

四、多项式回归与非线性回归

五、其他主流回归算法简介

六、正则化回归（岭回归、Lasso）（可选/拓展）

第五节：聚类任务详解——让AI自己“分群分组”

一、聚类任务的定义与生活化案例

二、聚类与分类的区别

三、聚类数据集结构及特征分析

四、聚类算法基础原理与常见类型

第六节：集成模型（Ensemble Models）——群体智慧的预测新高度

第七节：模型训练与评估——让AI更“聪明”

一、模型训练的标准流程

二、分类模型评估指标

三、回归模型评估指标

四、聚类任务评估指标

五、交叉验证与模型泛化能力

六、超参数调优方法（Hyperparameter Tuning）

七、训练过程监控与模型选择

分類
Classification

回歸
Regression

聚類
Clustering

模型評估
Clustering

機器學習算法
Scikit-learn



深度學習入門
Scikit-learn入門 → Torch

主流深度學習網絡
CNN、RNN、LSTM

前沿深度學習網絡
Transformer與Attention等

🤖 第六章：深度学习基础

🚀 前言：如何更简单地学习本章内容

🧠 第一节：深度学习发展简史与核心概念

🎨 第二节：（图解篇）神经网络工作原理图解

🧠 第三节：（图解篇）神经网络学习原理

🧠 第四节：（理论篇）多层感知机与神经元结构

🔧 第五节：（理论篇）正向传播、反向传播与训练机制

⚙️ 第六节：常用优化器与正则化

📝 第七节：手写数字识别（MNIST）动手实践（一）

📝 第七节：手写数字识别（MNIST）动手实践（二）

🏠 第七章：主流深度学习网络与自主搭建

📦 第八章：Transformer与自注意力机制

🎨 第九章：生成式AI与扩散模型（Diffusion）

🔧 第十章：常用AI工具与开源平台实践



1.2 正向传播的数学推导

我们用一个简单的两层神经网络（一个隐藏层）进行说明：

- 假设输入数据为向量 $\mathbf{x} = [x_1, x_2, \dots, x_n]^T$
- 隐藏层神经元数量为 m
- 输出层神经元数量为 k

则有：

输入层到隐藏层的计算公式为：

$$z^{(1)} = W^{(1)}x + b^{(1)}$$

- 其中， $W^{(1)}$ 是隐藏层权重矩阵，维度为 $(m \times n)$ ， $b^{(1)}$ 是偏置项，维度为 $(m \times 1)$ 。

再通过激活函数（如Sigmoid、ReLU）处理：

$$a^{(1)} = f(z^{(1)})$$

隐藏层到输出层的计算公式为：

$$z^{(2)} = W^{(2)}a^{(1)} + b^{(2)}$$

- 其中， $W^{(2)}$ 是输出层权重矩阵，维度为 $(k \times m)$ ， $b^{(2)}$ 是偏置项，维度为 $(k \times 1)$ 。

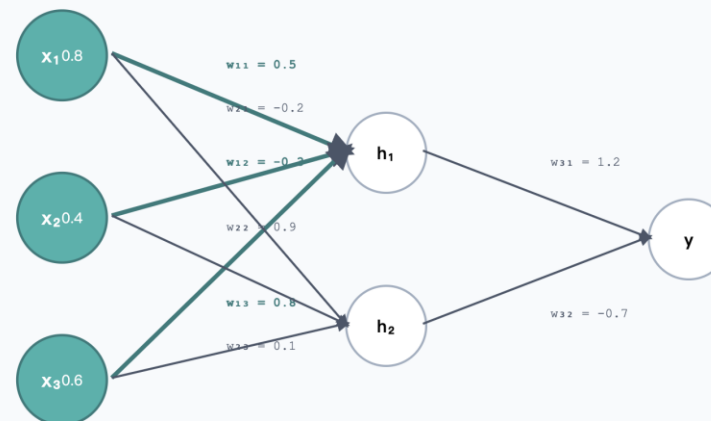
输出层激活函数后得到最终结果：

$$\hat{y} = f(z^{(2)})$$

此过程为一次完整的正向传播。



MLP 前向传播详细计算演示



h1 线性组合：
 $z_1 = 0.8 \times 0.5 + 0.4 \times (-0.3) + 0.6 \times 0.8 + 0.1 = 0.860$

上一步

下一步

從數學上加深對人工智能算法的理解



```

Python | 运行代码 | 复制代码
1  from sklearn.neural_network import MLPClassifier
2
3  mlp = MLPClassifier(
4      hidden_layer_sizes=(128, 64),      # 2层: 128、64
5      activation='relu',
6      solver='adam',
7      alpha=1e-4,
8      learning_rate_init=0.001,
9      max_iter=30,
10     early_stopping=True,
11     random_state=42
12 )
  
```

• 隐藏层结构如何选择?

- “逐层递减” (如128, 64) 能让模型的信息逐步抽象、去冗余, 最后只留下最关键的高阶特征, 有助于泛化、防过拟合。你可以尝试(128, 128)、(256, 128, 64)、(100,) , 甚至(64, 32)等不同组合, 然后比较准确率、loss曲线和训练耗时。
- 如果后面层比前面更宽, 模型会引入大量多余参数, 增加过拟合风险, 也更难训练。当然, 也可以用(128, 128)等结构做实验, 有时在大样本下宽层能提升表达力, 但要用验证集判定其合理性。

• 激活函数的作用?

- 激活函数决定网络的非线性能力。ReLU计算快、梯度不易消失, 是现代深度网络的首选; tanh和sigmoid历史悠久, 但层数多时易梯度消失。建议用'relu'和'tanh'各做一次实验对比。

• 正则化 (alpha) 为什么重要?

- L2正则化参数过小, 约束不够, 模型容易过拟合 (训练集效果好, 测试集差); L2正则化过大则抑制过头, 导致欠拟合 (模型表达力不足, 训练集和测试集都低)。需要结合训练/测试表现和loss曲线调优。你可以尝试alpha=1e-5、1e-4、1e-3, 观察不同强度对结果的影响。

• early_stopping=True自动早停有什么好处?

- 自动早停可以在模型在验证集上准确率不再提升时及时停止训练, 防止模型“背死书”, 大大降低过拟合风险, 也节省了训练时间。是神经网络工程中常用的自动防过拟合手段。

思考: 仅看准确率就够了吗? 为什么还要关注precision/recall/F1?

请大家回顾第五章第七节的相关部分内容:

二、分类模型评估指标

三、回归模型评估指标

你可以回顾:

- 准确率**: 整体预测正确的比例, 适合类别分布均衡时的主指标。
- 精度 (Precision)**: 关注被判为某类的有多少是真正的那一类, 查准率高说明“猜测”不会乱判。
- 召回率 (Recall)**: 关注某一类别实际样本中有多少被检出, 查全率高说明“没有漏掉太多”。
- F1分数**: 综合考虑精度和召回率的调和平均, 更能反映模型的均衡表现。
- 混淆矩阵**: 可以分析模型在哪些类别上容易混淆。

在本实验的多分类 (手写数字识别) 场景下, 这些指标配合使用可以帮助我们发现模型在不同数字类别上的表现差异, 发现模型“短板”, 指导下一步模型优化和调参。

类似地, 在回归任务、聚类任务中, 我们也有各自的评估指标 (如MSE、RMSE、MAE、轮廓系数等), 模型训练完成后都需要用“多种视角”评估, 不能只看一个数字。

內容體系完備，強調自主與個性化發展

課程在內容設計上系統梳理核心算法的實現細節, 深入闡釋各函數與參數的理論依據及工程意義。通過結構化、層級化的知識展開, 支持學生根據個人基礎個性化學習, 兼顧基礎性與創新性, 促進多層次能力提升與個性化發展。



信息系統

人工智能算法

開源硬件

研究方法

技能型

研究型

“科” 技” 並重的校本教材的建設

人才培養導向





三、课程结构与章节安排 📁

📖 第一章：认识研究

- ? 什么是科学研究？科学研究的本质与特点
- 🕒 实证研究 vs 探索研究 vs 理论研究
- 📦 “问题-过程-成果”三位一体结构

📖 第二章：选题与研究设计

- 🎯 如何提出一个好问题
- 🌱 研究问题、研究假设与研究目标
- 📋 研究框架与研究变量的设定
- 🛠️ 研究工具选择：问卷、访谈、实验、观察

📖 第三章：文献综述与资料收集

- 🔍 如何检索与筛选有价值的文献
- 📄 使用CNKI、Google Scholar等学术平台
- 📝 撰写高质量文献综述的结构与技巧
- 📌 引用格式与避免抄袭（APA、MLA简介）

📊 第五章：数据分析基础

- 📄 定量数据的录入与整理（Excel或SPSS）
- 📈 描述性统计指标：均值、中位数、标准差
- 📏 常用检验方法：T检验、卡方检验、相关分析
- 📊 数据可视化基础：图表选用与排版美学

📖 第六章：研究报告与论文写作

- 📄 项目研究报告结构：摘要、引言、方法、结果、讨论
- 📝 科技论文写作技巧：逻辑、图表、引用、附录
- 🖼️ 图表嵌入与统一规范排版技巧
- 💡 项目摘要的写法与申报技巧
- 🔄 反复打磨与查重修改的重要性

📖 第七章：项目展示与答辩

- 🗣️ 如何进行口头报告与PPT展示
- 🧠 演讲结构：开场、亮点、总结、提问应对
- 📄 项目海报与项目网页展示建议
- 🎭 模拟答辩训练与评审规则解析

本次调查共访问英语老师20位，问卷调查的数据如下👈

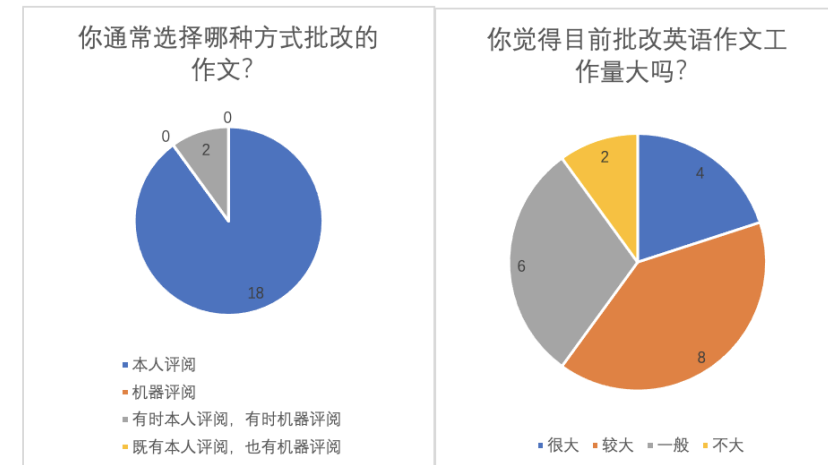


图2-2 调查问卷分析 (1)

图2-3 调查问卷分析 (2)

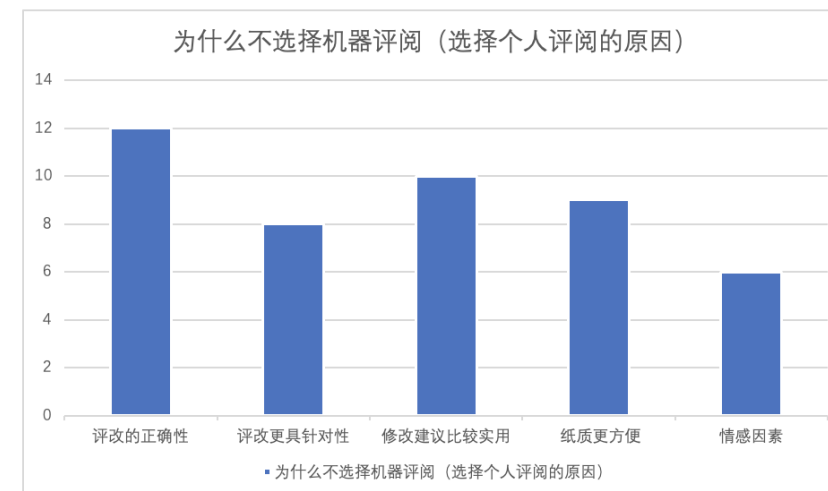


图2-4 调查问卷分析 (3)



愛上實驗課

(Scan the QR code with WeChat to follow our public account for content on experimental teaching.)





高一第一學期

- 搜索引擎的使用
- 網頁設計
- HTML
- CSS
- JavaScript

高一第二學期

網站（信息系統）開發

- HTML、CSS、JS (es6)
- jQuery、Layui、Vue
- Ajax
- Python Flask/Django
- MySQL-增刪改查

人工智能算法

經典算法

啟發式搜索

爬山、A*、模擬退火、大洪水
分類和聚類

機器學習Sklearn

線性回歸，決策樹，向量機，
Adaboost，GBDT等

人工神經網絡

DNN CNN RNN LSTM GAN

高二第一學期

- 開源硬件基礎
- 項目式學習
- 研究方法論
- 項目指導

構建“科”“技”並重的校本課程體系



紮實開展社團活動，形成有梯隊，可持續的社團發展模式。



2022年暑期集訓

左一為2021年“創新之星”何東毅暑期回校助教指導
左二為2022年“創新之星”高源泉



2023年一同前往國賽現場

左二為我校侯俊傑同學參與23年國賽現場



2024年交流

左一為2024年“創新之星”陳展濤與學長們進行交流

畢業生返校交流指導



積極探索跨學校和跨地區的合作模式，與香港基智學校、華南師範大學、重慶理工大學兩江人工智能學院、重慶理工大學AiRobot實驗室的師生進行交流。不僅擴大了“中山市小欖中學小平實驗室”的教育影響力，也促進了不同教育背景下學生和教師之間的知識與經驗交流，增強了科技教育的整體協同和創新能力。



與香港基智學校學生一起學習



與華南師範大學生參加教育創新成果博覽會



學生信息科技創新大賽

- 2020-2024年，創新開發連續4年參加全國活動，3次獲全國“創新之星”
- 近6年來，獲國家級獎項8項，省級獎項23項（省一等獎11項），市級獎項八十項。
- 在計算思維程序設計（創新應用開發、代碼編程）、3D創意設計、計算機動畫均獲得過國家級獎項以及廣東省一等獎。

2025年粵港澳學生信息科技創新大賽再獲廣東省一等獎



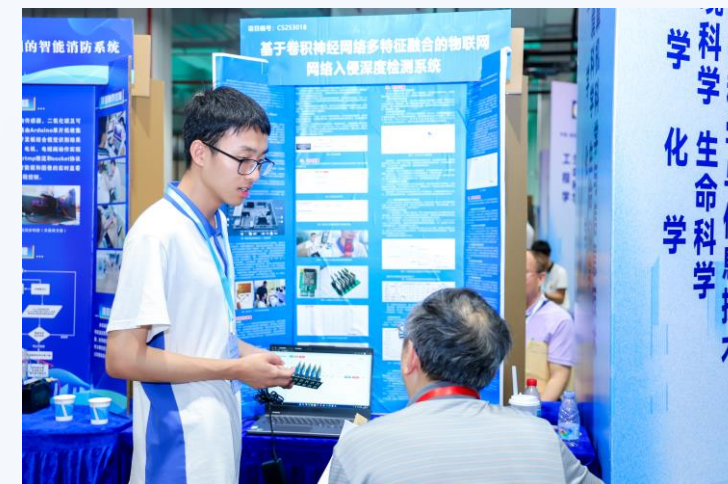




青少年科技創新大賽

- 近5年來，指導學生每年都獲得市一等獎榮譽。
- 指導學生獲省一等獎2人次，省二等獎3人次，市一等獎16人次。
- 其中2021年何東毅同學的“利用自然語言處理技術（NLP）開發的英語作文批改小程序”獲廣東省一等獎、中山大學逸仙英才獎、廣東工業大學校長創新獎。
- 獲得科技輔導員項目市一等獎

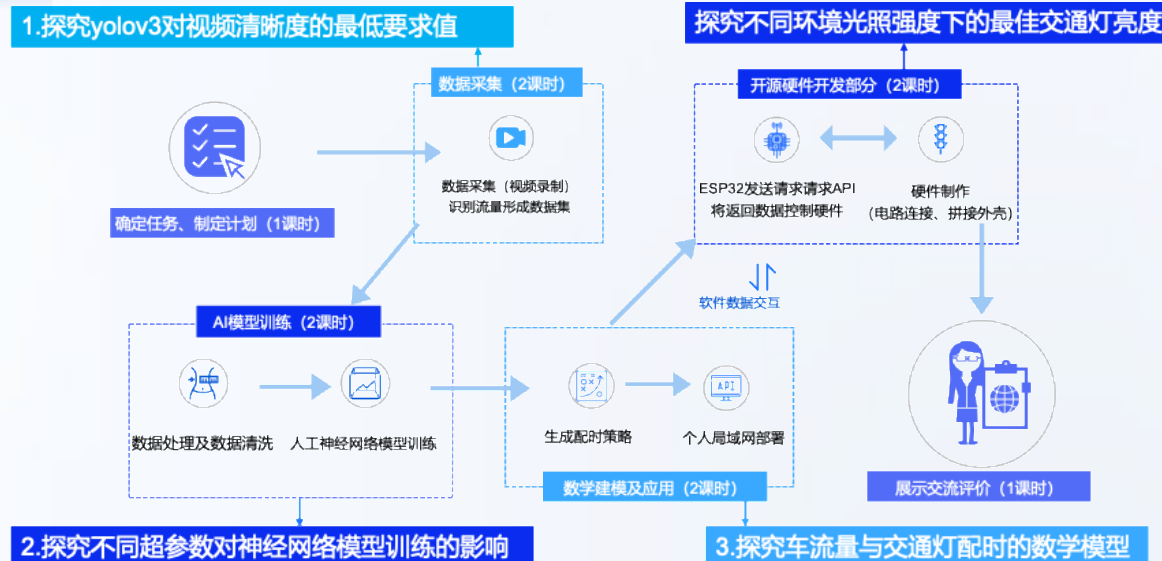






A Study on Intelligent Traffic Signal Timing Based on Traffic Flow Prediction Models

利用預測車流模型實現交通燈智能配時模型的研究



控制

ESP32

请求

{ "direction": "xx", "time": xx }

绿灯方向及时间

traffic_light_holiday					
time	east	west	south	north	
0	15	15	0	0	
15	0	0	21	21	
36	15	15	0	0	
51	0	0	21	21	
72	15	15	0	0	
87	0	0	21	21	
108	15	15	0	0	
123	0	0	21	21	
144	15	15	0	0	
159	0	0	21	21	
180	15	15	0	0	
195	0	0	21	21	
216	15	15	0	0	

搜索网格



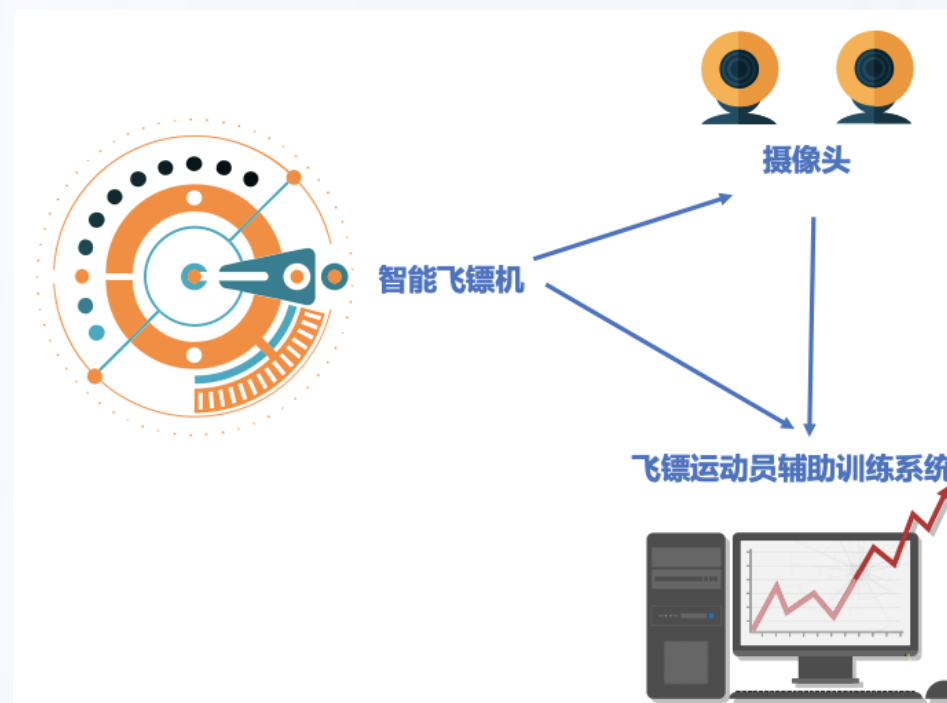
Human Pose Recognition Darts Training System(HPR-DATS) 基於人體姿態識別的飛鏢運動員輔助訓練系統



搭建飛鏢輔助訓練系統



Human Pose Recognition Darts Training System(HPR-DATS) 基於人體姿態識別的飛鏢運動員輔助訓練系統



飛鏢輔助訓練系統架構圖



Human Pose Recognition Darts Training System(HPR-DATS) 基於人體姿態識別的飛鏢運動員輔助訓練系統

信息系統開發

為每位運動員建立個人數據庫，個人數據模型獨立自訓練分析

HPR-DATS基于人体姿态识别的飞镖运动员辅助训练系统

运动员数据表

+添加运动员

ID	用户名	操作
1	陈展涛	总分析 训练数据 开始训练
5	杨风墨	总分析 训练数据 开始训练
17	彭步禹	总分析 训练数据 开始训练
18	高楚殇	总分析 训练数据 开始训练
19	谭自清	总分析 训练数据 开始训练
20	张少辞	总分析 训练数据 开始训练
21	胡见逸	总分析 训练数据 开始训练
22	张萧	总分析 训练数据 开始训练
23	孙浮衍	总分析 训练数据 开始训练
24	吕新少	总分析 训练数据 开始训练

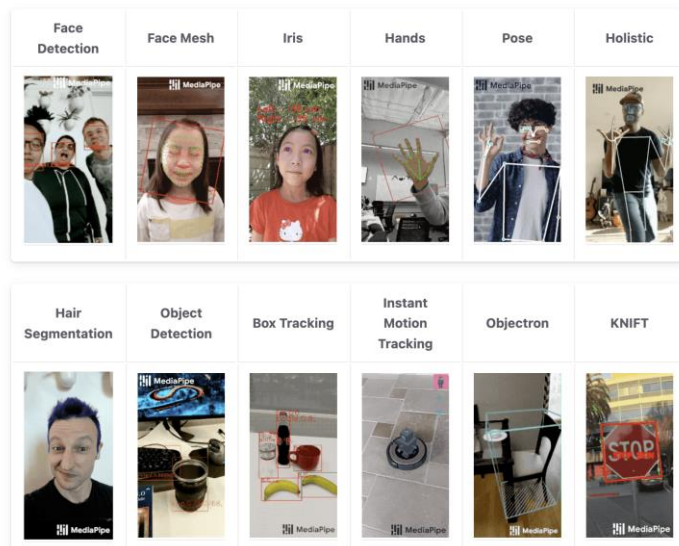
共 19 条 10 条/页





Human Pose Recognition Darts Training System(HPR-DATS) 基於人體姿態識別的飛鏢運動員輔助訓練系統

ML solutions in MediaPipe



	0	1	2
0	0.42099	0.17895	-0.00005
1	0.40627	0.13559	-0.01688
2	0.40386	0.13434	-0.01707
3	0.40106	0.13338	-0.01736
4	0.40354	0.13723	-0.05078
5	0.39959	0.13707	-0.05079
6	0.39537	0.13723	-0.05088
7	0.37047	0.15127	-0.05955
8	0.36838	0.15676	-0.20955
9	0.41280	0.22170	-0.00402
10	0.40911	0.22293	-0.04726
11	0.26402	0.34767	0.06575
12	0.38571	0.40126	-0.33877
13	0.26274	0.63885	0.28173
14	0.41907	0.76130	-0.27289
15	0.32386	0.75619	0.49017
16	0.45458	1.03425	-0.06771
17	0.33295	0.77985	0.53312
18	0.48433	1.09502	-0.07681
19	0.34588	0.76357	0.49977

去噪 (平滑)

Mediapipe Pose采集特定24個關節點

仿射變換 (平移、拉伸)

可用於後續姿態識別/分類/訓練 的數據

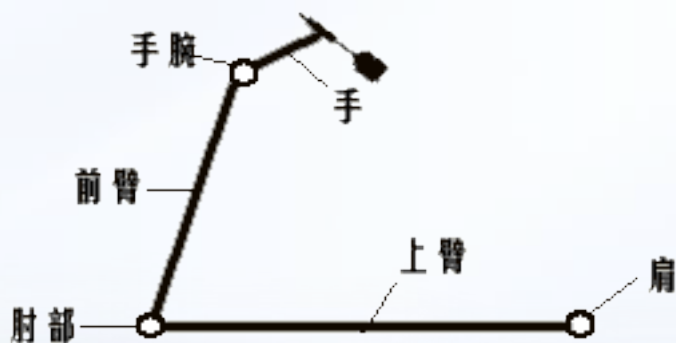
Human Pose Recognition Darts Training System(HPR-DATS)
基於人體姿態識別的飛鏢運動員輔助訓練系統

設計研究：影響飛鏢得分的因素

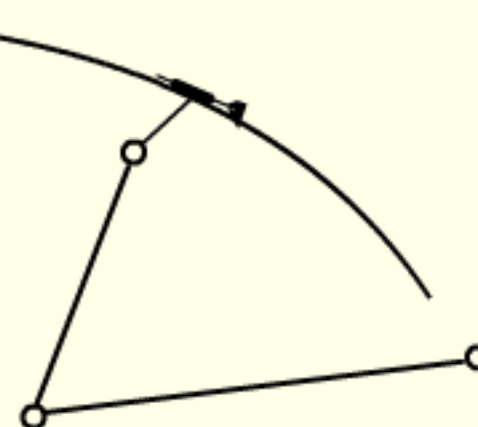
穩定性

發力

投擲軌跡



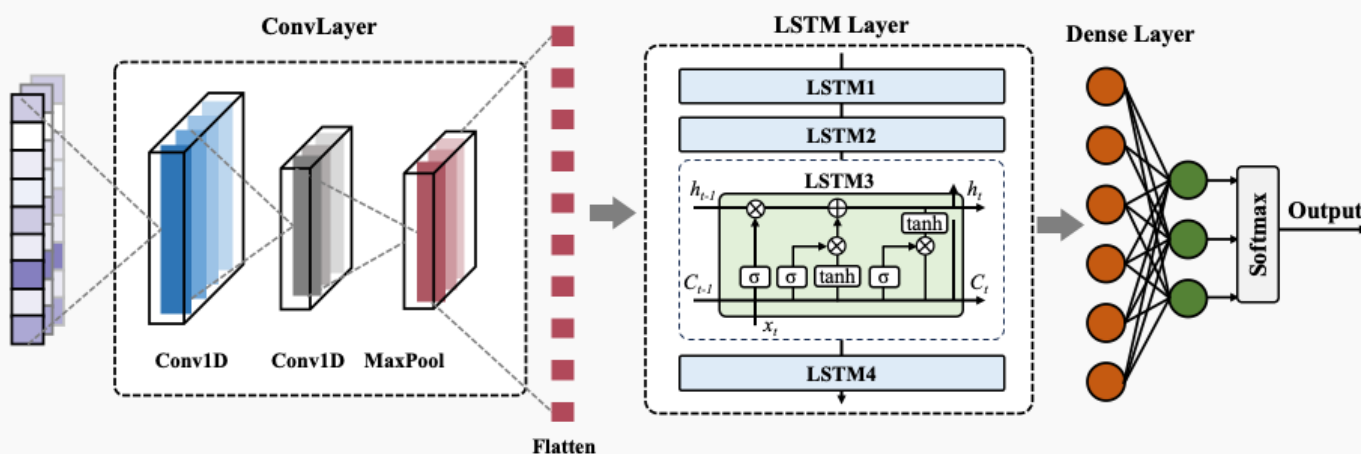
Aiming position



投擲飛鏢的“三連杆機構”以及投擲飛鏢過程

Human Pose Recognition Darts Training System(HPR-DATS)
 基於人體姿態識別的飛鏢運動員輔助訓練系統

穩定性


 自主改良的C-L-DNN神經網絡結構^[9-11]

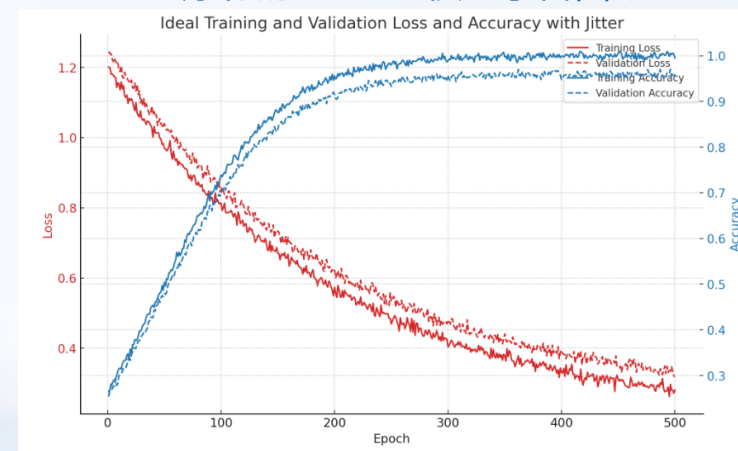
[9]巫兩環. 基於CLDNN混合模型的中文語音識別系統[D]. 電子科技大學, 2021. DOI:10.27005/d.cnki.gdzk u.2021.003113.

[10]孫藝聰,田潤瀾,王曉峰等. 基於改進CLDNN的輻射源信號識別 [J]. 系統工程與電子技術, 2021, 43 (01): 4 2-47.

[11]Emam, Ayman, et al. "A comparative study between CNN, LSTM, and CLDNN models in the con text of radio modulation classification." 2020 12th International Conference on Electrical Engineeri ng (ICEENG). IEEE, 2020.



原有的CLDNN模型學習曲線



改進後的CLDNN模型學習曲線



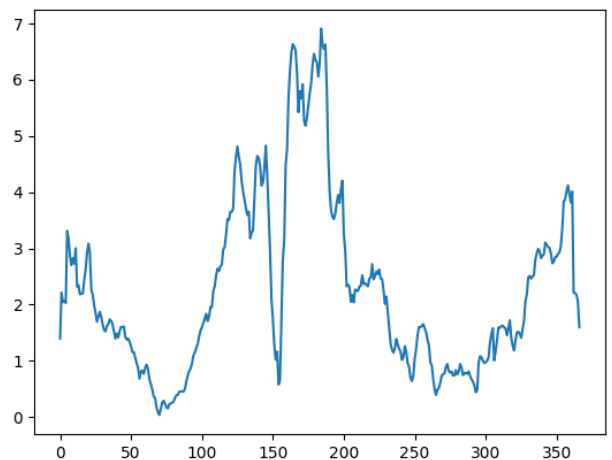
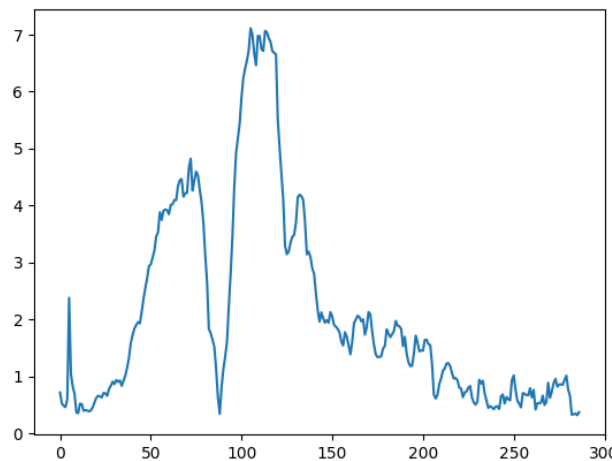
Human Pose Recognition Darts Training System(HPR-DATS)

基於人體姿態識別的飛鏢運動員輔助訓練系統

发力



運動員投擲飛鏢動圖



手腕速度變化圖

波穀到第二波峰的陡峭情況即為
加速度 (力度)。
 可有效幫助分析發力點情況



Human Pose Recognition Darts Training System(HPR-DATS) 基於人體姿態識別的飛鏢運動員輔助訓練系統

預測投擲軌跡

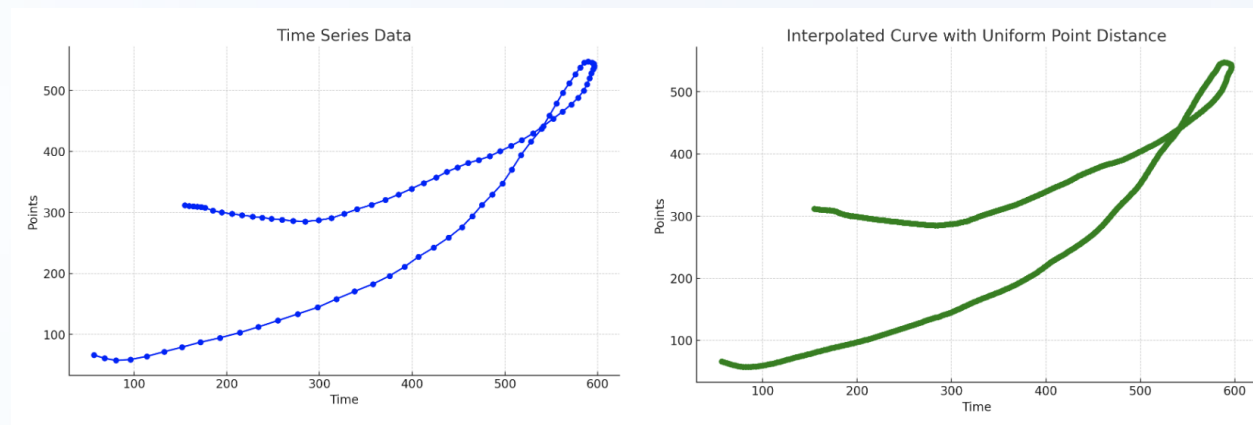
1. 軌跡重采樣 (插值)

采用**三次樣條插值 (Cubic Spline Interpolation)** 进行平滑连接离散数据点

2. 預測曲線

通過不同的機器訓練模型

選用**R2決定系數最高 (得分)** 的模型所預測的曲線。



線性回歸



決策樹



KNN



隨機森林



Adaboost



Bagging



ExtraTree
極端隨機樹

選得分最高的模型預測出曲線值



Human Pose Recognition Darts Training System(HPR-DATS) 基於人體姿態識別的飛鏢運動員輔助訓練系統

撰寫研究報告 (論文)

基于人体姿态识别的飞镖运动员辅助训练系统

【摘要】 本项目是利用智能基于人体姿态识别的飞镖运动员辅助训练系统,通过搭建软硬件系统设备,利用自主搭建的 C-L-DNN 卷积神经网络对运动员的动作姿态进行分析,同时建立个人运动数据库,能有效的辅助运动员进行飞镖运动的姿态调整,帮助他们更加科学地进行训练。

【关键词】 人体姿态识别; 飞镖; 辅助训练系统;

一、研究背景

(一) 飞镖项目的发展现状

飞镖在中国是一项新兴的体育运动,人们对它的喜爱程度并不亚于一些流行体育项目。中央电视台的“假日体育”栏目定期播放国际飞镖比赛的实况录像,给广大飞镖爱好者提供很好的学习机会。

飞镖运动在亚洲热门起来还是近几年的事,在中国,绝大多数的人对飞镖的认识还停留在硬式飞镖阶段,少数认识较深的人也是通过观看电视转播飞镖比赛来认识软式飞镖运动,尽管在外国领事馆及外国人聚集的高档酒吧可以看到电脑智慧飞镖机,但毕竟离大多数的市民还是太远,据国家体育总局飞镖协会预测,占世界人口 1/5 的中国,在不久的将来,飞镖爱好者将独占世界鳌头。

为了推动全民健身运动在全国的深入开展,进一步普及飞镖运动,国家体育总局社会体育指导中心,于 1999 年 5 月把飞镖运动列为正式体育项目,这标志着飞镖运动在中国得到了认可,成为一项名副其实的体育运动项目。飞镖已列为“2022 年课标”新兴体育类正式项目,学生参加比赛均为省学体联与省镖协主办,属于官方比赛。



图 1 2021 年广东省中学生飞镖锦标赛

目前,我校成为中山唯一一所开展飞镖项目的学校,得到了省镖协、学校的大力支持。为了更好地帮助我们学校开展飞镖项目的实践,在我们学校体育老

师和信息老师的共同指导下,我开发了一套基于人体姿态识别的飞镖运动员辅助训练系统。通过这个系统,我可以通过人体姿态识别帮助运动员矫正投掷飞镖的动作,同时为每个飞镖运动员建立个人训练数据库,他们能更好地了解自身的训练情况,也能得到体育老师更好的指导。

(二) 动作辅助系统研究现状

用计算机对动作进行识别研究从而辅助训练的系统已经得到了研究界许多专家学者的验证,市面上也有几款用来分析体育运动的实体系统,它们分别用到了影响观察法,图片分析法,便携式传感器法等实现系统。

(1) 影像观察法

图像观察法就是用肉眼观察一个人在运动中的动作,这种方法可以用来对这个人的动作提出改进建议。这种方法经常被用于体育训练和康复,例如,在高尔夫运动中,运动辅助系统可以用来回放和分析球员的挥杆动作,使教练能够观察到挥杆的细节,并向球员提出改进方案。然而,这种方法是主观的,依赖于教练的经验。为了克服这些局限性,开发一个运动姿势检测系统可以提供一个更客观和个性化的训练标准。通过分析用户的身体状况和运动模式,运动姿势检测系统可以提供个性化的反馈和训练建议,提高用户的训练效果和效率。

(2) 图片分析法

即对图片进行分析,使用图像或图片来分析动作,与其他动作捕捉系统类似,它需要事先为每个动作准备一个标准模板,并将模板中的参数与从图片中获得的参数进行比较,以评估该动作。还可以使用激光设备来获得训练者在训练过程中身体关键部位的坐标和运动速率。通过图片或激光设备等各种手段获得的身体参数,由专业人士进行分析,从而提出训练建议。还可以用单反相机捕捉训练者的正面姿态,这样可以避免身体部分被遮挡的情况,提高评估的准确性。图片分析法是动作捕捉的另一种形式,它为分析和评估动作提供了一种更客观和准确的方法。

(3) 便携式传感器法

便携式传感器可以用来收集训练者在运动过程中身体关键部位的物理参数,如可以测量加速度、方向和旋转的加速度或陀螺仪。这些传感器被设计成轻量级和便携式,使它们在运动中使用起来很方便。然而,使用传感器可能会干扰运动

实验结果如下图所示,当视频的分辨率在 450×338 时,视频的关键点识别准确率已经达到了比较好的状态,因此,在今后的数据采集过程中,我们都采用该分辨率进行视频数据的采集。

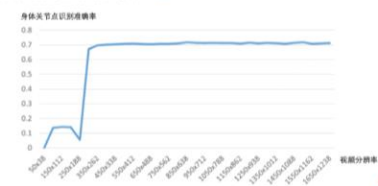


图 10 探究视频清晰度和关键点识别准确率的关系

(三) 数据预处理

首先进行数据标注与数据清洗工作,在学校专业飞镖教练的指导下将拍摄到的视频分为四组数据,这四组视频分别是,飞镖专业运动员与飞镖教练投掷飞镖的标准三连杆与稳定性视频共 440 个,毫无基础的同学投掷飞镖不标准三连杆与稳定性视频共 231 个。

接着,用 Mediapipe 对每个视频中的人物进行关键点提取。提取出全身 33 个不同位置的三维位置坐标。并将其保持成 numpy 数据方式以便后续数据处理。

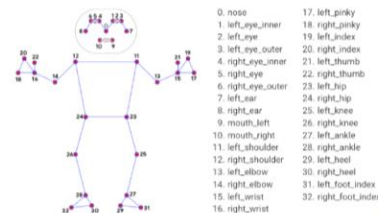


图 11 Mediapipe Pose 中的地标模型可预测 33 个姿势地标的位置

然后,将 Mediapipe 采集到的数据进行数据补齐,数据标准化,将视频采集



Human Pose Recognition Darts Training System(HPR-DATS) 基於人體姿態識別的飛鏢運動員輔助訓練系統

撰寫研究報告 (論文)

Human Pose Recognition Darts Training System (HPR-DATS)⁴²

Research Report⁴²

Zhantao Chen, Xiaolan Middle School, Zhongshan City⁴²

Advisor: Jin Fang⁴²

February 2024⁴²

Abstract: This project utilizes an intelligent darts training assistance system based on human pose recognition. By constructing a hardware and software system, the system conducts a stability analysis of the athletes' movements, fits the best-throwing trajectory for the athlete, and establishes an individual sports database. It effectively aids athletes in adjusting their posture for darts sports, enabling more scientific training methods.⁴²

Keywords: Human Body Pose Recognition; Stability Analysis; Trajectory Fitting; Personal Sports Database; Dart Training⁴²

I. Research Background⁴²

(A) Development Status of Darts Sport⁴²

Darts Sport has recently become popular in Asia, especially in China, where most people's understanding of darts is still at the stage of hard darts. A small number of people with a deeper understanding learn about soft darts through televised competitions. Although electronic smart dartboards can be seen in foreign consulates and upscale bars frequented by foreigners, they are still far from most of the public. According to predictions by the National Sports General Administration's Darts Association, China, with one-fifth of the world's population, will soon lead the world in the number of darts enthusiasts.⁴²

To further promote national fitness activities and popularize darts, the Sports Administration's Social Sports Guidance Center officially listed darts as a sports event in May 1999, signifying the official recognition and establishment of darts as a legitimate sport in China. Darts has been included as an "emerging sport" in the "2022 curriculum," with students participating in competitions organized by provincial student sports associations and dart associations, making them official competitions.⁴²

Currently, Xiaolan Middle School (Zhongshan Foreign Language School) is the only school in Zhongshan that offers a darts program, and it is supported enthusiastically by the provincial dart association and the school. To better support the practice of the darts program at our school, under the joint guidance of our school's information technology and physical education teachers, we developed the Human Pose Recognition Darts Training System (HPR-DATS). With HPR-DATS, we can assist athletes in analyzing their dart-throwing movements through human pose recognition, establish a personal training database for each dart athlete, help them better understand their training conditions, and receive better guidance from coaches.⁴²



Fig1 2021 Guangdong Secondary School Darts Championship⁴²

(B) Current Research Status of Motion Assistance Systems⁴²

Research on using computers to recognize actions for training assistance has been verified by many experts and scholars in the research community, and several physical sports analysis systems are available on the market. They employ observational methods, image analysis methods, and portable sensor methods, among others, to achieve system implementation.⁴²

1. Observational Method⁴²

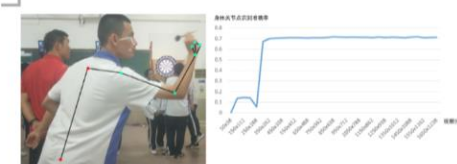
The observational method involves using the naked eye to observe a person's movements during sports activities and providing suggestions for improvement. This method is often used in sports training and rehabilitation. For example, in golf, training assistance systems can replay and analyze a player's swing action, allowing coaches to observe the details of the swing and suggest improvements. However, this method is subjective and depends on the trainer's experience. To overcome these limitations, developing a motion posture detection system can provide a more objective and personalized training standard. By analyzing the user's physical condition and movement patterns, the motion posture detection system can offer personalized feedback and training suggestions, improving the effectiveness and efficiency of training.⁴²

2. Image Analysis Method⁴²

This involves analyzing images or pictures to assess actions, similar to other motion capture systems. It requires a standard template for each action and compares the parameters from the pictures to the template to evaluate the action. Laser devices can also obtain the coordinates and motion rates of key body parts during training. Professionals analyze the body parameters obtained through pictures or laser devices, and other means to provide training suggestions. A single-lens camera can capture the trainee's frontal posture, avoiding the obstruction of body parts and improving the accuracy of the assessment. The image analysis method provides a more objective and accurate way to analyze and evaluate movements as another form of motion capture.⁴²

3. Portable Sensor Method⁴²

Portable sensors can collect physical parameters of key body parts during sports, such as accelerometers or gyroscopes that measure acceleration, direction, and rotational speed. These sensors are designed to be lightweight and portable for easy use during sports activities. However, using sensors may interfere with the athlete's full engagement in the activity and may prevent some key movements from being performed during the sport. Once the sensors collect the physical parameters, the data is transmitted to the training system,



Figures 9 & 10 demonstrate the recognition of key points using Mediapipe's Blaze-Pose and the relationship between video clarity and key point recognition accuracy, respectively.⁴²

(C) Data Preprocessing⁴²

The initial step involved data labeling and cleansing under the guidance of professional dart coaches. Videos captured were categorized into four groups, totaling 440 videos of professional athletes and coaches performing standard triple-hinge and stability throws and 231 videos from novices showing non-standard triple-hinge stability in their throws.⁴²

Subsequently, Mediapipe was utilized to extract key points from each individual in the videos, identifying 33 different three-dimensional position coordinates across the body, focusing on points 1 to 24. These data points were saved in numpy format for further processing. The gathered data underwent completion and standardization processes. Videos containing less than 80 frames were discarded, while those with more than 80 were normalized through an averaging frame extraction method.⁴²



Figures 11 & 12 display the landmark model in MediaPipe Pose, capable of predicting 33 pose landmark positions and showcasing some processed data, respectively.⁴²

- Noise Reduction (Smoothing):** The moving average method was employed to smooth the data by calculating the local average values around each data point. Specifically, a window (defined by the window_size parameter) was selected around each data point, the mean of all data points within this window was calculated, and this mean value was then assigned as the new value for the point. This



Human Pose Recognition Darts Training System(HPR-DATS) 基於人體姿態識別的飛鏢運動員輔助訓練系統

基於人體姿態識別的飛鏢運動員輔助訓練系統

項目成效



在HPR-DATS和專業飛鏢教練的指導下，我校飛鏢隊在2023年廣東省中小學飛鏢錦標賽取得了豐碩的成績。飛鏢隊8位同學不畏強敵、沉著應戰、勇於拼搏，在9個項目比賽中勇奪4項冠軍，並最終奪取了高中組團體總分第一名。HPR-DATS的訓練輔助效果得到飛鏢教練的充分肯定。





electronics MDPI

Article

Improving Short-Term Load Forecasting with Multi-Scale Convolutional Neural Networks and Transformer-Based Multi-Head Attention Mechanisms

Sheng Ding ^{1,*}, Dongyi He ^{2,†} and Guiran Liu ^{3,†}

¹ School of Software, Fudan University, Shanghai 200433, China
² School of Artificial Intelligence, Chongqing University of Technology, Chongqing 401135, China
³ San Francisco State University, College of Science & Engineering (CoSE), San Francisco, CA 94132, USA; gliu@sfsu.edu

* Correspondence: 19212010001@fudan.edu.cn
 † These authors contributed equally to this work.
 ‡ This author contributed to the data analysis and experimental design.

Abstract: This research introduces an original approach to time series forecasting through the use of multi-scale convolutional neural networks with Transformer modules. The objective is to focus on the limitations of short-term load forecasting in terms of complex spatio-temporal dependencies. The model begins with the convolutional layers, which perform feature extraction from the time series data to look for features with different temporal resolutions. The last step involves making use of the self-attention component of the Transformer block, which tries to find the long-range dependencies within the series. Also, a spatial attention layer is included to handle the interactions among the different samples. Equipped with these features, the model is able to make predictions. Experimental results show that this model performs better compared to the time series forecasting models in the literature. It is worth mentioning that the MSE score or mean square error of the model was 0.62, while the measure of fit R2 was 0.91 in predicting the individual household electric power consumption dataset. The baseline models for this dataset such as the LSTM model had an MSE of 0.784 and a R2 score of 0.79. [Abstract](#) [Full Article](#) [Full Article \(PDF\)](#)

[Check for updates](#)

Citation: Ding, S.; He, D.; Liu, G. Improving Short-Term Load

em Biomedical Signal Processing and Control Bin Jiang | Logout

Home Main Menu Submit a Manuscript About Help

← Revisions Being Processed for Author

Page: 1 of 1 (1 total revisions being processed) Results per page 10

Action	Manuscript Number	Title	Date Submission Began	Status Date	Current Status
View Submission Send E-mail	BSPC-D-25-02627R1	DSFRANet: A Lightweight Depthwise Separable Residual Attention Network for sEMG Gesture Recognition	Jun 23, 2025	Jun 23, 2025	"With Editor"

DATA Intelligence

Home Author Review

Author Dashboard

Author Dashboard

- 1 Submitted Manuscripts
- 3 Manuscripts I Have Co-Authored
- [Start New Submission](#)
- [Legacy Instructions](#)
- [5 Most Recent E-mails](#)

Submitted Manuscripts

STATUS	ID	TITLE
Contact Journal	DI-2025-0170	R2A-MAGNet: A Mamba-Based Attention Network for Reconstructing Central Arterial Pressure Waveforms
M-EIC: intelligence, data		
M-EIC: Liu, Fenghong		
		View Submission
Under Review		

em Neurocomputing

Home Main Menu Submit a Manuscript About Help

← Submissions Being Processed for Author

Page: 1 of 1 (1 total submissions)

Results per page 10

SPRINGER NATURE SNAPP scientific reports Account

MSDS-FusionNet: A Multi-Scale Dual-Stream Fusion Network for High-Accuracy sEMG-Based G

CURRENT STATUS

Your submission is in peer review

News about your peer review process

- The editor has invited more than 10 reviewer(s)
- There are 3 reviewer(s) that have accepted to review your manuscript
- The editor has received 3 reviewer report(s)

After the editor has collated and reviewed all the reports they need, which may involve seeking additional reviews, you'll be notified about their decision.

Progress so far [Show history](#)

- Submission received
- Technical check
- Editorial assignment
- With editor
- Peer review

[Learn about our submission process](#)

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Reg. No. 200404393R

8 July 2024

何东毅
Dongyi He
Chongqing University of Technology
Resident Identity Number: 442000200310298413

INVITATION LETTER
MACHINE LEARNING AND ITS APPLICATION
22 July to 26 July 2024
Nanyang Technological University, Singapore



FGITNet: A Two-Stream Gated Network for Enhanced Violence Detection in Surveillance Videos

Dongyi He¹, Yuanquan Gao¹, Jin Fang², He Yan¹, Wei Liu¹

¹School of Artificial Intelligence, Chongqing University of Technology
²Xiaoping Technology Innovation Lab, Zhongshan Xiaolan High School

DISE Conference
 December 20, 2024

Dongyi He¹, Yuanquan Gao¹, Jin Fang², He Yan¹ FGITNet: A Two-Stream Gated Network for I 1 / 18



A Two-Stream Gated Network for Violence Detection Using Inflated 3D ConvNet and Transformer

Authors Dongyi He, Yuanquan Gao, Jin Fang, He Yan, Wei Liu
 Pages 95 - 104
 DOI 10.3233/FAIA250257
 Category Research Article
 Series Frontiers in Artificial Intelligence and Applications
 Ebook Volume 405: Design Studies and Intelligence Engineering

Abstract

In this paper, we propose FGITNet, an innovative Two-Stream Gated Network designed for accurate violence detection in surveillance videos. Leveraging both Inflated 3D ConvNet and Transformer architectures, FGITNet extracts spatiotemporal features by processing RGB and Optical Flow frames simultaneously. A key contribution of our model is a novel background suppression mechanism that effectively filters out non-motion elements, enhancing motion feature detection while maintaining computational efficiency. This two-stream approach achieves state-of-the-art performance on the RWF-2000 and Hockey datasets, surpassing existing methods by 4% on RWF-2000, reaching an accuracy of 93.75%. Our method demonstrates significant potential for real-time applications in intelligent security systems, contributing to the field of design and intelligent engineering.

A Two-Stream Gated Network for Violence Detection Using Inflated 3D ConvNet and Transformer

Dongyi He^a, Yuanquan Gao^a, Fang Jin^a, He Yan^a and Wei Liu^{a,1}

^aSchool of Artificial Intelligence, Chongqing University of Technology, Chongqing, 400054, China.

^bXiaoping Technology Innovation Lab, Zhongshan Xiaolan High School, Zhongshan, Guangdong, 528415, China

Abstract. In this paper, we propose FGITNet, an innovative Two-Stream Gated Network designed for accurate violence detection in surveillance videos. Leveraging both Inflated 3D ConvNet and Transformer architectures, FGITNet extracts spatiotemporal features by processing RGB and Optical Flow frames simultaneously. A key contribution of our model is a novel background suppression mechanism that effectively filters out non-motion elements, enhancing motion feature detection while maintaining computational efficiency. This two-stream approach achieves state-of-the-art performance on the RWF-2000 and Hockey datasets, surpassing existing methods by 4% on RWF-2000, reaching an accuracy of 93.75%. Our method demonstrates significant potential for real-time applications in intelligent security systems, contributing to the field of design and intelligent engineering.

Keywords. Two-Stream Network, Violence Detection, Inflated 3D ConvNet, Transformer, Background Suppression

1. Introduction

The increasing demand for public safety has led to widespread use of surveillance cameras, which have significantly reduced crime rates in cities like Baltimore, Chicago, and Washington, D.C., as shown in Fig. 1. However, real-time violence detection remains a challenge due to the vast amounts of video data and the limitations of manual monitoring, often leading to delayed responses.

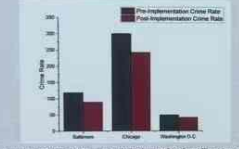


Figure 1. Crime rates before and after the implementation of surveillance cameras.

2. Related Work

2.1. Existing Methods for Violence Detection

The task of violence detection in video surveillance systems has garnered significant attention, with various methods developed to tackle this complex problem. Early techniques predominantly relied on single-modal approaches, which typically used visual data from RGB frames to identify violent actions. However, these methods have been criticized for their inability to effectively capture dynamic motion patterns, especially when background noise or occlusion is present, leading to reduced accuracy in real-world scenarios. Feil et al. (2019) highlighted that single-modal methods, despite their simplicity, often fail to generalize across different environments, making them unsuitable for large-scale surveillance systems that require robust performance in diverse conditions [1].

3. Proposed Approach

In this section, we describe the proposed FGITNet model, which encompasses several key stages: Overall Architecture, Data Preprocessing and Background Suppression.



4. Experiments and Results

We evaluated FGITNet using two benchmark datasets: RWF-2000 and Hockey. These datasets are widely used in violence detection tasks and provide a balanced representation of violent and non-violent actions.

4.1. Experimental Setup

The model was implemented using PyTorch and trained on an NVIDIA RTX 4070Ti GPU. Key hyperparameters included a batch size of 16 and a learning rate of 1×10^{-4} . We trained the model for 2000 epochs.

4.2. Evaluation and Comparison

FGITNet was compared to other state-of-the-art models. It achieved the highest accuracy on both datasets, as shown in Table 1.

Table 1. Performance Comparison.

Model	RWF-2000 (%)	Hockey (%)
Flow Gated Network	89.25	91
SeqConv-LSTM	86.75	90.5
FGITNet (Ours)	93.75	93.5

4.3. Ablation Study

The ablation study on RWF-2000 demonstrates the importance of background suppression, improving accuracy from 91.25% to 93.75% and enhancing precision and recall.

Table 2. Ablation Study on RWF-2000

Variant	Accuracy (%)	Precision (%)	Recall (%)
Without Background Suppression	91.25	93.10	93
With Suppression	93.75	96.3	91

5. Discussion and Future Work

5.1. Advantages and Limitations

FGITNet demonstrates several advantages in violence detection tasks, notably its effective background suppression mechanism, which helps the model focus on motion-related elements, significantly improving accuracy and robustness. The two-stream architecture, combining RGB and optical flow, enables FGITNet to capture both spatial and temporal features, enhancing its ability to detect violent actions in challenging real-world environments.

However, the model has some limitations. While it performs well on benchmark datasets, its generalization to more diverse real-world scenarios, such as varying lighting conditions, camera angles, or complex interactions in crowded scenes, remains to be fully tested. Additionally, real-time processing is not fully optimized, as the

6. Conclusion

In this work, we proposed FGITNet, a novel two-stream gated network for violence detection, integrating efficient background suppression with 3D and Transformer architectures. The model demonstrated superior performance on the RWF-2000 and Hockey datasets, achieving state-of-the-art accuracy in both cases. FGITNet's background suppression mechanism plays a crucial role in enhancing its focus on motion-related features, resulting in more reliable detection. While the results are promising, future work will aim to address the model's scalability and real-time processing capabilities, making it more applicable to diverse and real-world surveillance scenarios.





Let science and technology go hand in hand—empowering students to explore, innovate, and thrive in the age of artificial intelligence.

——讓「科學」與「技術」攜手並進，賦能學生在人工智能時代探索、創新、成長！

