

# 个人简介

能够独立开发各类具挑战性的软件系统  
尤其熟悉手写轨迹相关算法

陈颂光

chan@chungkwong

<https://chungkwong.cc/>

软件工程师  
机器学习工程师

2025 年 5 月 1 日

# 重点

- 扎实的理论基础
  - 持有基础数学专业的理学博士学位
  - 广泛自学计算机科学与技术的各个领域
  - 能够把理论知识应用于实践中
- 丰富的实践经验
  - 曾在与手写相关的国际性模式识别竞赛获得过冠军、亚军和季军
  - 具备从产品设计到落地的全流程项目经验
  - 非常可靠的程序设计和编程能力
- 良好的学习态度
  - 香港中学会考三优三良
  - 香港高级程度会考二优二良
  - 预科两年校内成绩均在理科生中排名全级第一
  - 大学四年总成绩在同专业中排名全级第一

# 教育背景和工作经历

# 教育背景

香港金文泰中学学生 (2004.9-2011.8)

2007-2008 香港数学竞赛初赛三等奖

中四数学科第一名、积极学习奖、进步生奖、守时奖

2008-2009 澳洲数学公开赛荣誉奖

中五数学科第一名、附加数学科第一名、商业科第一名、守时奖

2009 年香港中学会考三优三良 (数学 A、附加数学 A、生物 A、物理 B、化学 C、商业 C、中文 3、英文 3)

2009-2010 入选香港资优教育学院学员 (数学)

中六理组年考第一名、纯粹数学科第一名、物理科第一名、勤学生奖、守时奖

2010-2011 中七理组模拟考试第一名、纯粹数学科第一名、勤学生奖、积极学习奖、守时奖

港澳台联考 543 分 (语文 92/150、英语 127/150、数学 106/150、物理 80/150、化学 138/150, 自修)

中山大学对港独立招生考试 232 分 (语文 62/100、英语 72/100、数学 98/100, 自修)

香港高级程度会考二优二良 (高级程度纯粹数学 A、高级程度化学 A、高级程度物理 B、高级程度应用数学 C、高级补充程度英语运用 D)

# 教育背景

中山大学数学与计算科学学院数学与应用数学专业本科生 (2011.9-2015.7)

2011-2012 优秀学生二等奖学金

2014-2015 大学生创新训练计划校级优秀项目

2011-2015 必修课及专业选修课四年平均绩点为 4.3, 在本专业 110 人中排名第 1

# 教育背景

中山大学数学学院基础数学专业硕博连读研究生 (2015.8-2021.7)

- 2015-2020 获得“优生优培”资助计划支持
- 2017-2018 担任春季学期《实变函数》课程助教
- 2018-2019 担任春季学期《实变函数》课程助教
- 2019-2020 担任秋季学期《偏微分方程》课程助教
- 2019 获得国际模式识别学会 (IAPR) 文档分析与识别国际会议 (ICDAR) 联机手写数学公式识别竞赛 (CROHME) 脱机识别任务季军 (与学业无关)
- 2019 提出 Sauvola、Niblack、Bernsen 等局部自适应二值化方法的一种内存高效的快速实现, 在去掉挥霍空间的积分图像的同时还能稍为提高速度 (与学业无关)
- 2019-2020 重新探讨把脱机手写识别归结为联机手写识别的可行性 (与学业无关)
- 2021 毕业论文获全体评阅人和答辩委员一致通过, 内容与截面曲率下有界空间上的向量值单相自由边界问题及 Ricci 曲率下有界空间的奇异集的余维数估计问题有关

# 工作经历

泓宇星电子科技有限公司算法工程师 (2021.6-)

2022 突出贡献奖

2023 优秀员工奖

2023 国际模式识别学会 (IAPR) 文档分析与识别国际会议 (ICDAR) 联机手写数学公式识别竞赛 (CROHME) 全部三个任务 (轨迹、图片、双模态) 冠军 (优势显著)

2024 国际模式识别学会 (IAPR) 文档分析与识别国际会议 (ICDAR) 化学结构式识别竞赛 (CROCS) 亚军 (指标非常接近冠军且大幅领先季军)

2025 技术进步奖

# 职业技能

# 机器学习

## 问题建模

- 处理涉及时间序列、图、自然语言、图像等不同模态数据的问题
- 灵活运用神经网络（卷积、循环、图、全连接）、支持向量机、决策树、最近邻、聚类、概率统计方法和它们的结合来建模

## 数据工程

- 数据收集计划和工具设计
- 生成任务特定数据与设计数据增广策略
- 整合异构数据集

# 机器学习

## 模型训练

- 利用常见训练框架（TensorFlow、Keras、Pytorch）搭建网络结构和训练
- 设计辅助损失函数以改善模型的已知弱点
- 借助迁移学习加速收敛和利用其它模型学过的知识
- 调优提前停止、学习率调度、优化器、初始化方法、正规化层、梯度裁剪、激活函数其它元参数

## 端侧与服务器端模型部署

- 应用常见推理框架（ONNX Runtime、TensorFlow Lite、QNN）在端侧或云端部署模型
- 使用模型量化、算子融合和硬件加速等手段优化安装包大小、内存占用和速度

# 软件开发

## 知识储备

熟悉计算机系统、编译原理、操作系统、网络、离散数学、计算理论、算法分析、数据结构、数据库、信息安全、信息检索等知识并能灵活运用

## 开发工具

熟悉 Linux 命令行、版本管理 (SVN/git)、构建系统 (Maven/Gradle)、单元测试 (JUnit)、静态分析、调试、性能分析等工具来提高效率和质量

# 软件开发

## 多范式编程

**面向对象** 精通 Java、熟悉 Python

**面向过程** 掌握 C、BASH

**函数式** 掌握 Scheme

**声明式** 掌握 Prolog、SQL

## 跨平台交互设计

**移动前端** Android View、Kotlin Compose

**桌面前端** Swing、JavaFX、Kotlin Compose

**Web 前端** HTML、CSS、JavaScript

**Web 后端** Spring Boot

# 项目经验

# 数学公式识别

## 任务

把手写轨迹形式和图片形式（包括手写体和印刷体）的数学公式转换为 LaTeX 或 MathML 等便于计算机程序处理格式

## 应用

- 联机手写识别为科技排版软件和计算机代数软件提供一种方便的数学公式输入方式
- 脱机手写识别以便复用、重新编辑、搜索或批量索引现有的笔记、黑板或论文中的公式



$$\vec{AB} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

$$\vec{AB} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$



$$\begin{cases} 1+y^2 = (a-1)^2 + (b-4)^2 \\ -1 = a+1 \\ u = b+4 \end{cases}$$

$$0. \left\{ \begin{matrix} 1+y^2 = (a-1)^2 + (b-4)^2 \\ -1 = a+1 \\ u = b+4 \end{matrix} \right. \Rightarrow a = -2.$$

$$\begin{cases} 1+y^2 = (a-1)^2 + (b-4)^2 \\ -1 = a+1 \\ u = b+4 \end{cases} \Rightarrow a = -2.$$

$$1. \left\{ \begin{matrix} 1+y^2 = (a-1)^2 + (b-4)^2 \\ -1 = a+1 \\ u = b+4 \end{matrix} \right. \Rightarrow a = -2.$$

$$\begin{cases} 1+y^2 = (a-1)^2 + (b-4)^2 \\ -1 = a+1 \\ u = b+4 \end{cases} \Rightarrow a = -2.$$

# 数学公式识别

- 基于编码器 - 解码器方案实现触摸事件序列到 LaTeX 命令序列的转换，并定制损失函数以显式优化输入输出元素间的对应关系
- 对于轨迹模式使用多层双向 LSTM 作为编码器，而对于图片则使用轻量化的 DenseNet 变种作为主干网络
- 实现一个根据 LaTeX 代码合成手写轨迹或印刷体图片的工具以可控地批量生成有多样性的训练样本，从而大幅提高模型的准确性和方便扩大支持范围
- 对轨迹进行重排序、重采样和坐标正规化来提高对笔顺、设备采样率、书写位置和书写大小的鲁棒性
- 对数学公式的 LaTeX 进行规范化以提高惟一性，降低学习难度
- 通过把 LL(1) 语法的预测分析法和集成学习技巧结合到集束搜索解码以支持在运行时定制支持的识别范围和提高准确性

# 数学公式识别

2023 年文档分析与识别国际会议 (ICDAR) 联机手写数学公式识别竞赛 (CROHME) 成绩

	Structure + Symbol Labels			Structure Correct rate
	Correct	$1 \leq s.err$	$2 \leq s.err$	
<b>Task 1: On-line</b>				
Sumia*	<b>82.34</b>	<b>90.26</b>	<b>92.47</b>	<b>92.41</b>
YP_OCR	72.55	83.57	86.22	86.60
TUAT	41.10	54.52	60.04	56.85
DPRL_RIT	38.19	53.39	58.39	59.98
<b>Task 2: Off-line</b>				
Sumia*	<b>70.81</b>	<b>81.74</b>	<b>86.13</b>	<b>86.95</b>
YP_OCR	67.86	80.86	85.13	85.99
PERO	58.37	71.22	75.57	75.55
TUAT	51.02	63.17	67.48	64.59
UIT@AIClub_Tensor	14.83	22.87	28.57	32.19
<b>Task 3: Bi-modal</b>				
Sumia*	<b>84.12</b>	<b>91.43</b>	<b>93.70</b>	<b>94.13</b>
YP_OCR	72.55	83.57	86.22	86.60
TUAT	53.76	68.83	74.22	70.81

 图: 在全部三项任务 (联机/脱机/双模态) 任务上以压倒性优势夺冠

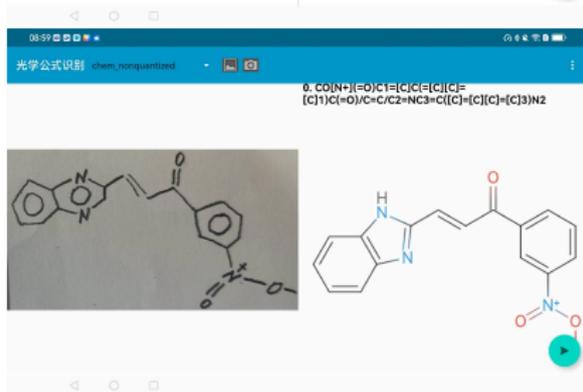
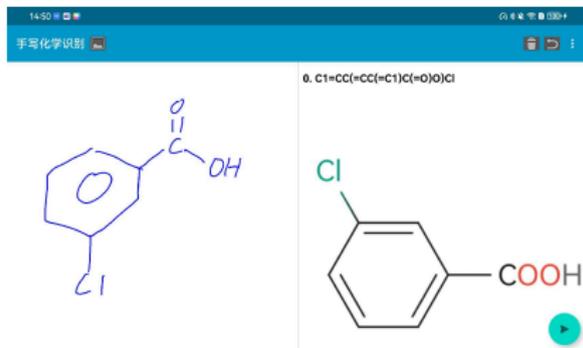
# 化学结构式识别

## 任务

把手写轨迹形式或图片形式（包括手写体和印刷体）的化学结构式转换为 SMILES 或其它方便化学信息系统的电子形式

## 应用

- 联机手写识别为科技排版软件和化学信息软件提供一种方便的结构式输入方式
- 脱机手写识别以便复用、重新编辑、搜索或批量索引现有的笔记、黑板或论文中的结构式



# 化学结构式识别

## 方法

- 使用尽可能惟一和简单的化学结构式表示以降低学习难度
- 利用 SMILES 代码生成印刷体和手写体化学结构式样本并随机化各种渲染参数以提高多样性
- 使用随机图像滤镜、噪声和几何形变等数据增强策略以提高泛化能力
- 仿照 MobileNetv2 用深度可分离卷积改造主干网络 DenseNet 以降低参数量和计算量
- 把卷积核从  $3 \times 3$  改为  $5 \times 5$  以扩大感受野
- 把批正规化换成实例正规化以提高训练与推理的一致性
- 使用计数损失函数作为辅助损失函数
- 使用类别加权的交叉熵以聚焦更关键的 token
- 使用语法约束确保输出的合法性
- 通过在集束搜索时结合多个模型的预测提高准确性

# 化学结构式识别

2024 年文档分析与识别国际会议 (ICDAR) 化学结构式识别竞赛 (CROCS) 成绩

**Competition Results**

> Final results of all teams on EDU-CHEM (Test Set B)

**Winner: CNKI\_AI**

Rank	Team Name	EM(1)	Structure EM(1)	Affiliations
1	CNKI_AI	70.66%	80.12%	Tongfang Knowledge Network Digital Publishing Technology Co., Ltd.
2	Sunia	70.05%	79.44%	Innovation department, Sunia Pte Ltd.
3	imucs	62.03%	72.72%	Inner Mongolia University
4	ocr-shenwiping	56.76%	67.56%	Huazhong University of Science and Technology
5	TSNUK	50.11%	67.16%	Taras Shevchenko National University of Kyiv

ICDAR 2024

图: 惜败给参数量大数百倍的模型 0.61 个百分点而大幅领先季军 8 个百分点

# 联机手写文字识别

## 任务

把手写轨迹转换为字符串：

- 支持中、日、韩、泰、缅、泰米尔、天城、阿拉伯、拉丁、西里尔、希腊、希伯来等不同的书写体系近 80 种语言且可以相互混合
- 提供横向书写、多方向书写和叠写等模式



## 应用

- 手写输入法/全局手写
- 笔记软件的手写转文本功能



# 联机手写文字识别

## 方法

- 利用有限的手写单字符/音节/单词样本根据文本语料生成大量仿手写的行样本（特别是缺乏真实样本的竖写和叠写），用于缓解数据不足和加强解码器隐含的语言模型
- 利用不同的预处理和数据生成策略来支持笔顺不敏感的文本行模式（中日韩文同时支持横写和竖写）和笔顺敏感的叠写模式
- 通过对标签进行规范化和逻辑顺序到物理顺序转换来降低模型的学习难度，从而提高对带内部结构字符（如韩文音节）和双向文字（例如右到左的阿拉伯文与左到右的英文混合）的泛化的能力
- 结合语言检测模型和跨语言识别模型以支持多种语言混写
- 可结合 n-gram 统计语言模型和词典来进一步提高识别的准确性
- 可利用注意力机制估计字符与笔画间的对应关系

# 联机手写文档分析

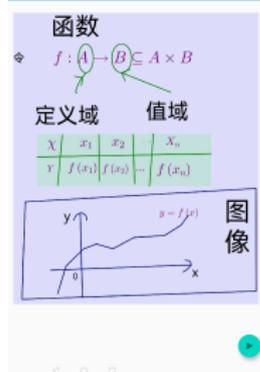
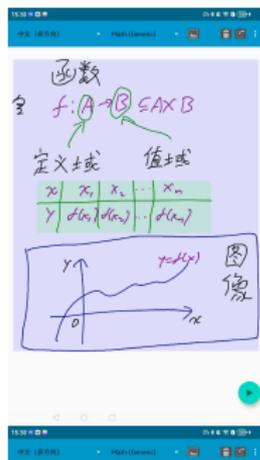
## 任务

对整页手写笔记中的笔画进行两个粒度的分组和分类：

- 粗粒度对象可能为段落、表格、图表或批注
- 细粒度对象可能为文本行、公式行、形状、绘图、化学结构式或误触笔画

## 应用

- 容许识别混合了多个不同类型对象的笔记，只要分别有识别模块可以处理各细粒度对象
- 容许对手写笔记进行某些编辑和美化操作



# 联机手写文档分析

## 方法

- 使用单一模型一举解决文本-非文本分类、文本行分割、数学公式检测、表格检测与识别、图表检测与识别等子任务
- 结合了利用循环神经网络和层次池化自动学习出来的特征以及人工特征，能够充分利用时空信息而不太依赖特征工程
- 利用图神经网络的模型以支持多粒度的分割与分类，其中每条笔画对应图中的一个顶点，两条笔画在时间或空间上相近时对应的顶点相邻，通过对顶点分类可得笔画所属对象的类型，而对边作分类可得两条笔画是否属于同一对象
- 搭建用户友好且可以利用上一版模型进行预标注的工具来提高标注效率
- 利用数据合成（如文档拼接）和数据增广策略（如多粒度的内容替换与删除、精细的顺序打乱、几何形变）以提高泛化到复杂版面的能力

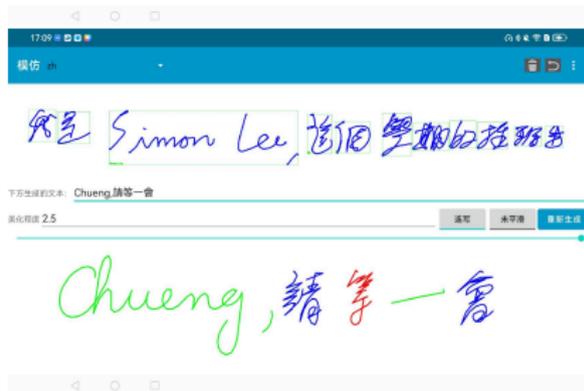
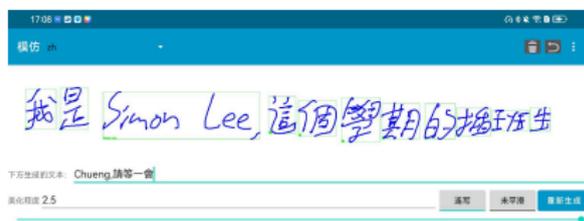
# 手写文字轨迹生成

## 任务

生成内容为指定字符串而风格与特定样本相似的文字的轨迹

## 应用

- 把印刷体文档转换为用户风格的手写以提高亲切性
- 用手写形式呈现手写内容的拼写纠正、自动补全和计算结果
- 生成笔迹鉴定用的假冒签名样本
- 制作特定风格的手写字体



# 手写文字轨迹生成

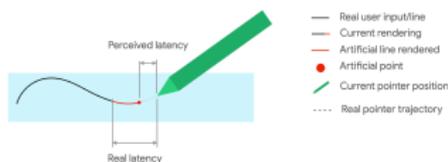
## 方法

- 把字符串按书写体系分解，分别转换为轨迹后再拼接，从而简化支持大量书写体系的难度
- 对于中日韩字符，利用 CNN 分别提取印刷体字符的内容特征和手写样本的风格特征，再用 Transformer 自回归地预测各点的位移和抬笔状态，从而容许生成训练时未见过的字符
- 对于其它书写体系的字符串，利用带单调注意力机制的循环神经网络和高斯混合模型，从而容许生成连写的单词
- 利用文字识别结果进行验证，在需要时触发重新生成或回退机制以降低生成错误的风险

# 笔迹补偿预测

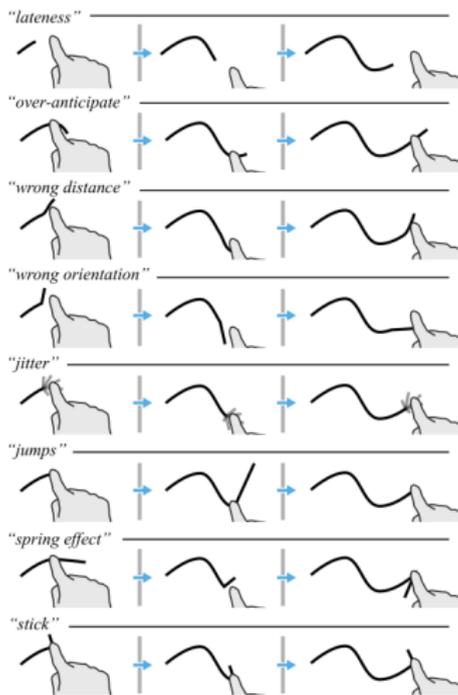
## 任务

通过利用过往的触摸事件预测未来的触摸事件，可以偷步绘制笔画，从而降低笔尖与落墨处间的延迟感觉



## 方法

通过让模型同时预测轨迹的二次参数方程和安全时长来实现变时长预测，从而在抑制离谱预测的前提下保障容易场景下的预测时长



图：不良预测算法常见的副作用

# 光学文档识别 (OCR)

## 任务

从文档图片（特别是包含手写内容的）中提取其中的文字、数学公式和化学结构式信息

## 方法

- 1 对象检测，即定位每行文字、每行公式和每个化学结构式
  - 使用基于 DBNet 的检测模型，这类基于分割的算法天然支持倾斜甚至弯曲的区域
- 2 对象分类，即区分每个对象的方向（是否需旋转  $180^\circ$ ）、类型（当作文字、公式还是化学结构式去识别）和书写体系（确定用哪个模型去识别文字）
  - 使用基于 MobileNetv3 的分类模型
- 3 对象识别，即把每个裁剪出来的对象图片转换为字符串
  - 使用基于 CRNN 的文本识别模型 (PPOCRv2)，这类基于 CTC 而非自回归解码的模型推理效率高和效果稳定

# 个人项目

- 基于笔划提取的脱机手写数学公式识别  
(<https://doi.org/10.1109/ACCESS.2020.2984627>)
- 局部自适应二值化的内存高效快速实现  
(<https://arxiv.org/abs/1905.13038>)
- 结构化数据搜索引擎 ViewFact (<https://www.viewfact.org/>)
- JVM 上的 Scheme 语言 R7RS 标准解释器 JSchemeMin  
(<https://github.com/chungkwong/JSchemeMin>)
- JVM 上的 Prolog 语言解释器 JPrologMin  
(<https://github.com/chungkwong/JPrologMin>)

# 谢谢大家