

# “英特尔杯”第一届中国研究生人工智能创新大赛

## 基于跨模态学习和可穿戴设备的手势输入法

### 项目文档

V20190721R2

2019.07.21

EmNets

技术创新

# 目录

## 1.1 1 项目概况 1

1.1	背景和基础.....	1
1.1.1	项目起因.....	1
1.1.2	已有工作基础.....	2
1.1.3	项目灵感.....	2
1.1.4	团队构成.....	3
1.2	场景和价值.....	3
1.3	所需支持.....	4
2	项目规划.....	4
2.1	整体目标.....	4
2.2	技术创新点.....	5
3	实施方案.....	5
3.1	技术可行性分析.....	5
3.2	技术细节.....	6
3.3	计划和分工.....	8
4	参考资料.....	8

## 记录更改历史

序号	更改原因	版本	作者	更改日期	备 注
1	初次编写	V20190620 R1	刘汶鑫	2019.06.20	撰写项目概况
2	添加已有工作基础	V20190620 R2	刘汶鑫	2019.06.20	撰写项目概况
3	添加项目灵感	V20190621 R1	刘汶鑫	2019.06.21	撰写项目概况
4	添加团队构成	V20190621 R2	刘汶鑫	2019.06.21	撰写项目概况
5	添加所需支持	V20190621 R3	刘汶鑫	2019.06.21	撰写项目概况
6	添加场景和价值	V20190621 R4	袁宇	2019.06.21	撰写项目概况
7	添加计划和分工	V20190622 R5	刘汶鑫	2019.06.22	撰写实施方案
8	添加整体目标	V20190715 R1	曾思钰	2019.7.15	撰写项目规划
9	添加技术创新点	V20190715 R2	曾思钰	2019.7.15	撰写项目规划
10	添加技术细节	V20190717 R1	曾思钰	2019.7.17	撰写实施方案
11	添加技术可行性分析	V20190721 R1	周寒	2019.7.21	撰写实施方案
12	添加应用场景图片	V20190721 R2	曾思钰	2019.7.21	撰写项目概况

# 1 项目概况

## 1.1 背景 and 基础

### 1.1.1 项目起因

文本输入是日常生活中人们使用移动设备时常见的一种交互方式，但是，在智能手表上输入文字具有挑战性。一方面，智能手表的外形和触摸屏都很小，用触摸输入的方式错误率较高，输入效率往往较低，用户的体验较差。另一方面，在智能手表上进行输入时，需要使用两只手来完成输入任务，一只手负责把手表移动并固定到自己的视线范围内，另一只手负责在触摸屏上进行输入，在某些特定场景下，这种输入方式会变得困难。例如，当用户拿着雨伞或购物袋走在路上的时候，这时在智能手表上输入文本是非常困难的，因为用户只有一个自由的手可以活动，另一只手需要拿着物品，当有重要消息需要回复时，用户不得不中断他们正在进行的任务（停止撑伞或提购物袋）来输入文本，显得十分麻烦。因此，如果用户能够仅仅通过一只手就可以完成输入任务的话，在面对这些特殊场景时，也能轻松应对了。

为了解决这一问题，一种解决方案是语音输入，但在某些情况下（例如，在会议或教室中）是不合适的，并且还可能暴露用户的隐私。另一种解决方案是使用智能回复助手事先定义好一些常见的语句（例如，“我正在开会，现在不方便回复”），当相应消息发来时，用户可以快捷地使用已定义好的语句进行回复，但是这些可定义的快捷回复语句数量有限，只适用于某些特定的场景，当遇到新的应用场景时，如果相应的快捷回复语句未定义，用户不得不自己根据需求自己手动输入文字进行回复。

基于上述问题，我们设计了一款基于跨模态学习的可穿戴设备上的单手 3D 手势输入法 **AirText**，通过该输入法，用户仅仅使用单只手与智能手表进行交互就可以实现文字的输入。具体使用方式是，用户利用佩戴智能手表的那只手的食指在空中书写拼音字母，智能手表在捕捉到用户手写字母的手势后，通过神经网络对手势进行分类来预测用户输入的字母，并将用户输入的字符串（按照输入顺序将用户输入的字母进行组合得到）放入英文词库中查询后，在手表屏幕的上、下、左、右四个方向上给出相应的预测单词，用户随后通过“向上/向下/向左/向右”挥动手腕的方式来选择对应方向上自己想要的单词，以完成相应语句的输入。



图 3-1 AirText 应用场景示意图

### 1.1.2 已有工作基础

团队在可穿戴设备应用方面的研究已积累了大量经验，在 ACM SenSys、ACM Ubicomp 等国际学术会议上投稿并发表了三篇论文，已受理国家发明专利三项，研究成果在世界物联网大会、阿里巴巴云栖大会、ACM SenSys 会议上进行了展示，得到了国内外学者、技术专家的一致认可。

团队成员自读研以来，一直在从事可穿戴设备应用方面的研究工作，所涉及到的技术包括但不限于超声波、毫米波、IMU（惯性传感器）、UWB、CSI、PPG、红外线和皮表肌电。团队从事的研究最多的是利用智能手表的 IMU 数据（包括加速度、角速度、磁场强度和方向）进行手势追踪，目前已利用这项研究成果在智能手表上开发了多款智能应用，包括手势追踪系统、体感游戏、健身助手、高尔夫教练、康复医师、多家电控制等。这些智能应用的核心技术就是通过智能手表采集用户的加速度计、陀螺仪、磁力计的数据，利用互补滤波、卡尔曼滤波、粒子滤波等算法对数据进行预处理，然后通过位姿估计算法计算出手表的位置，从而推测出用户肢体的位置，进而在连续的位置追踪中得到用户的手势，再结合事先对各种手势的定义触发相应的事件，以满足各种应用的需求。例如，在高尔夫教练的应用中，我们事先定义好各种标准的挥杆动作，然后在用户进行挥杆训练的时候，捕捉用户整个的挥杆动作，并与标准动作进行对比，计算两个动作之间的偏差，对用户的动作进行打分，同时给出相应的改进建议。

团队在掌握了处理 IMU 数据的白盒技术后，开始研究深度学习等黑盒技术来处理 IMU 数据，以实现相应的应用。例如，通过神经网络的方法对用户手势的 IMU 数据进行分类，这样，通过 IMU 数据就可以得到不同类别的用户手势，基于这些手势就可以进一步实现不同的应用了。

总之，团队在利用可穿戴设备实现智能应用方面取得了很多成果，并成功地将人工智能应用到了智能感知领域。因此，团队若想利用深度学习技术，在可穿戴设备上实现一种单手输入的 3D 手势输入法，就可以在已有工作的基础上，进一步地实现技术创新。

### 1.1.3 项目灵感

项目灵感来自于 CVPR'18 的一篇论文 Through-Wall Human Pose Estimation Using Radio Signals，该论文实现了可穿墙透视评估人体姿态的系统，即用跨模态学习技术来教无线设备感知人体的位姿和移动，甚至隔了一堵墙也能做到。

论文使用一个神经网络来分析人的身体反射的无线电信号，然后创建一个随人类动作而同步的动态图像，包括人行走、停下、坐下和移动时肢体的图像。由于无法对无线信号进行标注，所以通过同步的视频信号来提取人体的姿势信息，随后完成无线信号到姿势信息的训练。训练完成后，就可以直接由无线信号得到人体的姿势信息。训练数据需要通过视频数据来进行监督，但训练完成后，即使没有视频数据，人体的姿势信息也能被检测出来，甚至穿墙也能够检测出来，因为神经网络已经学会了如何用无线信号分析出人体的姿势信息。相对于从图像得到人体姿势信息的优势是，在不可见的情况下，人体的姿势信息也能被检测出来，比如在夜晚或者墙的另一边，但是利用图像信号就无法完成检测。

受到该论文的启发，我们团队在对用户的手写数据进行训练的时候，也利用了视频数据对训练过程进行监督，从而让神经网络学会如何用 IMU 数据分析出用户手写的字母，视频数据是通过 Leap Motion 采集得到的。训练完成后，神经网络已经学会了如何用 IMU 数据

去分析用户手写的字母,不再需要视频数据的监督了,只需要通过智能手表采集 IMU 数据,就可以分析得到用户输入的字母了。

### 1.1.4 团队构成

团队成员全部来自浙江大学,指导老师为董玮教授和高艺副教授,小组成员为周寒博士、曾思钰硕士、袁宇硕士和刘汶鑫硕士。

董玮是浙江大学计算机学院教授、博士生导师,浙江大学求是青年学者,IEEE 会员,ACM 会员,CCF 高级会员,CCF 互联网专委、物联网专委委员,浙江省计算机学会嵌入式与物联网技术专委秘书长。在 ACM MobiCom, SIGMETRICS, UbiComp, IEEE INFOCOM, ICNP 等著名国际会议以及 IEEE/ACM Trans on Networking, IEEE Trans on Mobile Computing, IEEE Trans on Computers 等著名国际期刊上发表论文 100 余篇,其中 CCF A 类期刊/会议论文 40 余篇,IEEE/ACM Trans 长文 30 余篇。

高艺是浙江大学计算机学院副教授、硕士生导师,2015 年入选中国计算机学会 CCF 青年人才发展计划,CCF 专业会员,CCF 青工委委员,ACM 会员,IEEE 会员。2009 年 9 月至 2014 年 12 月,于浙江大学计算机学院直接攻博,连续 5 年获得优秀研究生等荣誉,同时获得研究生国家奖学金等奖励。在国际顶级期刊 IEEE/ACM ToN, IEEE TC, IEEE TPDS, IEEE TMC, IEEE TII 以及国际顶级会议 ACM MOBICOM, ACM SIGMETRICS, IEEE ICNP, IEEE INFOCOM, IEEE ICDCS 等发表 40 余篇高质量论文。

周寒是浙江大学计算机学院在读博士,研究方向为物联网与智能感知。本科就读于西安电子科技大学,专业为电子信息工程,2016 年获工学学士学位。

袁宇是浙江大学计算机学院在读硕士,研究方向为物联网与智能感知。本科就读于东北大学,专业为物联网工程,2018 年获工学学士学位。

曾思钰是浙江大学工程师学院在读硕士,研究方向为深度学习与智能感知。本科就读于大连理工大学,专业为电子商务,2018 年获管理学学士学位。

刘汶鑫是浙江大学计算机学院在读硕士,研究方向为深度学习与智能感知。本科就读于东北大学,专业为计算机科学与技术,2017 年获工学学士学位。

## 1.2 场景和价值

简单阐述该项目适用的应用场景及潜在社会价值,包括市场调研、对比性分析等。

随着智能穿戴设备的发展,智能手表越来越成为年轻人时尚的选择,而如何与智能手表的小屏幕更好地进行交互,成为亟待解决的问题。当前智能手表主要采用与手机类似的输入法键盘,但由于智能手表键盘小,经常导致错误输入,严重影响用户体验。而且,当前智能手表无法进行单手输入,当用户在手持东西时无法进行输入操作及回复消息。

我们的产品 AirText 旨在提升用户与智能手表的交互体验,在一只手持快递等物品的情形下依然能够单手输入快速回复消息,在开会等不方便语音输入的场景下仍可实现交互。除此之外,对于单臂残障人士,AirText 也能够使他们依靠单臂输入消息实现与智能手表的交互,进而完成更丰富的指令控制。

AirText 通过运用 AI 算法训练用户独特的输入特征,改善用户与智能手表的小屏幕交互体验,将极大丰富智能手表的使用场景,促进智能穿戴行业的变革与发展。

## 1.3 所需支持

本项目实施过程中需要的算力有：

双精度浮点运算能力	4.7 teraFLOPS
单精度浮点运算能力	9.3 teraFLOPS
半精度浮点运算能力	18.7 teraFLOPS
NVIDIA NVLink 互联带宽	-
PCIe x16 互联带宽	32 GB/s
CoWoS HBM2 堆叠式显存容量	16 GB or 12 GB
CoWoS HBM2 堆叠式显存带宽	732 GB/s or 549 GB/s
提升使用页面迁移引擎编程的能力	√
ECC 保护助力实现可靠性	√
针对数据中心部署优化服务器	√

本项目实施过程中需要的硬件有：

1. 智能手表，包括 Huawei Watch 2 Pro (Android 8.0.0 Wear OS, sampling rate 200Hz)、Ticwatch Pro (Android 9 Wear OS, sampling rate 170Hz)、LG Watch Urbane (Android 7.1.1 Wear OS, sampling rate 200Hz)，用于采集 IMU 数据。
2. Leap Motion，最先进的手部跟踪设备。
3. 基于 PCIe 的服务器的 NVIDIA Tesla P100, 更快速地处理要求最严格的高性能计算。

本项目实施过程中需要的相关培训有：

1. Intel AI SW-HW Portfolio
2. Pytorch
3. Tensorflow
4. MxNet
5. AnalyticsZoo
6. BigDL
7. Devcloud
8. Movidius
9. FPGA & Open CL
10. OpenVINO
11. StarlingX

## 2 项目规划

### 2.1 整体目标

得益于 apple watch 4 领衔的智能手表向医用可穿戴设备转型的风潮，2018 年全球智能手表的出货量相比 2017 年增长 41%，2019 年第一季度又比去年同期增长了 48%。现在越来越多的智能手表都加入了 e-sim 功能，厂商也推出了手表端的应用市场，这使得智能手表

不再是手机的“附属品”，而是可以作为一个独立的智能设备使用。这样，作为交互手段之一的文字输入就是必不可缺的。

Wear os 内置了 Google 键盘和 Google 手写输入法，Google play 市场也存在搜狗输入法等各式键盘和手写输入法。这些输入法都需要另一只手在虚拟键盘上按键或直接在表盘上书写。考虑到大部分手表表盘的尺寸都在 2 英寸以下，Google 键盘在这之上布局的 26 键全键盘，误触率非常高；搜狗输入法的九宫格布局配合强大的字库，在中文拼音输入方面体验最佳；Google 手写输入采用了水平卷轴滚动的方式实现持续输入，但是在识别准确率和自动断句上的体验仍然差强人意。

表盘空间始终是有限的，很难兼容显示和输入。AirText 目标就是将输入拆分出来，使用用户在空中用手指书写字母时，智能手表所捕获的 IMU 数据，通过两阶段神经网络模型的学习和分类，识别出用户所书写的字母。通过迁移学习，可以针对每个用户的数据继续进行学习，生成每个用户专有的模型。结合单词预测来纠正单词中部分输入错误的字母，确保用户体验。

## 2.2 技术创新点

通过相关技术对比调研，简单阐述项目的主要技术创新点。

Through-Wall Human Pose Estimation Using Radio Signals 文中，直接对接收到的无线电信号进行卷积，得到人像热点图，再生成确切的人体姿态估计；WristText 在表带上部署了 12 个红外线感应器和 1 个压电振动传感器，高精度还原手腕动作以在表盘上实现单手对操纵杆的控制，重新设计了环形键盘，并实现了自动补全和自动校正；SignSpaeker 对 IMU 数据先进行了快速傅里叶变换，把时域信号变换为频域信号后，CTC 对单词进行切割，再利用 3 层双向 LSTM 进行分类和识别。

主要技术创新点：相比其他论文直接将数据放入模型进行分类，我们使用两阶段模型来提高分类准确率，第一阶段模型学习如何用 IMU 数据去还原手指轨迹，第二阶段根据学习得到的轨迹结果进行分类；其中第一阶段将光学传感器捕捉到的手指、手腕轨迹作为 ground truth，将 IMU 数据作为输入，经过插值放缩后直接放入神经网络进行训练，去学习写字时手部的运动轨迹；只利用市售的智能手表，不需要额外的硬件辅助，利用 IMU 数据实现单手输入字母。

## 3 实施方案

### 3.1 技术可行性分析

请阐述项目所需的数据如何采集、行业知识如何获取，是否有足够的算力、硬件支持保证项目顺利进行等可行性分析。

项目所需的数据集主要包含：

1) 用户佩戴可穿戴设备情况下，使用手势输入法时惯性传感器（加速度计，陀螺仪，磁力计）采集的传感器读数。我们使用商用智能手表 Huawei Watch 2 Pro, TicWatch Pro, LG Watch Urbane 等作为数据采集设备，这几款智能手表都基于安卓平台，是国内市场上占有率较高或用户评价较高的几款设备。同时，因为这几款设备基于安卓平台，我们已经开发了相



应的安卓应用采集传感器数据并记录，其中主要用到 `Android.Sensor API`。对于传感器数据的采集，我们设定了 `API` 所支持的最高采样率 200Hz，以最大程度的增加信息量，从而增加系统的精度。后期我们将会追加其它平台可穿戴设备上的数据采集和实验，如苹果手表和手环。

2) 用户佩戴可穿戴设备情况下，使用手势输入法时手部运动轨迹的记录。我们使用基于深度视觉的手部追踪设备 `Leap Motion` 记录各个手部关节的运动，轨迹以三维坐标的形式记录，作为阶段 1 训练时的 `ground truth`。`Leap Motion` 的追踪精度达到 0.01 毫米，其结果可以极好地表达不同手势手部运动的特征。

以上数据采集过程，除团队成员 4 人以外，另外将招募至少 20 人进行数据采集。招募的方式为有偿招募，支付每小时 120 元报酬。

项目需要的算力估计：目前团队已经收集 10 人的数据进行分阶段训练。总样本 23784 个，大小 1.45GB，阶段 1 训练集样本共 7510 个，评估集样本 939 个，测试集样本 939 个，阶段 2 训练集样本共 15830 个，评估集样本 3958 个，测试集样本 3996 个。团队使用阿里云深度学习服务器进行训练，以阶段 1 为例，使用 6 层深度神经网络迭代 12 次需要 4 小时时间，在相近的网络深度情况下，现有的服务器已经可以满足算力需求。

## 3.2 技术细节

请阐述项目相关技术细节，并结合可行性分析对核心技术进行论证及预期技术指标。

步骤 1，指尖轨迹学习。

将两组数据进行裁剪和时间对齐，压缩时相同时间长度，分别生成 `ground truth` 数据集 `G1` 和 `input` 数据集 `I1`。同时记录手写的字母作为阶段 2 分类模型使用的 `ground truth` 数据集 `G2`。分别计算掌骨的长度  $l_{metacarpal}$  和食指的长度  $l_{finger}$ ：

搭建卷积神经网络，各层从上到下逐层链接，分别是卷积层 `C1~C5`，扁平化 `F6`，全连接层 `F7`；优化器为 Adam 优化器；激活函数为 ReLU 函数；补偿系数 `k` 定义为 `G1` 中计算得到的食指长度  $l_{finger}$  和掌骨长度  $l_{metacarpal}$  的平均值与网络输出结果计算的到的食指长度  $l'_{finger}$  和掌骨长度  $l'_{metacarpal}$  的平均值的差；模型的损失函数设计为拟合轨迹（网络输出）与 `ground truth` 之间的欧式距离，再乘以补偿系数 `k`。

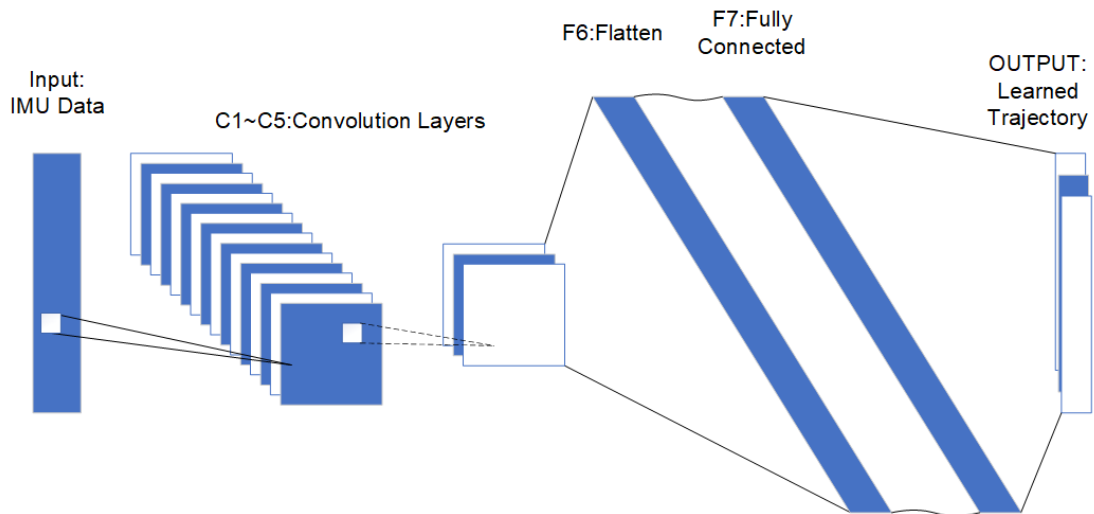


图 3-2 阶段 1 指尖轨迹学习网络模型

学习得到的手指轨迹与 ground truth 之间的距离差，训练集为 25.67mm，验证集为 35.16mm，测试集为 35.94mm，整体轨迹误差为 31.15mm。

#### 步骤 2，分类，迁移学习

4 层 CNN 网络，一层全连接层。将学习到的手指轨迹和提取的加速度计与陀螺仪特征作为阶段 2 分类学习的输入，输出结果为 26 个小写字母的识别概率。

迁移学习：针对多用户，我们采取迁移学习的方法，保持阶段 1 轨迹学习的模型不变，让新用户按照提示书写字母，在已学习到的阶段 2 分类模型基础上，将新用户的 IMU 数据放入模型持续进行学习，更新模型参数以提高该用户的识别准确率。

针对不同设备硬件不同的情况，实际实验表明，在不同的硬件上，使用相同的采样频率，同一用户保持原有的书写习惯，我们的模型不经过训练也能保持较好的识别率。

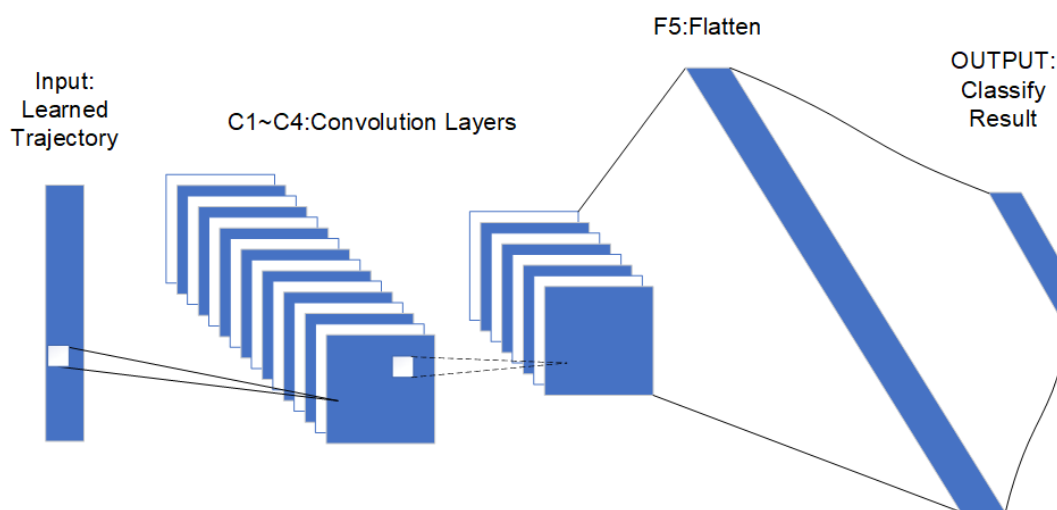


图 3-3 阶段 2 字母分类网络模型

提供阶段 1 轨迹数据的用户，在阶段 2 的字母识别率能达到 90.1~93.6%；不同用户由于字母书写速度、字母书写大小、手腕粗细等物理外因，在初始未经学习的分类模型上表现不佳，在经过迁移学习之后的识别率提升到 71.7~93.3%不等。与提供阶段 1 轨迹数据的用户书写方式相近的用户能获得几乎相同的最佳效果，而不同习惯的用户也能获得较好的体验。多设备方面，阶段 1 的训练数据集混合了 Huawei Watch 2 Pro、LG Watch Urbane 和 Ticwatch Pro 的 IMU 数据，阶段 2 训练则只使用了 LG Watch Urbane 的 IMU 数据，经过测试，三块手表的字母识别率分别为 87.3~89.5%、90.1~93.6%、88.0~91.2%。

#### 步骤 3，单词预测和纠错

通过字典树建立相应的单词查找库，逐一将每个单词的各个字母插入到字典树中；插入前先看前缀是否存在，如果存在，就共享该前缀字符串；否则创建对应的节点和边；字典树建立完成后，相邻节点之间的边代表一个字符，从根节点到某一节点的路径上经过的所有字符连接起来则代表该节点对应的字符串，并且每个节点对应的字符串都不相同；

对用户的当前输入字符串进行插入、删除、替换、转置相邻字符等四种字符串操作后，生成候选字符串集合。使用候选字符串集合去字典树中查询符合条件的单词。如果找到符合条件的单词，则根据词库中相应单词的出现频率，返回出现次数最多的前 N 个单词作为预测结果。如果未找到，则根据用户输入，使用相同的方法构建编辑距离为 2 的字符串集合，再次在字典树中进行查找并返回结果。

在经过单词预测后，单词错误率 WER 为 3.9~9.8%。多用户方面，基于他们不同的字母识别准确率，WER 为 6.3~18.9%。相比阶段 2 的字母识别率，WER 能进一步降低输入错误，

提升了用户体验。

### 3.3 计划和分工

结合参赛时间点，简单阐述本项目的整体计划和团队分工。

本项目的整体计划如下：

时间	计划
2019 年 6 月 17 日	完成报名
2019 年 6 月 18 日 - 2019 年 6 月 19 日	项目适用场景、潜在社会价值及可行性调研
2019 年 6 月 20 日 - 2019 年 6 月 21 日	根据调研结果撰写项目文档中的项目概况
2019 年 6 月 22 日	分配具体任务，撰写项目文档中的计划和分工
2019 年 6 月 23 日 - 2019 年 7 月 4 日	实现 AirText 基本功能组件
2019 年 7 月 5 日 - 2019 年 7 月 7 日	测试功能组件，调试并修正 BUG
2019 年 7 月 8 日 - 2019 年 7 月 9 日	采集不同用户的手写数据集
2019 年 7 月 10 日 - 2019 年 7 月 12 日	训练模型，并验证模型的可用性
2019 年 7 月 13 日 - 2019 年 7 月 15 日	根据实际应用情况对 AirText 进行改进
2019 年 7 月 16 日 - 2019 年 7 月 17 日	录制项目视频
2019 年 7 月 18 日 - 2019 年 7 月 19 日	撰写项目文档中的项目规划和实施方案
2019 年 7 月 20 日	根据项目文档撰写参赛作品简介
2019 年 7 月 21 日	根据项目文档撰写其他辅助材料
2019 年 7 月 22 日	提交作品
2019 年 7 月 23 日 - 2019 年 10 月 31 日	作品扩展（支持更多终端设备，如苹果手表、手环等；扩大实验，提高系统鲁棒性，提升多用户多设备友好度，降低 WER）

本项目的团队分工如下：

团队成员	具体分工
周寒	项目调研、撰写参赛作品简介和其他辅助材料
袁宇	项目调研、录制项目视频
曾思钰	实现 AirText 手势输入法、撰写项目文档
刘汶鑫	实现 AirText 手势输入法、撰写项目文档

## 4 参考资料

- [1] Zhao M, Li T, Abu Alsheikh M, et al. Through-wall human pose estimation using radio signals[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 7356-7365.
- [2] Gong J, Xu Z, Guo Q, et al. Wristext: One-handed text entry on smartwatch using wrist gestures[C]//Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. ACM, 2018: 181.
- [3] Hou J, Li X Y, Zhu P, et al. SignSpeaker: A Real-time, High-Precision SmartWatch-based Sign Language Translator[J]. 2019.

- [4] Gregory V Bard. 2007. Spelling-error tolerant, order-independent pass-phrases via the Damerau-Levenshtein string-edit distance metric. In Proceedings of the fifth Australasian symposium on ACSW frontiers-Volume 68. Citeseer, 117–124.
- [5] Eric Brill and Robert C Moore. 2000. An improved error model for noisy channel spelling correction. In Proceedings of the 38th Annual Meeting on Association for Computational Linguistics. Association for Computational Linguistics, 286–293.
- [6] Alex Graves and Navdeep Jaitly. 2014. Towards end-to-end speech recognition with recurrent neural networks. In International conference on machine learning. 1764–1772.
- [7] Mu Li, Yang Zhang, Muhua Zhu, and Ming Zhou. 2006. Exploring distributional similarity based models for query spelling correction. In Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics. Association for Computational Linguistics, 1025–1032.
- [8] Ji S, Xu W, Yang M, et al. 3D convolutional neural networks for human action recognition[J]. IEEE transactions on pattern analysis and machine intelligence, 2012, 35(1): 221-231.
- [9] Taktak M, Triki S, Kamoun A. 3D Handwriting characters recognition with symbolic-based similarity measure of Gyroscope signals embedded in smart phone[C]//2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA). IEEE, 2017: 319-326.
- [10] Amma C, Georgi M, Schultz T. Airwriting: a wearable handwriting recognition system[J]. Personal and ubiquitous computing, 2014, 18(1): 191-203.
- [11] Amma C, Schultz T. Airwriting: bringing text entry to wearable computers[J]. XRDS: Crossroads, The ACM Magazine for Students, 2013, 20(2): 50-55.
- [12] Islam R, Mahmud H, Hasan M K, et al. Alphabet recognition in air writing using depth information[C]//The Ninth International Conference on Advances in Computer-Human Interactions, At Venice, Italy. 2016.
- [13] Drotar P, Mekyska J, Rektorová I, et al. Analysis of in-air movement in handwriting: A novel marker for Parkinson's disease[J]. Computer methods and programs in biomedicine, 2014, 117(3): 405-411.
- [14] Shen S, Wang H, Roy Choudhury R. I am a smartwatch and i can track my user's arm[C]//Proceedings of the 14th annual international conference on Mobile systems, applications, and services. ACM, 2016: 85-96.
- [15] Vikram S, Li L, Russell S. Handwriting and Gestures in the Air, Recognizing on the Fly[C]//Proceedings of the CHI. 2013, 13: 1179-1184.
- [16] Pathak D, Krahenbuhl P, Darrell T. Constrained convolutional neural networks for weakly supervised segmentation[C]//Proceedings of the IEEE international conference on computer vision. 2015: 1796-1804.
- [17] Kim D W, Lee J, Lim H, et al. Efficient dynamic time warping for 3D handwriting recognition using gyroscope equipped smartphones[J]. Expert Systems with Applications, 2014, 41(11): 5180-5189.
- [18] Lee J Y, Wick M, Tristan J B, et al. Enforcing constraints on outputs with unconstrained inference[J]. arXiv preprint arXiv:1707.08608, 2017.
- [19] Xu C, Pathak P H, Mohapatra P. Finger-writing with smartwatch: A case for finger and hand gesture recognition using smartwatch[C]//Proceedings of the 16th International Workshop on Mobile Computing Systems and Applications. ACM, 2015: 9-14.

- [20] Meli L, Barcelli D, Baldi T L, et al. Hand in air tapping: A wearable input technology to type wireless[C]//2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN). IEEE, 2017: 936-941.
- [21] Tran D, Bourdev L, Fergus R, et al. Learning spatiotemporal features with 3d convolutional networks[C]//Proceedings of the IEEE international conference on computer vision. 2015: 4489-4497.
- [22] Yoon H, Park S H, Lee K T. Lightful user interaction on smart wearables[J]. Personal and ubiquitous computing, 2016, 20(6): 973-984.
- [23] Chen M, AlRegib G, Juang B H. 6dmg: A new 6d motion gesture database[C]//Proceedings of the 3rd Multimedia Systems Conference. ACM, 2012: 83-88.
- [24] Georgi M, Amma C, Schultz T. Recognizing Hand and Finger Gestures with IMU based Motion and EMG based Muscle Activity Sensing[C]//Biosignals. 2015: 99-108.
- [25] Adib F, Hsu C Y, Mao H, et al. Capturing the human figure through a wall[J]. ACM Transactions on Graphics (TOG), 2015, 34(6): 219.
- [26] Lin X, Chen Y, Chang X W, et al. SHOW: Smart handwriting on watches[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2018, 1(4): 151.
- [27] Kumar P, Saini R, Roy P P, et al. Study of text segmentation and recognition using leap motion sensor[J]. IEEE Sensors Journal, 2016, 17(5): 1293-1301.
- [28] Jing L, Dai Z, Zhou Y. Wearable handwriting recognition with an inertial sensor on a finger nail[C]//2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR). IEEE, 2017, 1: 1330-1337.