CUHK Jockey Club AI for the Future Project
Secondary School Think and Create Competition 2021

「眼」明「手」快

AEye

Project Proposal

St. Paul's Co-educational College
AI Project Group

0. Team Particulars

School Name:

St. Paul's Co-educational College

Project Title:

AEye「眼」明「手」快

Teacher Advisor:

Mr Tse Wai Tak

Team Members:

1	CHEUNG Chee Tung Dione	張芷澎	S.3C
2	WONG Ping Kuen Angus	王炳權	S.5E
3	CHAN King On Anthony	陳璟安	S.5I
4	SEE Chak Sum Matthew	施澤琛	S.6I
5	TANG Justin Kit Hang	鄧杰恆	S.6I

1. Interpretations on the Competition Theme

Definition

The competition theme, 'AI for the New Normal', could be interpreted on a broad spectrum. In a time when nobody can evade the impacts of the COVID-19 pandemic, it is necessary to give a clear definition to the term "New Normal". Therefore, presupposing a societal state in the post-COVID-19 era, we could define the term specifically as,

A state in which people adopt behaviours that match the requirements for their sustained well-being under COVID.

Upon this, we attempt to answer the questions: what 'new behaviours' do we find the most frustrating? What are the problems that arise from these 'new behaviours'?

Evaluating Topics

Based on our interpretation of the theme, we proceed to outlining the critical criterion for a proposed topic. We believe that a good topic should:

A. Solve an existing real-life issue
B. Do not overcomplicate, that it must be the simplest/most effective method to solve the issue
C. Leverage the power of existing technologies or knowledge related to AI
D. Can be practically implemented and cost-effective
E. Induce long-term societal change, such as raising awareness towards a certain issue

Direction of Exploration

Our team decided to approach the devise our topic based on a stakeholders' approach, that is, to analyse the COVID-related frustrations faced by various stakeholders in the society. Examples of such include children, the elderly, and the disabled. The aim is to apply AI in addressing to solve such frustrations.

2. Motivation

A News Article

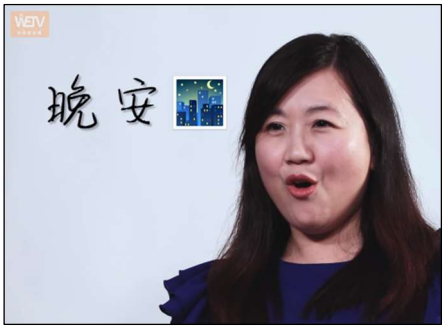
While discussing the effects of COVID on the disabled community, a news article came into our minds: A US university student, Ashley Lawrence, noticed the difficulties the deaf and mute face when reading facial cues with masks on, thus sewed her own transparent masks to help the community.[1]

This sparked our interest in the issue: with many in the deaf-mute community lacking access to cheap and protective transparent masks, their communication effectiveness is greatly hindered. Thus, we decided to see how we could solve this issue with AI.

Background Research

It is estimated that more than 70 million people use sign languages to communicate[2]. While sign language consists of multiple variants based on locale and language, *facial cues* are one indispensable part of sign languages. To better illustrate their importance, we generalized the role of facial cues into three categories: to imitate certain characteristics (Img. 1)[3], to show the pronunciation of a word (Img. 2)[4], or to deliver a sentiment associated with a word (Img. 3).

Img. 1 Facial cue of the word 「風」



Img. 2 Facial cue of the word 「安」

Img. 3 Facial cue of the phrase「唔明白」

[1] Coyne, M. (2020a, April 4). This See-Through Mask Lets The Deaf Communicate While Staying Safe. Forbes. https://www.forbes.com/sites/marleycoyne/2020/04/04/this-see-through-mask-lets-the-deaf-communicate-while-staying-safe/?sh=1606a8fd5257

[2] Without Sign Language, Deaf People Are Not Equal. (2020, October 28). Human Rights Watch. https://www.hrw.org/news/2019/09/23/without-sign-language-deaf-people-are-not-equal

[3] *The Three Little Pigs in Sign Language | Kong Wan Ki 江韻琪 | TEDxYouth@HongKong*. (2017, May 11). [Video]. YouTube. https://www.youtube.com/watch?v=E-sZjwQTk2g

[4] 香港手語教學 1：問候語篇｜Angel 手語世界. (2017, February 10). [Video]. YouTube. https://www.youtube.com/watch?v=c-PN62v1xH8

Therefore, it could be easily seen how, in the absence of facial cues, it would be extremely difficult for the mute and deaf to understand each other.

Necessity in Hong Kong

With the project based in Hong Kong, it is important to validate its necessity in Hong Kong. According to a headline article published in a 2016 CUHK Newsletter[5], out of the 15 thousand deaf people in Hong Kong, only 4,000 know sign language due to misconceptions that learning sign language would impede oral language development. Thus, based on the SLCO Programme led by Professor Gladys Tang aimed at promoting sign language to the deaf community, we believe that a higher social awareness towards sign language is needed. Together with the challenges to sign language brought by mask-wearing under COVID, this provides a strong justifies for our topic selection.

3. Existing Approaches on the Issue

Avoiding a Solution

The first approach is a 'primitive' yet undesirable solution under the 'new normal'. By this approach, sign language users would simply 'guess' what others are saying given the limited hand and facial cues (such as the movement of eyes). While this is possible, it hinders communication as it is hugely frustrating to make continuous guesses. However, this is the most prevalent solution given the regrettable lack of alternatives.

Design Approaches

The second approach is in line with what is mentioned previously. Individuals and companies design transparent masks for the deaf and mute community, such that they could see the mouth movements of each other. One example is the ClearMask™. Once again, while this is possible, disadvantages include:

A. Costly: each FDA-approved ClearMask™ costs approximately $20 HKD and is intended for single use only. Moreover, anti-fog technology must be incorporated in each mask. In comparison, surgical masks can be as cheap as $1 HKD/unit.

[5] Information Services Organization, CUHK. (2016, November). 培育手口同心下一代. 香港中文大學資訊處 Information Service Office, CUHK. https://www.iso.cuhk.edu.hk/chinese/publications/CUHKUPDates/article.aspx?articleid=1825

B. Unsafe: If sign language users choose to sew their own transparent masks, while it is cheaper, it compromises disease protection
C. Comfort: Transparent masks, made from plastic, are stuffy and are hard to breathe in
D. Personal acceptance: if the deaf-mute are the only ones who wear transparent masks in public areas, people around may be skeptical about their level of protection or pay excessive attention to those wearing these masks.

Legislative Approaches

Under the 'new normal', many regions require everyone to wear masks in public areas. However, certain legislation allows sign language users to not wear one. Taking Hong Kong's Cap. 599 as an example, "*a person has a reasonable excuse for not wearing a mask if – (b) the person is accompanying, or providing assistance to, another person who relies on lip reading to communicate with the person*"[6]. The disadvantage of this approach is obvious – sign language users are unprotected against diseases.

4. Proposed Solution

Overview

The aim of our proposed solution - AEye is to allow sign language users to understand each other, with regular masks on, using a pair of smart glasses and their mobile devices.

Technical Details

We propose a solution to automatically translate sign language (involving hand gestures together with mask motion arising from mouth movements) into comprehensible Cantonese sentences.

The mouth movements associated with different words would generate unique mask motion. Using the example of two commonly used Cantonese words, 「我」 and 「你」;「我」, when enunciated, would cause the mask to stretch vertically while 「你」would cause the mask to stretch horizontally. Our goal is to capture such minute mask movements to assist the interpretation of 'pandemic' sign

[6] "Requirement to wear mask in specified public place during specified period": Prevention and Control of Disease (Wearing of Mask) Regulation, Cap. 599 (2020)

language.

A small camera would be attached to a pair of specially designed smart glasses, utilized to capture both the hand gesture and mask movements of the speaker. The images captured by the camera are transmitted to a wirelessly linked phone and analysed by a trained AI model. Finally, the AI model will translate the images into written Cantonese dialogue and display the translated sentences in a phone app.

Advantages over Existing Solutions

AEye is effective in enabling communication between deaf-mute people who are using conventional masks. As mentioned, the most promising solution is the adoption of transparent masks. Comparing it with our solution,

Transparent Masks	AEye
• Costly • Not widely available	• Users can wear conventional masks • High benefits-cost ratio as a one-time investment • Sustainable; could be reused • Could be used offline

In addition, not only can our solution be used between deaf-mute people exclusively, but it also empowers communication for those that have no knowledge of sign language. They can make use of our solution to understand sign languages of the mute behind masks. Thus, with our solution, we hope that we can also raise social awareness towards the difficulties faced by the deaf-mute in Hong Kong.

5. Prototype Development

Hardware and Costs

Item	Description	Price (HKD)
Raspberry Pi	For wireless communication and powering the camera	$700
compatible camera (e.g., Rasp Cam 2)	For capturing the sign language and mask movements	$270
Wirelessly enabled smartphone	For the AI processing of captured videos	N/A
Glasses frame	For modifying into smart glasses	$30
	Grand Total:	$1000

Software

- Language: Python
- Training of AI: Google Colaboratory, TensorFlow
- Mobile phone application: CV2 (and for manual data-collection)
- Hand and mask tracking: Hand pose estimation networks

Training of the AI

Due to the novelty of our proposal, in tandem with the small database available for the Cantonese sign language system, we would need to create our own training dataset. As it is virtually impossible to cover all the words given the limitations of this high-school project, we would choose a dictionary of approximately 50 words, and manually record team members performing the words with sign languages and masks on. This process is recorded by a CV2-based mobile application and stored in a central database.

Then, the AI model is trained (supervised) on each video. While this depends on initial training results, we expect to track the hand gestures and mask movements by separate pre-trained networks.

Operation Mechanism

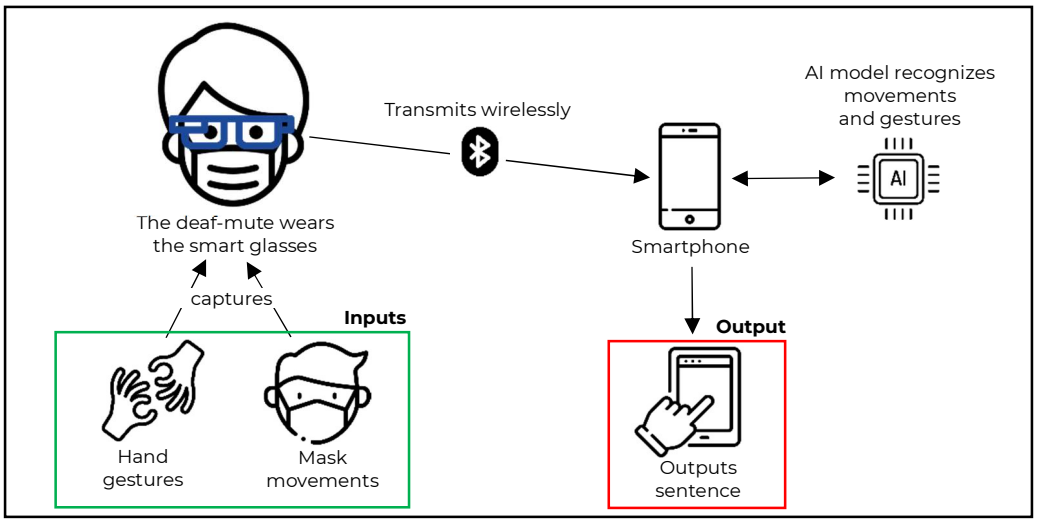


Figure 1. Operation Mechanism

The prototype design is shown in the above figure. It consists of two main parts: (1) the smart glasses, and (2) a smart mobile device. The camera on the glasses records the hand/masks movements and instantaneously transmits the video feed to the smartphone.

The AI model, pre-downloaded in the device, recognizes the corresponding hand gesture using computer vision and estimates the mouth location/movement based on the wearer's facial features. It searches in the dictionary for a word or vocabulary that fits both hand and face criteria, then displays the most probable result on the mobile device.

The wearer would only need to glance on his device if he is unsure about any part of the communication. The entire process is kept offline.

Discussions on Practicality

While the prototype smart glasses require connection to a relatively heavy mini-computing device, in theory, as all processing is completed at the smartphone's end, there is only a need for a power source (small battery) and a camera to be on the glasses. This proves our solution practical in real-life.

6. Linkages to Theme

Under the COVID 'new normal', we are unsure of when the health threats would subside. It may be possible that people need to wear masks for several years into the future. With our solution, the deaf-mute would no longer need to worry about the communication difficulties in such situations. Moreover, under globalization where people are more connected online, the smart glasses could also act as a 'sign language translator', allowing people to understand variants of sign language.

7. Addressing Existing AI Sign Language Detectors

Recognizing sign language using AI is not a novel idea. In fact, we came to found out that one participating team last year, backed by the HKU, worked on a seemingly similar topic[7]. However, upon close inspection, not only are the focuses and aims fundamentally different from each other, but it is also apparent that our solution remains novel as previous works have focused on pre-COVID sign languages, in which, following our discussion, cannot be applied under the 'new normal'. There are words that share highly similar signs that will confuse the original AI. In addition, the practicality and accuracy of our proposal is advantageous due to the introduction of smart glasses and use of a pre-trained body pose detector instead of overgeneralized convolutional neural networks that does not target any

[7] Z. Zhou, Y. Neo, K.S. Lui, V. Tam, E. Lam, N.Wong. 2020. A Portable Hong Kong Sign Language Translation Platform with Deep Learning and Jetson Nano. In The 22nd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '20). Association for Computing Machinery, New York, NY, USA, Article 89, 1–4. DOI:https://doi.org/10.1145/3373625.3418046

specific part of the body.

8. Ethical and Social Considerations

Social Benefits

The United Nations Convention on the Rights of Persons with Disabilities calls to promote the use of sign languages with the goal to ensure that people with disabilities can enjoy their rights on an equal basis with others.

Through adapting AI as to address COVID-related inequalities in society, we are imploring those around us to use the tools they are skilled and expert at to lend a hand to all those around them in need.

Privacy Concerns

There might be concerns for privacy since the fundamental prerequisite is for the AI to record one's communication. The system will be developed as a closed loop and will not connect to the Internet. For the initial development stages, only the team members' video and audio footage will be used to train the AI.

9. Conclusion

Our solution, AEye, effectively addresses the issue of impaired communication between the deaf-mute by leveraging the ability of AI into our daily lives. The system brings long-term social change in people's attitudes towards the disabled and encourages the use of technology to assist the disadvantaged.

Not only does it bear high social value, but it is also in accordance with the United Nations Sustainable Development Goals of reduced inequalities as well as sustainable cities and communities. These are goals that we find lacking in Hong Kong's society and potential future development.

Recalling the criteria to evaluate topics, we believe that AEye indeed fulfils the 5 criterion and would certainly bring a foreseeable improvement to the quality of lives of the disadvantaged.

THE END