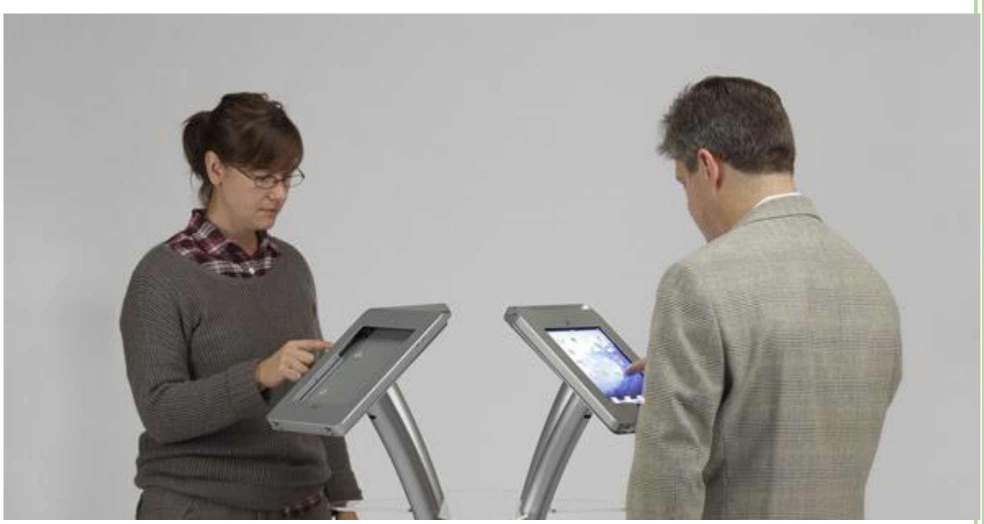
MIE1624 Project Report: Automated Sign language Interpretation via Gesture Recognition



Group 18

Adheesh Boratkar	1004641701
Alok Deshpande	999859309
Bibin Sebastian	1003752691
Nishank Thakrar	1004633079
Shaily Patel	1004218740

Table of Contents

Introduction

The United Nations’ new 2030 Agenda for Sustainable Development contains a number of goals that relate to Smart Cities. Many of these goals focus on inclusivity. This project is inspired by the goal of making cities and human settlements inclusive, safe, resilient and sustainable. In particular, we focus on inclusivity for the community with hearing and/or speaking impairments.

According to a 2007 report by the World Health Organization (WHO) [1], an estimated 360 million people around the world have some degree of hearing loss. Loss of productivity, due to unemployment and premature retirement among people with hearing loss, is conservatively estimated to cost and $105 billion worldwide annually [1].

In Canada, hearing and speaking disabilities are also significant issues affecting more than 800,000 people [2]. Hearing loss poses a significant challenge in their lives Unless suitable care is made available, the inevitable social exclusion can cause feelings of loneliness, isolation and frustration; in elderly individuals it may also lead to cognitive decline [1].

In this report, we propose that action be taken on the issue of hearing and speech disability in Canada. The report begins by first describing the key issues related to these disabilities, and the target populations that could benefit from a solution. A technical solution is then proposed that will reduce the communication barriers between Canadians with and without sensory impairments. Finally, the report demonstrates the business viability of this solution and its competitiveness on the market. For brevity, in this report we will use the word *disabled* to refer to those with hearing and speech disabilities, unless otherwise specified.

Key Issues

Work

While it is difficult to quantify all the difficulties that arise from disability, one can focus on employment, as this factor directly impacts most aspects of lifestyle. According to Statistics Canada, there is also an issue of underemployment: many people with disabilities work in low-income industries. In Canada, the average personal income is $32,600 for the former as opposed to $48,300 for the latter [3]. We moreover analyzed raw data from the US census for comparison [3], and discovered that the average income was 27.7 % higher for those without disabilities. A t-test was used to confirm that the difference was statistically very significant ($p<0.01$). In brief, there is an unmet need in the population with disabilities.

According to the survey in [4], people with disabilities most frequently mentioned three barriers that impede their ability to gain meaningful employment. By the percentage of people citing them, these are: “limited in ability to work” (63.4%), “few jobs available locally” (22.5%), and “training and/or experience not adequate” (22.2%). This shows that there are two general areas in which improvements can be made: improving education, and providing modifications to enable those with disability to perform work that they otherwise cannot.

With regards to work modifications, there are 15 major job aids that are provided to workers with all types of disabilities [5]. This includes “communication aids”, which we considered to be the most relevant for the deaf and mute. Based on the 2012 Statistics Canada survey containing this data, we found that 81.1% of workers needing this type of aid lack it.

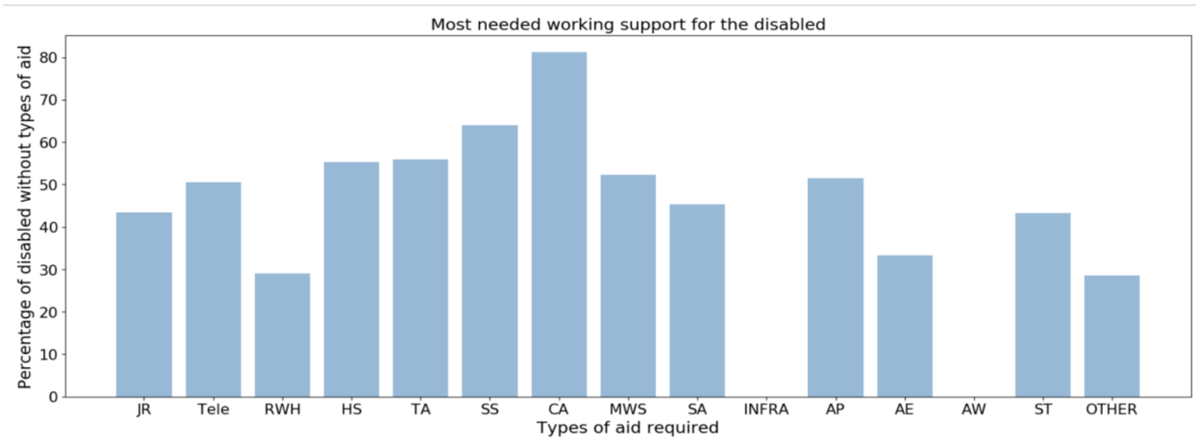


Figure 1 Most needing working supports for the disabled

As shown in Figure 1, this is the highest value amongst all 15 types. (see appendix for the list of acronyms)This shows that deaf and mute office workers are a neglected group who would could consider a product to meet their needs.

Education

Similarly, with regards to education, we discovered that disabled students are also a group that could benefit from a solution. This conclusion is based on two findings. Firstly, in Canada, there is a correlation between one’s level of education and income, as one can see from Figure 6 [6]. Clearly, increases in level of education are generally related with higher income.

Moreover, people with disabilities generally have lower levels of educational attainment: whereas 11% of Canadians without disabilities lack a high school diploma, about 22% of people with severe disability lack a diploma [2]. This implies that eliminating any factors which lead to students dropping out of school would improve their ability to improve their job prospects, as expected.

Secondly, we have inferred that the factors leading to students with disabilities dropping out of school occur at an earlier age. Based on US census data collected during the years 2009-2017 [3], we obtained counts of both deaf-or-mute students and hearing students in each grade in the American school system. We tracked the students who began in kindergarten over the course of the 10 years, and analyzed the year-over-year change in their numbers to determine trends. Figure 2 shows this change in the numbers of students for both disabled and regular students.

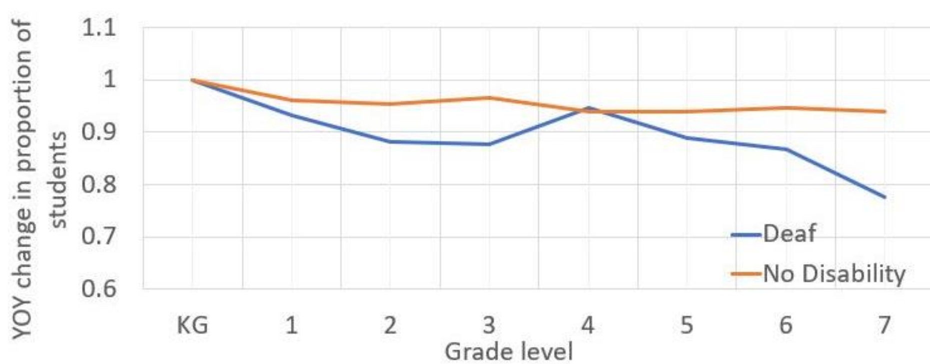


Figure 2 Student retention trends, for students entering in 2010

Although we only show results for 2010 for clarity, we observed the following trend for other years, too: there is a general drop in the numbers of disabled students compared to the regular ones over the same period. One can conclude that younger disabled students might need some form of support related to their condition.

Access to interpretation services

A final point one should consider in the case for helping the deaf-or-mute population is that they currently have limited access to resources. Aside from obtaining communication aids at work, the best way for such an individual to overcome his/her condition would be in hiring an interpreter, as there is currently no substitute for the human component of translation.

Unfortunately, there are also a limited number of sign language interpreters in Canada, all registered under the Association of Visual Language Interpreters of Canada (AVLIC). After web-scraping the AVLIC registry [7]to determine a count of the number of interpreters (847), we mapped their locations in Canada, as shown in Figure 3. This shows that almost all translators live in large cities, which poses a geographic issue for the disabled who live in more remote areas.

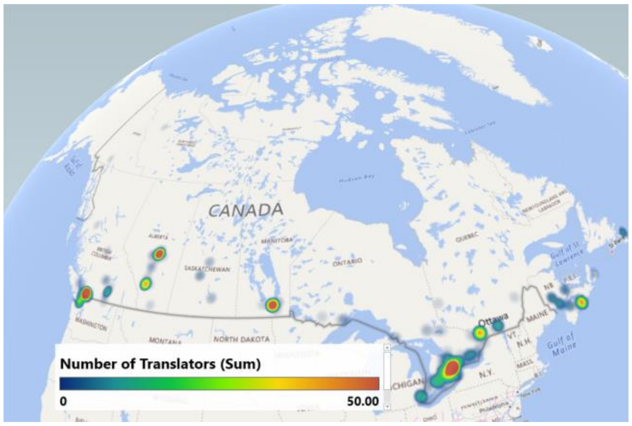


Figure 3 Map of Interpreters in Canada

Using the 2012 census data, we can obtain their numbers per province [8]. We used this to determine the ratio of interpreters to individuals, as shown in Figure 4. The labels for each province also show the number of interpreters in the province. As indicated by the red bars, in the case of PEI and Nunavut, there are actually no interpreters at all. For these cases, it is clear that an alternate solution must be found to support these Canadians.

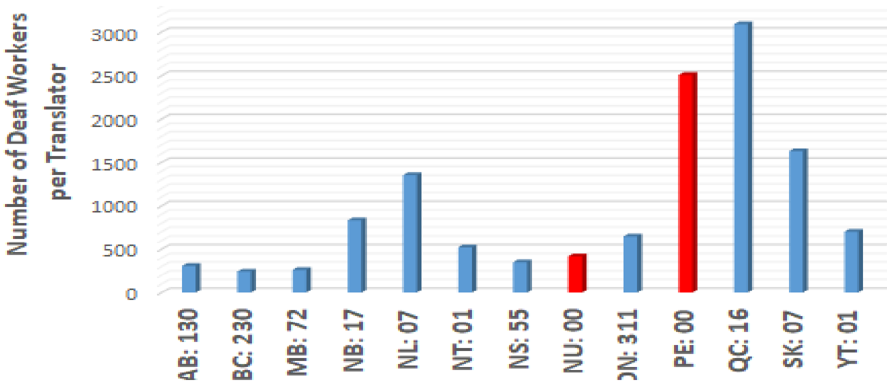


Figure 4 Disabled-to-Interpreters Ratio

The Solution

Hearing disability severely restricts an individual’s ability to communicate. To bridge this gap, we have developed an automated gesture recognition system called Otikos. In this section, we described the proposed solution.

Product Implementation

The system may be operated in two modes. In the Peer-to-Peer Mode, the system is deployed as a mobile application. Both the deaf/mute and their hearing/speaking colleagues, (hereafter referred to as colleagues) have the app installed on their smartphones. When the colleague speaks, the app on the deaf person's phone converts that speech to text. When the deaf/mute person expresses in sign language, the app on the colleague's phone interprets the sign language gestures from a video stream and displays the translation as text.

In the Enterprise Mode, a kiosk is deployed at enterprise locations like banks, ticketing counters, etc. It consists of screens on both sides and a camera facing the deaf person/customer. With this change, the operation works likewise. Additional hardware cost is comparable to that of two tablet computers.

Technology Overview

The system is implemented by the means of a deep learning algorithm called convolutional neural networks (CNNs). IBM Watson text to speech and speech to text conversion APIs are also deployed.

Initially, a CNN architecture called "InceptionV3" was set up, trained and tested. The primary limitation of this architecture is that it cannot detect objects within a frame and can only classify entire frames. It therefore could not detect the hand within the frame and had difficulty in correctly classifying the gestures based on the relative position of hand on the frame. The current system is based on the "YoloV3" architecture. This is able to detect a hand within the frame and then classify the gesture. While this yields much better results, it requires a larger, and more diverse training set.

For the purpose of this project, a dataset was manually create consisting of 4000 annotated images per character for letters, space, and two commands ("confirm" and "delete"). Models were trained on Google Cloud VM instance with Nvidia - Tesla K80 GPU (12GB GPU Memory) and the accuracy of frame classification from the video feed was >99% for both the Training and Validation sets.

Justification

Having described our solution, we now justify its financial viability. We first estimate the cost of the current market solution: sign language interpreters. We then present the cost of the proposed solution and show it is competitive.

Cost of Alternative Solution

As previously explained, the human component of sign language interpretation is invaluable for the those who are completely deaf or mute, of which there are approximately 150,000 in Canada [9]. However, this also makes their services expensive.

As a result, not every disabled person may hire an interpreter due to the hourly cost. However, we assumed that at least 20% will choose to buy the proposed service full-time if it has a sufficiently low cost. This meant that we had to estimate the costs that a market of 30,000 users *would* currently pay for interpreters, to fairly show that that our costs are much lower than this alternative and put an upper bound on proposed costs.

To determine the costs that individuals pay to interpreters, we first obtained the distribution of interpreter salaries in Canada. We initially used translator data from 2006 Canadian census, as visualized in Figure 7 in Appendix 1. This graph provides shows the range of salaries that individuals pay for translators. However, because this dataset is for translators in general (not just interpreters).

Secondly, we used this distribution to estimate the amount paid by a market of 30,000. As long as our total sales value is less than the total costs paid to interpreters, one can say with some probability that our product will cost less to a given person who is deaf/mute than an interpreter. We assumed that the number of people from the population spending a given amount on interpreters is proportional to the number of interpreters in each bin of Figure 8 in Appendix 1. Using this as a probability distribution, a Monte Carlo simulation was done to estimate the probability distribution of the total amount spent by the deaf population.

By converting the result into a probability distribution and then dividing by the number of interpreters, we obtained an estimate for hourly individual spending on interpreters.

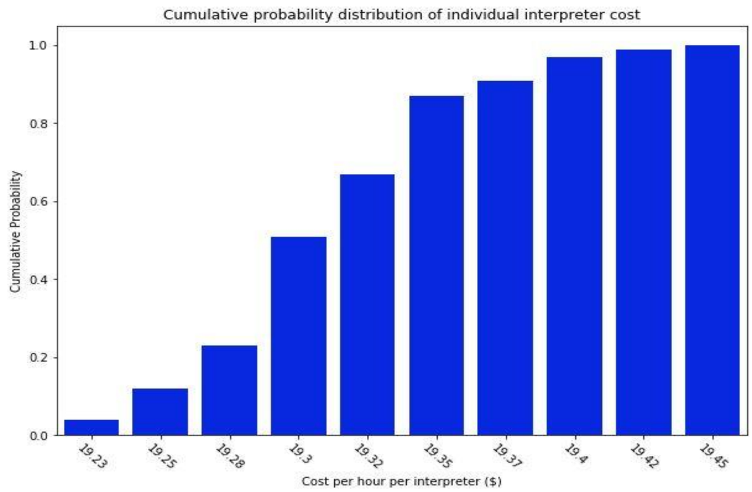


Figure 5 CDF of interpreter cost per individual

Figure 5 in Appendix 1 shows that our product will be preferable to hiring an interpreter when it costs less than approximately $19.20/hr. If it costs more, the above plot shows the probability that it will cost more for a given person who is deaf compared to hiring an interpreter.

Cost of Proposed Solution

Having presented the estimated cost of employing interpreters, we now present the cost of the Otikos system. We considered the costs involved in operating this service for one year. The capital investment and the non recurring engineering costs have not been considered for this analysis. They are assumed to have been paid for by the investors of the company. The product is proposed to be offered to cities/individuals only in the SaaS (software as a service) mode.

To maintain the product across the Android and iOS platforms after completion of the initial development, we estimated a workforce consisting of the positions shown in Table 1 of Appendix 1. The salaries for these positions were budgeted from the median salaries for these roles from Indeed.ca [9]. We further estimated the cost of office space. The annual operating cost for this service is $1.6 million. Adding a 30% operating profit margin and expecting to capture 20% of the market share amongst the severely deaf population, the **cost for a user $4.84/month**. This is clearly much lower than the cost of the main alternative solution, and so the proposed product is competitive.

Conclusion

Most issues faced by hearing disabled individuals across their education, work and social life arise from the communication barrier. Translator services are not adequately accessible. Otikos bridges this gap, with an affordable, automatic gesture recognition system.

References

[1] "Global costs of unaddressed hearing loss and cost-effectiveness of interventions: a WHO report, 2017," World Health Organization, Geneva, 2017.

[2] "Canadian Survey on Disability 2012," [Online]. Available: https://open.canada.ca/data/en/dataset/4ec6a8c9-b0b6-4f0e-9f4d-b144bf31decc.

[3] [Online]. Available: https://www2.census.gov/programs-surveys/acs/tech_docs/pums/data_dict/PUMS_Data_Dictionary_2008-2012.pdf.

[4] "Limitations and barriers to employment for adults with disabilities," Open Data Canada.

[5] " Modifications for labour force participation for adults with disabilities," Open Data Canada.

[6] [Online]. Available: https://www12.statcan.gc.ca/census-recensement/2016/as-sa/98-200-x/2016024/98-200-x2016024-eng.cfm.

[7] [Online]. Available: http://www.avlic.ca/about/members-directory.

[8] "Labour force status for adults with disabilities by disability type," Open Data Canada.

[9] "Android-Developer-Salaries," Indeed Canada, [Online]. Available: https://ca.indeed.com/salaries/Android-Developer-Salaries.

Appendix 1

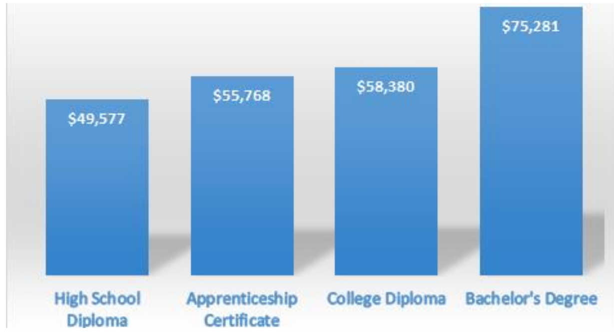


Figure 6 Income by Education Level bar graph

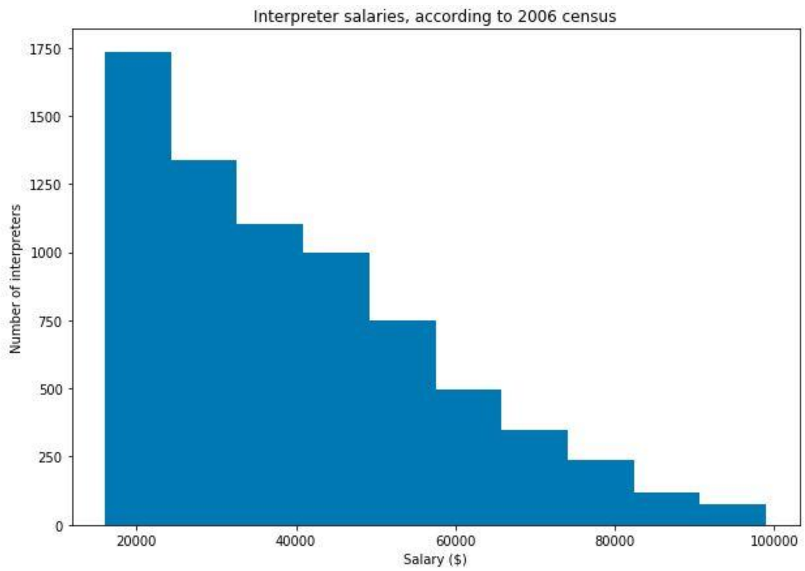


Figure 7 Interpreter salaries, according to 2006 census

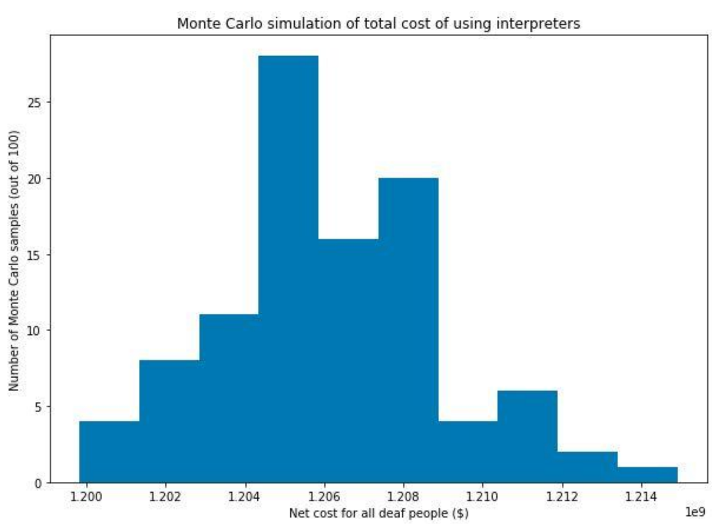


Figure 8 Monte Carlo simulation of total cost of using interpreters

Table 1 Operating costs for 1 year

Skill set/ Commodity	Unit Annual Expenditure	Number of workers/ units	Total Annual Cost
Android Developer	$85,258.00	2	$170,516.00
IOS Developer	$87,084.00	2	$174,168.00
Product Manager	$85,258.00	1	$85,258.00
Full Stack Developer	$84,463.00	2	$168,926.00
Data Scientist	$97,575.00	2	$195,150.00
Customer Service Representative	$31,200.00	5	$156,000.00
Office Space	$360	1800 sq feet	$648000.00
Total Cost			**$1,598,018.00**