HRI Project

FER – HRI
André Santos
Sérgio Estêvão

Motivation

- Study **Emotive Feedback** as an HRI investigation task
- Study its appliance in a **Learning Environment**

- **Sign language** is the primary language of communication for groups of people with disabilities.
- Amount of new users has been stagnating (6.000 users in the UK from 2011 to 2021 [1]);
- Potential to be used in various scenarios: interacting with deaf people, scuba diving, communicating in quiet areas, international dialogue;

1 - https://www.signature.org.uk/census-2021-british-sign-language/

Other Work

duolingo

L2TOR

Training the Model

Dataset

- 87000 images
- 26 letters
- 3 actions(space, delete, nothing)

Images too similar or in the same light conditions. The different experimented models have good accuracy on the test set but poor on new data inputs.

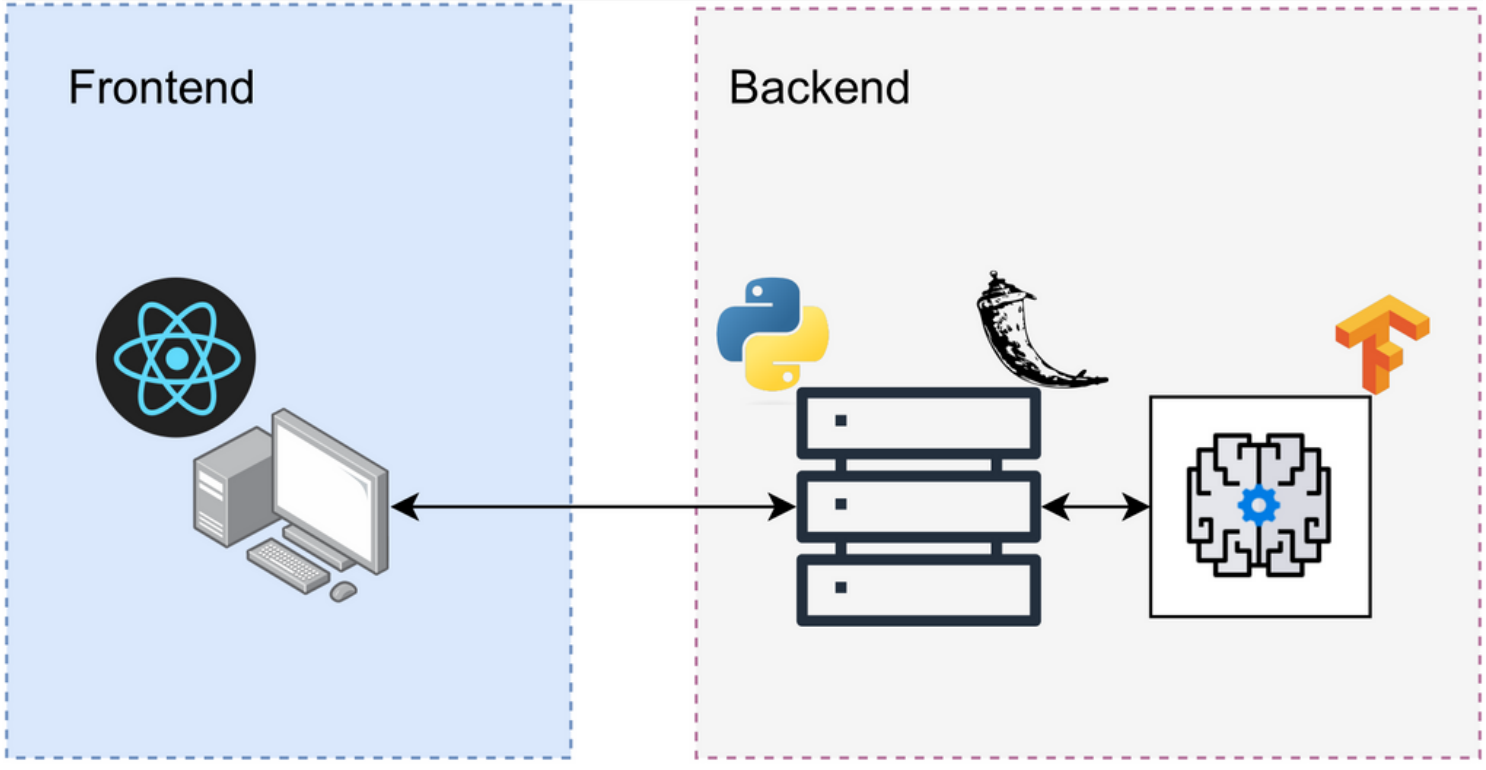
Model Overfiting

Sign Language

Some characters are too similar so the model has high uncertainty in the output

Validate various outputs

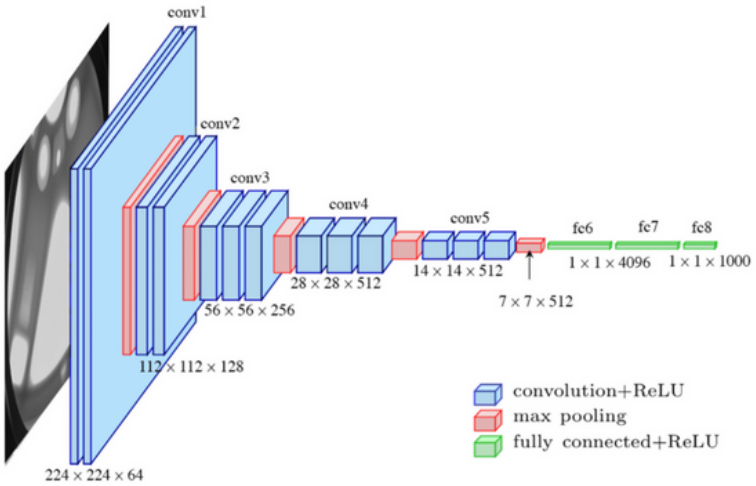
Architecture



Training the Model

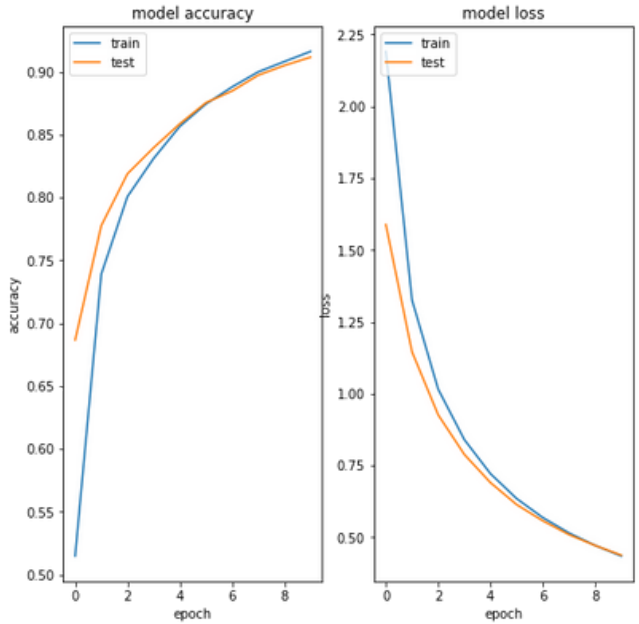
VGG16 + DenseLayer

- CNN'S are the best in image classification tasks
- Sequence of Convolutions help extracting features in the image
- Outputs likelihood for each of the classes



Performance

- 80/20 Train/Test Split
- Optimizer: Adam
- Learning Rate: 0.0001
- Number of Epochs: 10
- Loss function: Categorical cross-entropy



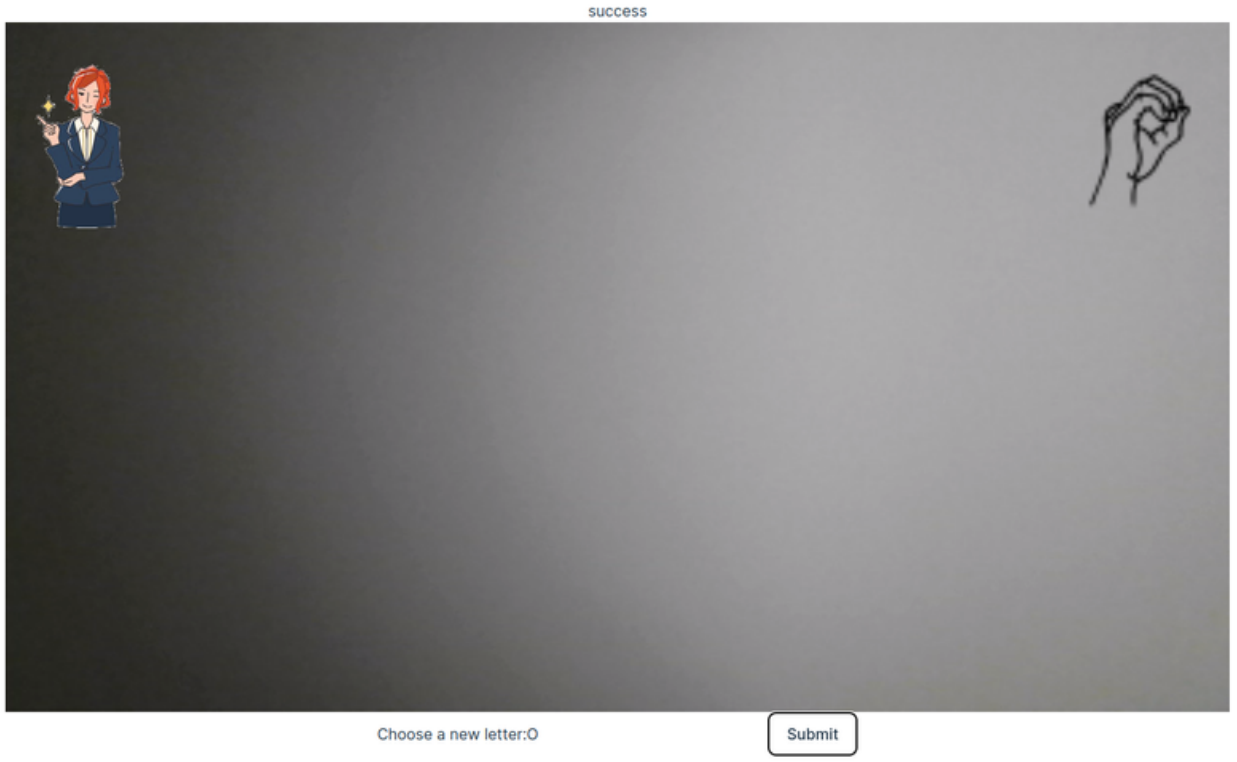
Interface

Feedback

- Audio Cues
- Cartoon Animation
- Emoji Prompt
- Human-like Figure

Goal

- Correct hand sign for the chosen letter



Methodology

Experiments

- User Studies (6 participants):
  - Duration of 10/15 minutes;
  - Ages between 18 to 22
  - 3 male + 3 female

Evaluation

- Qualitative:
  - Users evaluate their experience generally by sharing their opinion on the experience and effect.

- Quantitative:
  - User Surveys;
  - 1 to 5 grading;

Results

- Q1: How did you find your experience? (No Feedback)
- Q2: How did the emoji prompt affect your experience?
- Q3: How did the audio feedback affect your experience?
- Q4: How did the presence of a human-like figure affect your experience?
- Q5: How did the appearance of cartoons affect your experience?
- Q6: How easy was learning sign language? (No Feedback)
- Q7: How easy was learning sign language? (Generally w/Feedback)

Questions	Subjects					
	A	B	C	D	E	F
Q1	4	4	3	4	5	3
Q2	3	4	3	3	3	3
Q3	2	4	3	4	4	3
Q4	3	4	3	3	4	3
Q5	3	4	3	3	3	3
Q6	1	3	2	3	2	2
Q7	1	3	2	3	3	2

Table 1. Survey on Subjects Perception of the Application

'Wizard of Oz' simulating success influences the emotion perception but not learning

Conclusions

User feedback:

- Hard to evaluate
- Almost impossible to remove bias

Possible Interactions:

- Robot vs Software tradeoff
- Interaction vs Accessibility

Sign Language Evaluation Limitations were a hard to solve issue.

Overall improved the user experience in a sign language learning environment.

Related Work

[1] [n. d.]. ASL Alphabet. https://doi.org/10.34740/KAGGLE/DSV/29550
[2] Kerstin Dautenhahn and Aude Billard. 1999. Bringing up robots or—the psychology of socially intelligent robots: from theory to implementation. In Proceedings of the Third Annual Conference on Autonomous Agents (Seattle, Washington, USA) (AGENTS '99). Association for Computing Machinery, New York, NY, USA, 366–367. https://doi.org/10.1145/301136.301237
[3] Duolingo Inc. [n. d.]. Duolingo. https://www.duolingo.com/
[4] ITU AI for Good Global Summit. [n. d.]. The Future of Educational Robotics: Enhancing Education, Bridging the Digital Divide, and Supporting Diverse Learners. https://aiforgood.itu.int/the-future-of-educational-robotics-enhancing education-bridging-the-digital-divide-and-supporting-diverse-learners/
[5] Yueru Lang, Ke Xie, Shaoying Gong, Yanqing Wang, and Yang Cao. 2022. The Impact of Emotional Feedback and Elaborated Feedback of a Pedagogical Agent on Multimedia Learning. Frontiers in Psychology 13 (2022). https://doi.org/10.3389/fpsyg.2022.810194
[6] Khuyen Le. 2021. An overview of VGG16 and NiN models. https://medium.com/mlearning-ai/an-overview-of-vgg16-and-nin-models-96e4bf398484
[7] Manav Mandal. 2023. Introduction to Convolutional Neural Networks (CNN). https://www.analyticsvidhya.com/blog/2021/05/convolutional-neural-networks-cnn/
[8] Masahiro Mori. 2012. Uncanny Valley. https://spectrum.ieee.org/the-uncanny-valley
[9] Sharon Oviatt, Björn Schuller, Philip R. Cohen, Daniel Sonntag, Gerasimos Potamianos, and Antonio Krüger (Eds.). 2017. The Handbook of Multimodal-Multisensor Interfaces: Foundations, User Modeling, and Common Modality Combinations - Volume 1. Vol. 14. Association for Computing Machinery and Morgan & Claypool.
[10] Paul Timothy Mooney. 2018. Interpret Sign Language with Deep Learning.
[11] Anna Rowe, Julie Fitness, and Leigh Wood. 2014. The role and functionality of emotions in feedback at university: A qualitative study. Australian Educational Researcher 41 (07 2014), 283–309. https://doi.org/10.1007/s13384-013-0135-7
[12] Paul Vogt, Rianne van den Berghe, Mirjam de Haas, Laura Hoffmann, Junko Kanero, Ezgi Mamus, Jean-Marc Montanier, Cansu Oranç, Ora Oudgenoeg-Paz, Daniel H. Garcia, Fotios Papadopoulos, Thorsten Schodde, Josje Verhagen, Chris D. Wallbridge, Bram Willemsen, Jan de Wit, Tony Belpaeme, Tilbe Göksun, Stefan Kopp, Emiel Krahmer, Aylin C. Küntay, Paul Leseman, and Amit K. Pandey. 2019. Second Language Tutoring using Social Robots. A Large-Scale Study. In Proceedings of the 2019 ACM/IEEE International Conference on Human-Robot Interaction(HRI 2019)