## Introduction

Mission: Bring vision to the blind or people with visual impairment through a conversational platform.

The three goals of our product are:
- Tell users what is contained in an image
- Give gender, age, emotion, number information of people in an image
- Understand the image scene, and be able to describe it

## Interesting Points

- **Social Impacts**
  Help people with visual impairment through AI
- **Technical Novelty**
  Integrate conversation with vision and language tasks

## Use Cases

- Know what’s happening in image
  Q: Tell me what happens.
  A: Well... It is likely to be a young man riding a skateboard.

- Know the information of people
  Q: How many people are there?
  A: Interesting... There are three. More questions?
  Q: How old is the left woman?
  A: Well... I consider: she is about 47.
  Q: Is she happy?
  A: Yes, she is. Any more questions?

## Convision

Convision is a smart conversation AI implemented with the Clinc platform. This tool will allow users to talk with it and understand what’s happening based on the image input. They can also understand the information of people in the image through follow-up dialogues. It provides a starting point of bringing vision to the people with visual impairment through conversations.

![Fig 1. Architecture Graph of Convision](image)

![Fig 2. Use Cases](image)

![Fig 3. State Graph of Convision](image)

Linyi Jin, Ray Cao, Shengyi Qian, Chung-Hsun Wang, Yi Wen, I-Le Wu, Xinyi Zheng. Instructor: Prof. Jason Mars

## Impact

Future impacts of Convision include:
- Help the blind or people with visual impairment perceive the world
- Connect classical vision tasks with conversational AI features

Send feedback and ideas or get more info: [convision@umich.edu](mailto:convision@umich.edu)