

W19 EECS 498 Conversational AI: Principles and Practice 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

Bring Vision to the Blind through Conversations

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 bring vision to the people with visual impairment through conversations.

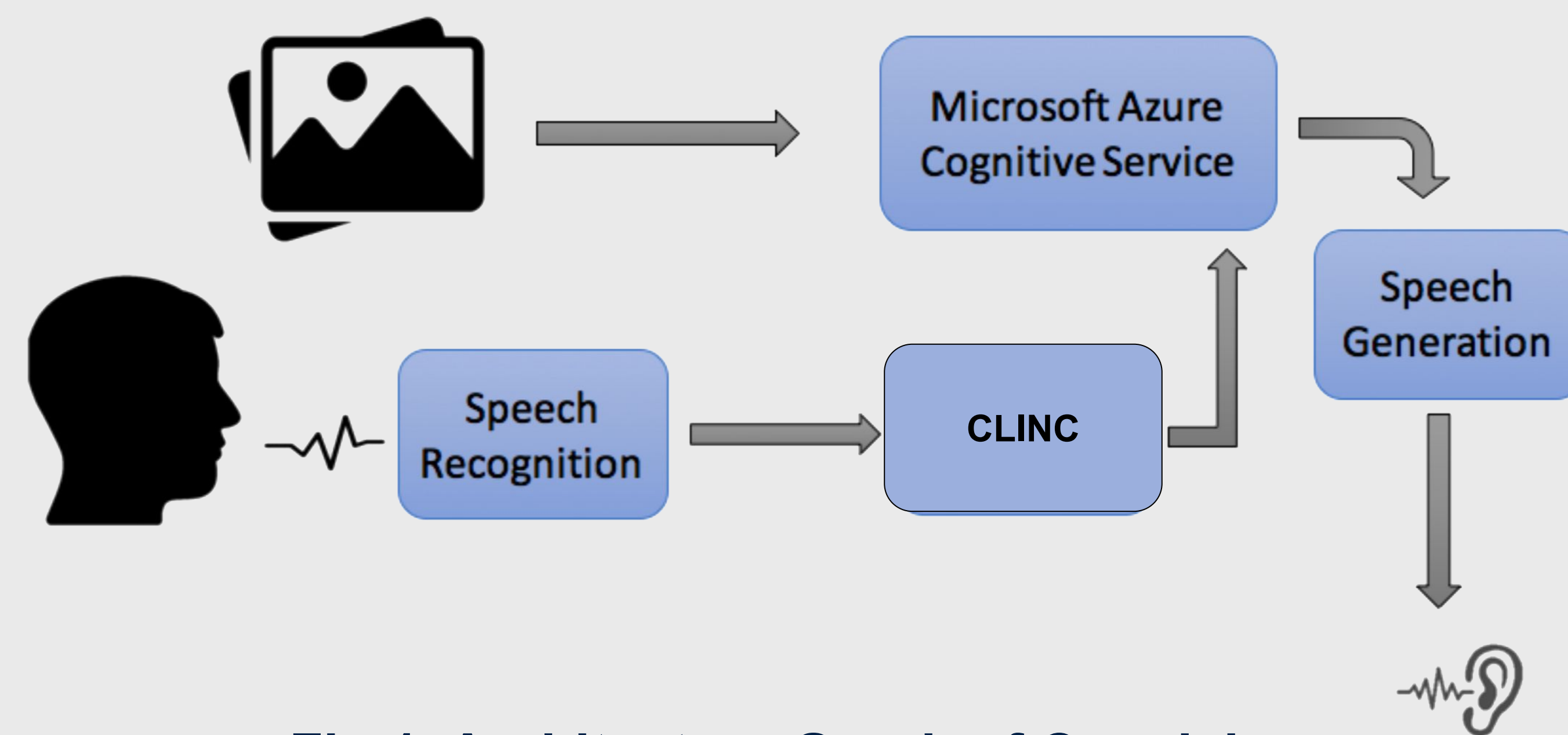


Fig 1. Architecture Graph of Convision

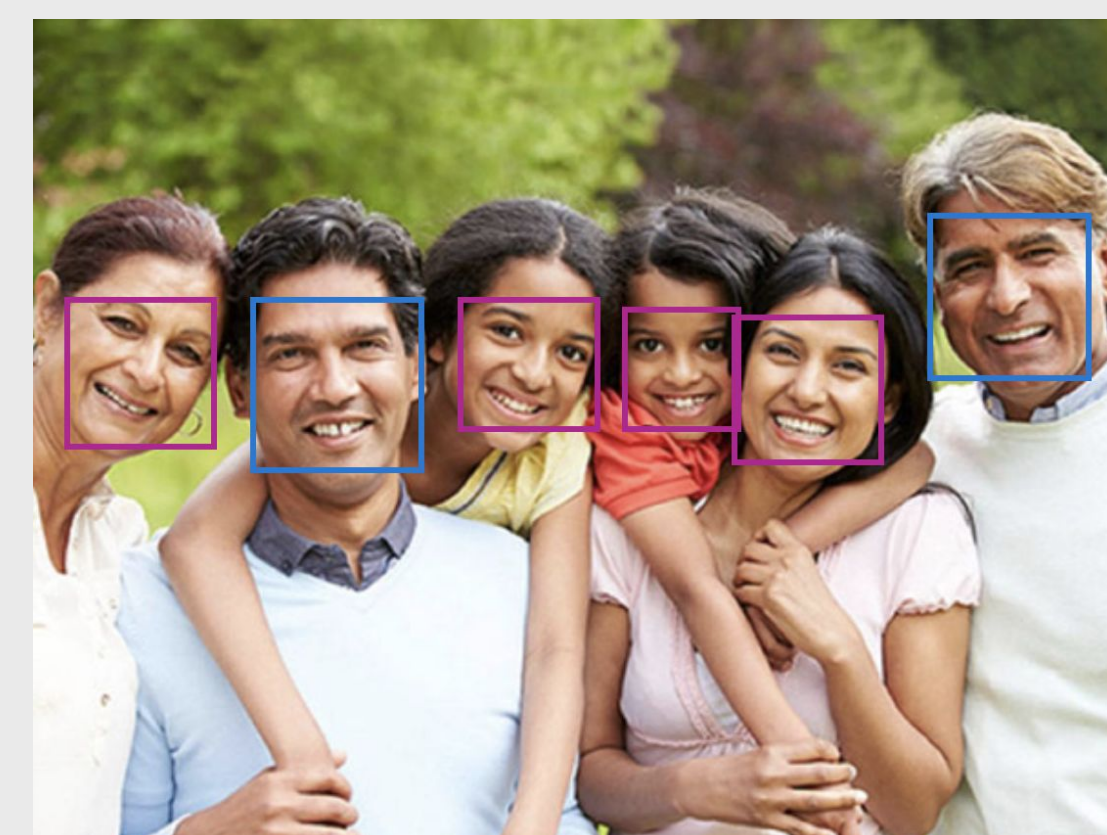
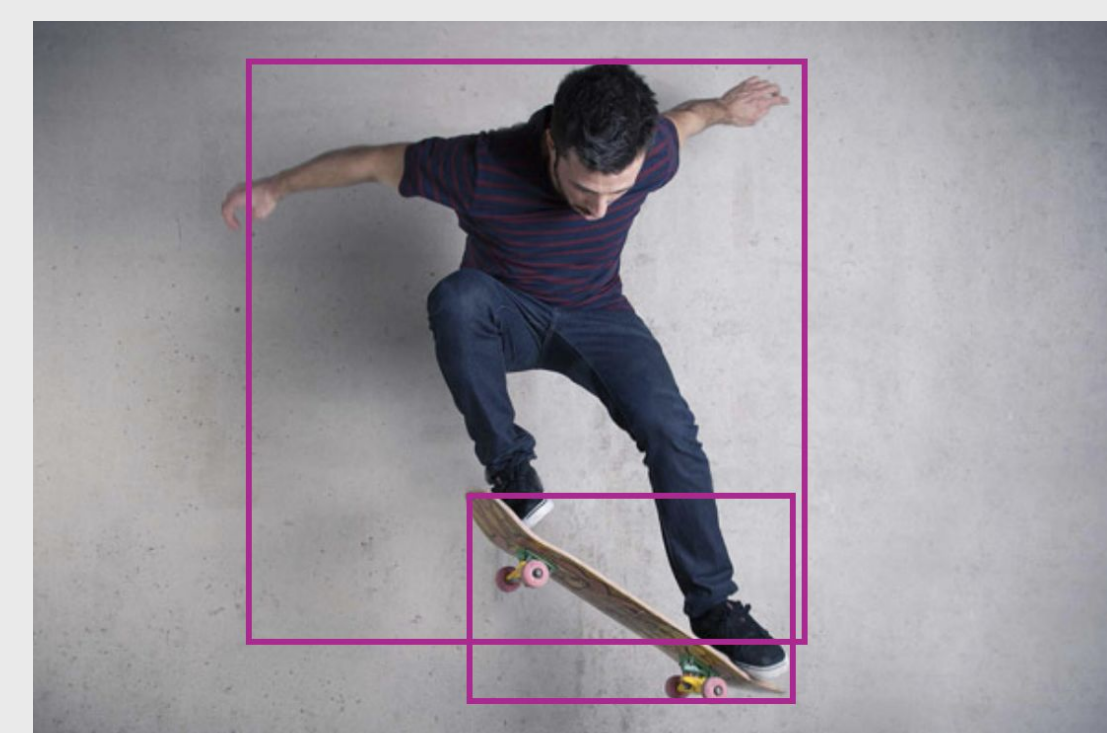


Fig 2. Use Cases

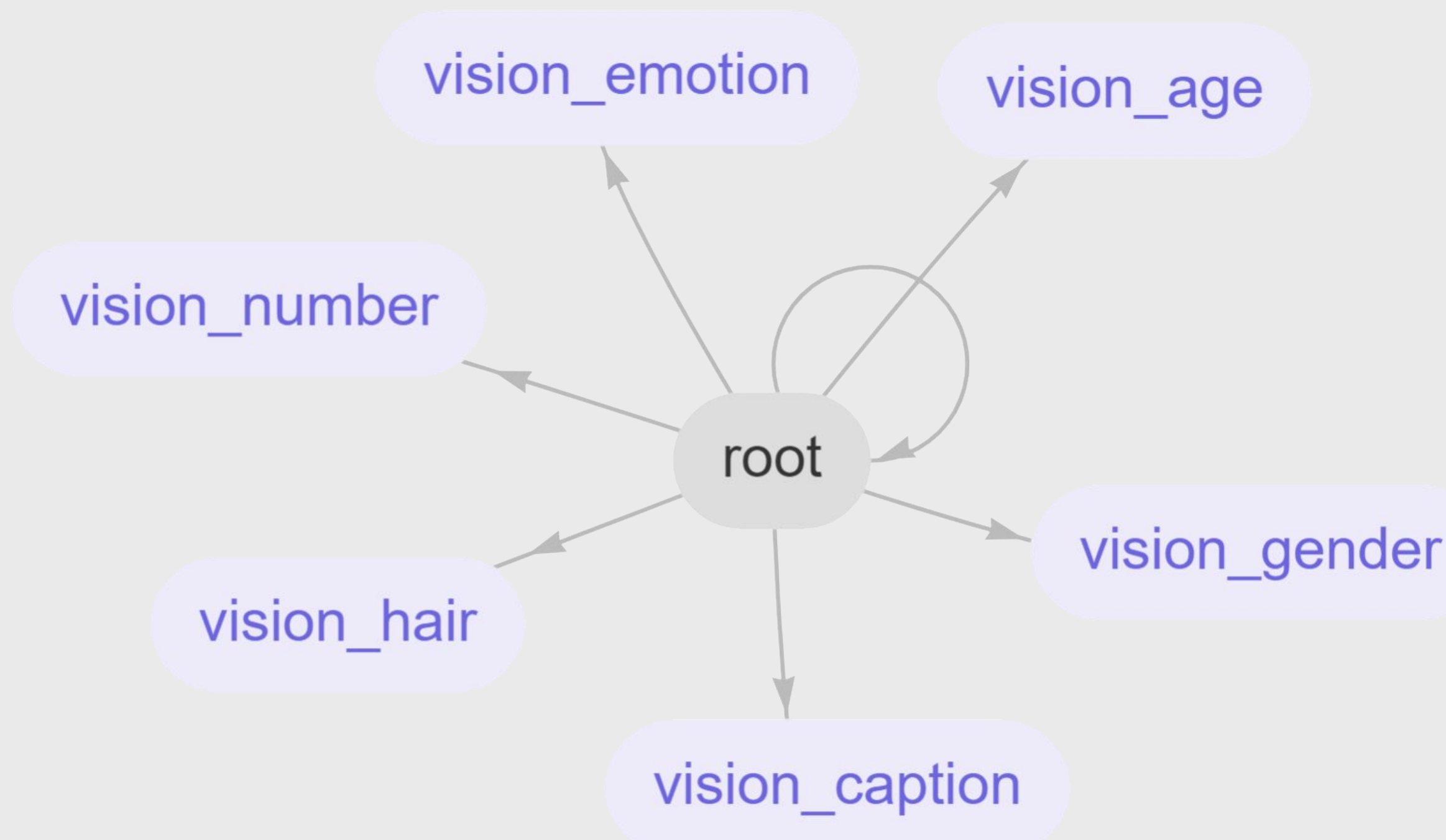


Fig 3. State Graph of Convision

Class & SVP Scope

- Slot Names
 - Gender
 - Emotion
 - Age
 - Number of people
 - Position
 - Order
 - Color
 - Subject
- Example Utterances
 - Are there **two angry man** in the picture? (vision_number)
 - What's the feeling of the **old female**? (vision_emotion)
 - Is the person in the picture a **male**?(vision_gender)
 - How many **happy female** are there? (vision_emotion)
 - Can you tell me the gender of the **first** people from the **left**? (vision_gender)
 - Is the **second** people from **right** **unhappy**? (vision_emotion)
 - Tell me the hair color of that **happy female**. (vision_hair)
 - Does the **woman** have **black** hair? (vision_hair)
 - Is **she** **glad**?(vision_emotion)

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