



Grand Challenge:
Answering Visual Questions from Blind People

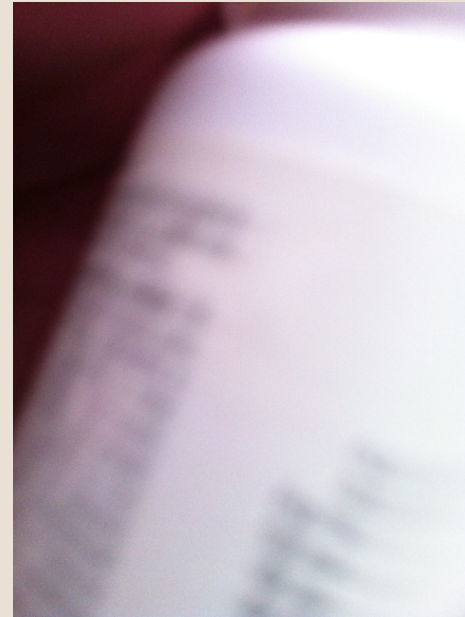
Motivation: Answer Visual Questions



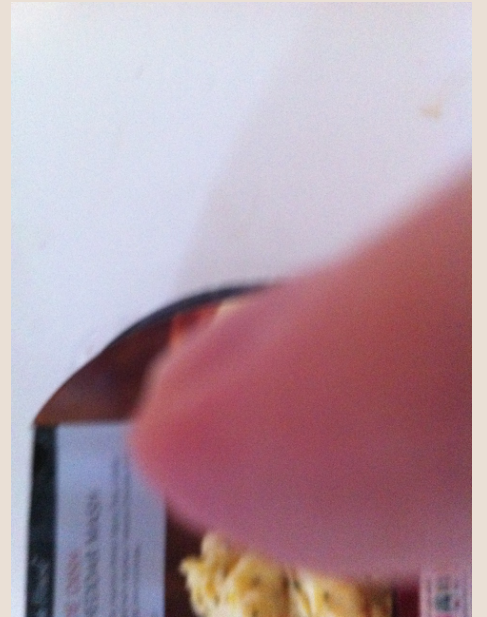
Is this shirt
clean or dirty?



Hi there can you
please tell me what
flavor this is?



What type of pills
are these?



What is this?

Motivation: Assistive Technology for Blind People



Workshop Goals

- What are current research and application issues?
- What are other meaningful challenges [beyond VQA]?
- What are meaningful evaluation protocols?
- How can we accelerate converting cutting edge research into market products that have social impact?
- How to maintain a community for this work?

Invited Speakers



Jeffrey P. Bigham
Carnegie Mellon University



Kris Kitani
Carnegie Mellon University



Devi Parikh
Georgia Tech



Saqib Shaikh
Microsoft (Seeing AI)



Yonatan Wexler
Orcam



Roberto Manduchi
UC Santa Cruz

Schedule – Morning

- 9-9:10am: Opening remarks
- 9:10-9:30am: Jeffrey Bigham ("VizWiz: From Visual Question Answering to Supporting Real-World Interactions")
- 9:30-9:50am: Kris Kitani ("Wearable Sensing for Understanding, Forecasting and Assisting Human Activity")
- 9:50-10:10am: Devi Parikh ("Forcing Vision and Language Models to Not Just Talk But Also Actually See")
- 10:10-10:30am: Break
- 10:30-10:50am: Overview of challenge, winner announcements, & results analysis
- 10:50-11:20am: Talks by challenge winners
- 11:20-12:30pm: Poster session
- 12:30-1:45pm: Lunch

Schedule – Afternoon

- 1:45-2:05pm: Saqib Shaikh ("Seeing AI: Leveraging Computer Vision to Empower the Blind Community")
- 2:05-2:25pm: Yonatan Wexler ("OrCam: Life-Changing Wearable AI")
- 2:25-2:45pm: Roberto Manduchi ("Finding and reading scene text without sight")
- 2:45-3:15pm: Break
- 3:15-3:45pm: Panel discussion
- 3:45-4pm: Open discussion
- 4-4:10pm: Closing remarks

Let's Begin!

- 9-9:10am: Opening remarks
- 9:10-9:30am: Jeffrey Bigham ("VizWiz: From Visual Question Answering to Supporting Real-World Interactions")
- 9:30-9:50am: Kris Kitani ("Wearable Sensing for Understanding, Forecasting and Assisting Human Activity")
- 9:50-10:10am: Devi Parikh ("Forcing Vision and Language Models to Not Just Talk But Also Actually See")
- 10:10-10:30am: Break
- 10:30-10:50am: Overview of challenge, winner announcements, & results analysis
- 10:50-11:20am: Talks by challenge winners
- 11:20-12:30pm: Poster session
- 12:30-1:45pm: Lunch

Next Up...

- 9-9:10am: Opening remarks
- 9:10-9:30am: Jeffrey Bigham ("VizWiz: From Visual Question Answering to Supporting Real-World Interactions")
- 9:30-9:50am: Kris Kitani ("Wearable Sensing for Understanding, Forecasting and Assisting Human Activity")
- 9:50-10:10am: Devi Parikh ("Forcing Vision and Language Models to Not Just Talk But Also Actually See")
- 10:10-10:30am: Break
- 10:30-10:50am: Overview of challenge, winner announcements, & results analysis
- 10:50-11:20am: Talks by challenge winners
- 11:20-12:30pm: Poster session
- 12:30-1:45pm: Lunch

Next Up...

- 1:45-2:05pm: Saqib Shaikh ("Seeing AI: Leveraging Computer Vision to Empower the Blind Community")
- 2:05-2:25pm: Yonatan Wexler ("OrCam: Life-Changing Wearable AI")
- 2:25-2:45pm: Roberto Manduchi ("Finding and reading scene text without sight")
- 2:45-3:15pm: Break
- 3:15-3:45pm: Panel discussion
- 3:45-4pm: Open discussion
- 4-4:10pm: Closing remarks

Finally...

- 1:45-2:05pm: Saqib Shaikh ("Seeing AI: Leveraging Computer Vision to Empower the Blind Community")
- 2:05-2:25pm: Yonatan Wexler ("OrCam: Life-Changing Wearable AI")
- 2:25-2:45pm: Roberto Manduchi ("Finding and reading scene text without sight")
- 3:10-3:25pm: Break
- 3:25-3:50pm: Panel discussion
- 3:50-4pm: Open discussion
- 4-4:10pm: Closing remarks