Intuitive Physics in Virtual Reality

Yu-Yu Chen
American University
Data Science Practicum Project
Adviser: Dr. Bei Xiao

When playing Jenga...

All the players are trying to make the tower stay stable...

But how do we know?

This Study

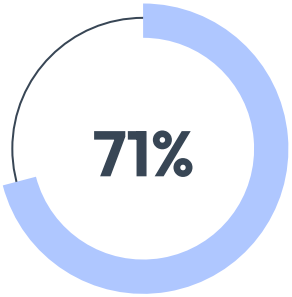
- **Human Intuitions about Tower Stability in VR**
  - Evaluate various tower designs in an immersive virtual reality environment
- **Predictive Modeling**
  - Image-based models to predict whether towers will fall or not

Experiment Design

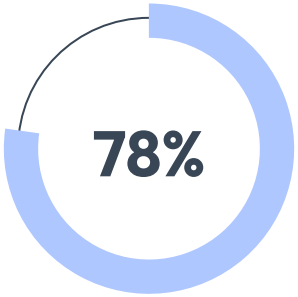
- **50 unique** tower design
- For each tower:
  - Rate **stability** on 1 – 7 scale
  - Predict **likely fall direction**
- 3 evaluation rounds, 150 total trials

Demo

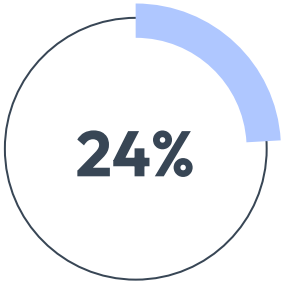
Human Behavior Results



Accuracy in fall prediction



Consistency in repeated judgments



Random Guess

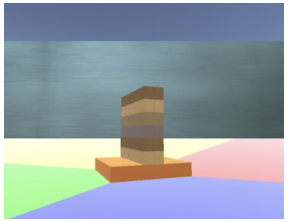
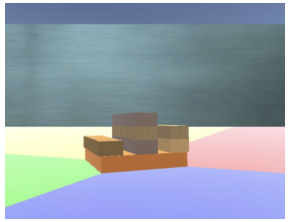
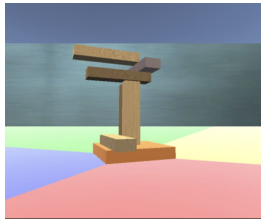
Observers rely on well-developed intuitive skills and knowledge of physical principles rather than simple guessing.

Tower Difficulty

The portion of participants that correctly judged the towers' falling direction

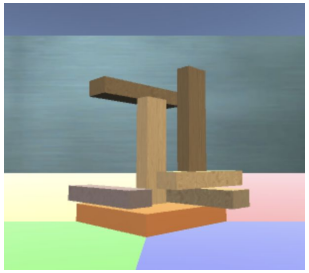
Easiest

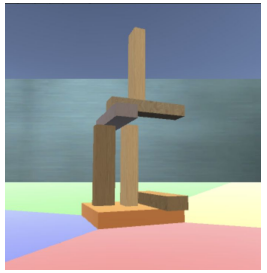
(> 90% correct)

Medium

(~ 70%)





Hardest

(< 40%)

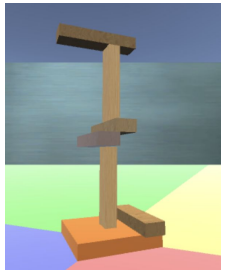
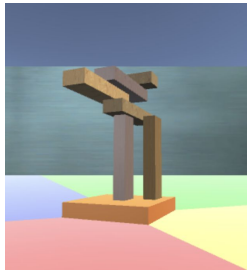
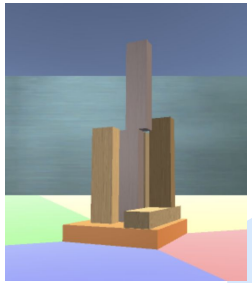




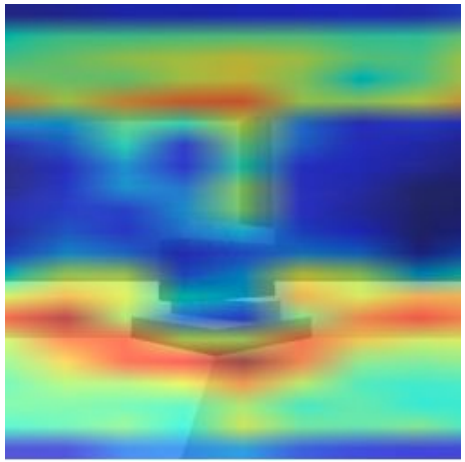
Image-based models Performance

Dataset: Images of different viewing angles of 50 towers, total 164 unstable & 144 stable towers' photo

- Over 95% accuracy
- Precision drop from training:
  - Potential overfitting
  - Sensitivity to variations

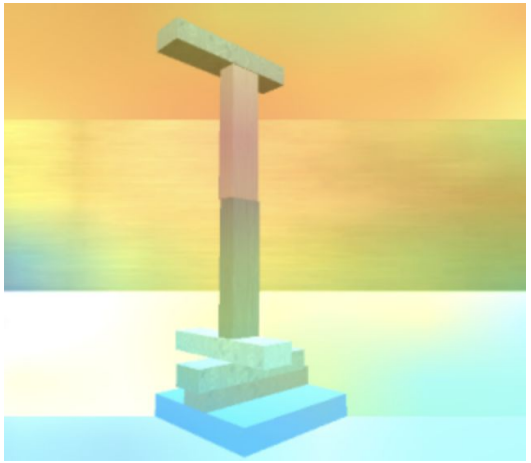
	ResNet50		InceptionV3	
	Training	Testing	Training	Testing
Accuracy	**95.98**	**95.45**	**96.43**	**95.45**
Precision	95.45	**85.71**	95.49	**85.71**
Recall	97.67	100.00	98.45	100.00
F1 Score	96.55	92.31	96.95	92.31

Model Explainability



ResNet50

Focuses on top/bottom



InceptionV3

Considers full structure

Spectrum from warm to cool.

Red =

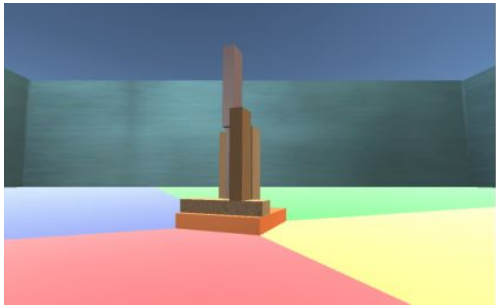
Areas that contribute most to the decision making process

Blue =

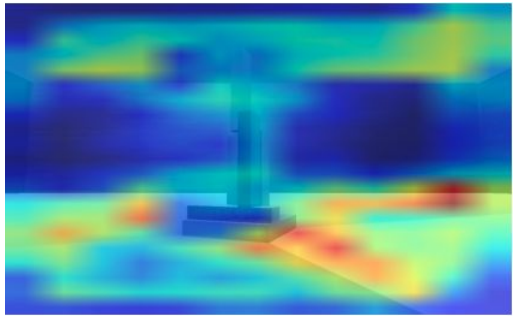
Areas that contribute least to the decision making process

Interesting Case

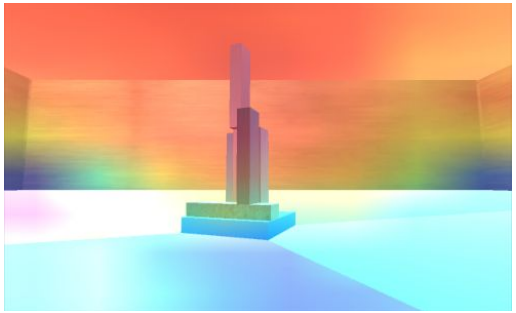
Same tower, different viewing angles



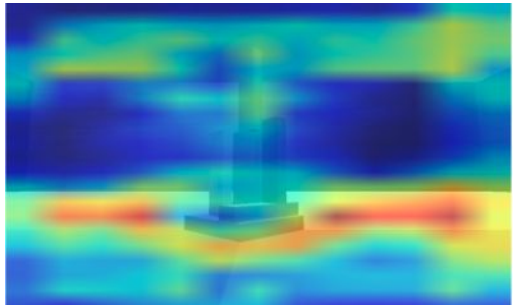
Human: Stable (38%)



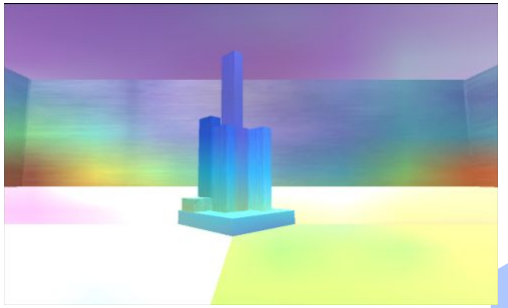
ResNet50: Stable



InceptionV3: Stable



ResNet50: Unstable



InceptionV3: Unstable

Summary

Human Behavior Results

Humans show robust intuitive grasp of physical principles with **high accuracy** (71%) and **significant consistency** (78%)

Image-Based Models

Deep learning models show the potential to reach human capabilities but more work is needed to understand **the gap between human and machines' intuitive physics**

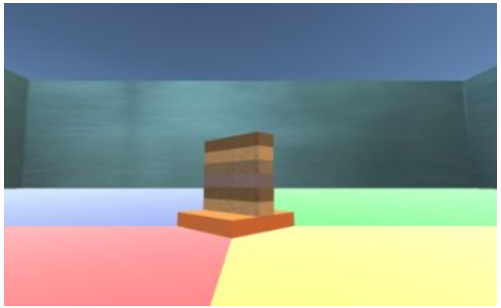
Thanks!

Dr. Bei Xiao
Jesse Schwartz (Lab alumni)
Michael Reinisch (MS student)
Chenxi Liao (PhD student)

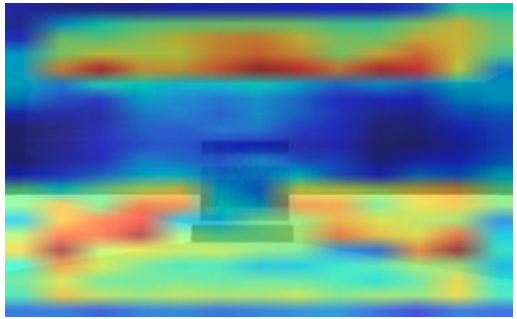
Yu-Yu Chen, MS candidate in Data Science
yc0688a@american.edu

Interesting Case

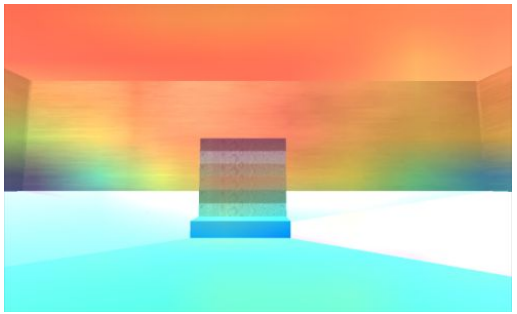
Same tower, different viewing angles



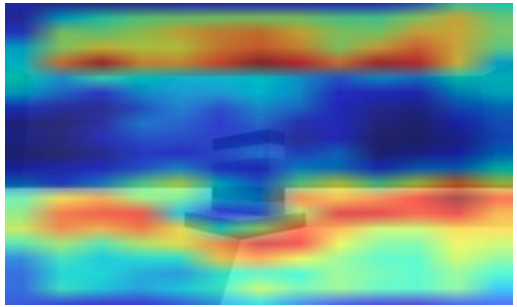
Human: Stable (96%)



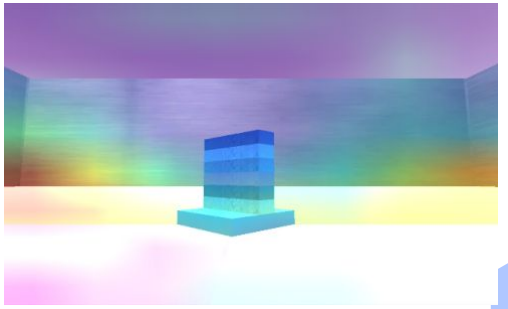
ResNet50: Stable



InceptionV3: Stable



ResNet50: Stable



InceptionV3: Unstable