Building IMU-based Gesture Recognition!!!

Jennifer Wang

jewang.net
Hello world!

Jennifer Wang
Software Engineer & Hardware Prototyper
I <3 sensors

- Phones
- Wearables
- Robots
- Telescopes

jewang.net / jen@jewang.net
Preamble

Thank you to these lovely people:

- Tim ‘mithro’ Ansell
- Kat Fang
- Cynthia Gan
- Samy Kamkar
- Sophi Kravitz
- Jinna Lei
- Jen Tong
- Tony Valderrama
- Ruth Grace Wong

Please go vote!
That door in front of you is magically locked. Cast Alohomora to open it.
Wingardium (Leviosa) Levitation

Flippendo Bread and butter
Final Product

https://github.com/jewang/gesture-demo
How to Gesture Recognition

1. Define & scope the problem
2. Propose machine learning model / algorithm
3. Collect data
4. Train model & iterate
5. Productionize the model
IMUs + Machine Learning = Lots of fun!

- Smartwatch gestures/input/stealing ATM Pins
- Music Control Gloves (Mimu Gloves used by Imogen Heap)
- Real firebending (Allen Pan)
- Gait-based auth
- Android Camera Control Gestures
- Fitness / Sleep Tracking "actigraphy" (FitBit)

Jewang.net // Hackaday Supercon '18
What will the final product look like?

1. Sample sensor data
2. Generate Features
3. Machine Learning 🧠
4. Post-process result

Output!!!
What will the final product look like?

Sample sensor data → Generate Features → Machine Learning → Post-process result → Output!!!

What is a feature?
Summary. Captures what is interesting.
Sensor Data
Sensor Data
Sensor Data
Sensor Data
Sensor Data
Sensor Data
Sensor Data
Sensor Data
Sensor Data

![Sensor Data Chart](image_url)
Sensor Data
How to Gesture Recognition

1. Define & scope the problem
2. Propose machine learning model / algorithm
3. Collect data
4. Train model & iterate
5. Productionize the model
How to Gesture Recognition

1. Define & scope the problem
2. Propose machine learning model / algorithm
3. Collect data
4. Train model & iterate
5. Productionize the model
## Define & scope the Problem

<table>
<thead>
<tr>
<th>Question</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>What are your resources? Time? $?</td>
<td></td>
</tr>
<tr>
<td>Who/when/where will it be used?</td>
<td></td>
</tr>
<tr>
<td>Power, Latency, Memory Req’s?</td>
<td></td>
</tr>
<tr>
<td>When is your model good enough?</td>
<td></td>
</tr>
</tbody>
</table>

---

Jewang.net // Hackaday Supercon '18
Define & scope the Problem

<table>
<thead>
<tr>
<th>Question</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>What are your resources? Time? $?</td>
<td>Me. 1 Month! &lt;$200?</td>
</tr>
<tr>
<td>Who/when/where will it be used?</td>
<td></td>
</tr>
<tr>
<td>Power, Latency, Memory Req’s?</td>
<td></td>
</tr>
<tr>
<td>When is your model good enough?</td>
<td></td>
</tr>
</tbody>
</table>
## Define & scope the Problem

<table>
<thead>
<tr>
<th>Question</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>What are your resources? Time? $?</td>
<td>Me. 1 Month! &lt;$200?</td>
</tr>
<tr>
<td>Who/when/where will it be used?</td>
<td>Me. Halloween evening! Outdoors, in my neighborhood.</td>
</tr>
<tr>
<td>Power, Latency, Memory Req’s?</td>
<td></td>
</tr>
<tr>
<td>When is your model good enough?</td>
<td></td>
</tr>
</tbody>
</table>
Machine Learning on an Embedded System: Pick 2

- Memory-efficient (Model Size)
- Real-time Fast Detection (Latency)

Accuracy
Machine Learning on an Embedded System: Pick 2

- Memory-efficient (Model Size)
- Real-time Fast Detection (Latency)
- Accuracy

Ship as fast as possible
<table>
<thead>
<tr>
<th>Define &amp; scope the Problem</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>What are your resources? Time? $$?</strong></td>
<td>Me. 1 Month! &lt;$200?</td>
</tr>
<tr>
<td><strong>Who/when/where will it be used?</strong></td>
<td>Me. Halloween evening! Outdoors, in my neighborhood.</td>
</tr>
<tr>
<td><strong>Power, Latency, Memory Req’s?</strong></td>
<td></td>
</tr>
<tr>
<td><strong>When is your model good enough?</strong></td>
<td></td>
</tr>
<tr>
<td>Define &amp; scope the Problem</td>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td><strong>What are your resources? Time? $?</strong></td>
<td><strong>Me. 1 Month! &lt;$200?</strong></td>
</tr>
<tr>
<td><strong>Who/when/where will it be used?</strong></td>
<td><strong>Me. Halloween evening! Outdoors, in my neighborhood.</strong></td>
</tr>
<tr>
<td><strong>Power, Latency, Memory Req’s?</strong></td>
<td><strong>Don’t worry about it for now. Ship as fast as possible.</strong></td>
</tr>
<tr>
<td><strong>When is your model good enough?</strong></td>
<td>****</td>
</tr>
</tbody>
</table>

Jewang.net // Hackaday Supercon ’18
<table>
<thead>
<tr>
<th>Define &amp; scope the Problem</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>What are your resources? Time? $?</strong></td>
<td><strong>Me. 1 Month! &lt;$200?</strong></td>
</tr>
<tr>
<td><strong>Who/when/where will it be used?</strong></td>
<td><strong>Me. Halloween evening! Outdoors, in my neighborhood.</strong></td>
</tr>
<tr>
<td><strong>Power, Latency, Memory Req’s?</strong></td>
<td><strong>Don’t worry about it for now. Ship as fast as possible.</strong></td>
</tr>
<tr>
<td><strong>When is your model good enough?</strong></td>
<td><strong>When Halloween happens When kids are happy :)</strong></td>
</tr>
<tr>
<td>What sensors?</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td></td>
</tr>
<tr>
<td>What is the basic architecture?</td>
<td></td>
</tr>
</tbody>
</table>

Define & scope the hardware
Common sensors for gesture recognition

- Camera?
Common sensors for gesture recognition

- **Camera?** → No, it’s dark outside
Common sensors for gesture recognition

- **Camera?** → No, it’s dark outside
- **IR/Wireless Beacons?**
Common sensors for gesture recognition

- **Camera?** → No, it’s dark outside
- **IR/Wireless Beacons?** → No, I can’t mount these all over my neighborhood
Common sensors for gesture recognition

- **Camera?** → No, it’s dark outside
- **IR/Wireless Beacons?** → No, I can’t mount these all over my neighborhood
- **Sonar / Audio?**
Common sensors for gesture recognition

- **Camera?** → No, it’s dark outside
- **IR/Wireless Beacons?** → No, I can’t mount these all over my neighborhood
- **Sonar / Audio?** → No, the environment will be noisy because I’ll be outside! Walking around!
Common sensors for gesture recognition

- **Camera?** → No, it’s dark outside
- **IR/Wireless Beacons?** → No, I can’t mount these all over my neighborhood
- **Sonar / Audio?** → No, the environment will be noisy because I’ll be outside! Walking around!
- **Magic E/M Sensing?**
Common sensors for gesture recognition

- **Camera?** → No, it’s dark outside
- **IR/Wireless Beacons?** → No, I can’t mount these all over my neighborhood
- **Sonar / Audio?** → No, the environment will be noisy because I’ll be outside! Walking around!
- **Magic E/M Sensing?** → No, I am not an academic :)

Jewang.net // Hackaday Supercon ’18
Common sensors for gesture recognition

- **Camera?** → No, it’s dark outside
- **IR/Wireless Beacons?** → No, I can’t mount these all over my neighborhood
- **Sonar / Audio?** → No, the environment will be noisy because I’ll be outside! Walking around!
- **Magic E/M Sensing?** → No, I am not an academic :)
- **Inertial Measurement Unit (IMU)?**
Common sensors for gesture recognition

- **Camera?** → No, it’s dark outside
- **IR/Wireless Beacons?** → No, I can’t mount these all over my neighborhood
- **Sonar / Audio?** → No, the environment will be noisy because I’ll be outside! Walking around!
- **Magic E/M Sensing?** → No, I am not an academic :)
- **Inertial Measurement Unit (IMU)?** → Yes!
What is an inertial measurement unit (IMU)?

✨ Orientation! ✨

Accelerometer + Gyroscope + Magnetometer

3 Axis
<table>
<thead>
<tr>
<th>What sensors?</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>What is the basic architecture?</td>
<td></td>
</tr>
</tbody>
</table>
### Define & scope the hardware

<table>
<thead>
<tr>
<th>What sensors?</th>
<th>IMU (Bosch BNO055)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>What is the basic architecture?</th>
</tr>
</thead>
</table>
Parts!

- BNO055 IMU
- UART
- Raspi Zero W
- USB Power
- Phone battery
- USB Audio
- Speaker

Glue this to a cosplay wand.

Jewang.net // Hackaday Supercon '18
Speaker goes up your sleeve ;}
## Define & scope the hardware

<table>
<thead>
<tr>
<th>What sensors?</th>
<th>IMU (Bosch BNO055)</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is the basic architecture?</td>
<td></td>
</tr>
</tbody>
</table>
### Define & scope the hardware

<table>
<thead>
<tr>
<th>What sensors?</th>
<th>IMU (Bosch BNO055)</th>
</tr>
</thead>
</table>

| What is the basic architecture? | Raspi Zero W glued to a wand. Embedded Linux. |
How to Gesture Recognition

1. Define & scope the problem
2. Propose machine learning model / algorithm
3. Collect data
4. Train model & iterate
5. Productionize the model
How to Gesture Recognition

1. Define & scope the problem
2. Propose machine learning model / algorithm
3. Collect data
4. Train model & iterate
5. Productionize the model
## What algorithm should I use?

<table>
<thead>
<tr>
<th></th>
<th>Deep learning!</th>
<th>Signal processing</th>
<th>Traditional Machine Learning</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Domain knowledge</strong></td>
<td>Not much</td>
<td>Lots!</td>
<td>Medium</td>
</tr>
<tr>
<td><strong>Amount of data</strong></td>
<td>Lots!</td>
<td>Not as much</td>
<td>Medium</td>
</tr>
<tr>
<td><strong>Easy to debug</strong></td>
<td>Harder</td>
<td>Easier</td>
<td>Medium</td>
</tr>
<tr>
<td></td>
<td>Deep learning!</td>
<td>Signal processing</td>
<td>Traditional Machine Learning</td>
</tr>
<tr>
<td>------------------------------</td>
<td>----------------</td>
<td>------------------</td>
<td>----------------------------</td>
</tr>
<tr>
<td>Domain knowledge</td>
<td>Not much</td>
<td>Lots!</td>
<td>Medium</td>
</tr>
<tr>
<td>Amount of data</td>
<td>Lots!</td>
<td>Not as much</td>
<td>Medium</td>
</tr>
<tr>
<td>Easy to debug</td>
<td>Harder</td>
<td>Easier</td>
<td>Medium</td>
</tr>
</tbody>
</table>
### What algorithm should I use?

<table>
<thead>
<tr>
<th>Domain knowledge</th>
<th>Amount of data</th>
<th>Easy to debug</th>
<th>Deep learning</th>
<th>Signal processing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not much</td>
<td>Lots!</td>
<td>Harder</td>
<td>Traditional Machine Learning</td>
<td>Medium</td>
</tr>
<tr>
<td>Lots!</td>
<td>Not as much</td>
<td>Easier</td>
<td>Medium</td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
<td></td>
</tr>
</tbody>
</table>

1. Ship your MVP
2. Get more users
3. More users = more data
4. Switch to deep learning

[Source](Jewang.net // Hackaday Supercon '18)
Software Used

- **Language:** Python3
- **Numerical Libraries:** Pandas, Numpy
- **Data Notebook:** Jupyter Notebook
- **Data Visualization:** Plot.ly
- **Machine learning library:** scikit-learn
scikit-learn algorithm cheat-sheet
How to Gesture Recognition

1. Define & scope the problem
2. Propose machine learning model / algorithm
3. Collect data
4. Train model & iterate
5. Productionize the model
How to Gesture Recognition

1. Define & scope the problem
2. Propose machine learning model / algorithm
3. Collect data
4. Train model & iterate
5. Productionize the model
Planning data collection

- How much data do you need?
- Look for pre-existing data
- More data = More $ / 🕒
  Design collection procedure, manage data collectors, QA...
Doing data collection

Time to code!

Lock mechanical design!

Match & record data for all expected use cases!

Who × Posture × When × Where
Record data for all expected use cases!

<table>
<thead>
<tr>
<th></th>
<th>Wingardium</th>
<th>Flippendo</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sitting Couch</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Sitting Chair</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Sitting Floor</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Standing</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Walking</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
Record data for all expected use cases!

<table>
<thead>
<tr>
<th></th>
<th>Wingardium</th>
<th>Flippendo</th>
<th>Negative*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sitting Couch</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Sitting Chair</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Sitting Floor</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Standing</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Walking</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
Data collection is an exercise in diligence
Final data collected for magic wand

`collect_data.py` → CSV

286 1.5s traces
98 Wingardium
99 Flippendo
89 Negative

~7 minutes of data
257,000+ data points
How to Gesture Recognition

1. Define & scope the problem
2. Propose machine learning model / algorithm
3. Collect data
4. Train model & iterate
5. Productionize the model
How to Gesture Recognition

1. Define & scope the problem
2. Propose machine learning model / algorithm
3. Collect data
4. Train model & iterate
5. Productionize the model
Train model & iterate

- Explorations.ipynb
- Mostly data cleaning :P
Features

- max_accel
- min_accel
- Range_accel
- mean_accel
- std_accel

- max_gyro
- min_gyro
- range_gyro
- mean_gyro
- std_gyro
Train model & iterate

- Iteration #1 → Didn’t work well. 102 traces.
Train model & iterate

- Iteration #1 → Didn’t work well. 102 traces.
  - Clean Data! Collect new data!
Train model & iterate

- **Iteration #1** → Didn’t work well. 102 traces.
  - Clean Data! Collect new data!
- **Iteration #2** → Didn’t work well. 167 traces.
Train model & iterate

- **Iteration #1** → Didn’t work well. 102 traces.
  - Clean Data! Collect new data!
- **Iteration #2** → Didn’t work well. 167 traces.
  - Clean Data! Collect new data!
Train model & iterate

• Iteration #1 → Didn’t work well. 102 traces.
  ○ Clean Data! Collect new data!
• Iteration #2 → Didn’t work well. 167 traces.
  ○ Clean Data! Collect new data!
• Iteration #3: → Worked OK, distinguish gestures. 286 traces.
Train model & iterate

- **Iteration #1** → Didn’t work well. 102 traces.
  - Clean Data! Collect new data!
- **Iteration #2** → Didn’t work well. 167 traces.
  - Clean Data! Collect new data!
- **Iteration #3**: → Worked OK, distinguish gestures. 286 traces.
  - Feature Design
Wingardium has 2 accel z-axis peaks while flippendo has 3!
Wingardium has 2 accel z-axis peaks while flippendo has 3!

Hack a 75% Accurate Peak Counter

[Graphs showing accel and gyro data for Wingardium and Flippendo]
Features

- max_accel
- min_accel
- Range_accel
- mean_accel
- std_accel

- max_gyro
- min_gyro
- range_gyro
- mean_gyro
- std_gyro
Features

- max_accel
- min_accel
- Range_accel
- mean_accel
- std_accel
- accel_z_peaks
- max_gyro
- min_gyro
- range_gyro
- mean_gyro
- std_gyro
Train model & iterate

- **Iteration #1** → Didn’t work well. 102 traces.
  - Clean Data! Collect new data!
- **Iteration #2** → Didn’t work well. 167 traces.
  - Clean Data! Collect new data!
- **Iteration #3:** → Worked OK, distinguish gestures. 286 traces.
  - Feature Design
Train model & iterate

- **Iteration #1** ➔ Didn’t work well. 102 traces.
  - Clean Data! Collect new data!
- **Iteration #2** ➔ Didn’t work well. 167 traces.
  - Clean Data! Collect new data!
- **Iteration #3:** ➔ Worked OK, distinguish gestures. 286 traces.
  - Feature Design
- **Iteration #4** ➔ Good enough. It’s Halloween! Ship it!
How to Gesture Recognition

1. Define & scope the problem
2. Propose machine learning model / algorithm
3. Collect data
4. Train model & iterate
5. Productionize the model
How to Gesture Recognition

1. Define & scope the problem
2. Propose machine learning model / algorithm
3. Collect data
4. Train model & iterate
5. Productionize the model
Productionize the model

- `gesture_detector.py`
- <100 lines of code
- Might need a little / lot of tuning ;)

Jewang.net // Hackaday Supercon '18
How to Gesture Recognition

1. Define & scope the problem
2. Propose machine learning model / algorithm
3. Collect data
4. Train model & iterate
5. Productionize the model
How to Gesture Recognition

1. Define & scope the problem
2. Propose machine learning model / algorithm
3. Collect data
4. Train model & iterate
5. Productionize the model

Build more cool stuff!
IMUs + Machine Learning on a large scale...

I <3 sensors

- City walkability
- Better health phenotypes
- Depression treatments
- Census data
- Earthquake detection
IMUs + Machine Learning = Lots of fun!

- Real firebending (Allen Pan)
- Gait-based auth
- Smartwatch gestures/input/stealing ATM Pins
- Music Control Gloves (Mimu Gloves used by Imogen Heap)
- Android Camera Control Gestures
- Fitness / Sleep Tracking “actigraphy” (FitBit)

Jewang.net // Hackaday Supercon ‘18
How to Gesture Recognition

1. Define & scope the problem
2. Propose machine learning model / algorithm
3. Collect data
4. Train model & iterate
5. Productionize the model

Build more cool stuff!

jewang.net / jen@jewang.net
Appendix
How does a Linear SVC model work?