The author(s) shown below used Federal funding provided by the U.S. Department of Justice to prepare the following resource:

Document Title:	**Automatic Acquisition and Identification of Footwear Class Characteristics**
Author(s):	**Susan Vanderplas**
Document Number:	**309346**
Date Received:	**July 2024**
Award Number:	**2019-MU-MU-4096**

National Institute of Justice

Research and Development in Forensic Science for Criminal Justice Purposes

Award: 2019-MU-MU-4096

Automatic Acquisition and Identification of Footwear Class Characteristics

Award Period: January 1, 2020 - June 1, 2024

[redacted]

Submitted: July 12, 2024

EIN: 42-6004224

PI:
Susan Vanderplas
Email: svanderplas2@unl.edu
Phone: 515-509-6613

[redacted]

Table of Contents

1 Executive Summary

1.1 Project Goals and Accomplishments

The overarching goal of this project was to enable footwear examiners to make empirical statements about the frequency of class characteristics in the local geographic population. To support this overarching goal, we created the following objectives for this project, which would push research in this area forward and eventually enable quantitative evaluation of class characteristic frequency and similarity.

1. Develop robust, weather-resistant scanning equipment which can passively gather images of shoe soles and uppers from local populations.
2. Develop automated software to automatically identify relevant class characteristic information from images collected by the scanning equipment.
3. Collect local population footwear data over multiple seasons, weather conditions, days of the week, and times of the day. Assess the changes in identified class characteristic frequency associated with temporal and weather-related variability.

Of these project objectives, the first has been completed successfully. The second objective has been partially completed, but the results suggest that the hypothesis implicit in the objective was flawed. The third objective was affected by both the COVID-19 pandemic and IRB restrictions, but should still be feasible as written should sufficient data be collected to support it.

While progress on the overall goal has been slower than anticipated, the work completed as part of this grant provides a meaningful step forward - it is now possible to collect footwear information from the local population using the MANTIS scanner. Methodology for automatically analyzing this data has been developed as well, though there is substantial room for improvement in classification accuracy as well as in region proposals.

2 Research Objectives

2.1 Development of a Shoe Scanner to Collect Class Characteristic Information

Over the course of this project, the MANTIS scanner was developed, tested, and deployed in a range of weather conditions. MANTIS successfully captures images of shoe soles and uppers when the scanner is traversed.

We overcame a number of obstacles during the development of the scanner: - optical configurations to ensure data captured more than 6” from the surface of the scanner is out of focus, which is necessary to preserve privacy. - glare reduction - reduction of calibration time after scanner deployment - reduction of equipment thermal profile, reducing equipment wear - self-cleaning functionality to reduce maintenance required for outdoor data collection.

In addition, while working to reduce calibration time during scanner deployment, we developed a lighter-weight scanner which can be deployed indoors but which is still compatible with the weatherproof shell developed as part of the initial scanner iteration. The weatherproof shell is 400 lbs and requires a truck hoist and two individuals to move; in most data collection scenarios, the ability to withstand -20° F temperatures, hail, wind, torrential rain, and snow is not necessary. However, the heavy shell has an additional advantage with respect to security - theft or even damaging the equipment is difficult because of the robustness necessary to withstand in-ground freeze-thaw cycles while minimizing water intrusion.

2.2 Software for Automatic Identification of Class Characteristics

We have developed several computer vision models which identify relevant class characteristic information. Each model has flaws which limit the effectiveness of the model, but we are confident that this objective is tractable - our approaches, while unsuccessful, suggest viable paths forward in the future. The hypothesis underlying our initial approach to and formulation of this objective was that if neural networks can differentiate between African and Asian elephants, that identifying geometric features would be an even easier task. The work performed in support of this objective suggests that this underlying hypothesis is false, with implications for different approaches to this task that may prove more successful.

2.2.1 Convolutional Neural Network for Object Recognition (CoNNOR)

An initial model (model code: https://github.com/srvanderplas/CoNNOR-paper, paper: https://github.com/srvanderplas/CoNNORFSI) used transfer learning with the pre-built VGG16 model to classify square sections of shoe soles, applying one or more labels to each 256 x 256 pixel image created by slicing up an image of a shoe sole. This model had an accuracy of around 80-90% across the 9 shape classes used: circle, quadrilateral, triangle, bowtie, lines, chevrons, polygon, text, and a catch-all “other” class, as shown in Figure 1.

One notable factor which became apparent when training CoNNOR is that the 9 classes used were not entirely distinct: text contains many different circles which confuse the model, resulting in confusion between the text and circle classes, as shown in Figure 2. Figure 3 shows several examples of predictions made where circles and text are confused, as well as examples where they are correctly identified.

2.2.2 Object Detection Models

As we implemented CoNNOR, there were many changes in the software available for computer vision tasks, and many new developments in network types and capabilities. In an attempt to both improve model accuracy and reduce preprocessing, we transitioned from a classification model (which applies labels to the entire image) to an object detection approach, which takes a whole image, identifies regions of interest within the image, and then applies a label to each region.

Figure 7 shows the difference between classification, object detection, and image segmenta-

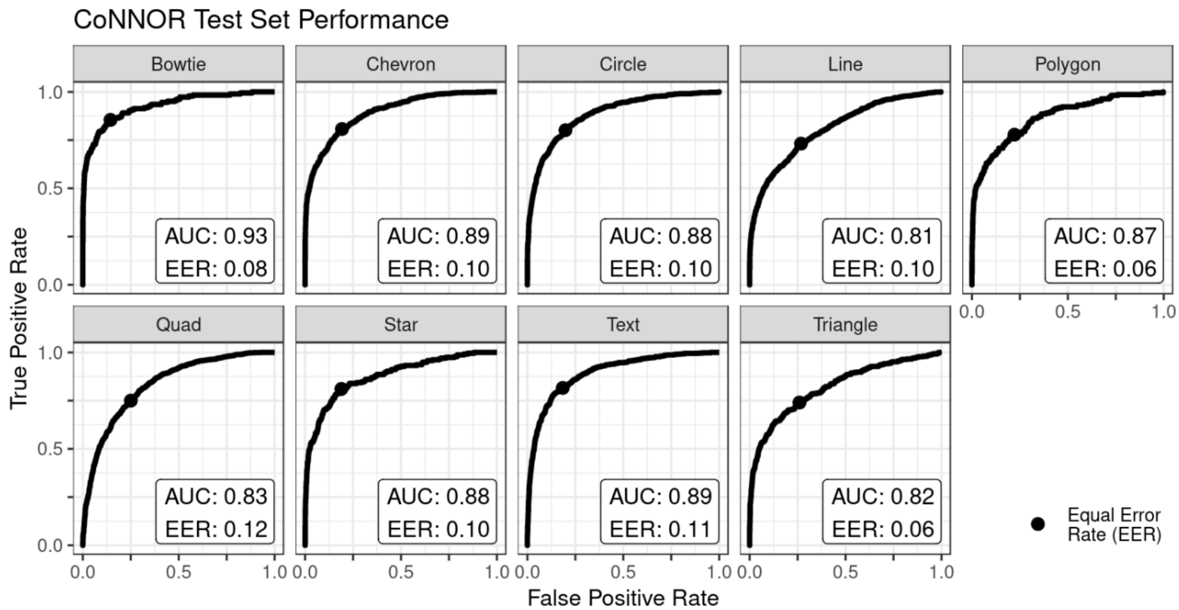


Figure 1: Class-by-class ROC curves. Area under the curve (AUC) is a measure of overall model performance. Equal error rates (EER) are marked, indicating the position at which there is an equal probability of false positive and false negative error rates.

tion problems. All of these approaches can credibly be used to identify shoe class characteristics, but each requires a different type of labeled data. As this project progressed, we created labeled data in both classification and object detection formats, but re-creating the data for each type of problem is labor-intensive and extremely slow.

This approach ended up requiring an entirely different data format, and, during the data collection process, the toolkit we were using to build the model was deprecated, requiring a rewrite of all previously developed model code. The resulting model did not accurately identify regions with features, either in groups or single; large numbers of features were missed. In subsequent investigations, we discovered that the region proposal algorithm included assumptions about the density of features - specifically, that areas are present that do not have relevant features (“negative” samples). Most athletic shoe soles (which made up a large proportion of our available data) do not contain large regions devoid of geometric shapes, which makes off-the-shelf object detection models unsuitable for this purpose. Unfortunately, this discovery came at the end of the grant period. Initial investigation into model modifications suggests that one possible way to overcome this issue is to use dense networks which can identify exemplar objects (these networks are often used for counting e.g. the number of fish in a school) to develop proposal regions for each feature. Unfortunately, it is likely that separate networks would need to be trained to identify each class of feature. This is a promising area of future investigation which could be used to solve this underlying problem.

Figure 2: Confusion matrix, showing the correct classification rate on the diagonal and classification errors on the off-diagonal. In multi-label images, correct off-diagonal labels have been excluded from the calculation of false positives. Quadrilaterals are generally over-predicted across true labels, while polygons and triangles are under-predicted. Circles and text are often confused.

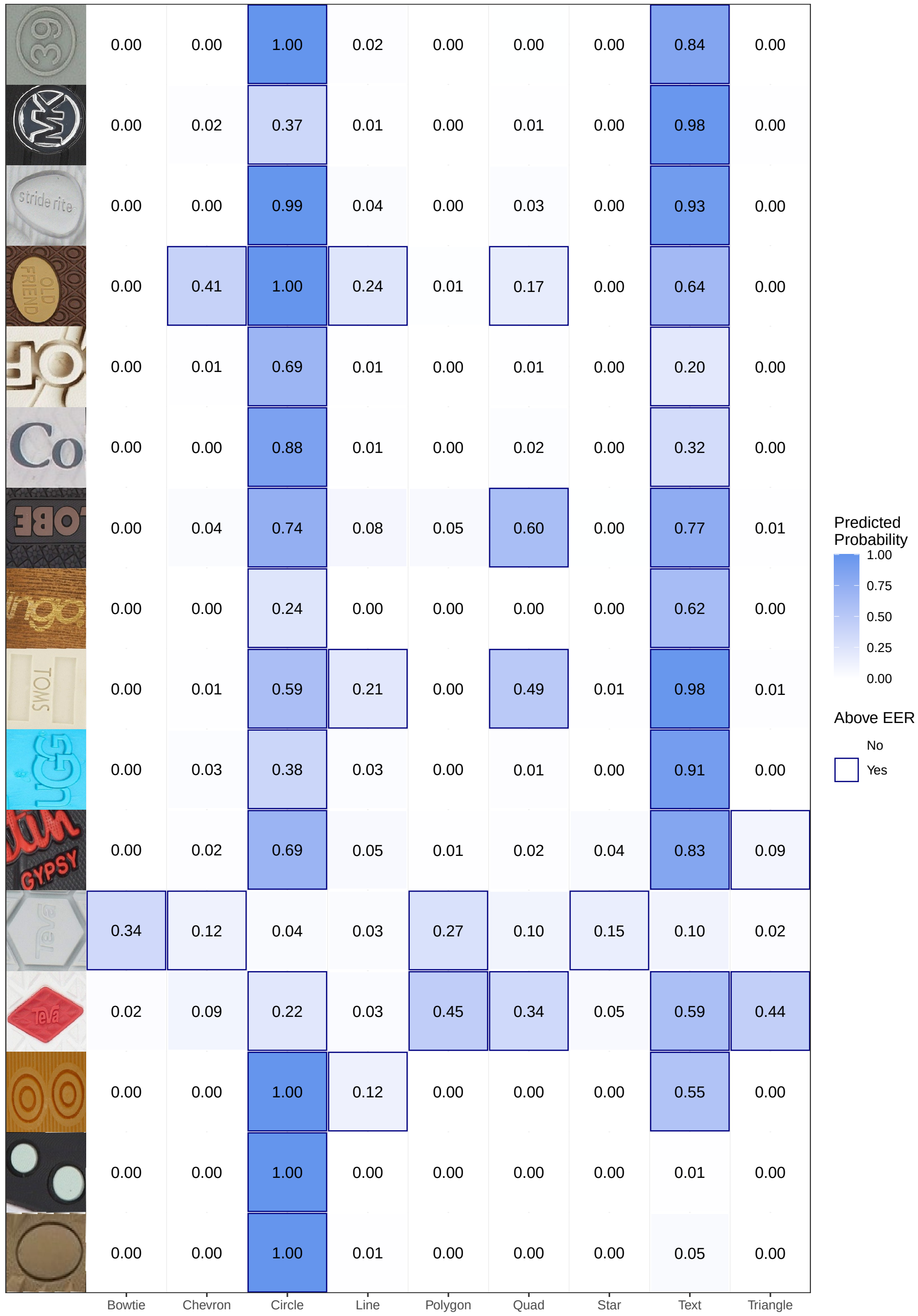


Figure 3: Model predictions of images containing circles and/or text. The model correctly identifies circles in all images where they are present, but also predicts moderate probabilities of circles in many images where text is present, particularly when there are curved letters, such as ‘G’. Text is correctly identified in most images, with the exception of the false negative prediction in the Teva hexagon image, which has low contrast, and in the concentric circles near the bottom, where the text prediction is a false positive.

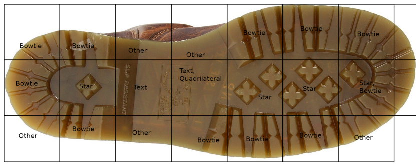


Figure 4: Classification: same-size regions labeled with one or more classes

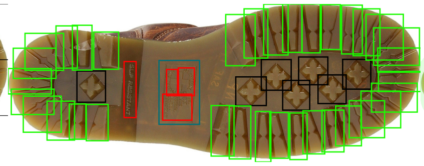


Figure 5: Object Detection: Propose a bounding box and label for each object in an image

Figure 6: Image segmentation: find regions of the image and label each region

Figure 7: Different types of computer-vision problems which require different types of labeled data in different formats.

2.2.3 Discussion: Computer Vision Models and Shoe Soles

One interesting finding that underlies the difficulties in implementing feature recognition software beyond the CoNNOR model is that shoe soles are fundamentally different from many other image recognition tasks in a way that makes transfer learning challenging. Models like AlexNET, VGG16, and RESNET are all trained on millions of photographs of natural scenes, which are much richer in context than the photographs of artificial shoe soles. In addition, shoe soles are the product of human design, composed primarily of geometric objects, with minimal depth information. Contrast this to a picture of a city street, where there are multiple layers of information present, objects of different sizes with corresponding perspective information, and where each object in the scene is much more complex. Even medical imaging data, which is not a "natural scene", contains considerably more variation and complexity than the composition of geometric objects which make up a typical shoe sole. The fundamental flaw in our hypothesis, then, was the assumption that because our data was simpler than natural scenes, models would be easier to train and use than the corresponding natural scene models. After all, circles are easy to identify, while project personnel cannot differentiate between African and Asian elephants, so models should be similarly skilled. This assumption was incorrect. The feature detectors necessary to correctly operate with natural features seem to be suboptimal for detecting geometric shapes that are not particularly natural in configuration.

Our initial, small scale experiments generating data to train a computer vision model from scratch (discarding the transfer learning approach) suggest that this approach has promise. Unfortunately, training computer vision models from scratch requires orders of magnitude more data than transfer learning, and this data should be generated programmatically in order to reduce the impact of mis-labeling that is inevitable when data is labeled by humans. In some cases, data isn't even mis-labeled so much as different people identify different features that are present; conversations with examiners at IAI in 2022 suggest that this is common even among practitioners. There are significant challenges in developing models that use human-recognizable features and labels, in part because this is a moving target: humans are not consistent when labeling data, which makes it hard to train a consistent model.

2.3 Class Characteristic Spatial and Temporal Trends

In the project proposal, we planned to place the scanner on the sidewalk (using ramps) or embed it into the sidewalk (e.g. zero-entry configuration) once the scanner design was fixed. Initial IRB screening suggested this was reasonable, however, when we went to get approval for the data collection phase of the project, the university risk assessment required that we not obstruct any sidewalk, and instead allowed us to place the scanner adjacent to the path. Unfortunately, this meant that participants had to walk out of their way to go across the scanner, and most potential participants declined to do so. When the scanner was placed in a new location, it was initially novel, and people would walk across it. As the novelty wore off, the decline in data collection rates was steep: we could not collect sufficient data to perform inference. This portion of the project was complicated by the COVID pandemic, supply chain issues which made it difficult to get replacement parts for optics and manufacturing of the scanner body, and ultimately, the resulting significant decrease in foot traffic in common areas around campus, which persisted through the 2022-2023 academic year due to an increased student preference for online classes.

It seems likely that another entity could still collect data using the scanner: Ames, IA does not have high population density and is not overly walkable, but data collection in an urban, walkable area might be considerably more successful, particularly if the scanner is incorporated into the environment more thoroughly than was feasible in our location. The scanner's ability to make use of a hard-wired power connection allows for relatively low maintenance collection of forensic intelligence data that would also facilitate the completion of analyses proposed as part of this project, but due to the combination of structural and pandemic-induced complications, this objective was essentially not feasible for us to complete during the project period.

3 Outcomes

3.1 Activities/Accomplishments

- Development and refinement of a scanning machine which can be used outdoors (day and night)
- Development of optics which protect the privacy of anyone walking over the machine.
- Labeling of "clean" (commercial) data with shoe size, brand, and texture information
- Develop a statistical model for class characteristic proposal and classification
- Develop a method to improve the model through the creation of synthetic data
- Developing a strategy for creating synthetic data with a sufficient variety of realistic characteristics

3.2 Training & Skill Acquisition

3.2.1 Graduate Students

Graduate Statistics student Jayden Stack received his MS while working on the project. Jayden developed skills in machine learning, computer vision, and python programming.

Graduate Statistics student Muxin Hua received her MS while working on the project and is a current Ph.D. student. She is expected to graduate with a Ph.D. in 2026.

Graduate IMSE student Colton Fales received his Ph.D. while working on the project.

Graduate IMSE student Brayden Westby is currently working on his Ph.D.

3.3 Dissemination of Results

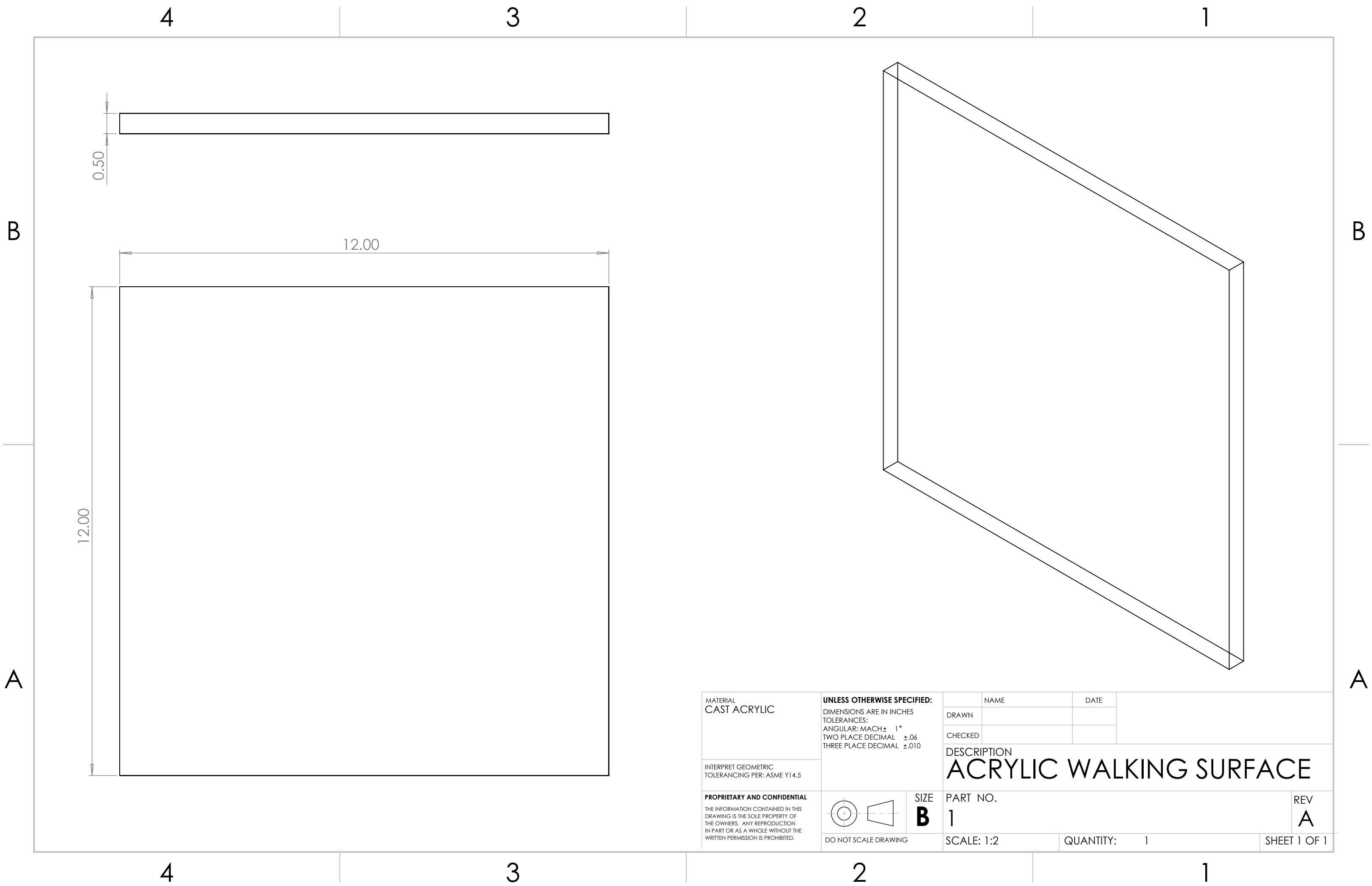
3.3.1 Outreach Events

The scanner was present at the CSAFE booth at AAFS in 2022 as well as at IAI in 2023.

The scanner has also been used at 5-10 outreach events at Iowa State.

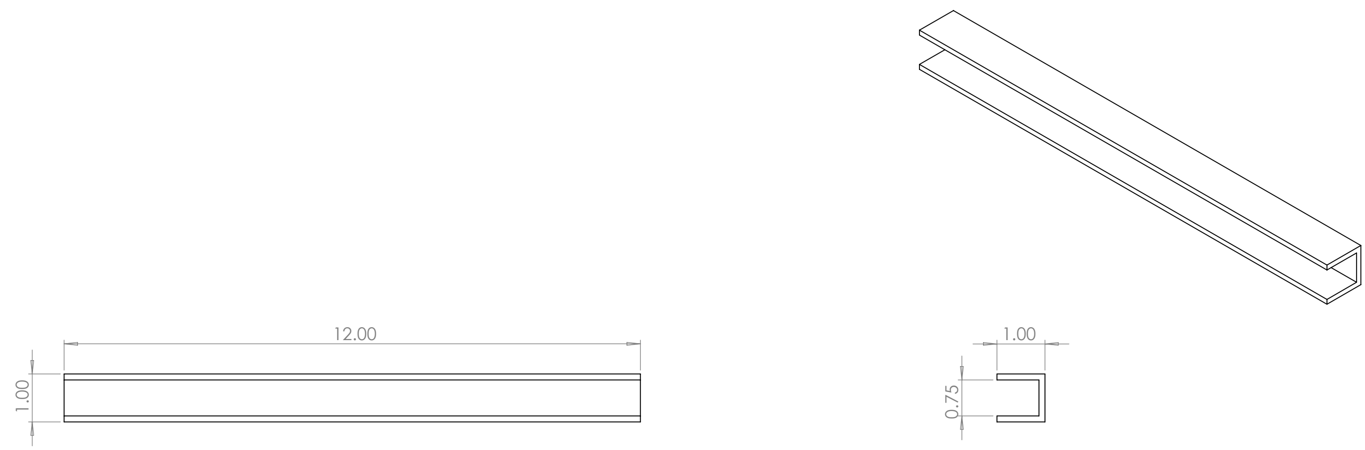
3.3.2 Products

The MANTIS (the footwear scanning device) is functional, and design schematics are included below. Additional information can be obtained by contacting Dr. Richard Stone (rstone@iastate.edu).



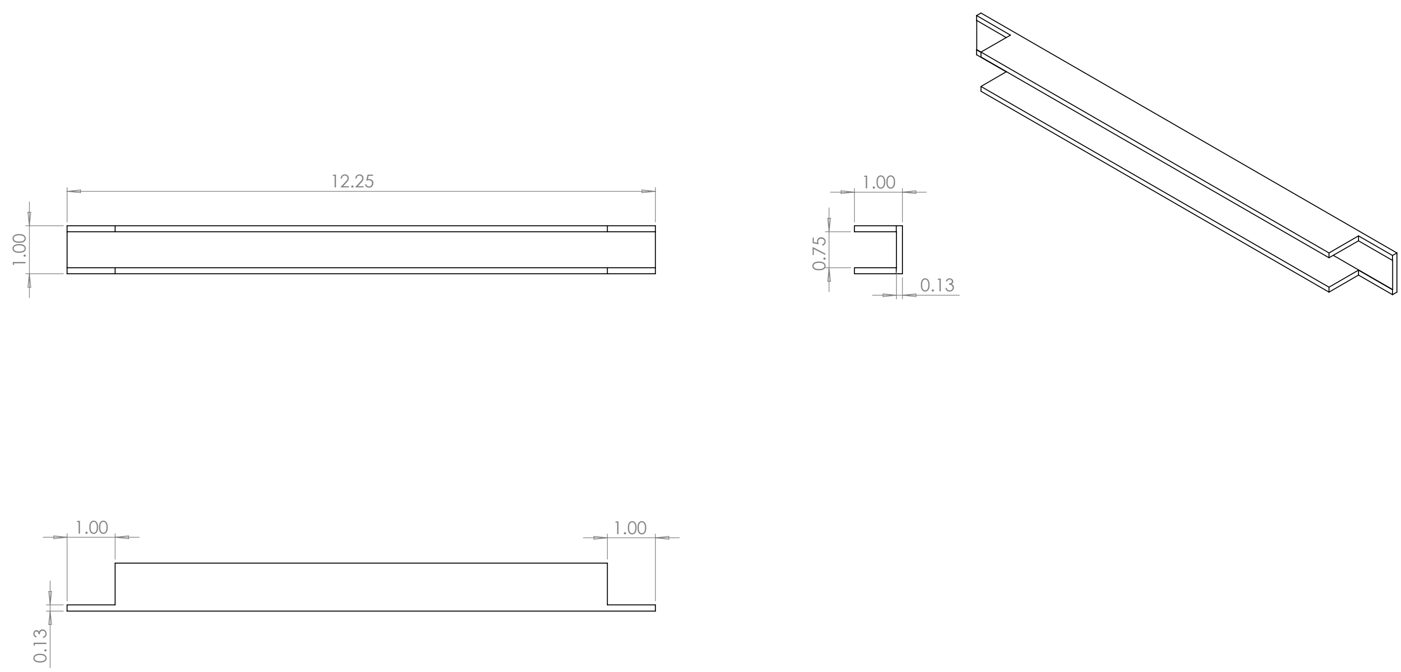
MATERIAL CAST ACRYLIC	UNLESS OTHERWISE SPECIFIED: DIMENSIONS ARE IN INCHES TOLERANCES: ANGULAR: MACH ± 1° TWO PLACE DECIMAL ±.06 THREE PLACE DECIMAL ±.010		NAME / DATE
		DRAWN	
		CHECKED	
INTERPRET GEOMETRIC TOLERANCING PER: ASME Y14.5		DESCRIPTION ACRYLIC WALKING SURFACE	
PROPRIETARY AND CONFIDENTIAL THE INFORMATION CONTAINED IN THIS DRAWING IS THE SOLE PROPERTY OF THE OWNERS. ANY REPRODUCTION IN PART OR AS A WHOLE WITHOUT THE WRITTEN PERMISSION IS PROHIBITED.	SIZE B	PART NO. 1	REV A
	DO NOT SCALE DRAWING	SCALE: 1:2 QUANTITY: 1	SHEET 1 OF 1

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

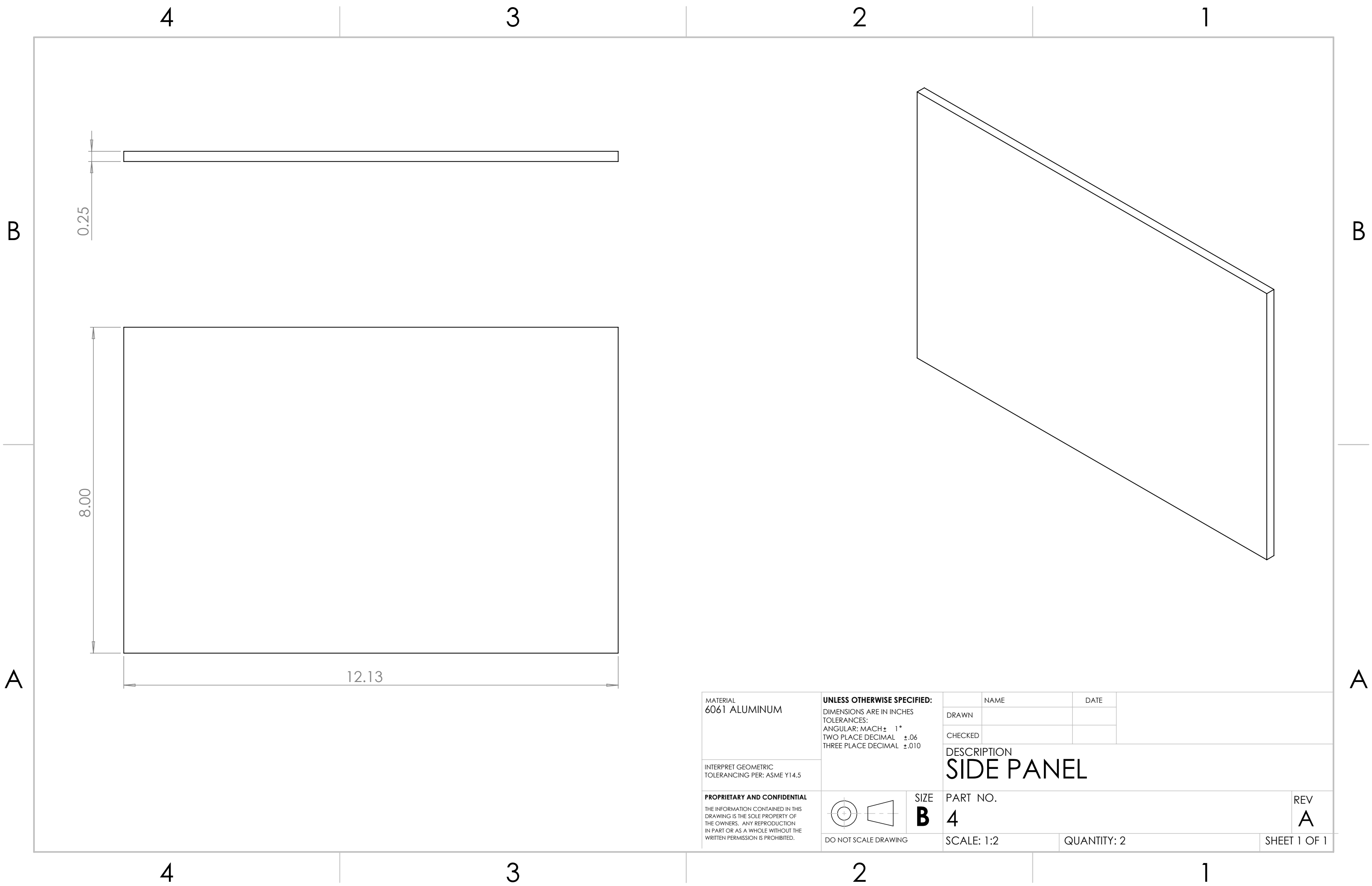


MATERIAL 6061 ALUMINUM	UNLESS OTHERWISE SPECIFIED: DIMENSIONS ARE IN INCHES TOLERANCES: ANGULAR: MACH± 1° TWO PLACE DECIMAL ±.06 THREE PLACE DECIMAL ±.010		NAME	DATE
		DRAWN		
		CHECKED		
INTERPRET GEOMETRIC TOLERANCING PER: ASME Y14.5		DESCRIPTION U-CHANNEL		
PROPRIETARY AND CONFIDENTIAL THE INFORMATION CONTAINED IN THIS DRAWING IS THE SOLE PROPERTY OF THE OWNERS. ANY REPRODUCTION IN PART OR AS A WHOLE WITHOUT THE WRITTEN PERMISSION IS PROHIBITED.	SIZE **B**	PART NO. 2		REV A
	DO NOT SCALE DRAWING	SCALE: 1:2	QUANTITY: 2	SHEET 1 OF 1

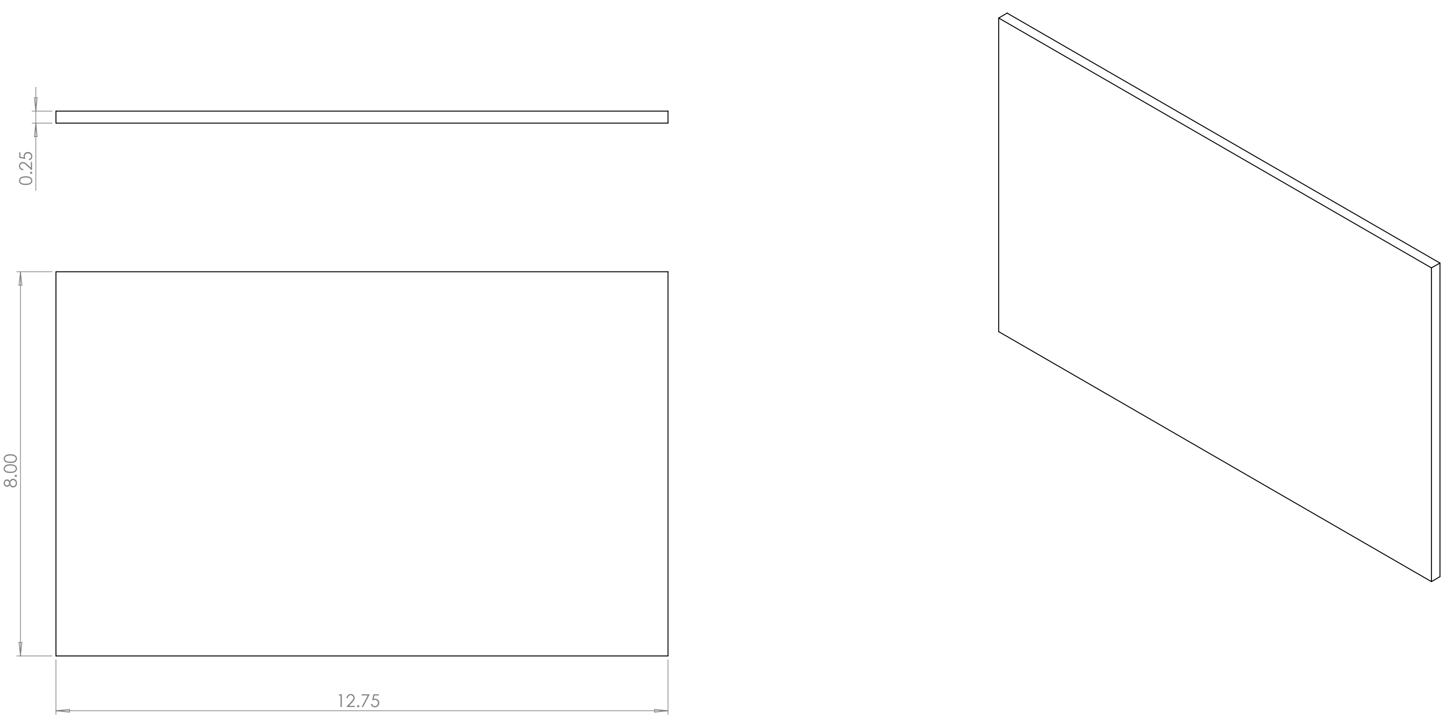
This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.



MATERIAL 6061 ALUMINUM	UNLESS OTHERWISE SPECIFIED: DIMENSIONS ARE IN INCHES TOLERANCES: ANGULAR: MACH± 1° TWO PLACE DECIMAL ±.06 THREE PLACE DECIMAL ±.010		NAME	DATE
		DRAWN		
		CHECKED		
INTERPRET GEOMETRIC TOLERANCING PER: ASME Y14.5		DESCRIPTION SLOTTED U-CHANNEL		
PROPRIETARY AND CONFIDENTIAL THE INFORMATION CONTAINED IN THIS DRAWING IS THE SOLE PROPERTY OF THE OWNERS. ANY REPRODUCTION IN PART OR AS A WHOLE WITHOUT THE WRITTEN PERMISSION IS PROHIBITED.	SIZE B	PART NO. 3		REV A
	DO NOT SCALE DRAWING	SCALE: 1:2	QUANTITY: 1	SHEET 1 OF 1

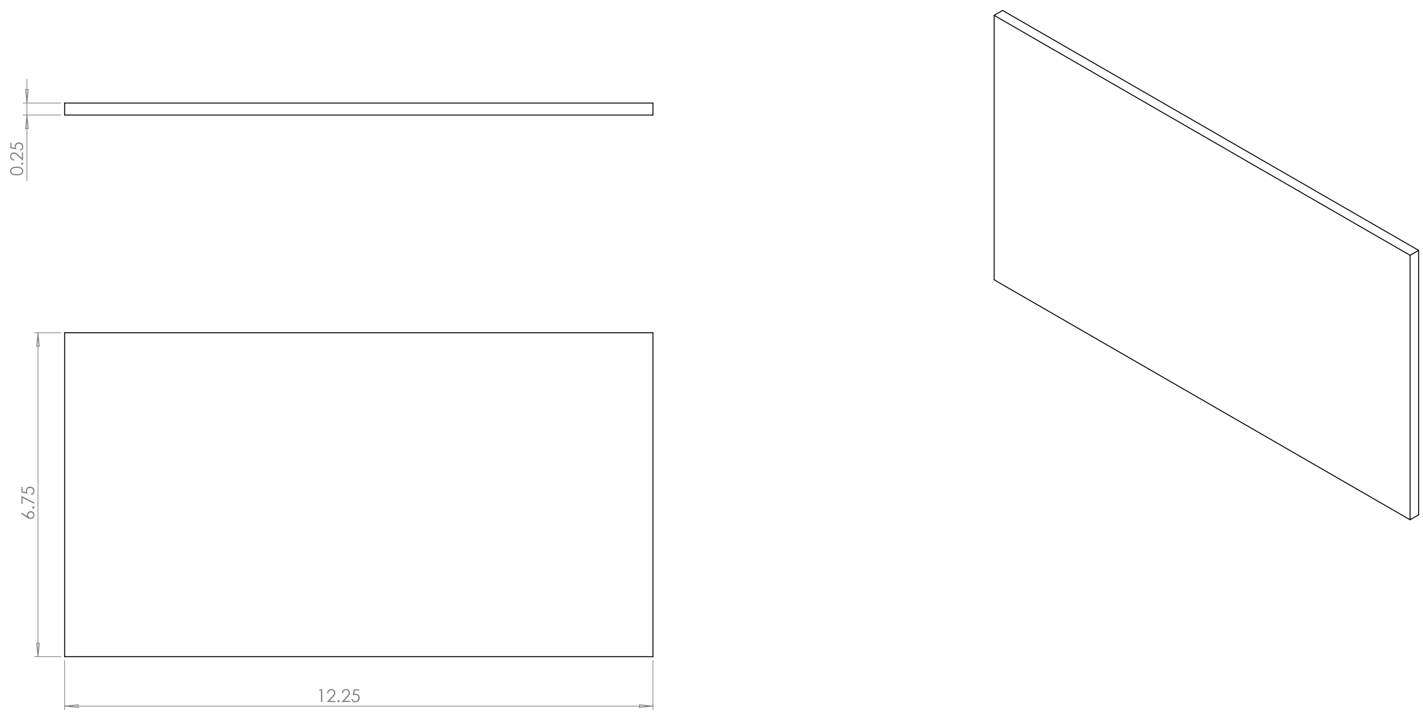


This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.



MATERIAL 6061 ALUMINUM	**UNLESS OTHERWISE SPECIFIED:** DIMENSIONS ARE IN INCHES TOLERANCES: ANGULAR: MACH ± 1° TWO PLACE DECIMAL ±.06 THREE PLACE DECIMAL ±.010		NAME / DATE DRAWN CHECKED
INTERPRET GEOMETRIC TOLERANCING PER: ASME Y14.5			DESCRIPTION BACK PANEL
PROPRIETARY AND CONFIDENTIAL THE INFORMATION CONTAINED IN THIS DRAWING IS THE SOLE PROPERTY OF THE OWNERS. ANY REPRODUCTION IN PART OR AS A WHOLE WITHOUT THE WRITTEN PERMISSION IS PROHIBITED.	DO NOT SCALE DRAWING	SIZE **B**	PART NO. 5 REV A
			SCALE: 1:2 QUANTITY: 1 SHEET 1 OF 1

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

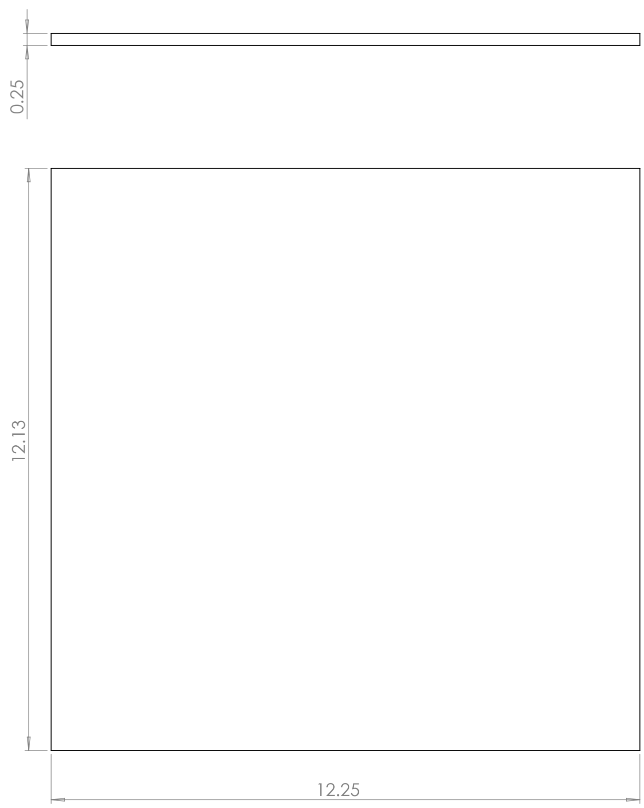


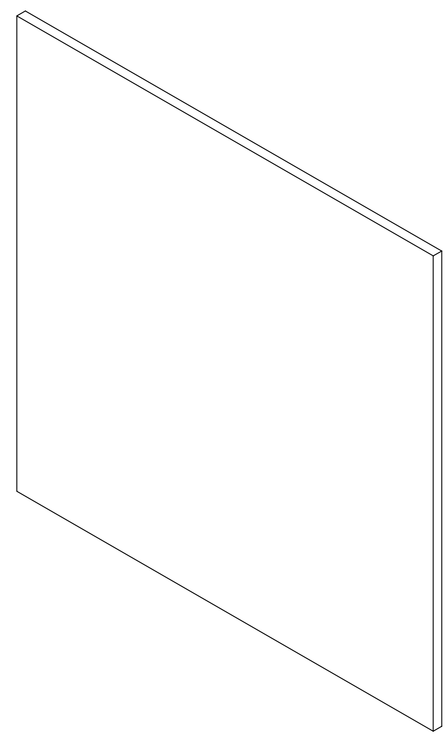
MATERIAL 6061 ALUMINUM	UNLESS OTHERWISE SPECIFIED: DIMENSIONS ARE IN INCHES TOLERANCES: ANGULAR: MACH ± 1° TWO PLACE DECIMAL ±.06 THREE PLACE DECIMAL ±.010		NAME	DATE
		DRAWN		
		CHECKED		
INTERPRET GEOMETRIC TOLERANCING PER: ASME Y14.5		DESCRIPTION FRONT PANEL		
PROPRIETARY AND CONFIDENTIAL THE INFORMATION CONTAINED IN THIS DRAWING IS THE SOLE PROPERTY OF THE OWNERS. ANY REPRODUCTION IN PART OR AS A WHOLE WITHOUT THE WRITTEN PERMISSION IS PROHIBITED.	SIZE B	PART NO. 6		REV A
	DO NOT SCALE DRAWING	SCALE: 1:2	QUANTITY: 1	SHEET 1 OF 1

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

4 | 3 | 2 | 1

B



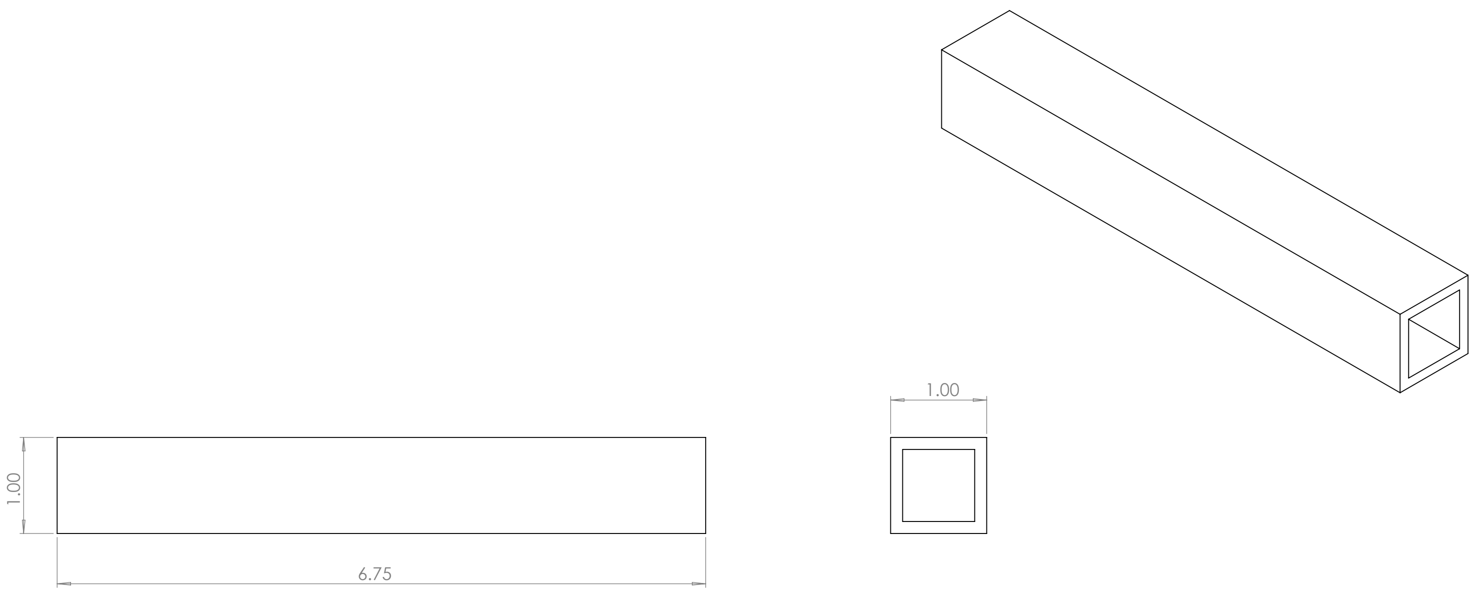


A

MATERIAL 6061 ALUMINUM	UNLESS OTHERWISE SPECIFIED: DIMENSIONS ARE IN INCHES TOLERANCES: ANGULAR: MACH ± 1° TWO PLACE DECIMAL ±.06 THREE PLACE DECIMAL ±.010		NAME	DATE
		DRAWN		
		CHECKED		
INTERPRET GEOMETRIC TOLERANCING PER: ASME Y14.5		DESCRIPTION BOTTOM PANEL		
PROPRIETARY AND CONFIDENTIAL THE INFORMATION CONTAINED IN THIS DRAWING IS THE SOLE PROPERTY OF THE OWNERS. ANY REPRODUCTION IN PART OR AS A WHOLE WITHOUT THE WRITTEN PERMISSION IS PROHIBITED.	SIZE **B**	PART NO. 7	REV A	
	DO NOT SCALE DRAWING	SCALE: 1:2	QUANTITY: 1	SHEET 1 OF 1

4 | 3 | 2 | 1

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.



MATERIAL 6061 ALUMINUM	**UNLESS OTHERWISE SPECIFIED:** DIMENSIONS ARE IN INCHES TOLERANCES: ANGULAR: MACH ± 1° TWO PLACE DECIMAL ±.06 THREE PLACE DECIMAL ±.010	NAME / DATE DRAWN CHECKED
INTERPRET GEOMETRIC TOLERANCING PER: ASME Y14.5		DESCRIPTION VERTICAL SUPPORT LEG
PROPRIETARY AND CONFIDENTIAL THE INFORMATION CONTAINED IN THIS DRAWING IS THE SOLE PROPERTY OF THE OWNERS. ANY REPRODUCTION IN PART OR AS A WHOLE WITHOUT THE WRITTEN PERMISSION IS PROHIBITED.	SIZE **B** DO NOT SCALE DRAWING	PART NO. 8 — REV A SCALE: 1:1 — QUANTITY: 4 — SHEET 1 OF 1

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

MATERIAL
6061 ALUMINUM

INTERPRET GEOMETRIC TOLERANCING PER: ASME Y14.5

PROPRIETARY AND CONFIDENTIAL

THE INFORMATION CONTAINED IN THIS DRAWING IS THE SOLE PROPERTY OF THE OWNERS. ANY REPRODUCTION IN PART OR AS A WHOLE WITHOUT THE WRITTEN PERMISSION IS PROHIBITED.

UNLESS OTHERWISE SPECIFIED:

DIMENSIONS ARE IN INCHES
TOLERANCES:
ANGULAR: MACH ± 1°
TWO PLACE DECIMAL ±.06
THREE PLACE DECIMAL ±.010

DO NOT SCALE DRAWING

	NAME	DATE
DRAWN		
CHECKED		

DESCRIPTION
SIDE BOTTOM SUPPORT LEG

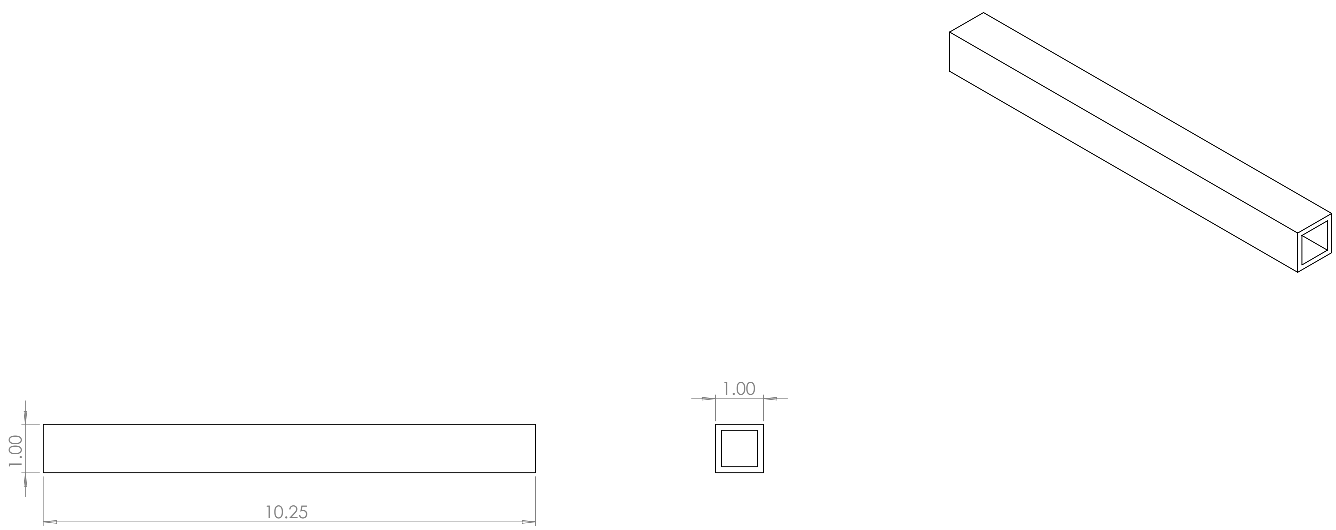
SIZE **B**

PART NO. 9

REV A

SCALE: 1:2 | QUANTITY: 2 | SHEET 1 OF 1

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.



MATERIAL
6061 ALUMINUM

INTERPRET GEOMETRIC
TOLERANCING PER: ASME Y14.5

PROPRIETARY AND CONFIDENTIAL
THE INFORMATION CONTAINED IN THIS DRAWING IS THE SOLE PROPERTY OF THE OWNERS. ANY REPRODUCTION IN PART OR AS A WHOLE WITHOUT THE WRITTEN PERMISSION IS PROHIBITED.

UNLESS OTHERWISE SPECIFIED:
DIMENSIONS ARE IN INCHES
TOLERANCES:
ANGULAR: MACH ± 1°
TWO PLACE DECIMAL ±.06
THREE PLACE DECIMAL ±.010

DO NOT SCALE DRAWING

	NAME	DATE
DRAWN		
CHECKED		

DESCRIPTION
BOTTOM SUPPORT LEG

SIZE **B**

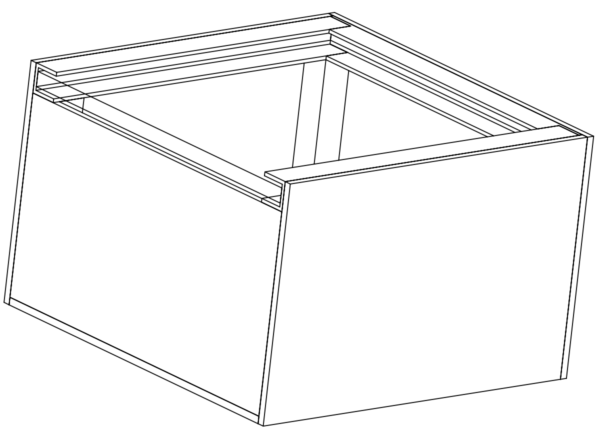
PART NO. 10

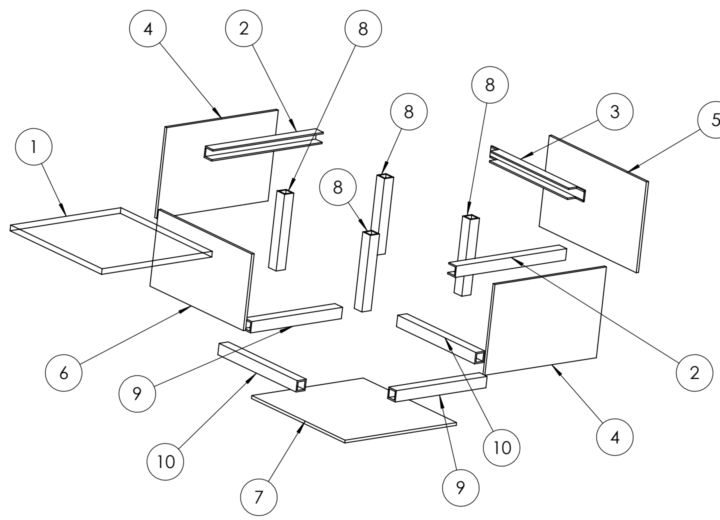
REV A

SCALE: 1:2 QUANTITY: 2 SHEET 1 OF 1

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

ITEM NO.	PART NAME	DESCRIPTION	QTY.
1	Acrylic Walking Surface	Clear Scratch- and UV-Resistant Cast Acrylic Sheet	1
2	U-Channel	Multipurpose 6061 Aluminum Rectangular Tube	2
3	Slotted U-Channel	Multipurpose 6061 Aluminum Rectangular Tube	1
4	Side Panel	Multipurpose 6061 Aluminum Sheet	2
5	Back Panel	Multipurpose 6061 Aluminum Sheet	1
6	Front Panel	Multipurpose 6061 Aluminum Sheet	1
7	Bottom Panel	Multipurpose 6061 Aluminum Sheet	1
8	Vertical Support Leg	Multipurpose 6061 Aluminum Rectangular Tube	4
9	Side Bottom Support Leg	Multipurpose 6061 Aluminum Rectangular Tube	2
10	Bottom Support Leg	Multipurpose 6061 Aluminum Rectangular Tube	2





MATERIAL 6061 ALUMINUM, CAST ACRYLIC INTERPRET GEOMETRIC TOLERANCING PER: ASME Y14.5	**UNLESS OTHERWISE SPECIFIED:** DIMENSIONS ARE IN INCHES TOLERANCES: ANGULAR: MACH ± 1° TWO PLACE DECIMAL ±.06 THREE PLACE DECIMAL ±.010	NAME / DATE DRAWN CHECKED	DESCRIPTION SMALL BOX
PROPRIETARY AND CONFIDENTIAL THE INFORMATION CONTAINED IN THIS DRAWING IS THE SOLE PROPERTY OF THE OWNERS. ANY REPRODUCTION IN PART OR AS A WHOLE WITHOUT THE WRITTEN PERMISSION IS PROHIBITED.	SIZE **B** DO NOT SCALE DRAWING	PART NO. A1 SCALE: 1:20	REV A QUANTITY: 1 SHEET 1 OF 1

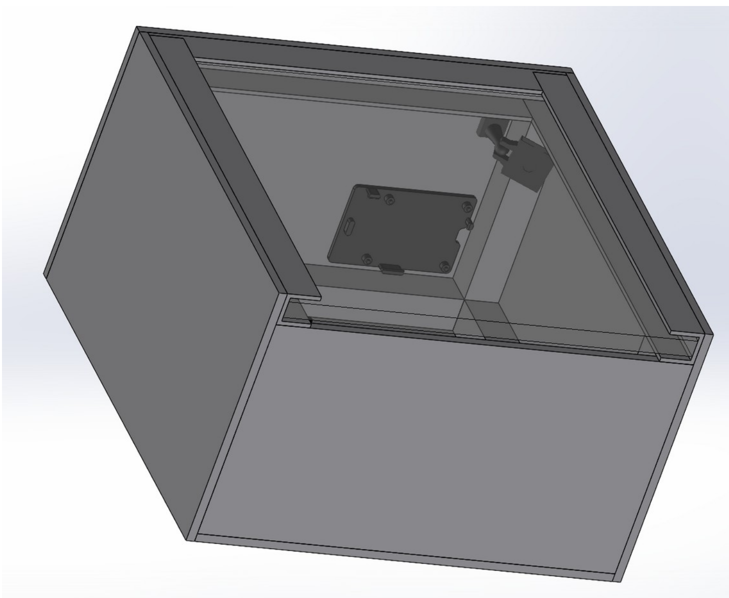
This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Small Box Usage Instructions

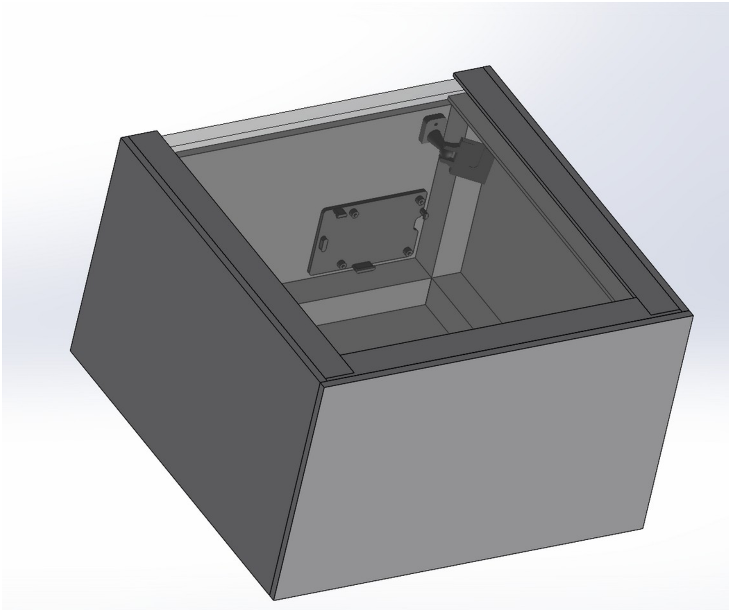
1. Slide the acrylic walking surface out of its slot to access box internals.
2. Connect both Raspberry Pis to a portable power bank via USB-C cable. The system is fully operational once both Pis are powered.
3. Replace the acrylic walking surface into its slot.
4. Step one foot diagonally onto the acrylic walking surface, in line with the distance sensor gangplank. The images will be captured automatically.
5. When you are done collecting data, slide the acrylic walking surface out of its slot to access internals.
6. Disconnect both Raspberry Pis from power.
7. To remove the data from the Raspberry Pi SD cards, either connect each Pi to a monitor setup to extract the images via the Raspberry Pi Desktop using a flash drive, or remove the SD card from the Pi and use a microSD reader to extract the images onto a computer. It is recommended to remove data after every data collection session.

Small Box Building Instructions

The attached PDF drawing package details the structure to be used for a vacuum former mold to create a small box. Additional 3D printed parts are required, and placement for those parts can be seen in the images below. The setup is identical on each of the two walls that have 3D printed parts mounted on them. 3D printed parts are affixed to the box with adhesives. All 3D printed parts were sourced from Thingiverse.com, and links are given below:
https://www.thingiverse.com/thing:5183183 use "pi_wall_mount_blank_4mm_clearance.stl"
https://www.thingiverse.com/thing:1156296 use both files



This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.



Software repositories are available containing all code written in support of the project.

- https://github.com/srvanderplas/fastai-test contains code for object recognition based models
- https://github.com/srvanderplas/ShoePatterns contains code for generating data to train future models

Presentations

- August 12, 2021: How Do you Define a Circle? Perception and Computer Vision Diagnostics. Susan VanderPlas, University of Nebraska Lincoln.

The synthetic data approach to footwear labeling was presented at a CSAFE footwear working group meeting on November 6, 2023.
Slides: https://srvanderplas.github.io/2023-Presentations/11-CSAFE-Shoe/.
Recorded Presentation
Abstract:

> Neural Networks are very complicated and very useful models for image recognition, but they are generally used to recognize very complex and multifaceted stimuli, like cars or people. When neural networks are used to recognize simpler objects with overlapping feature sets, things can go a bit haywire. In this talk, we'll discuss a model built for applications in statistical forensics which uncovered some very interesting problems between model-based perception and human perception. I will show visual diagnostics which provide insight into the model, and talk about ways we might address the discrepancy between human perception and model perception to produce more accurate and useful model predictions.

- February 25, 2022, 8:45-9:00 am
  Automatic Class Characteristic Recognition in Shoe Tread Images
  Jayden Stack (presenter) and Susan Vanderplas, University of Nebraska Lincoln.
  Abstract:

  > One of the fundamental problems in footwear forensics is that the distribution of class characteristics in the local population is not currently knowable. Surveillance devices for gathering this data are just half of the battle – it is also necessary to process the data gathered using these devices and identify relevant features.
  >
  > This presentation will describe progress made in automatic identification of relevant footwear features - brand, shoe size, and tread pattern elements, as well as complications which arise when combining machine learning algorithms with human-friendly features. Using transfer learning to connect pre-trained neural networks to newly gathered and labeled training data, this method bridges the gap between unfriendly numerical features and descriptors used by examiners in practice. Leveraging both clean training data and "messy" data gathered from the local community using newly developed footwear surveillance devices, the

authors will present developments in footwear forensics which will enable examiners to testify as to the frequency of class characteristics in the local population in the very near future.

- March 24, 2022, 11-12 am
  MANTIS: The Development, Deployment, and Application of An Active Footwear Data Collection System
  Richard Stone and Susan Vanderplas
  Recording: https://www.youtube.com/watch?v=iILj0pEIhio
- August 3, 2022: IAI workshop presenting the scanner and the potential for automatic collection of footwear information.
- November 6, 2023, CSAFE footwear working group meeting.
  Slides: https://srvanderplas.github.io/2023-Presentations/11-CSAFE-Shoe/
- December 6, 2023, International Association of Statistical Computing - Asian Regional Section in Macquarie Park, Australia on December 6, 2023.
  Slides: https://srvanderplas.github.io/2023-Presentations/12-IASCARS/#/title-slide

4 Participants and Other Collaborating Organizations

4.1 Individuals

Name	Project Role	Person Months Worked
Susan Vanderplas	PI	2.5
Richard Stone	Collaborator	2.5
Jayden Stack	Graduate Student	3.5
Muxin Ha	Graduate Student	5.5
Colten Fales	Graduate Student	2.5
Colton Fales	Hourly	1.5
Laura Paul	Hourly	0
Braden Westby	Hourly	3.5
Ayuush Mehta	Hourly	.25
Foster Moon	Hourly	2

4.2 Organizational Partners

4.2.1 Center for Statistics and Applications in Forensic Evidence

- Cooperative Agreement 70NANB20H019 between NIST and Iowa State, which includes activities carried out at Carnegie Mellon University, Duke University, University of California Irvine, University of Virginia, West Virginia University, University of Pennsylvania, Swarthmore College and University of Nebraska, Lincoln.
- Location: United States

- Role: advisory organization, provides opportunities for dissemination of research results to practitioners