# Computer Vision Introduction

INSTITUTE FOR ARTIFICIAL INTELLIGENCE

# The Summer Vision Project

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

PROJECT MAC

Artificial Intelligence Group
Vision Memo. No. 100.

July 7, 1966

THE SUMMER VISION PROJECT

Seymour Papert

The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition".

Figure 1. The Summer Vision Project by Seymour Papert

# The DARPA Image Understanding Program Began in 1975

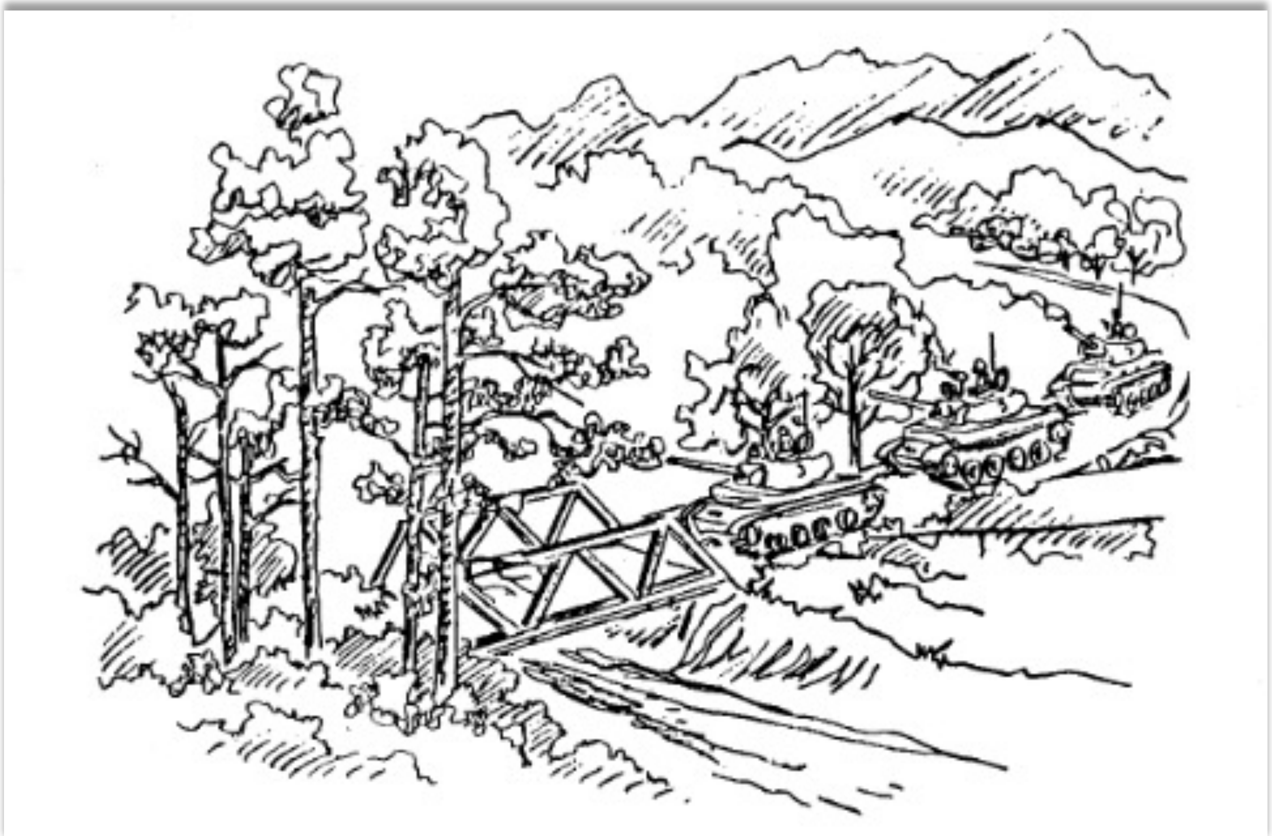


Figure 2. A human use knowledge of the world to understand what he sees in a two-dimensional image.

Tank Convoy Is Approaching Bridge
Over Elbe River On Route #410.
The Elbe River Is the Border
Betwe2n East and West Germany.
Route #410 Leads From Neuhaus.

Figure 3. Application of knowledge such as that contained in a map might produce a succinct summary of the scene in Fig. 2.

# ACRONYM – The First Vision System

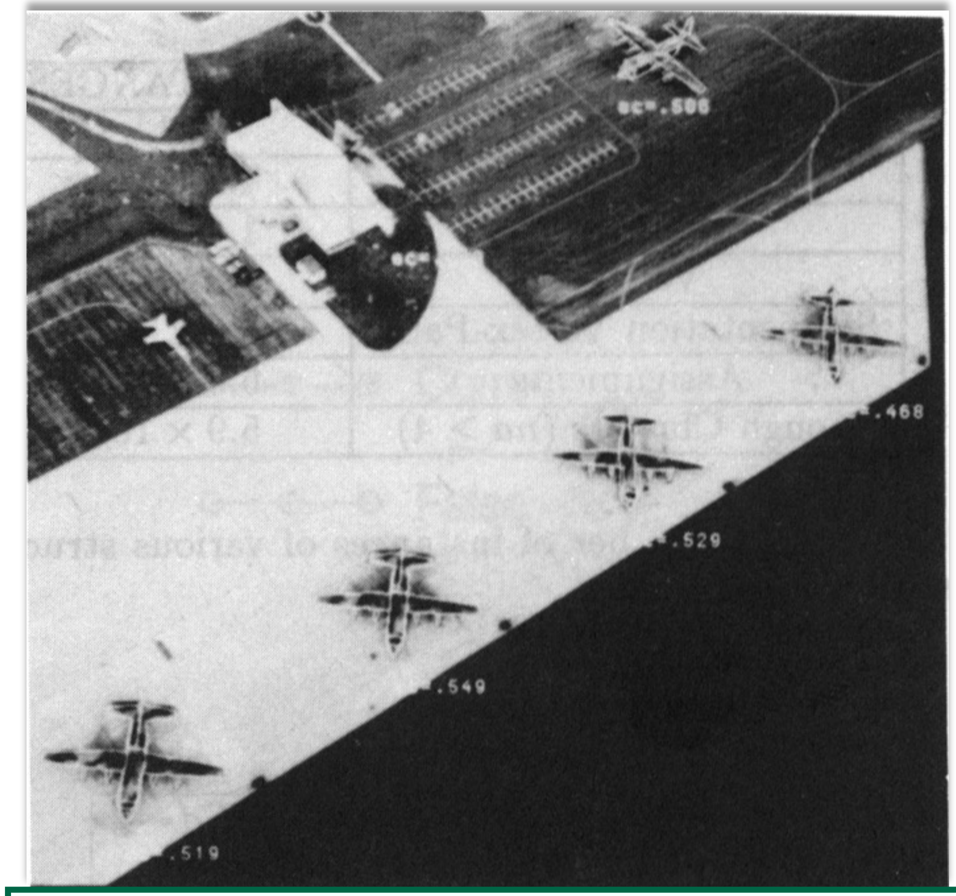



Figure 4. Satellite image of an airport. Taken from Brooks, R. A., Creiner, R., & Binford, T. O. (1979, August). The ACRONYM model-based vision system. In Proceedings of the 6th international joint conference on Artificial intelligence-Volume 1 (pp. 105-113). Morgan Kaufmann Publishers Inc..

# Complexity of Human Vision

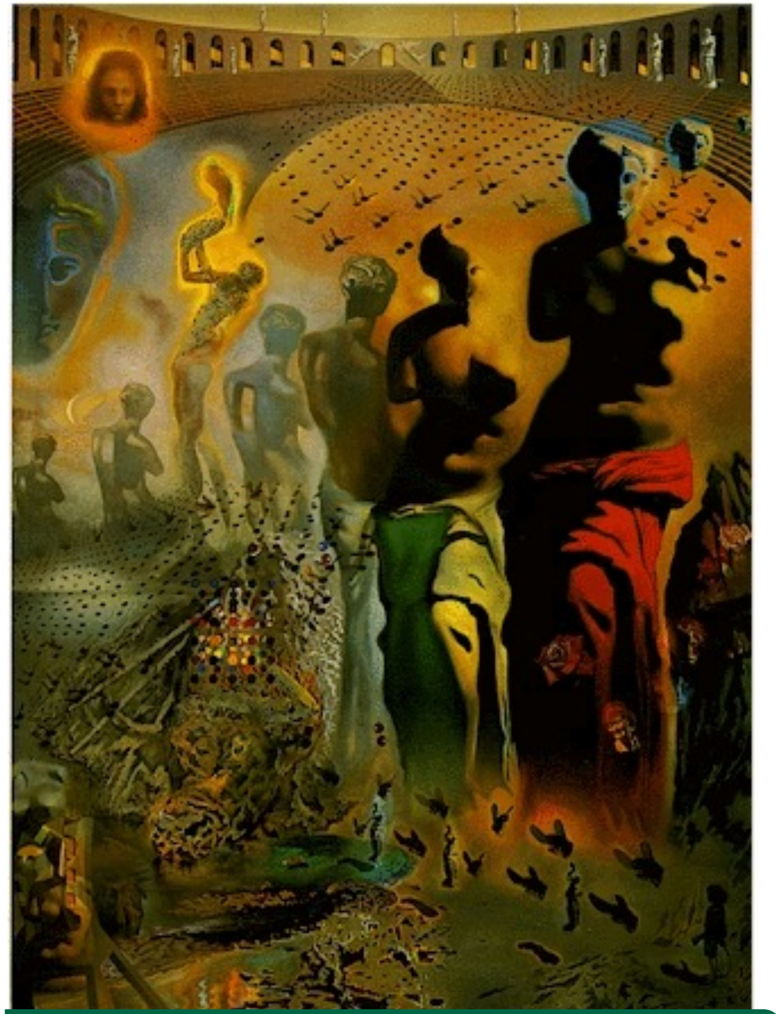
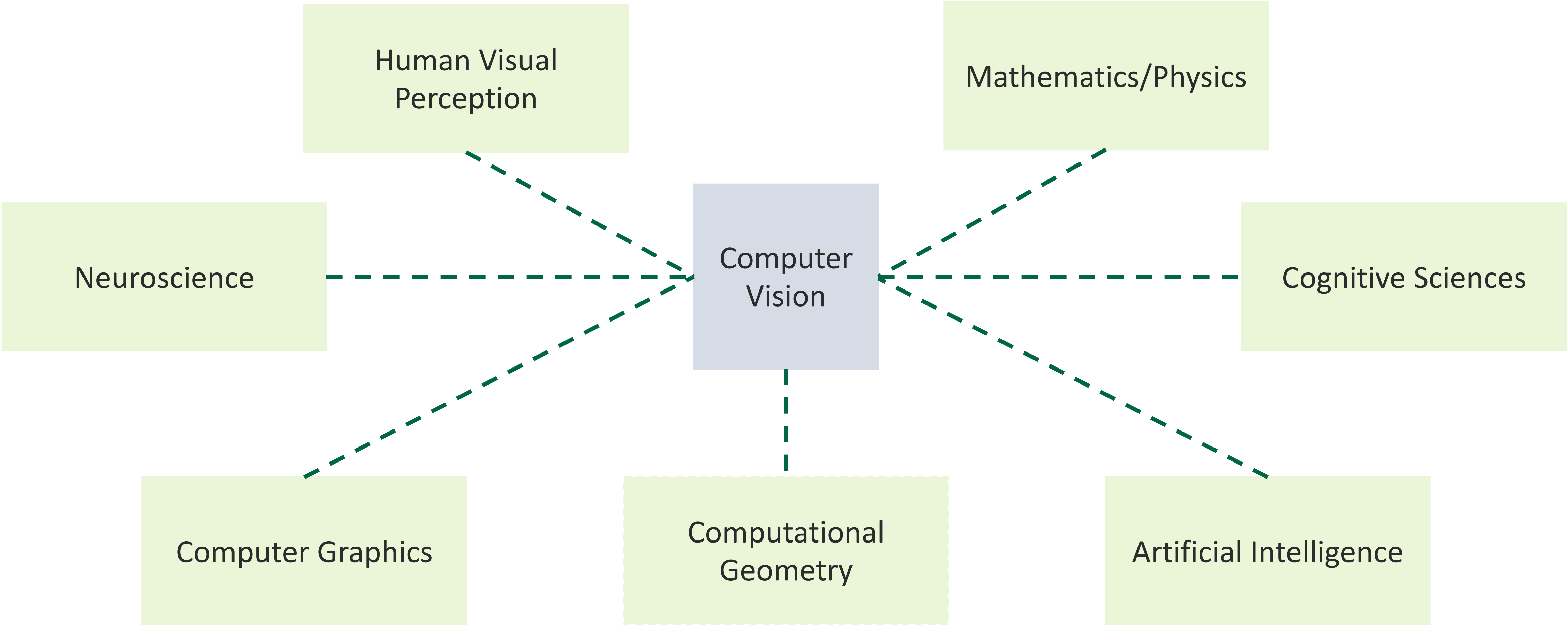


Figure 5. The Hallucinogenic Toreador, Salvador Dali

- 1/3 of the brain devoted to perception
  - Eye's just the front end
- Most of the processing not available for introspection
- Most reliable sensory mode
- Human vision
  - Collection of slapped together modules doing various kinds of tasks
    - Face recognition

# CV Draws from Many Disciplines



# AI + X

AI+X

INSTITUTE FOR ARTIFICIAL INTELLIGENCE

*“Unleashing human potential! Artificial intelligence to synergistically enable people to excel at what they do.”*

Source: USF Institute for Artificial Intelligence

# Event Understanding from Video Data

- Formulate a computer vision-based event understanding algorithm that operates in a self-supervised, streaming mode.
- There are many applications of this technology, such as building assistive robotics or smart spaces for independent living or monitoring wildlife.
- This research couples self-supervised learning process with prior knowledge, moving the field towards open-world algorithms, and needing little or no supervision.
- University of South Florida, Florida State University, and Oklahoma State University.
- Prof. Sudeep Sarkar
- National Science Foundation

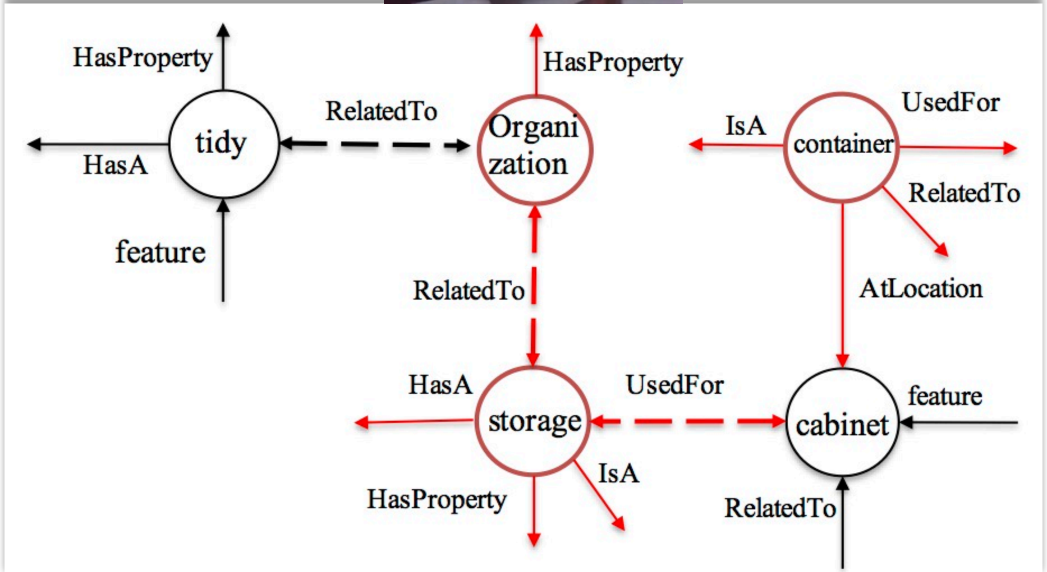



Figure 6. Interpretation of content in the video above.

# Recognizing Human Intentions

- Achieving Autonomy by Learning from Sensor-Assisted Control in a Wheelchair-Based Human-Robot Collaborative System.
- The goal of this project is to improve independence and quality of life by creating an adaptive human-robot collaborative system (a wheelchair mounted robotic arm) that learns from example to assist its user perform instrumental activities in a way that requires minimal user guidance.
- Rajiv Dubey, Redwan Alqasemi, Kyle Reed (Robotics) and Sudeep Sarkar (Computer Vision)
- National Science Foundation (NSF) project

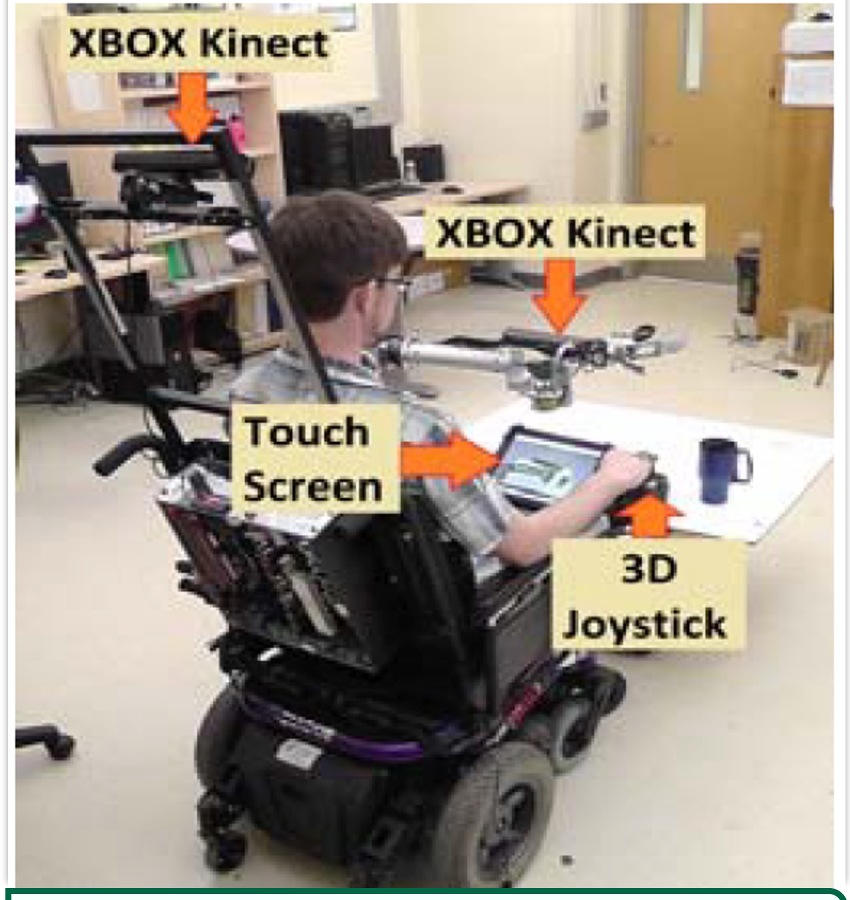



Figure 7. Instrumented wheelchair

# AI + Citizen Science for Mosquito Surveillance

- Large-scale automated identification of mosquito genera and species.
- Profs. Ryan Carney and Sriram Chellappan
- National Science Foundation

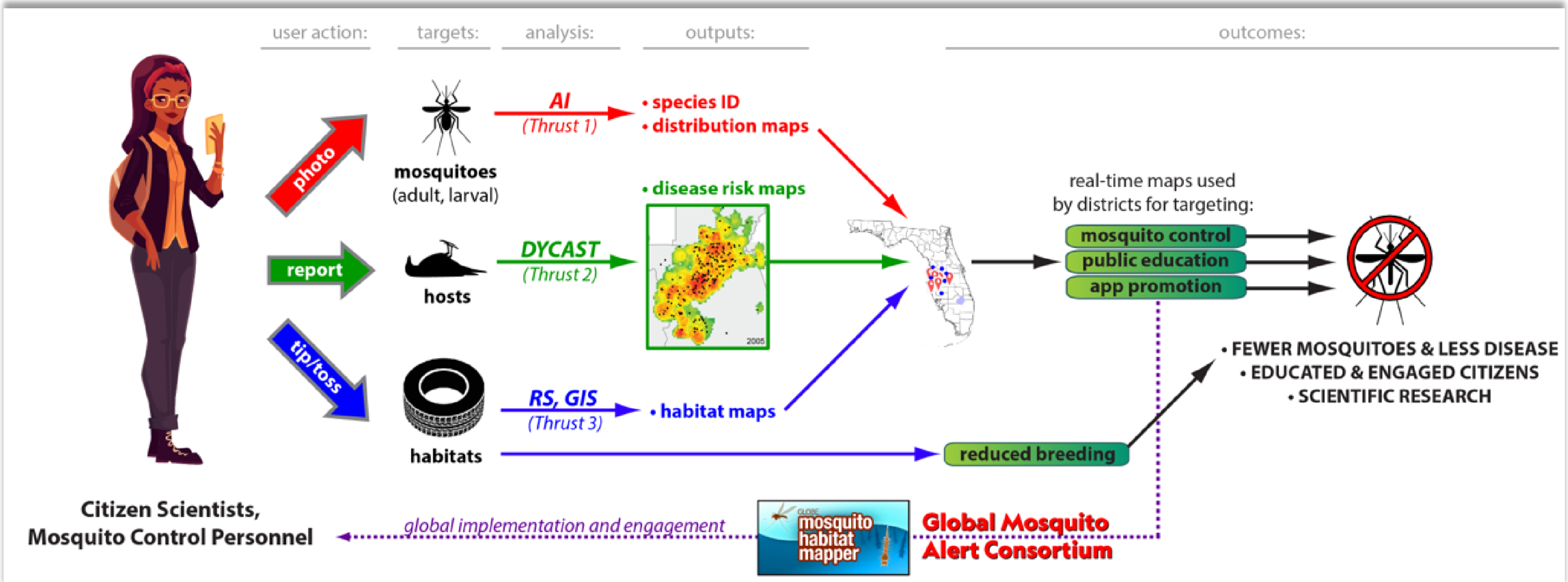



Figure 8. Framework for citizen science based mosquito surveillance

# Multimodal Assessment of Neonatal Pain

- Assessing neonatal pain is difficult because the current standard for assessment is subjective, inconsistent, and discontinuous.
- The continuous monitoring of pain, using affordable, non-invasive, and easily integrable devices, provides immediate pain detection and intervention, and therefore, contribute to improved long term outcomes; i.e., reduce the outcomes of under- and over-treatment.
- Profs. D. Goldgof, R. Kasturi, Y. Sun, T. Ashmeade
- National Institutes of Health

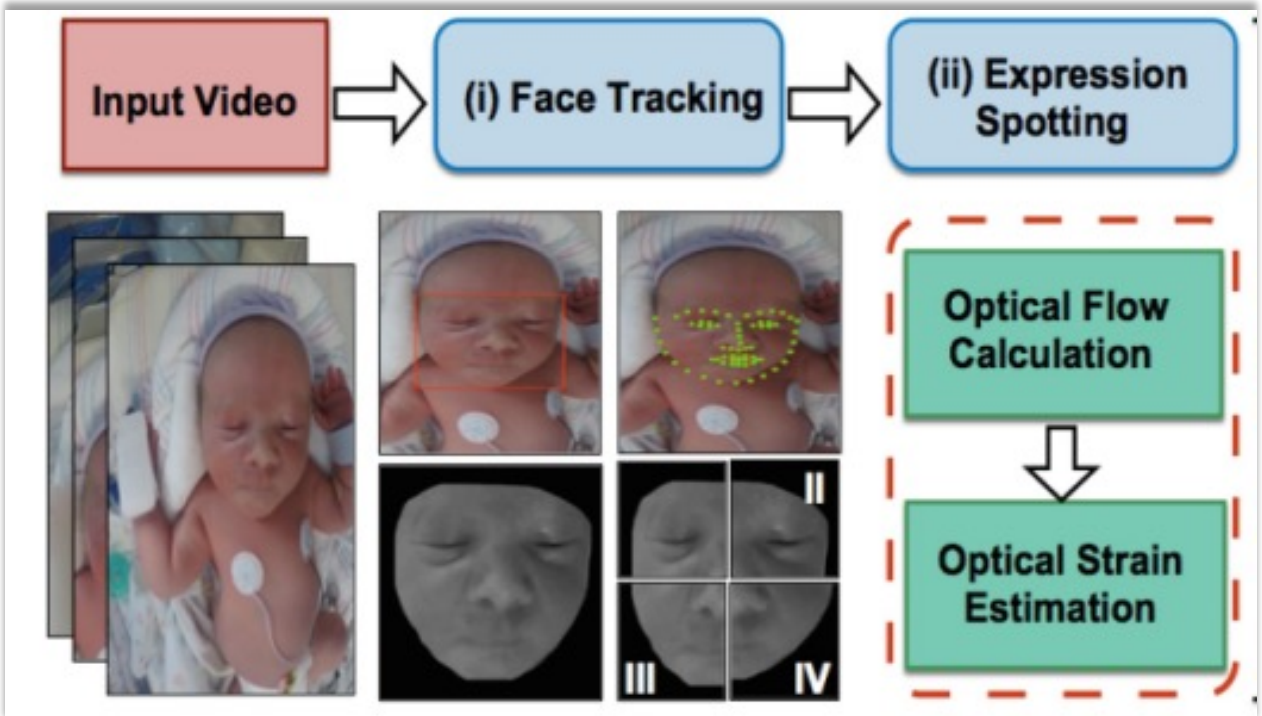



Figure 9. Detecting facial features for expression recognition in infants.

# Economic Activity from Satellite Images

- Remote sensing can help to understand their effects on human and economic activities by recording changes in human behavior over time by monitoring parking lots.
- Air traffic is of particular interest to unveil human and economic activities (e.g., travel, tourism, cargo) and track disease spread due to in-flight transmission.
- Solution adopted by European Space Agency. Winner of top prize Upscaling Euro Datacube
- Prof. Sudeep Sarkar
- USF, the University of California, Berkeley, and the global satellite company Maxar Technologies based in Colorado.

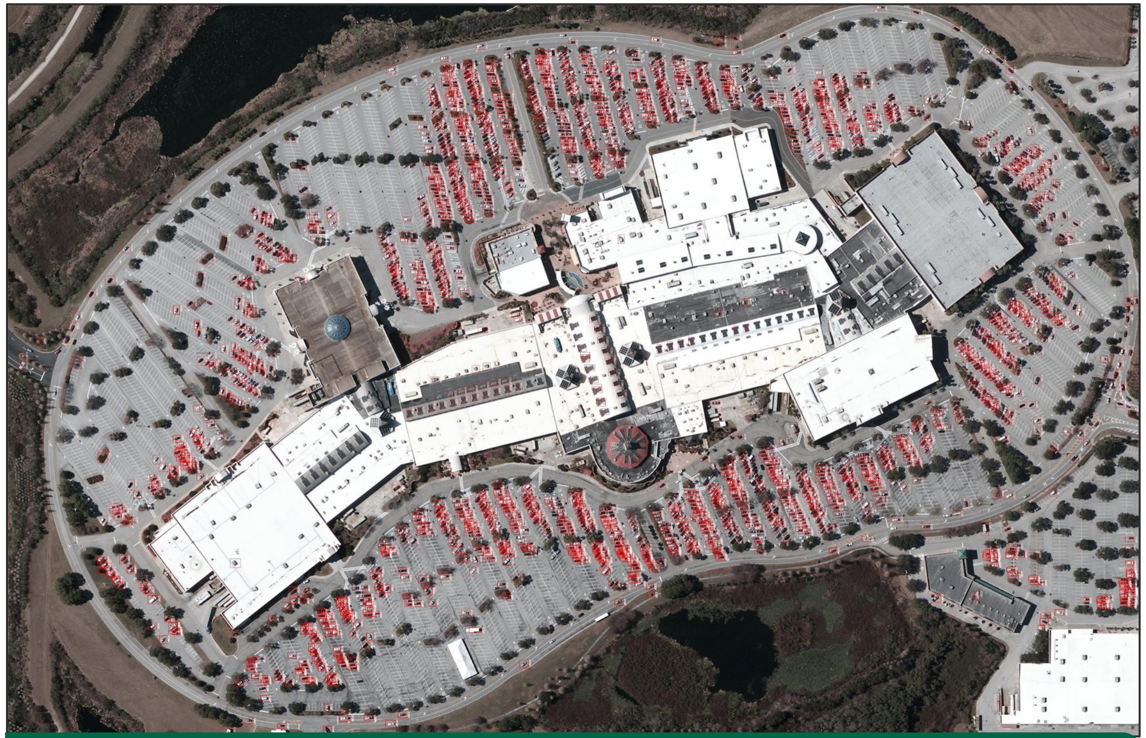


Figure 10. Cars detected (in red) in the parking lot of a mall in Tampa.

# Interpretable AI for COVID-19 Diagnosis

- The use of CT imaging enhanced by artificial intelligence to effectively diagnose COVID-19, instead of or in addition to reverse transcription-polymerase chain reaction (RT-PCR), can improve widespread COVID-19 detection and resource allocation.
- We developed, trained, validated, and tested an object detection model which detects features in three categories: ground-glass opacities (GGOs) for COVID-19, GGOs for non-COVID-19 diseases, and features that are inconsistent with a COVID-19 diagnosis.
- These collected features are passed into an interpretable decision tree model to make a suggested diagnosis.
- Prof. Shyam Mohapatra

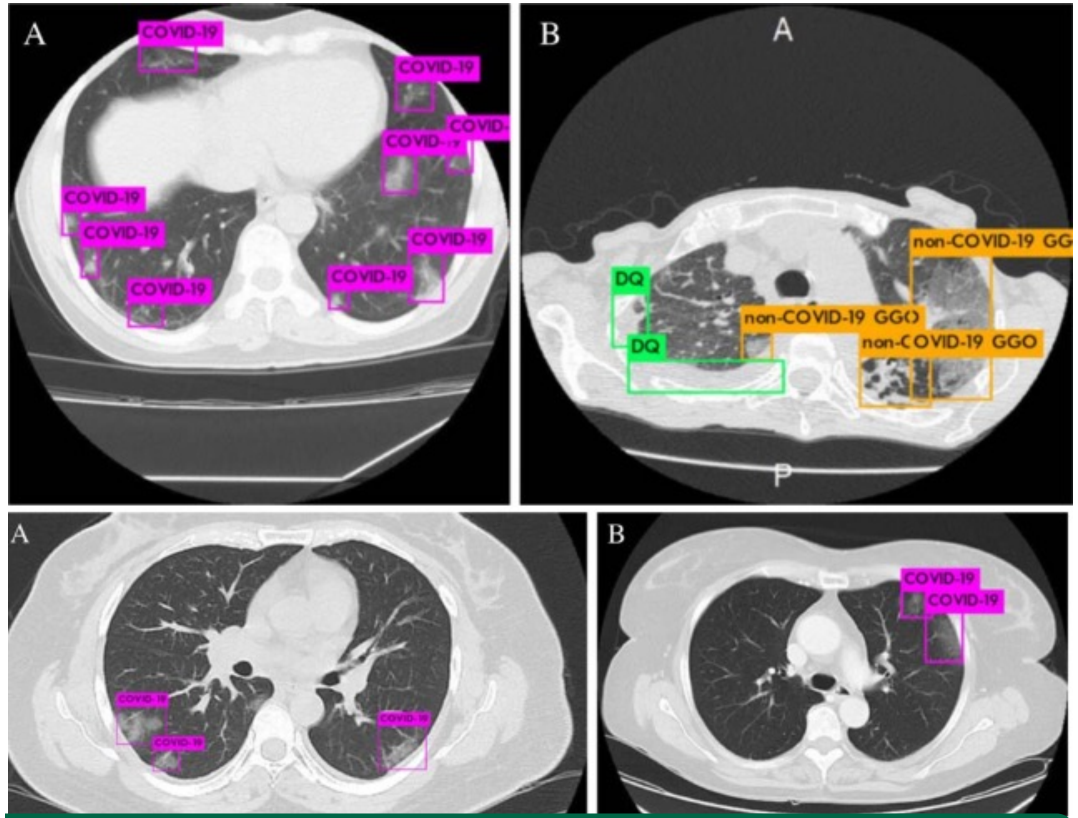
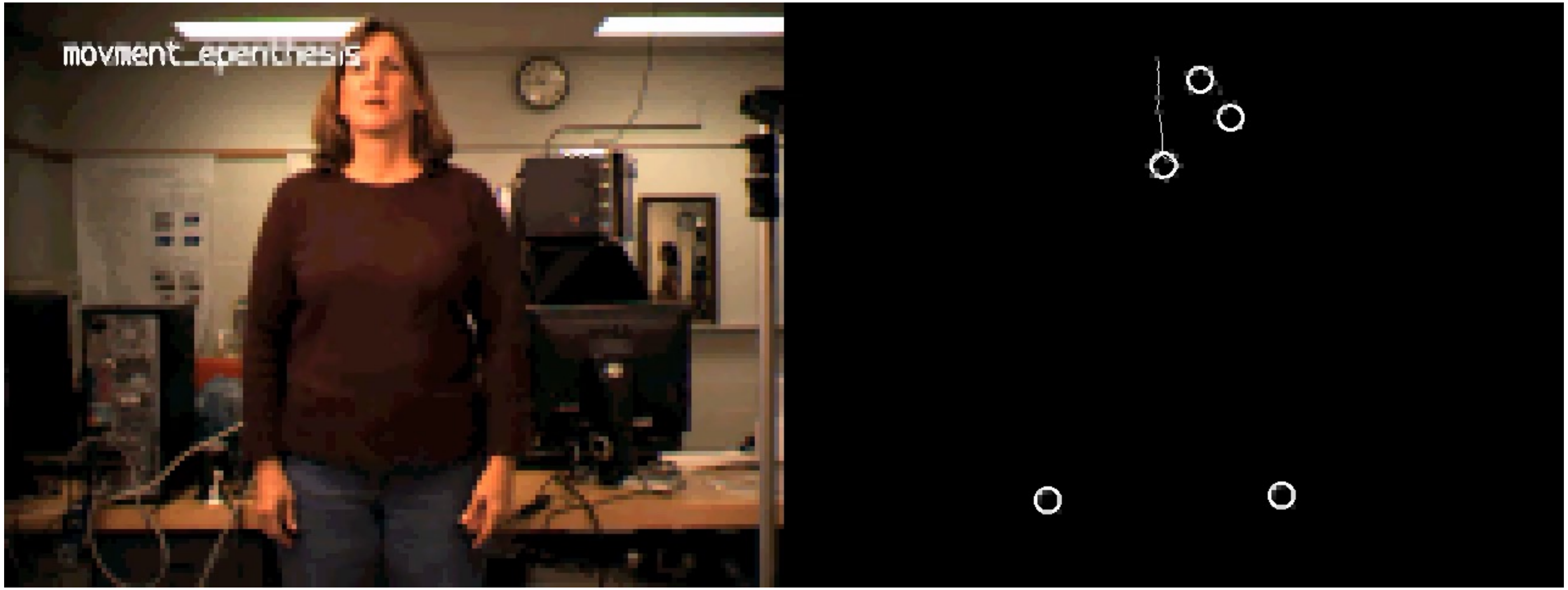



Figure 11. Computed tomography images of the lungs along with labels.

# Security from a Distance

The HumanID gait challenge problem: Data sets, performance, and analysis, S Sarkar, PJ Phillips, Z Liu, IR Vega, P Grother, KW Bowyer, IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI) 27 (2), 162-177

# Communicating with the Deaf




Handling movement epenthesis and hand segmentation ambiguities in continuous sign language recognition using nested dynamic programming.
R Yang, S Sarkar, B Loeding,
IEEE transactions on pattern analysis and machine intelligence 32 (3), 462-477

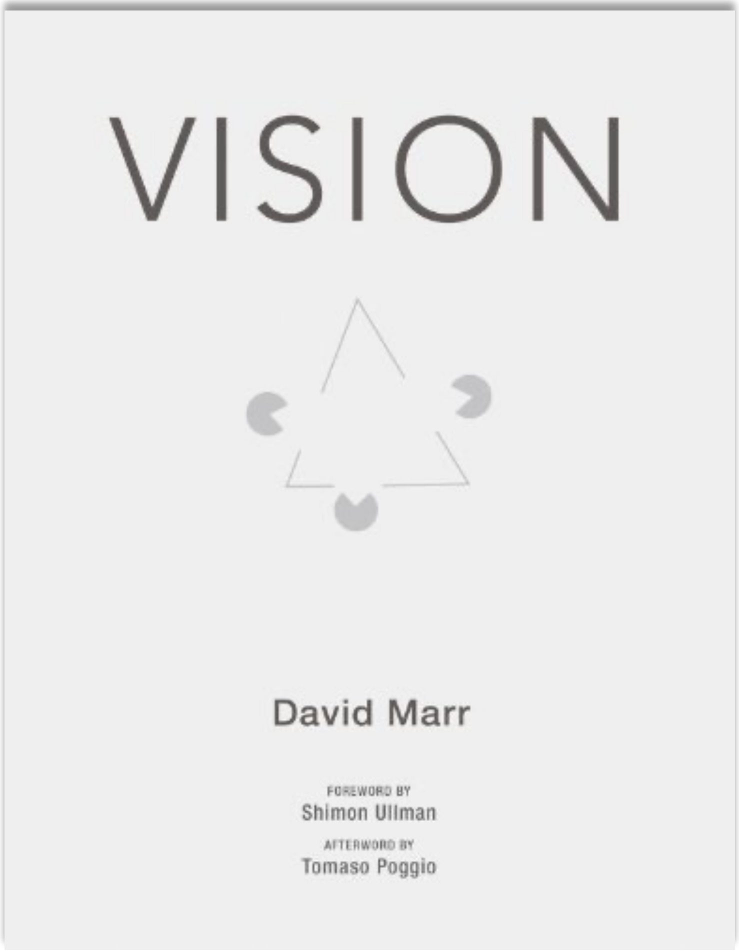
# NAE Grand Challenges for Engineering


Source: National Academy of Engineering

# Problems, Core Vision Problem Solved, Mathematical Methods

| Problems (NAE Grand Challenges) | Core Vision Problem Solved | Mathematical Methods |
|---|---|---|
| ▪ Advanced personalized learning<br>▪ Enhanced virtual reality<br>▪ Reverse engineering the brain<br>▪ Engineering better medicine<br>▪ Advanced health informatics<br>▪ Improve urban infrastructure<br>▪ Secure cyberspace<br>▪ Engineer the tools of scientific discovery | ▪ Segmentation<br>▪ 3D Motion estimation<br>▪ 3D Shape estimation<br>▪ Symbolic description of content<br>▪ Recover material property | ▪ Probabilistic methods<br>▪ Graph methods<br>▪ Linear optimization<br>▪ Non-linear optimization<br>▪ Linear algebra (SVD)<br>▪ Deep learning |

# Vision by David Marr




Source: The MIT Press

| Computational theory | Representation and algorithm | Hardware implementation |
|---|---|---|
| What is the goal of the computation, why is it appropriate, and what is the logic of the strategy by which it can be carried out? | How can this computational theory be implemented? In particular, what is the representation for the input and output, and what is the algorithm for the transformation? | How can the representation and algorithm be realized physically? |
| The three levels at which any machine carrying out an information-processing task must be understood. | | |

# You have reached the end of the lecture.

# Image/Figure References


Figure 1. The Summer Vision Project by Seymour Papert. https://people.csail.mit.edu/brooks/idocs/AIM-100.pdf
Figure 2. A human use knowledge of the world to understand what he sees in a two-dimensional image. Druffel L.E. (1982) Summary of the DARPA Image Understanding Research Program. In: Kittler J., Fu K.S., Pau LF. (eds) Pattern Recognition Theory and Applications. NATO Advanced Study Institutes Series (Series C — Mathematical and Physical Sciences), vol 81. Springer, Dordrecht
Figure 3. Application of knowledge such as that contained in a map might produce a succinct summary of the scene in Fig. 2. Druffel L.E. (1982) Summary of the DARPA Image Understanding Research Program. In: Kittler J., Fu K.S., Pau LF. (eds) Pattern Recognition Theory and Applications. NATO Advanced Study Institutes Series (Series C — Mathematical and Physical Sciences), vol 81. Springer, Dordrecht
Figure 4. Title. Brooks, R. A., Creiner, R., & Binford, T. O. (1979, August). The ACRONYM model-based vision system. In *Proceedings of the 6th international joint conference on Artificial intelligence-Volume 1* (pp. 105-113). Morgan Kaufmann Publishers Inc..
Figure 5. The Hallucinogenic Toreador, Salvador Dali. Retrieved from: https://en.wikipedia.org/wiki/File:The_Hallucinogenic_Toreador.png
Figure 6. Interpretation of content in the video above. Aakur, S., de Souza, F., & Sarkar, S. (2019). Generating open world descriptions of video using common sense knowledge in a pattern theory framework. Quarterly of Applied Mathematics, 77(2), 323-356.
Figure 7. Instrumented wheelchair. Carney RM, Mapes C, Low RD, Long A, Bowser A, Durieux D, Rivera K, Dekramanjian B, Bartumeus F, Guerrero D, Seltzer CE, Azam F, Chellappan S, Palmer JRB. Integrating Global Citizen Science Platforms to Enable Next-Generation Surveillance of Invasive and Vector Mosquitoes. Insects. 2022; 13(8):675. https://doi.org/10.3390/insects13080675
Figure 8. Framework for citizen science based mosquito surveillance
Figure 9. Detecting facial features for expression recognition in infants. G. Zamzmi, C. -Y. Pai, D. Goldgof, R. Kasturi, T. Ashmeade and Y. Sun, "An approach for automated multimodal analysis of infants' pain," 2016 23rd International Conference on Pattern Recognition (ICPR), 2016, pp. 4148-4153, doi: 10.1109/ICPR.2016.7900284.
Figure 10. Cars detected (in red) in the parking lot of a mall in Tampa. R. Minetto, M. P. Segundo, G. Rotich and S. Sarkar, "Measuring Human and Economic Activity From Satellite Imagery to Support City-Scale Decision-Making During COVID-19 Pandemic," in IEEE Transactions on Big Data, vol. 7, no. 1, pp. 56-68, 1 March 2021, doi: 10.1109/TBDATA.2020.3032839.
Figure 11. Computed tomography images of the lungs along with labels. Warman A, Warman P, Sharma A, Parikh P, Warman R, Viswanadhan N, Chen L, Mohapatra S, Mohapatra S, Sapiro G. Interpretable Artificial Intelligence for COVID-19 Diagnosis from Chest CT Reveals Specificity of Ground-Glass Opacities. medRxiv [Preprint]. 2020 May 18:2020.05.16.20103408. doi: 10.1101/2020.05.16.20103408. PMID: 32511545; PMCID: PMC7274226.