Bias in Computer Vision—It's Bigger than Facial Recognition!

Susan Kennedy, PhD
Assistant Professor of Philosophy
Santa Clara University

The Optimistic View

INTELLIGENT MACHINES

Will Smart Machines Be Less Biased Than Humans?

Robert D Atkinson Information Technology And Innovation Foundation

September 19, 2016

Bias in Computer Vision: Facial Recognition

Another Arrest, and Jail Time, Due to a Bad Facial Recognition Match

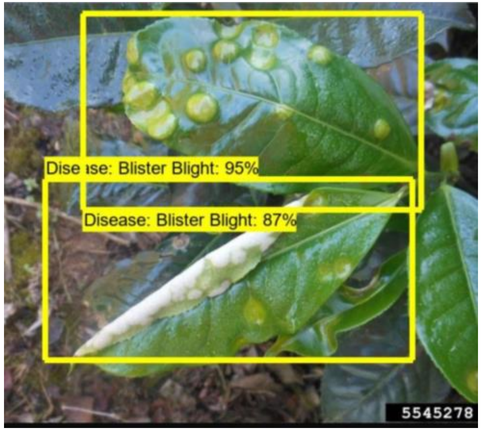
A New Jersey man was accused of shoplifting and trying to hit an officer with a car. He is the third known Black man to be wrongfully arrested based on face recognition.

By Kashmir Hill

Published Dec. 29, 2020   Updated Jan. 6, 2021

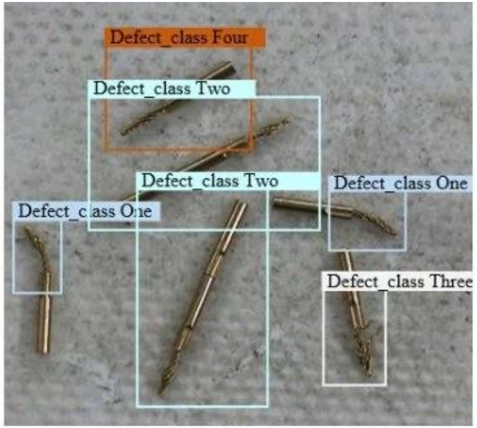
Bias in CV: Looking Beyond Facial Recognition

embedded VISION SUMMIT



Agriculture

Plant Disease Detection



Manufacturing

Quality Inspection

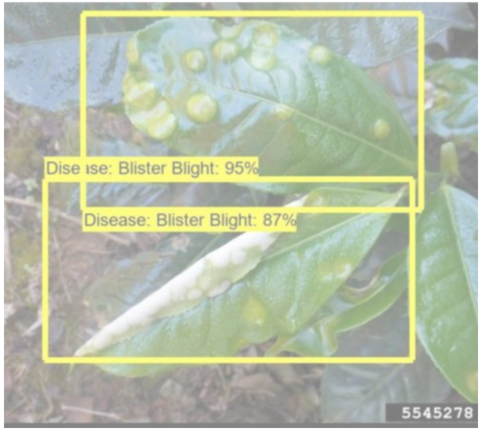
Transportation

Pothole Detection

Santa Clara University

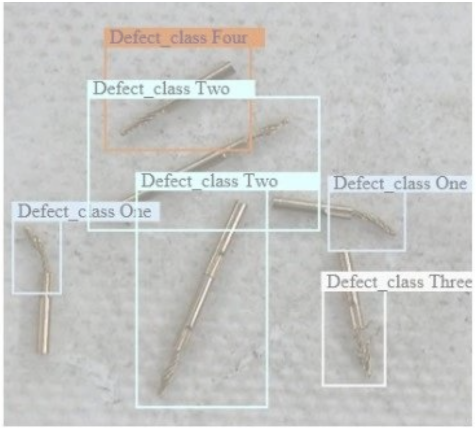
Bias in CV: Looking Beyond Facial Recognition

Bias can pose an ethical challenge, even without sensitive data!



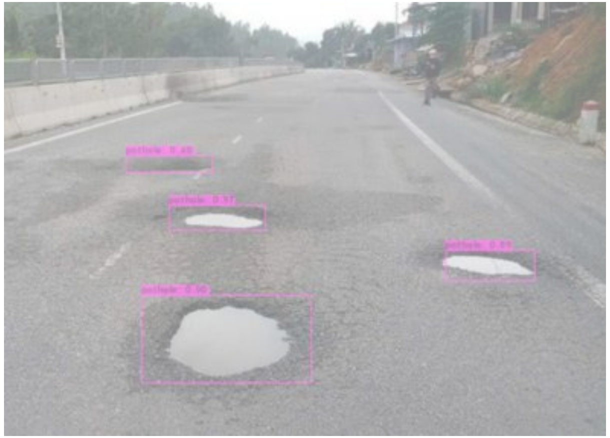
Agriculture

Plant Disease Detection



Manufacturing

Quality Inspection

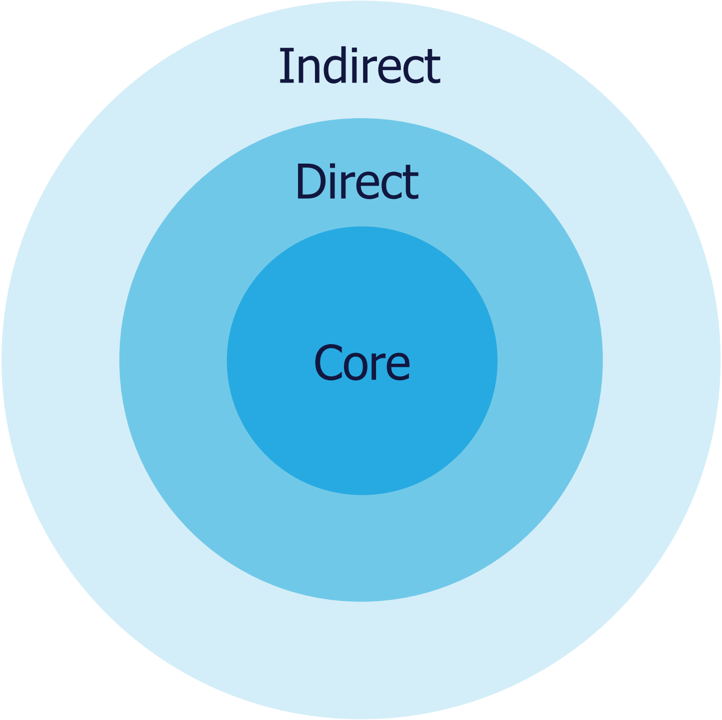


Transportation

Pothole Detection

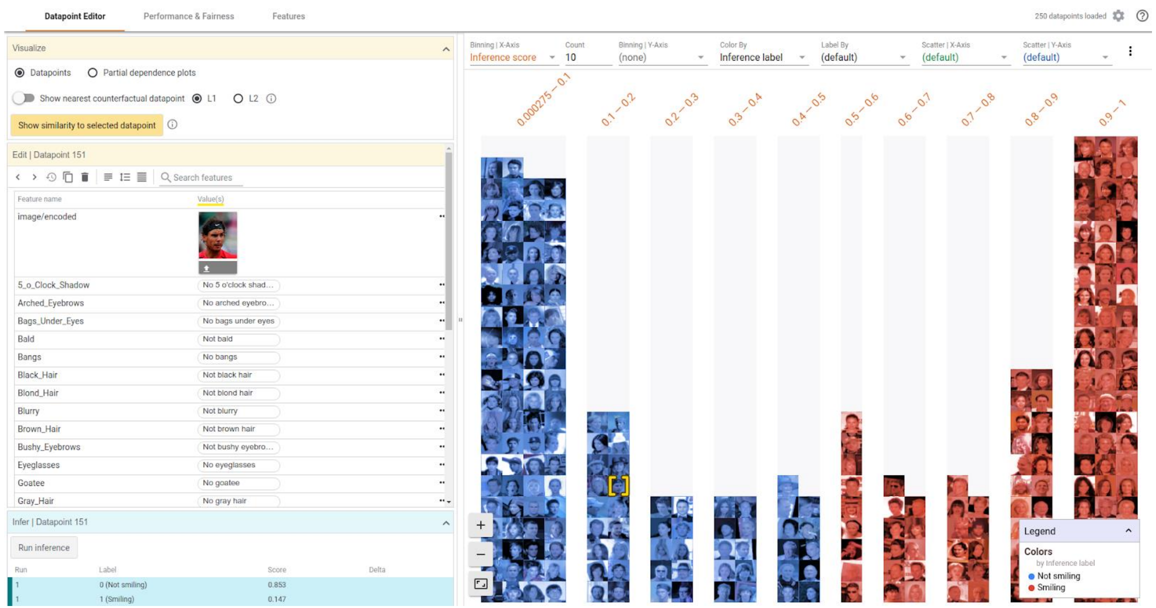
Human *Subjects* ≠ Human *Impacts*

- Expand the ethical circle to take into account the full range of stakeholders
  - *Who* does it not work for?
  - *What* does it not work for?
  - *When* does it not work?



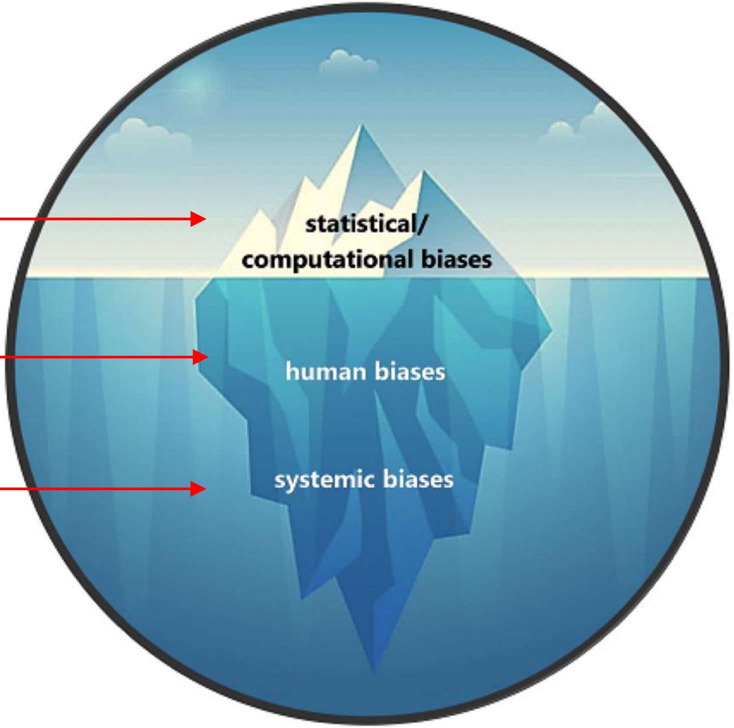
Mitigating Bias – Technical Solutions

- Toolkits
  - Google What-If
  - IBM Fairness 360
  - Microsoft FairLearn



The Tip of the Iceberg

- Systematic errors stemming from bias in the datasets and algorithmic processes used
- Human bias present across the AI lifecycle and in the use of AI once deployed
- Present in the datasets used in AI, and the institutional norms and practices across the AI lifecycle and in broader society



Bias-Aware not Bias-Free!

- Bias-free AI is an unachievable goal

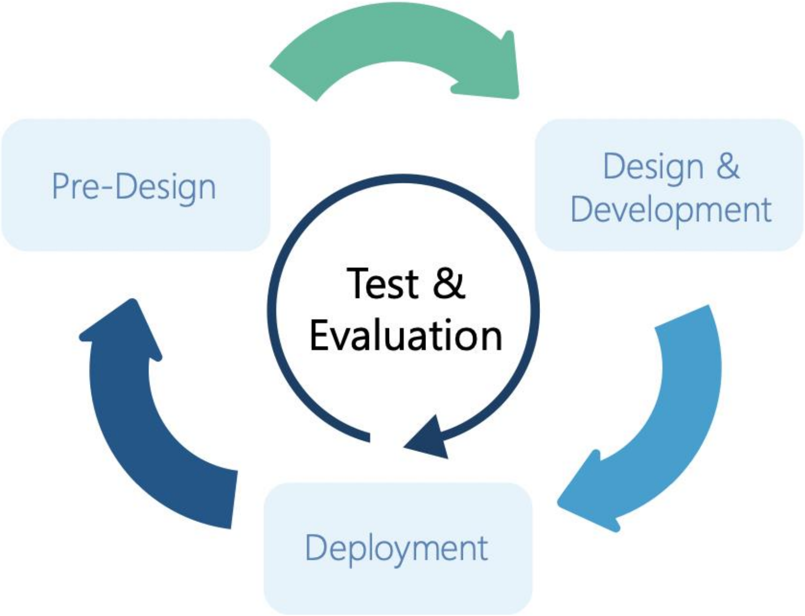
  → Mitigating bias requires a **bias-aware** approach

- AI exists within a larger social system

  → Mitigating bias requires a **sociotechnical** approach

Responsible AI – Ethical Pre-Mortems

- Instead of waiting for bias to strike, build a habit of anticipating preventable causes so they can be mitigated
- Across the entire AI lifecycle – From pre-design to deployment
- 3 key problem areas:
  - Datasets
  - Testing and evaluation
  - Human factors



Strategies to Employ

Dataset	• Interrogating decisions about who/what gets counted and how • Statistical methods to mitigate representation issues • Culture, context, & stakeholders in terms of dataset suitability
Testing and Evaluation	• Fairness metrics (context specific!) • Monitoring performance after deployment • Periodic model updates, test and recalibrate model parameters
Human Factors	• Multistakeholder engagement and diverse perspectives • Model and procedural transparency • Algorithmic impact assessments, iterative process

Thermal Imaging for Human-Wildlife Conflict - Arribada Initiative

Dataset

- Hardware testing – optimizing data collection methods across species
- Suitability – adjusting for variations in the environment (dirt, grass, snow)

Human Factors

- Stakeholder engagement – WWF team in Tezpur advised on locations for field testing
- Commitment to model transparency

Key Takeaways

1. Bias poses an ethical challenge, regardless of the stakes involved
2. Reframing the goal - bias-aware not bias-free
3. An effective mitigation strategy requires a combination of technical and social considerations

Resources

1. NIST Special Publication 1270 – Towards a Standard for Identifying and Managing Bias in Artificial Intelligence
https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1270.pdf

2. (In Progress) NIST Mitigation of AI/ML Bias in Context
https://www.nccoe.nist.gov/projects/mitigating-aiml-bias-context

3. An Ethical Toolkit for Engineering Design Practice – Markkula Center for Applied Ethics
https://www.scu.edu/ethics-in-technology-practice/ethical-toolkit/