



AI for Good Global Summit

Trust in AI: opportunities and challenges

May 16, 2018

Trust in AI by educating engineers to ethically aligned design

Prof. Hagit Messer Yaron

messer@eng.tau.ac.il



TEL AVIV אוניברסיטת
UNIVERSITY תל אביב

Faculty of Engineering

The Kranzberg Chair in Signal Processing



IEEE

The IEEE Global
Initiative on Ethics
of Autonomous and
Intelligent Systems



The World Commission
on the Ethics of
Scientific Knowledge
and Technology
COMEST* is an advisory
body and forum of
reflection that was set
up by UNESCO in 1998.

Introductory notes

- My talk is about: “educating engineers to ethically alighted design of autonomous intelligent systems (AIS)”
 - “**Engineers**” include anyone who is involved in the design and developments of AIS (i.e., EE, CS etc.)
 - **AIS** vs. **AI** : AI is a tool, as well as sensors, mechanical parts, etc., used in building AIS

Ethics and Trust

- Ethics is fundamental to fostering trust in AIS technologies, and it is crucial for current and future engineers to be educated on ethically aligned design. The need for ethics in AIS design has been emphasized at the international level by **UNESCO's World Commission on the Ethics of Scientific Knowledge and Technology**
- The challenge: The curriculum of most Engineering programs DO NOT include developing tools for raising awareness to ethical considerations in AIS ([Wendell Wallach](#) 2018)
- The opportunity: Joint efforts of International organizations can facilitate and accelerate a change

This PDF is available at <http://nap.edu/21889>

SHARE



Infusing Ethics into the Development of Engineers: Exemplary Education Activities and Programs

- The US Example: **Infusing** Ethics...(2016)
- Educating **ENGINEERS** to ethical thinking needs to bridge over a cultural gap between technology and **ART** (Accountability, Responsibility, Transparency)



Trust in AI and the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems **MISSION:**

To ensure every stakeholder involved in the design and development of autonomous and intelligent systems [“engineer”] is educated, trained, and empowered to prioritize **ethical considerations** so that these technologies are advanced for the benefit of humanity.



The IEEE Global Initiative brings together several hundred participants from six continents, who are thought leaders from academia, industry, civil society, policy and government in the related technical and humanistic disciplines to identify and find consensus on timely issues.

The ethical design, development, and implementation of these technologies should be guided by the following General Principles:

- **Human Rights**: Ensure they do not infringe on internationally recognized human rights
- **Well-being**: Prioritize metrics of well-being in their design and use
- **Accountability**: Ensure that their designers and operators are responsible and accountable
- **Transparency**: Ensure they operate in a transparent manner
- **Awareness of misuse**: Minimize the risks of their misuse

EXAMPLE: the Role of the Engineer in applying general ETHICAL principles

People Counting in High Density Crowds from Still Images

Ankan Bansal, and K S Venkatesh (2015)

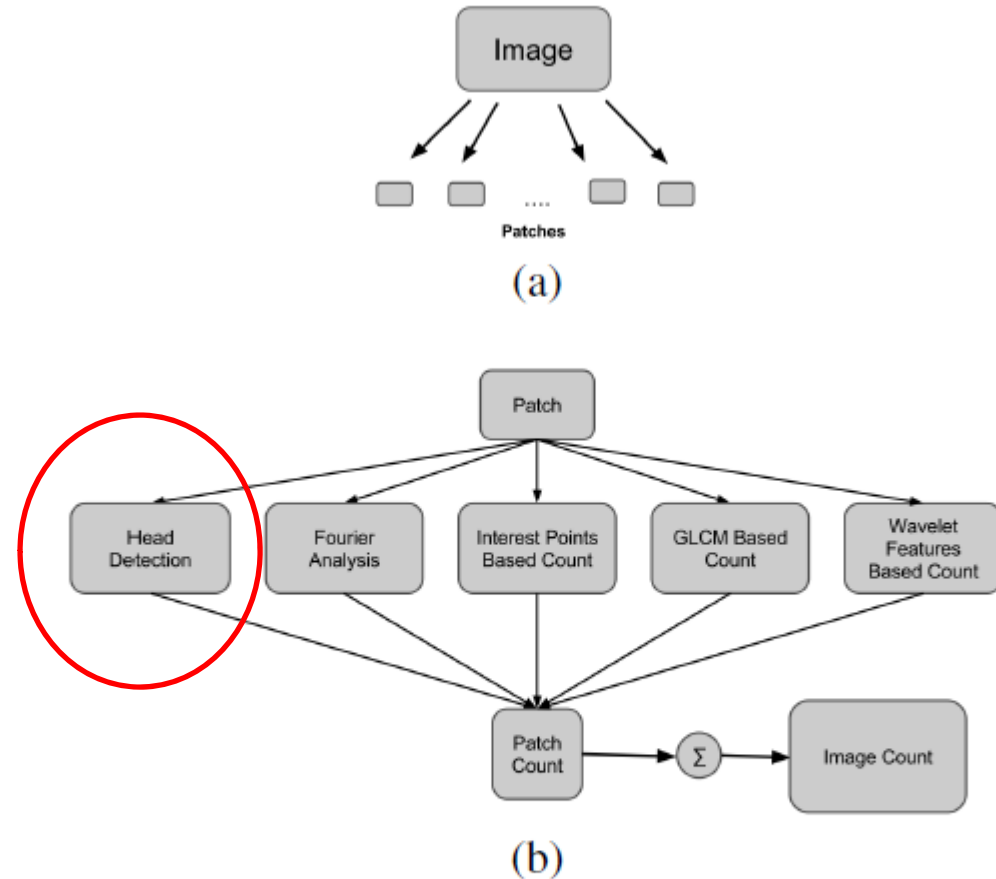


Fig. 2. A flow chart illustrating the methodology adopted in this paper. (a) The image is first divided into small cells in a grid. (b) Count is estimated for each cell by fusing estimates from five different methods. The final count estimate for the image is the sum of all cell counts.

Awareness of **misuse**

- **Recognize an Ethical Issue** – potential violation of privacy
- **Get the Facts** – do I really need the details?
- **Evaluate Alternative Options** – add image preprocessing for edge identification (contours)
- **Quantify the trade-offs** (processing time, memory use, performance, potential misuse, bias, etc.)
- **Decisions**

Summary

- Trust in AIS is based on EAD
- Regulation is not sufficient! The role of an individual engineer in building trust in AIS is extremely important
- EDUCATION, and in particular “Infusing Ethics into the Development of Engineers” is necessary to guarantee sustainable EAD of AIS, and therefore TRUST IN AI

Thank you!