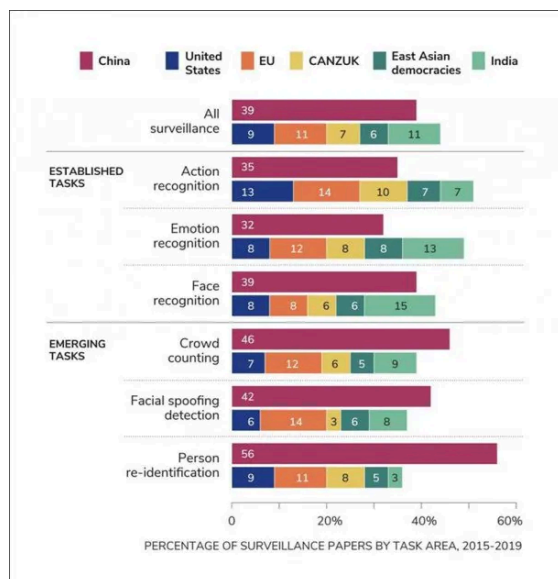


CSET: China Leads the World in Computer Vision Surveillance Research

Originally published at <https://www.unite.ai/cset-china-leads-the-world-in-computer-vision-surveillance-research/>



A new report from the Center for Security and Emerging Technology (CSET) has found that China's research sector produces 'a disproportionate share' of research into three core AI-related surveillance technologies, and that the CCP's more general contribution to computer vision technologies is growing at the same rate, and notably overtaking western rates of publication.



China's clear lead in research initiatives into more controversial sub-sectors of computer vision research, chiefly related to surveillance. Source: <https://cset.georgetown.edu/wp-content/uploads/Surveillance-in-the-CV-Literature.pdf>

The three key areas in which China has a massive lead are person re-identification (REID), crowd counting and spoofing detection (i.e. technologies that aim to expose attempts to subvert identification technologies).

Additionally, as indicated in the graph above, China's research community publishes a notably higher percentage of papers on human-facing computer vision tasks which, the paper argues, represent supporting technologies for wider surveillance solutions that use machine learning. These tasks include emotion recognition, face recognition, and action recognition.

The authors comment:

'These algorithms are often applied for benign, commercial uses, such as tagging individuals in social media photos. But progress in computer vision could also empower some governments to use surveillance technology for repressive purposes.'

On a less sinister note, the authors have found that papers related to visual surveillance account for under 10% of all computer vision research undertaken in the study period, and that the broader tranche of research is quite evenly distributed across countries.

However, China's dominance is clear, the researchers contend*:

'Researchers with Chinese institutional affiliations were responsible for more than one third of publications in both computer vision and visual surveillance research.'

'This makes China by far the most prolific country in both areas. Chinese researchers' share of global visual surveillance research is growing at a similar rate to their share of computer vision research.'

The new [report](#), titled *Trends in AI Research for the Visual Surveillance of Populations*, represents the application of Natural Language Processing (NLP) approaches to a dataset of published papers covering the years 2015-2019, and is written by Ashwin Acharya, Max Langenkamp and James Dunham.

English Language Bias

The authors of the paper observe that their study only touches on English language scientific papers, and that extending it to non-Anglophone publications could reveal a deeper iceberg of academic endeavor from China in these sectors. Further, the researchers believe that augmenting the data with ancillary information, such as patent data, camera deployment and pertinent government policies, could increase this statistical lead.

Naturally, the paper concedes, analyzing public and openly published papers cannot account for private corporate or state research, and classified research, but is a workable index of sector activity in the absence of these hidden data points.

Architecture and Data

The authors derived core data by training a [SciREX](#) document-level information extraction model on data from Papers With Code, with the framework deriving the relevance of papers by identifying references to tasks related to computer vision, and particularly to surveillance-centric projects and initiatives.

The model was then applied to an aggregated CSET body of scholarly literature containing more than 100 million individual publications across six academic datasets. The publishing platforms involved were Dimensions, Web of Science, Microsoft Academic Graph, China National Knowledge Infrastructure, arXiv, and Papers With Code.

Trained on Arxiv preprints, a [SciBERT](#) classifier was then tasked with identifying computer vision papers across the corpus.

The fact that SciREX and SciBERT are trained on English language documents prevented the researchers from extending the reach of the study beyond English. Of this, the researchers comment: *'This means that in national comparisons it underestimates non-English research output, and in particular it likely underrepresents China's share of*

Findings

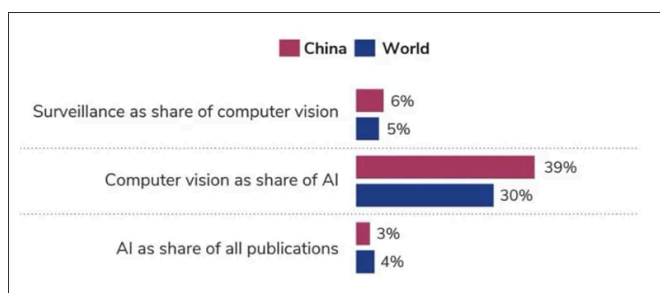
Within the visual surveillance sector, the study finds that face recognition was the most recurrent task, appearing in more than a thousand papers for the year 2019. However, the authors note that crowd-counting and face-spoofing recognition are 'fast growing' fields of pursuit.

TASK AREA	2019 PAPERS	TOP 10 MOST FREQUENT TERMS IN THE AREA
General*	30,994	Classification, Segmentation, Detection, Computer vision, Image processing, Object detection, Recognition, Feature extraction, Image segmentation, Image classification
Visual surveillance**	3,116	Face recognition, Action recognition, Face detection, Facial expression recognition, Video surveillance, Security, Human action recognition, Re-ID, Surveillance
Medical	2,493	Diagnosis, Medical imaging, Breast cancer, Medical image analysis, Diabetic retinopathy, Fusion, Medical image segmentation, Computer-aided diagnosis, Treatment
Remote sensing	1,263	Remote sensing, SAR, Target detection, Remote sensing images
Autonomous vehicles	724	Autonomous driving, Autonomous vehicles
Other	1,988	SR, Pose estimation, Hyperspectral image classification, Super-resolution, Pedestrian detection, Agriculture, Image super-resolution, Depth estimation

From the paper, the most frequently recurring tasks individuated for the years studied. Cited source is 'CSET merged corpus. Results generated July 22, 2021'

The authors of the paper consider that even the apparently more 'neutral' and less politically incendiary computer vision pursuits related to surveillance can also contribute to repressive control systems. For 'Action recognition', they observe that this can be used to identify 'abnormal behavior' in crowded public spaces; for face spoofing, they comment 'While sometimes used in biometric login systems or to prevent fraud, it may also prevent journalists and activists from hiding their identity'; and with regards to emotion recognition, the paper comments that 'In addition to its non-security-oriented and commercial purposes, some researchers, firms, and government agencies propose applying emotion recognition to identify security threats in crowded public areas'.

In general, the findings seem to show that China is above-averagely interested in computer vision research, compared to the global average.



The authors conclude:

'[The] share of both computer vision and visual surveillance from China increased over time. The United States, together with its allies and partners, published a similar amount of research in these areas as China published alone. However, these other regions' share of global surveillance research was stable or declined while China's grew.'

**The paper authors' bold emphasis.*

First published 6th January 2022.

Writer on machine learning, domain specialist in human image synthesis. Former head of research content at Metaphysic.ai, until its dissolution into DNEG's Brahma.ai.

Portfolio site: martinanderson.ai

Contact: martin@martinanderson.ai



Discover More