

Research Assessment Exercise 2020

Impact Case Study

University: [The Chinese University of Hong Kong]

Unit of Assessment (UoA): [13 CSIT - Computer Studies/Science (incl. information Technology)]

Title of case study: Deep Learning in Computer Vision

(1) Summary of the impact (indicative maximum 100 words)

The Multimedia Lab of CUHK (MMLab), led by Professor Xiaoou Tang, is a major innovator in deep learning and computer vision. MMLab was among the earliest to apply deep learning to solve computer vision problems, thus leading to technological breakthroughs in a number of areas, including the first face recognition system that surpasses human performance and the first super-resolution method ready for commercial use. MMLab's innovations have also resulted in the founding of SenseTime, which is one of the most valued AI companies in the world. SenseTime is now working with over 700 corporations to promote AI technologies in various industrial sectors and is recognized as the China's National Open Innovation Platform for Next-Generation AI on Intelligent Vision.

(2) Underpinning research (indicative maximum 500 words)

MMLab at CUHK is a pioneer in inventing innovative deep learning methods for tackling practical computer vision problems. Underpinning this case study, MMLab's research between 2013 and 2019 has made significant contributions in substantially pushing forward the envelope in a number of key tasks:

Face and Person Recognition

MMLab has always been a leader in face and person recognition. The lab developed a ground-breaking technique, GaussianFace, which won the **Outstanding Student Paper Award** [R3] in the most prestigious conference on artificial intelligence, AAAI 2015. On LFW, a large-scale benchmark widely used to test face recognition techniques, GaussianFace has attained an impressive accuracy of 98.52%, which, for the first time in the world, surpassed human performance (97.53%).

On the other hand, we actively explored the use of deep learning for face and person recognition, and developed the DeepID-series models [R1] (DeepID1, DeepID2, DeepID3, etc.) for face recognition and the DeepReID [R2] model for person recognition. These are among the earliest use of deep neural networks for face and person recognition, which achieved a nearly perfect accuracy (99.53%) on LFW. The DeepID and DeepReID series have been the state-of-the-art of face and person recognition techniques and are the foundation of many real-world face and person recognition systems.

Object Detection

Our lab pioneered the research of deep learning-based object detection with several impactful works. In 2016, our lab proposed the DeepID-Net system [R5], which is capable of robustly detecting more than 200 classes of objects despite from complex videos. It won the championship in multiple detection tasks of ImageNet Challenge 2016 and 2015 (one of the important challenges in computer vision).

In recent years, we actively advance the object detection paradigm to simultaneously detecting and parsing the images in pixel level. To fulfill these goals, we developed the Hybrid Task Cascade

(HTC), which won the first place in the MS-COCO challenge in 2018, the most prestigious competition in the detection and parsing community.

Action Recognition

MMLab actively explored deep learning in video analytics. We developed Temporal Segment Networks (TSN) [R6] and UntrimmedNet, which is a very efficient deep neural network for video analytics and allows very deep networks to be applied to long videos in a single GPU. With these techniques, we won the first place in the untrimmed classification task of ActivityNet competition in 2016, the most prestigious contest in the video analytics.

In our recent work, we further developed Structured Segment Networks (SSN) based on TSN for activity detection. We won the first place in the activity detection task, on the ActivityNet 2016 benchmark. Both TSN and SSN have been widely adopted in practical systems.

Image Super-resolution

In 2014, we addressed the challenging problem of image super-resolution through a novel deep neural, which demonstrates state-of-the-art restoration quality, and achieves fast speed for practical usage. Our work is widely acknowledged as the pioneering study that shows the potential of deep learning on low-level vision problems. Our paper [R4] was selected as the ‘Most Popular Article’ by IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI) from March 2016 to August 2016 and remains a top-10 popular article to date. Our method won the NTIRE 2017 Challenge, the world’s largest competition in image super-resolution, and has led to a widespread adoption of deep learning for solving the image super-resolution problem.

(3) References to the research (indicative maximum of 6 references)

- [1] Yi Sun, Xiaogang Wang, and Xiaoou Tang. "Hybrid Deep Learning for Face Verification," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 38 No. 10, pp. 1997-2009, 2016.
- [2] Wei Li, Rui Zhao, Tong Xiao and Xiaogang Wang. "DeepReID: Deep filter pairing neural network for person re-identification," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2014.
- [3] Chaochao Lu and Xiaoou Tang. "Surpassing human-level face verification performance on LFW with GaussianFace," Twenty-ninth AAAI conference on artificial intelligence (AAAI). 2015 (AAAI 2015 Outstanding Student Paper Award).
- [4] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. "Image super-resolution using deep convolutional networks," IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), Vol. 38, No. 2, pp. 295-307, 2016.
- [5] Wanli Ouyang, Xingyu Zeng, and Xiaogang Wang et al. "DeepID-Net: Object Detection with Deformable Part Based Convolutional Neural Networks," IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), Vol. 39, No. 7, pp. 1320-1334 2017.
- [6] Limin Wang, Yuanjun Xiong, Zhe Wang, Yu Qiao, Dahua Lin et al. " Temporal Segment Networks for Action Recognition in Videos," IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), Vol. 41, No.11, pp.2740-2755, 2018.

(4) Details of the impact (indicative maximum 750 words)

The research work by MMLab of CUHK has facilitated the launch of various products of SenseTime Group Limited [1]. The algorithms by MMLab cover almost all SenseTime’s realms of sales. According to public media coverage, SenseTime is the world’s most valuable artificial-

intelligence startup, with valuation breaching US\$7.5 billion in 2019, and has raised US\$2.6 billion in total.

AI for Smartphones

Smartphones have become the most widely adopted computational devices. The major functionality “face unlock” aims at using users’ faces to unlock their phones rapidly, accurately and safely. SenseTime’s face verification product, SenseID, was based on the algorithm by MMLab. It was adopted by Qualcomm’s Snapdragon 845 chipset to enable face unlock on Android smartphones [2] and has been sold for millions globally.

Smartphones also become the most popular devices for taking photos. The world’s leading smartphone manufacturers (e.g., OPPO, VIVO, etc.) adopted SenseTime’s mobile imaging and AR software, which are based on various algorithms developed by MMLab. SenseAR is SenseTime’s Augment Reality engine adopted by OPPO [3]. App developers can develop augment reality-based apps based on the OPPO ARKit, which can be ran on all OPPO smartphones. OPPO has sold more than 78 million smartphones globally in 2018 [4].

SensePhoto, also backed up by algorithms by MMLab, provides functionalities of dual-camera photo background bokeh, photo super-resolution and 3D face reconstruction. The dual-camera photo background bokeh functionality was widely adopted by all VIVO’s 2019 on-sale models (including the NEX series, X series, S series and Y series) [5]. VIVO has sold more than 102 million smartphones in 2018 [4]. The SensePhoto super-resolution functionality is adopted by VIVO’s X23, X27, X27 Pro models [6], which were sold for more than 5 million units in the first half of 2019. SensePhoto’s 3D face reconstruction functionality help changing the lighting effects of persons’ faces and was adopted by the VIVO NEX [5].

AI for Autonomous Driving and Driver Assistance

The SenseDrive system, powered by algorithms from MMLab, aims to achieve autonomous driving in the near future, has been adopted by many top-ranking automakers in the world, including Honda Motor Co., Shanghai Automotive Industry Co. (a top automaker in China), and Zhengzhou Yutong Bus Co. (a world’s top bus manufacturer). Specially, SenseTime has established a five-year strategic agreement with Honda to deploy SenseDrive to Honda’s tens of thousands vehicles to achieve L4 autonomous driving vehicles [7]. It is planned for mass production in 2026 and will have major impact to the autonomous driving industry. SenseTime is also cooperating with China’s domestic emerging electrical-vehicle automakers, including WM Motor Co., Hozon Motor Co., etc., to develop a new advanced driving assistance system (ADAS).

The SenseDrive system also helps monitoring drivers’ behaviors to enhance driving safety. Shanghai International Automobile City Group has adopted the SenseDrive system for real-time monitoring its bus drivers’ fatigue levels and detecting distraction, which substantially enhances driving safety on the road [8].

AI for Web-scale Video Analysis

SenseMedia is SenseTime’s AI-assisted platform for web-scale video analysis, tagging, and search. It has been adopted by many entertainment and multimedia providers, including Alibaba Cloud, Suning, Hisense, Shanghai Star Shine, etc. The world’s largest transportation hub, the new Beijing Daxing International Airport, has adopted SenseTime’s Smart Passenger Security Check System powered by SenseMedia and SenseID [9]. It automatically conducts super-fast person-ID-ticket

verification to make sure passengers carry eligible identity documents and flight tickets for departure. This system is expected to facilitate 72 million passengers yearly by 2025.

SenseKitchen, also powered by SenseMedia, is specifically adapted to the surveillance of restaurant kitchens in China to ensure food safety and hygiene, where cooks are required to wear hats, facial masks and chef clothes. The system is able to automatically recognize if all persons in kitchens follow the wearing regulations. It has been adopted by the governments of Shanghai and Jiangsu province of China, monitoring hundreds of kitchens every day to enhance food safety.

AI for Healthcare & Diagnosis

MMLab's algorithms on medical image analysis, have been successfully integrated into SenseCare - the AI-aided diagnosis platform developed by SenseTime, which provides assistance to doctors and helps them improve the efficiency and accuracy of medical diagnosis.

SenseCare has been adopted by many top-ranking hospitals in China, including Shanghai No. 1 People's Hospital, Shanghai No. 9 People's Hospital, Shanghai Ruijin Hospital, and the very first independent accredited pathology diagnosis center in China, HISTO Pathology Inc. [10]. HISTO has integrated the SenseCare platform into their pathology diagnosis system, which has been adopted by more than 73 hospitals, benefitting more than 5000 patients and 100 doctors in China.

(5) Sources to corroborate the impact (indicative maximum of 10 references)

[1] Attached supported letter 1 to certify the use of MMLab algorithms in SenseTime products

[2] Media coverage on Qualcomm adopting SenseTime's SenseID product:
<https://beebom.com/qualcomm-multiple-face-unlock-sensetime/>

[3] Media coverage on OPPO adopting SenseTime's SenseAR engine:
<https://yicai.com/news/oppo-and-sensetime-jointly-build-an-ar-developer-platform>

[4] Global smartphone sale research: <https://www.counterpointresearch.com/global-smartphone-share/>

[5] VIVO NEX smartphone using SenseTime's SensePhoto solution:
<https://www.sensetime.com/en/news/view/id/75.html>

[6] VIVO NEX and X23 using SenseTime's image super-resolution solution:
<https://www.sensetime.com/en/news/view/id/73.html>

[7] Honda's official announcement on 5-year strategic agreement with SenseTime:
<https://hondanews.eu/en/fi/corporate/media/pressreleases/124192/honda-and-china-based-sensetime-to-pursue-joint-research-and-development-in-ai-technologies-for-auto>

[8] SAICG using SenseTime's SenseDrive to improve traffic safety:
<https://www.sensetime.com/en/news/view/id/110.html>

[9] Beijing Daxing Airport adopting SenseTime's Person Verification system:
<https://www.sensetime.com/en/news/view/id/126.html>

[10] Attached support letter 2 to certify the use of SenseTime's SenseCare system in HISTO Pathology Inc.