PAIR

Pervasive Artificial Intelligence Research Labs

2019 Annual Report

國立交通大學 National Chiao Tung University

MOST 科技部 Ministry of Science and Technology

Contents

Dr. Yu-Chee Tseng
Director

The PAIR Labs (Pervasive AI Research Labs) was founded in January 2018. Its goal is to develop the core AI technologies and tools and further apply these new technologies to all types of intelligent services. We envision target applications that are closely related to domestic industries, such as finance, management, healthcare, manufacturing, communication networks, and agriculture.

As its name suggests, PAIR Labs target AI services in a pervasive way from all aspects of hardware and software approaches. The labs now grow from 17 teams to 21 teams across 10 universities around Taiwan. We have seen significant progress among our participating teams in areas such as smart agriculture, smart carriers, AI-enabled 5G services, and AI-applied education. Our teams also publish significant results in several AI core conferences and journals. During 2019, with the joining of several members, we have extended our coverage to the areas of affective computing, AI safety, and FinTech. Social issues of AI have also received attention from our new members. I am very pleased to see that our members have extended their work to 5G communications and demonstrated several new applications that involve both 5G and AI, which have got lots of press attention recently.

Looking toward the year 2020, I like to point out several directions that we will work on further.

The first driving force is a capstone project to encourage multi-team collaboration. A new capstone project in the area of traffic safety has been kicked off since September 2019. Field trials shall be done for proof-of-concept purposes. Second, some cross-team workshops will be held, in topics such as mobile carrier and drone platform. Through regular workshops, we hope to facilitate the sharing of new development and knowledge and build external collaborations. Third, international collaboration remains a high priority work. Over the past two years, we have signed several MOUs with major international partners. We have held workshops with these partners. We will continue these efforts and increase personnel exchanges. Last but not least, we have set 2020 as the year of industrial collaboration for PAIR Labs. Making industry impact is a top priority for 2020.

We thank our funding source MOST. I look forward to seeing the continuous impact made by our members in academic excellence, international collaboration, and the AI industry. With PAIR Labs, let us work together, go together, and eventually succeed together. At the end of 2019, I wish you a healthy and very fruitful year of 2020.

Dr. Ted Kuo
Deputy Director

As we are completing the second year of our mission, I like to congratulate all PAIR members for your dedication and efforts to make 2019 yet another fruitful one. Together we continue to win on many fronts that we set our mind to.

In our 2018 year-end review, we set the year 2019 to be the Year for International Outreach and to collaborate with at least ten (10) foreign research institutes at the center-to-center level. Well, thanks to all of you, we did it! For each partner, we either host a joint workshop here in Taiwan or on the partner-hosted premises to further explore topics for joint research. Thanks to our quality work and academic publications, which continue to be highly regarded by the research communities. Also, we hosted the first AI international grant challenge with IEEE MMSP 2019. The results were so well received that we got invited to host the IEEE International Conference on Multimedia and Expo (ICME) 2020, which has much broader participants. Our members have participated in domestic and international competitions and won twice the medals and honors than those of 2018. All these activities are gradually putting PAIR on the map of the global research communities. It's a good start for all of us, but the journey continues.

We built the first 5G Testbed in Taiwan with a local mobile operator. We successfully demonstrated multimodal AI and autonomous carriers that take advantage of 5G uRLLC, eMBB, and mMTC features to complete various surveillance and rescue missions in October. This work leads us to form a larger scale AI+5G V2X based Smart City Capstone project to address unique Taiwan traffic safety issues.

In 2020, our priority is to innovate and collaborate with industrial partners through joint industrial projects like the Smart City Capstone one. As most of our teams across the half-time marks of their projects, it's time for us to apply and adjust our technologies to meet the industrial needs. Working with the industry to solve their problems seems to be a mutually beneficial approach. To that end, in September, we have hired a new CEO, Dr. Po-Chi Hu, who comes from the industry to help the teams with industrial engagement.

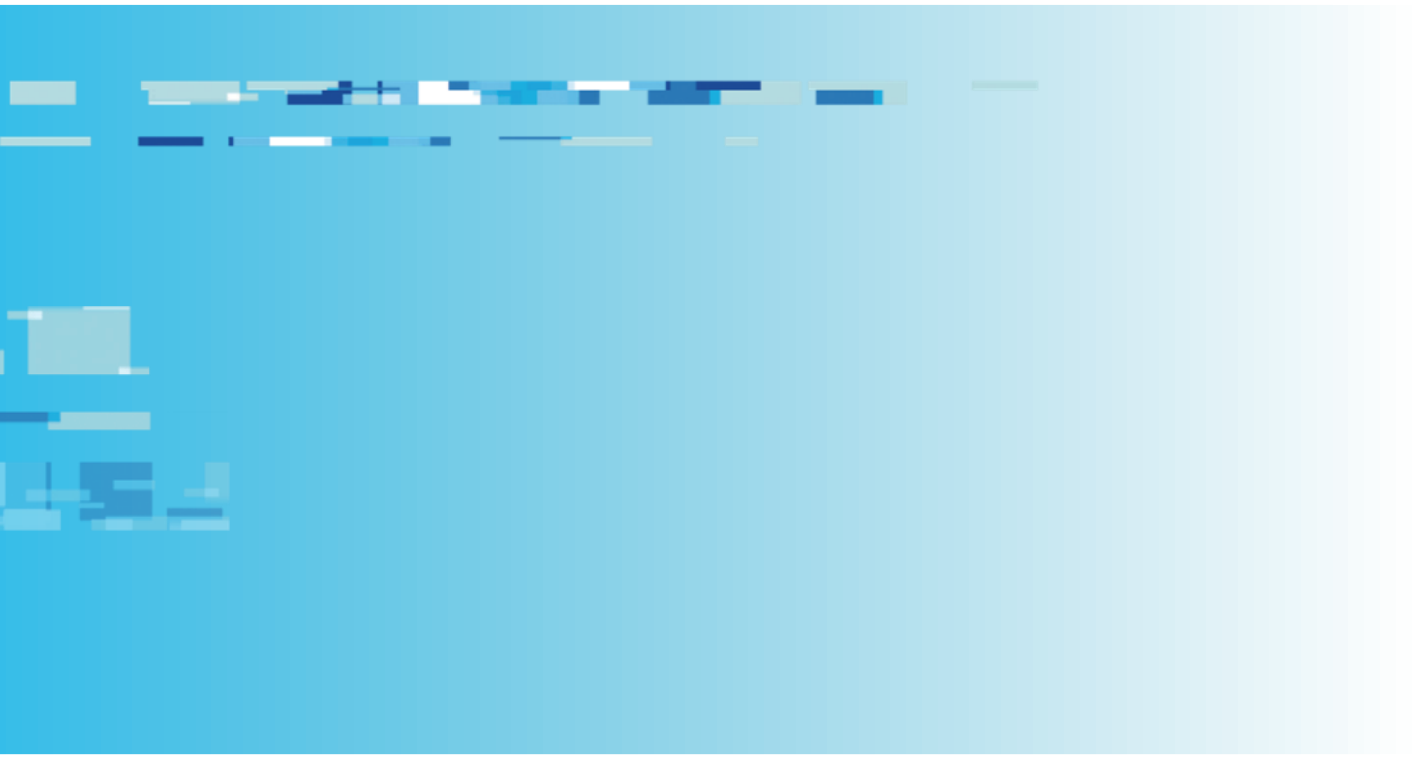
By now, I am sure all members are fully aware that every year, though, we designate a theme as the priority of the year, it does not mean we only focus on this one aspect of work. Together, I have no doubt PAIR will soon be one of the recognized AI research centers in the global research community. Keep up the excellent work!

Dr. Po-Chi Hu
CEO

In the year 2015, when I was serving in a well-known ODM manufacturer as a senior member of the strategy center, we did intensive studies on the Internet of Things (IoT) trying to find great opportunities in this hot topic of that time. We realized that IoT will bring in tremendous amounts of data and big data analysis inevitably will become crucial to the business. After we dig further, we found that Artificial Intelligence (AI) had emerged as an important technology to utilize data and successfully classified patterns that would be impossible by traditional programming approach. Soon later we predicted that the next decade will be the era of AI. A few months after, Amazon launched its famous Echo product and Alexa AI assistant. One year later, Deepmind's AlphaGo beat professional Go player 4-1, surprising the whole world.

Today, AI is everywhere. Google applied AI technologies to filter emails and provide smart reply suggestions. Facebook using AI proactively detects posts with self-harm thinking. Uber deployed an AI program to analyze if a rider has been drinking. Nevertheless, we have not fully utilized AI power to improve our daily life. Many areas could be benefited by current AI technologies but still waiting for it to be happening. This is why this Pervasive Artificial Intelligence Research (PAIR) labs was funded, to apply AI technologies in a pervasive way, to all areas of our daily lives.

The PAIR lab takes a two-track promotion approach. One is aimed at actual business applications, innovating or optimizing current service models or product designs by adopting AI technologies. On the other hand, by participating in international competitions and exhibitions, we guide the professional research teams to concentrate on cutting- edge technologies. Through the PAIR lab, we encourage industry and academia collaborations to develop practical solutions for real scenarios and generate commercial values with daily life applications. Let the AI technologies being everywhere, and in the end, it will lead the country to an AI revolution and leap forward to a highly productive and competitive new era.



The PAIR lab has outlined the following development strategies:

1. Technology breakthrough, focusing on research capability, market demands, building test fields, talents build-up and engineering supports

2. Commercialization, building up a patent portfolio, extensive technology transfer and licensing, start-up incubation and providing value-added services

3. International cooperation, through value promotion, impacts, mutual visits and joint researches.

4. Social responsibility, promoting right AI knowledge and concepts, discussion of social issues, and applying AI to solve problems.

There are many ongoing hot social topics in a discussion already. For example, how to make AI safe and explainable? On 23 Sep. 2019, Ministry of Science and Technology (MOST) of Taiwan established "Guidelines for the Development of Artificial Intelligence Research" to Improve the development environment of AI research, in order to reduce the potential doubts and risks of AI technology research and development, eliminate AI bias, discrimination, and exclusion. While we develop and extend the applications of AI technologies, we have to make sure that these AI technologies are properly used for the benefit of all mankind. The PAIR lab is not only an AI lab for research and technology commercialization, but also an organization to make people understand AI and ensure that it will meet social expectations.

Intelligent Communications and Networking Technologies for Drone-Cells

Principal Investigator

Prof. Li-Chun Wang

Summary

Our project proposes an "Artificial Intelligence (AI) Drone-Cruiser" base station for helping 5G mobile communication systems rapidly recover the network after the disaster and handle the instant traffic of the flash crowd. The drone-cruiser base station can overcome the communications problem for three types of flash crowds, such as in stadiums, parades, and large plaza so that an appropriate number of aerial base stations can be precisely installed to meet the huge and dynamical traffic requirements. Artificial intelligence can solve such problems by analyzing the collected data, and then adjust the system parameters to achieve the goals of self-configuration, self-optimization, and self-healing under the framework of Self-Organizing Network (SON). With the help of AI technologies, the 5G network can become even more intelligent. This project is to provide a novel service, On-Demand Aerial Base Station as a Service, including the following three technical challenges: (1) rapid air-to-ground 3D wireless channel learning technology; (2) optimal 3d placement for aerial base stations; (3) big data analysis and AI technologies for the automatic UAV management; and (4) innovational design for the long-time hovering drone-cruiser. It is hoped that the outcome of this project can help open up another emerging opportunity for Flying Base Station (Flying Access Point) in the post-5G era for Taiwan's information and communications industry, following the success story of Wi-Fi in Taiwan.

Keywords

Aerial Base Station, Drone-Cruiser, Artificial Intelligence, Self-Organizing Network, 3D Placement, Flying Access Point.

Innovations

- We design and develop a new type of UAV, Drone-Cruiser [1], to provide base station service with dynamic mobility for a long hovering time.
- We analyze the characteristics of the 3D air-to-ground wireless channel and design a new learning mechanism for UAV-BS to provide stable communication services.
- We propose a multi-agent deep reinforcement learning network for the crowd estimation and prediction on the UAV-BS.
- With the consideration of the above estimation information, we proposed some fast 3d placement algorithms of UAV-BSs for guaranteeing the allocated data rate [2] of each served user and improving the system sumrate [3].
- We integrate the latest LTE, WiGig, and mmWave technologies on the UAV-BS for guaranteeing the backhaul connections while providing network services.
- We propose a predictive on-demand placement of UAV-BSs using Echo State Network [4] for providing seamless communication services.
- We proposed a machine learning based interference management for multiple UAV-BSs [5].
- To guarantee the privacy protection on Internet of Drones, we propose two novel approaches using network coding and recurrent neural network (RNN)-assisted network coding [6][7].

Benefits

- The new type of UAV, Drone-Cruiser, can be one of an important product for the UAV-assisted communications industry.
- We collect the real 3D wireless channel data and construction a 3D+radio dataset/map to help the prediction of communication link quality.
- We develop a simulation platform to recommend the appropriate 3d placement configuration for improving the performance of UAV-assisted communications systems.
- We propose a data-driven UAV base station placement for serving arbitrarily distributed crowds and the proposed method can effectively improve the system sum-rate by almost 100%.

Fig. 1
The target scenario of the UAV-assisted cellular communications.

Fig. 2
The overview of the research architecture.

Fig. 3
3D radio map labeling.

Fig. 1

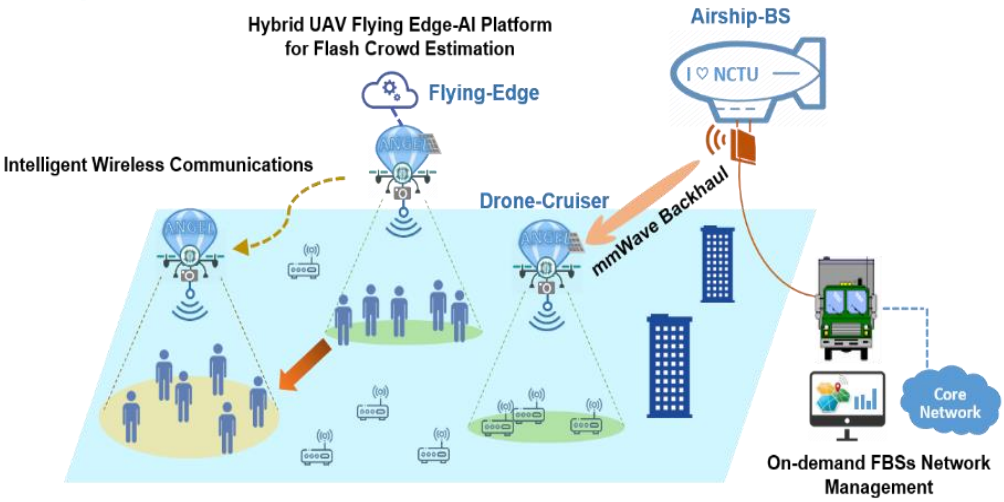


Fig. 2

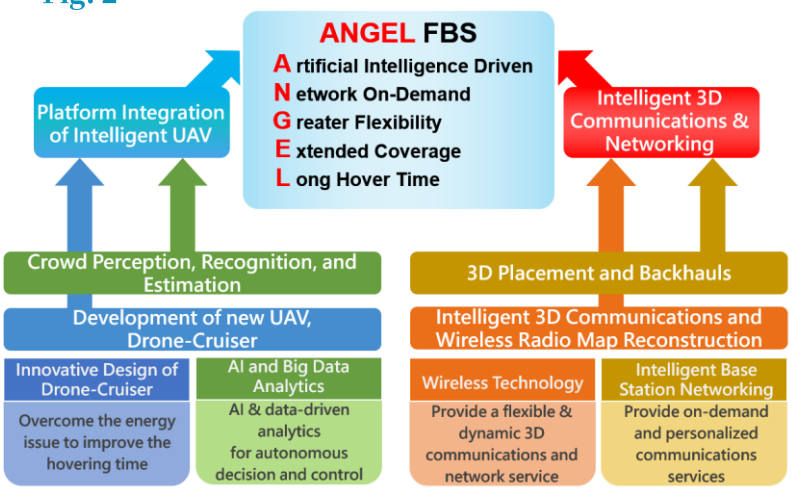
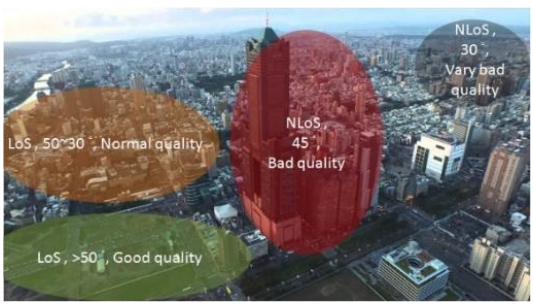


Fig. 3



Selected Publication

1. Li-Chun Wang, Chuan-Chi Lai, Hong-Han Shuai, Hsin-Piao Lin, Chi-Yu Li, Teng-Hu Cheng, and Chiun-Hsun Chen. "Communications and Networking Technologies for Intelligent Drone Cruisers." To appear in *IEEE Globecom Wkshps*, Waikoloa, HI, USA, 2019.
2. Chuan-Chi Lai, Chun-Ting Chen, and Li-Chun Wang. "On-Demand Density-Aware UAV Base Station 3D Placement for Arbitrarily Distributed Users with Guaranteed Data Rates." *IEEE Wireless Communications Letters*, vol. 8, no. 3, pp. 913-916, Jun. 2019.
3. Chuan-Chi Lai, Li-Chun Wang, and Zhu Han, "Data-Driven 3D Placement of UAV Base Stations for Arbitrarily Distributed Crowds." To appear in *IEEE Globecom*, Waikoloa, HI, USA, 2019.
4. Haoran Peng, Chao Chen, Chuan-Chi Lai, Li-Chun Wang, and Zhu Han. "A Predictive On-Demand Placement of UAV Base Stations Using Echo State Network." In *IEEE/CIC ICCC*, Changchun, China, 2019.
5. Li-Chun Wang, Yung-Sheng Chao, Shao-Hung Cheng, and Zhu Han, "An Integrated Affinity Propagation and Machine Learning Approach for Interference Management in Drone Base Stations," to appear in *IEEE Transactions on Cognitive Communications and Networking*.
6. Yu-Jia Chen and Li-Chun Wang, "Privacy Protection for Internet of Drones: A Network Coding Approach," in *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 1719-1730, Apr. 2019.
7. Chen-Hung Liao, Hong-Han Shuai, and Li-Chun Wang, "RNN-assisted Network Coding for Secure Heterogeneous Internet of Things with Unreliable Storage," to appear in *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 7608-7622, Oct. 2019.

Video Compression based on Generative Model

Principal Investigator

Prof. Hsueh-Ming Hang

Summary

The deep learning schemes recently developed provide a new direction for constructing a high-efficiency compression image/video coding system. Our research focuses on the essential elements of a video compression system including image compression, video frame predictor, compression residual artifacts reduction, saliency map estimation, and compression-oriented multi-task training. These topics are either the backbone of an image/video compression system or they can enhance the compression efficiency.

Keywords

Image Compression, Video Compression, Depth Learning, Video Predictor, Saliency Detection, Auto-encoder (AE), Generative Adversarial Network (GAN).

Innovations

- **End-to-end image compression:** One key issue in the image compression system design is that the Latent Variables (in an autoencoder) need to have low entropy rate and can be easily coded using a small number of bits. We use the autoencoder (AE) together with Principal Component Analysis (PCA) to design the end-to-end image compression systems. AE is able to generate condensed representatives and reduce dimensionality of an input image. PCA can further compact the energy of Latent Variable (feature maps) into fewer dimensions.
- **Restoration and compression:** We develop a framework for image compression and restoration together. The framework consists of a standard image compression codec and several post-processing networks. We aim to make the network learn the prior knowledge for image reconstruction, to understand the fundamental problems that what essential information is needed for transmission, and what prior information needs to be provided by the deep convolutional neural network and dataset.
- **Frame predictor:** We propose a reinforcement learning based video compression framework. It combines the conventional motion compensation algorithm with neural networks. Unlike most of the learned based approaches relying on the black-box-like model, which is often unexplainable and requires huge computations, we train a neural network to estimate the positions of key pixels and their motion vectors for generating the predicted video frame, a key for video compression.

Benefits

- **End-to-end image compression:** In addition to PCA, our design includes a soft-bit quantizer in the training phase, which solves the vanishing gradient problem and improves the rate estimation. Our compressed images have comparable MSE with BPG and have better visual quality (MS-SSIM).
- **Restoration and compression:** A unified framework for image compression and restoration consists of a BPG codec, an artifact removal network, a super-resolution network and a colorization network as shown in Fig. 1. Based on three learned networks embedded with prior knowledge, a good quality image can be reconstructed by the given compressed partial information of an image. Given a compressed low-resolution grayscale image and very few bits of color hints, the decoder can recover high quality images, achieving comparable or even better compression efficiency than the BPG codec.
- **Frame predictor:** We obtain promising results of our reinforcement learning based frame predictor. The objective metric, MSE score, is nearly equal to the best existing scheme, and its SSIM outperforms all the others (Tab. 1). The visual samples in Fig. 2 show that the predicted frames produced by our method is closest to the original frame. Also, the network structure becomes much simpler with few convolutional layers because the neural networks only estimate the key pixels and the motion vectors.

Fig. 1

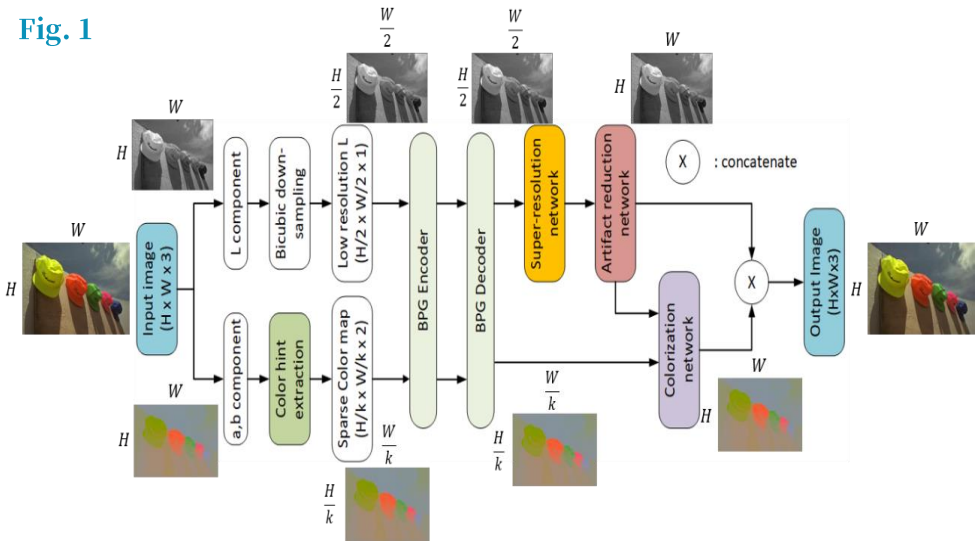


Fig. 1
The overall architecture of image compression and quality enhancement.

Tab. 1
Performance comparison of various Frame Predictors on the CaltechPed dataset.

Fig. 2
Samples of Frame Predictor results.

Tab. 1

Method	MSE	SSIM	#parameters	Type
CopyLast	0.00476	0.878	-	-
MCNet	0.00250	0.879	14M	direct
DualGAN	0.00241	0.899	113M	dense
Ours	0.00252	0.923	1M	sparse

Fig. 2

Researches on Applications with Deep Reinforcement Learning

Principal Investigator

Prof. I-Chen Wu

Summary

Recently, Deep Reinforcement Learning (DRL) has been applied to many AI applications. One of the successful achievements is the AlphaZero, called the Zero method in this project, was presented to learn without human knowledge and surprisingly surpass all the human players and all the AI programs.

This project focus on three classes of DRL applications: 1) Lightweight model, e.g., Go program CGI, exact methods, and general Zero methods. 2) Complex model, e.g., AI bot of video games. 3) Real-word model, e.g., DRL applications of robotics.

Keywords

Deep Reinforcement Learning, Reinforcement Learning, Deep Learning, Monte-Carlo Tree Search, AlphaGo Zero, Computer Games, Go, Video Games, Car Racing, Robotics, Random Bin Picking.

Innovations

- We propose a novel value network, the multi-labeled value network (ML-VN). Fig. 1 shows its extension, the board evaluation multi-labelled value network (BV-ML-VN).
- We propose the multiple policy value MCTS (MPV-MCTS). As shown in Fig. 2, it combines multiple policy value neural networks of various sizes to retain advantages of each network.
- We propose an approach to strength adjustment for MCTS-based programs, and perform a theoretical analysis to prove that the adjusted policy exceeds a lower bound in strength. Fig. 3 shows the derived strength, which is close to the empirical data of Go.
- We investigate whether the Zero method can also learn theoretical values and optimal plays for non-deterministic games, and develop the 2x4 Chinese Dark Chess Zero program.
- We propose a new weighted cross entropy method (WCEM), which can achieve a success rate of 96% in grasping tasks for robotic arms, while DDPG can only achieve 70%.
- We propose a new end-to-end hybrid action space DRL method, the Parameterized Proximal Policy Optimization (P3O). Fig. 4 illustrates its network architecture. P3O can greatly improve the accuracy of robotics grasping and pushing tasks.

Benefits

- We develop the world's first Go Zero program that can play under different komis by using BV-ML-VN. The proposed method has also been published on IEEE Transactions on Games.
- The proposed MPV-MCTS has been accepted by the top conference IJCAI-19 (acceptance rate is only 850/4,752 = 17.8%).
- The proposed MCTS-based strength adjustment method has been accepted by the top conference AAAI-19. (acceptance rate is only 1,150/7,095 = 16.2%).
- The computer Go lifelong learning system we developed is the world's first Go system that is able to provide different strengths from beginners to super-humans. This achievement is selected for the show of the 2018 Future Tech.
- The 2x4 Chinese Dark Chess Zero program we developed is the first Zero program for stochastic games in the world. A paper for this also won the best paper award in TAAI 2018.
- We propose distributed end-to-end DRL algorithms for training AI bots, which has been successfully applied to industrial-university joint projects, and also outperform top human players.
- The paper of P3O has been accepted by the Infer2Control workshop at NeurIPS 2018.
- We proposed a learning rate decay method, the hyperbolic-tangent decay (HTD), which has been accepted by IEEE WACV 2019.

Fig. 1

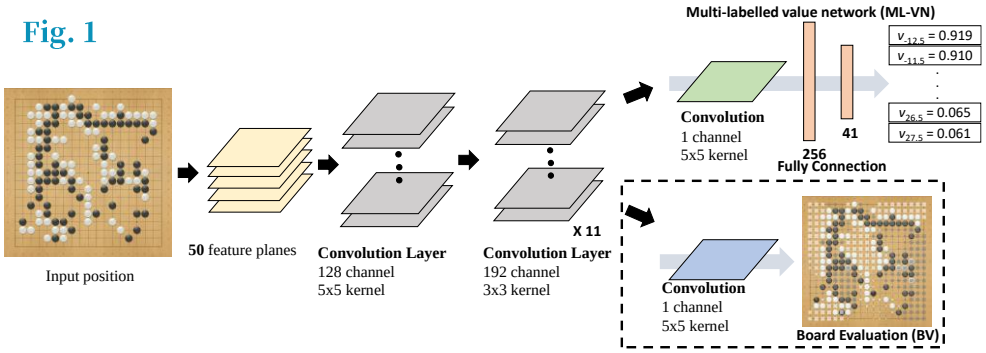


Fig. 1
The network architecture of BV-ML-VN, which is an extension of ML-VN. There are two outputs. One outputs the multi-labeled value, and the other outputs the board evaluation which indicates the ownership of each point for the given position.

Fig. 2

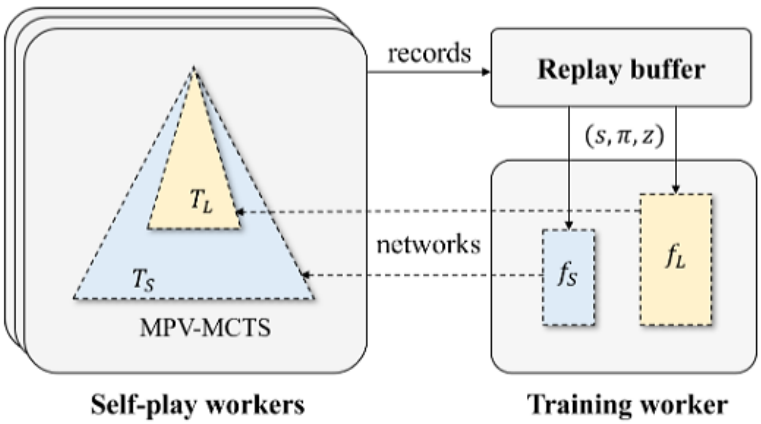


Fig. 2
MPV-MCTS combines smaller network f_S and larger network f_L. The smaller network with faster inference provides efficient exploration, while the larger network provides more accurate evaluation of key points.

Fig. 3

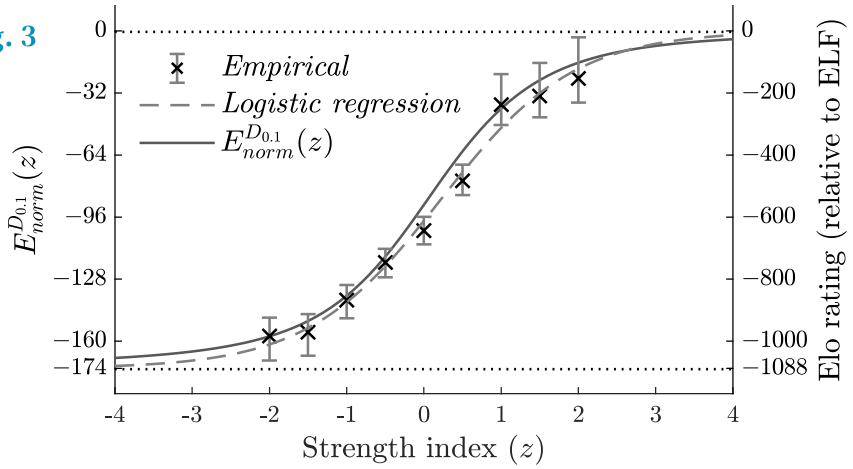


Fig. 3
The curve of $E^{D_{0.1}}_{norm}(z)$ and the empirical data for the strength adjustment of Go programs. The derived strength is also close to the empirical strength with regression error 40.45 Elo, and to a logistic function with regression error 10.51 Elo.

Fig. 4

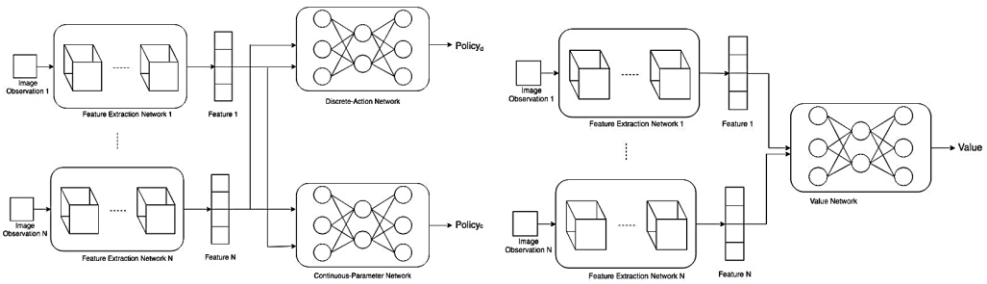


Fig. 4
The network architectures of P3O. The actor (left) consists of two policy networks, the discrete policy network, and the continuous policy network. Both policy networks take multiple images as the input. The critic (right) is a value network which takes multiple images as inputs and predicts the value of the state.

Development of AI Platform for Smart Drone - Intelligent Flight Control

Principal Investigator

Prof. Jen-Hui Chuang

Summary

The objective of the research is to develop an autonomous flight control (AFC) system for drones, or unmanned aerial vehicle (UAV), for innovative, real world applications such as smart agriculture and smart city (see Fig. 1). We have developed the control system over a computer simulator and physical drones, which is designed to avoid obstacles in the flying path, complete a flight mission, and land safely. To create a cost-effective training system, we built a flight simulator, which can receive somatosensory signal from the environment and react to somatosensory feedbacks of a pilot. Moreover, we have created 3D models for the simulator based on the real environments for the training via the flight control system. In addition to AFC, we also study the control for precise localization, navigation, and landing. The plan is to train the AFC system in the simulated environments, then transfer it over real drones when it is ready. We anticipate significant differences between simulated environments and the real ones, which will be one of main challenges of this research.

Keywords

UAVs, Smart Drones, Smart Sensing, Smart Control and Smart Simulation, Deep Learning (DL)

Innovations

- A new network architecture is developed for the analysis of aerial images of parking lots, to quickly identify available parking spaces with size smaller than 15×15 pixels.
- DL-based obstacle detection and depth prediction are merged into a single CNN architecture to attain real-time performance (see Fig. 2), with ground truth provided by our world leading stereo system (see Fig. 6).
- With flight control training based on cross-domains and cross-tasks learning, the training results in the simulated environment are expected to be directly applicable to real-world situations (see Fig. 3).
- The state-of-the-art end-to-end algorithm for static/dynamic obstacle avoidance and navigation using LIDAR is developed for smart sensing of drones.
- A wearable haptic device is developed to convert remote UAV's flight and sensor information to somatosensory signals to serve as flight control assistance for a pilot (see Fig. 5). Furthermore, the device can send messages triggered through gesture change to control the simulated UAV.
- A very efficient 3D modeling method is developed via the simplification of building shapes (see Fig. 4), while several sharing datasets of UAV are built and posted for easy access.

Benefits

- We have developed a real-time obstacle detection and avoidance system for drones that can effectively avoid static objects or moving pedestrians.
- We are developing a state-of-the-art drone flight simulator which will combine remote control, AirSim simulator, and VR technology to provide professional/amateur pilots a good environment to collect flight data without worrying about power consumption, equipment damage, etc.
- We have developed a prototype of the Remote Tactile and Sensing Suit (RETASENS), which allows the pilot to perceive the flight attitude and sensor signals from the drone, and to send control signals to a UAV.
- The sharing datasets can help researchers of UAV's techniques to accelerate the process of research and development.
- The simplified 3D building modeling method can effectively describe a complicated building model with minimum vertices, e.g., 99% less, in simulated scenario.
- We have published two international journal papers and 27 international conference papers (including several top conference papers, e.g., those in CVPR).

Fig. 1

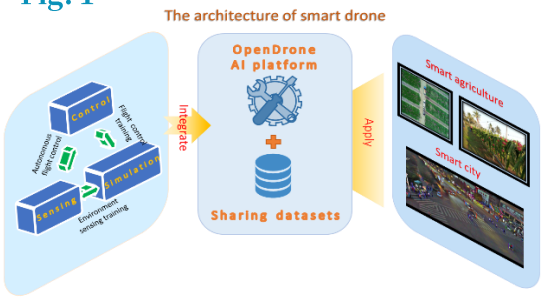


Fig. 2

Fig. 3

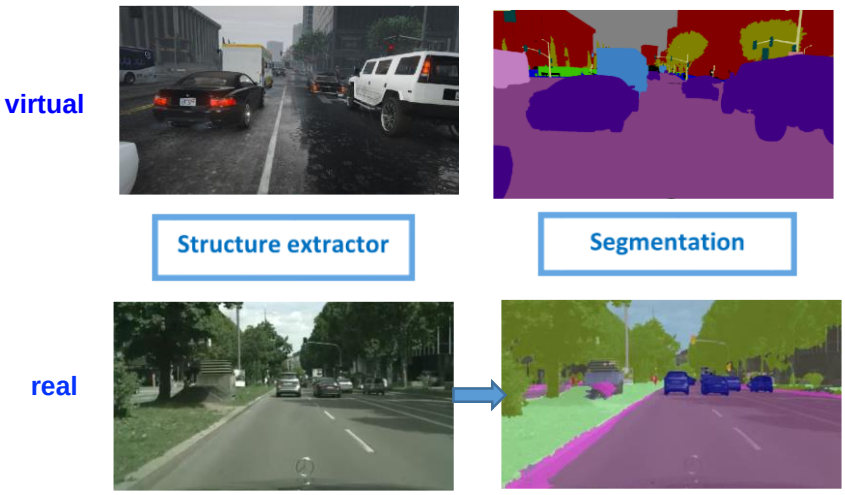


Fig. 4

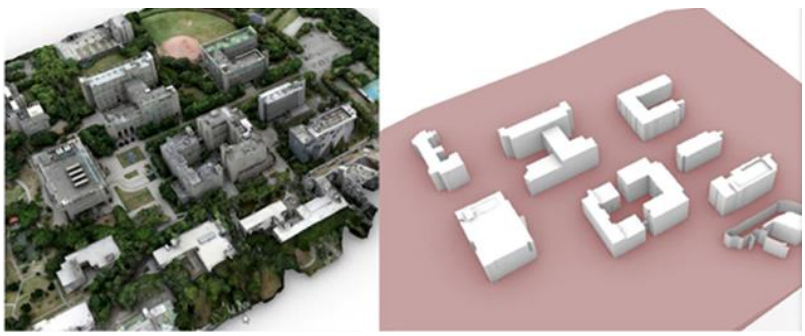


Fig. 5

Fig. 6

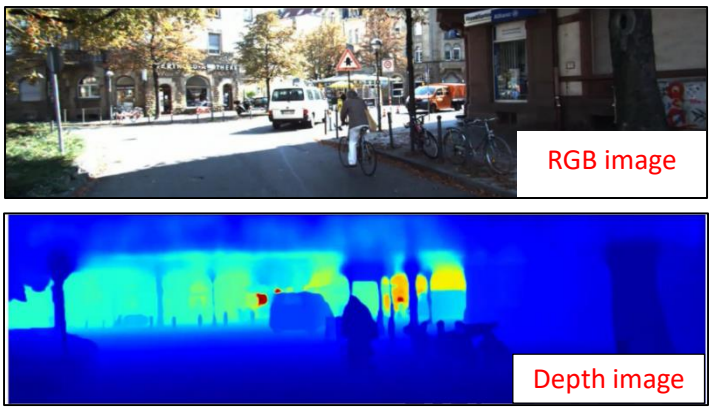


Fig. 1
Developing AI Platform for Smart Drones.

Fig. 2
Real-time obstacle detection and avoidance (ICIP 2019).

Fig. 3
Adapting structural info for boosting cross-domain semantic segmentation (CVPR 2019).

Fig. 4
Simplifying 3D building models.

Fig. 5
The RETASENS suit.

Fig. 6
A depth estimator -- ranked 1st on KITTI leaderboard for 4 months (CVPR 2018).

Embedded AI Deep Learning Technology for ADAS/Niche Self-Driving Applications

Principal Investigator

Prof. Jiun-In Guo

Summary

To assist Taiwan's automotive industry to break through the barrier established by the existing ADAS patent portfolio in the world, we propose the idea of developing embedded AI deep learning technology for ADAS/niche self-driving applications so that Taiwan's industry possesses the chance to be part of the eco-system of ADAS/Self-driving products in the world. Our research focuses on three areas: (1) Automatic data labeling tool for deep learning applications and labeled ADAS/Self-driving datasets; (2) Embedded deep learning object detection/behavior analysis algorithm and model development; (3) Real-time deep learning computing platform/ASIC/SoC development. Applications of our research include autonomous driving systems with level-1/2 functions like lane keeping, automatic emergency braking (AEB), and automatic lane changing and level-3/4 functions like high automation driving systems in city/urban/highway environments. In addition, we are also developing related technology required in certain niche self-driving applications, e.g., non-tracking AGV highly demanded in the Industry 4.0, to satisfy the needs of local industry.

Keywords

Embedded deep learning technology, fast data labeling, object detection and behavior recognition, ADAS, autonomous driving vehicle, industry 4.0

Innovations

- We have developed a world first fast labeling tool, ezLabel 2.3 (shown in Fig. 1), to speed up 10-15x labeling efficiency as compared to the existing manual video labeling tools, which won two prizes in 2018 AUDI Innovation Award Taiwan. We support the bounding box labeling, pixel-based segmentation labeling, as well as behavior labeling of moving objects.
- We have established a dataset with 20M+ of samples for deep learning object detection and behavior analysis for ADAS/Self-driving applications (as shown in Fig. 2). Out of the collected 20M+ samples, we have opened and shared 102K samples of them, most of which are related to vehicles, pedestrians, and cyclists under various weather conditions.
- We have developed two embedded SSD models (NCTU SSD lite 512x512/512x256) that is suitable for using the TI TDA2X/iCatch V37 chipsets for real-time object recognition. At 30fps, we obtained 72% mAP@Pascal VoC.
- We have developed a latest real-time embedded deep learning model called Multi-task Semantic Attention Model (MTSAN) that can support both object detection and scene segmentation at the same time that detect vehicles as far as 100-150m, which outperforms the Yolo v3 model (max. 100m) and Mobilenet+SSD model (max. 50m) and owns 4.5% higher in mAP as compared to Mobilenet+SSD model (as shown in Fig. 3).
- We have developed a deep learning technology based on 3D convolutional neural network to predict whether rear vehicles will overtake in the next three seconds with a 95% or more accuracy (as shown in Fig. 4). It is able to be applied to front camera applications to detect pedestrian crossing, vehicle/motorcycle cut-in, and emergency breaking and achieve over 90% accuracy.
- We have developed a Taiwan first smart self-driving wheelchair for indoor mobility in places like airports and hospitals (as shown in Fig. 5).
- We have developed a world first bit accurate dynamic fixed-point deep learning model training and inferencing tool (ezQUANT) for CNN hardware accelerators to ensure less than 2% mAP quality drop for most deep learning models.
- We have developed a hybrid fixed-point/binary CNN model training flow (ezHybrid-M) for CNN hardware accelerators (as shown in Fig. 6) to ensure no quality drop in Mobilenet +SSD model.

Benefits

- Have won two prizes in 2018 AUDI Innovation Award Taiwan.
- Have won an outstanding prize in AISlander 2018 with ezLabel and embedded AI technology.
- Have won a Technology Breakthrough Award in MOST 2018 Future Tech.
- Have won a Bronze award in 2018 MXIC design contest.
- Have conducted 18 cases of industrial collaboration projects with total grant about NT$14,580K (2018) and have conducted 13 cases of industrial collaboration project with total grant about NT$14,960K (2019).

Fig. 1
Fast labeling tool, ezLabel 2.3.

Fig. 2
Profiling of 20M labeled datasets for ADAS.

Fig. 3
The embedded MTSAN technology.

Fig. 4
The rear overtake warning system with object detection and behavior recognition.

Fig. 5
Self-driving electric wheelchair.

Fig. 6
Hybrid fixed point/binary CNN model training flow (ezHybrid-M).

Fig. 1

Fig. 2

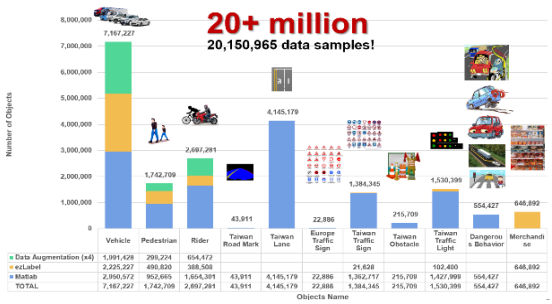


Fig. 3

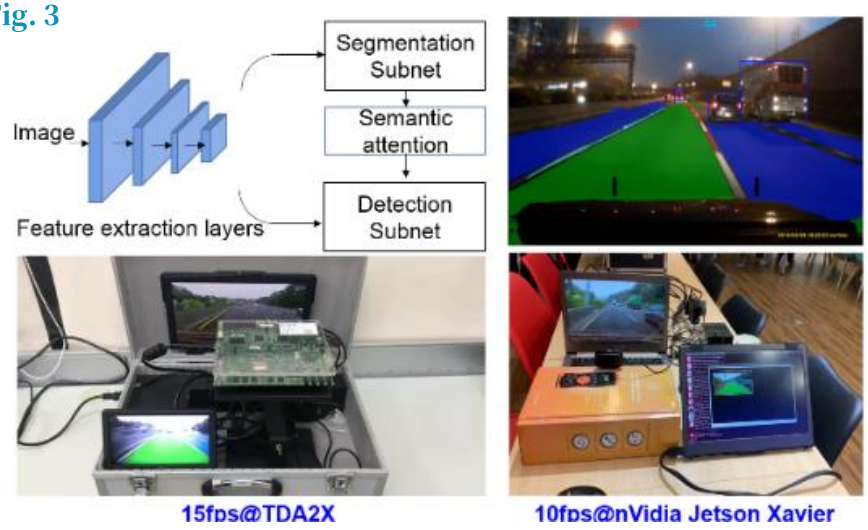


Fig. 4

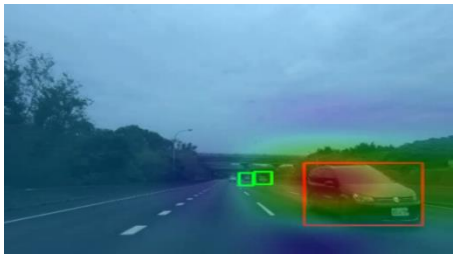


Fig. 5

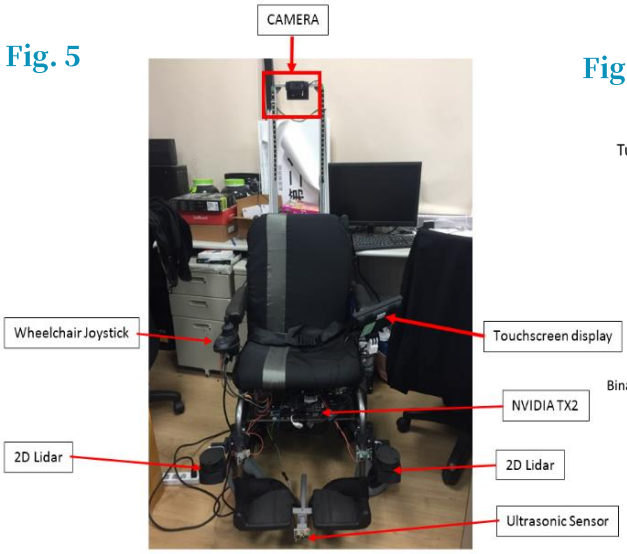
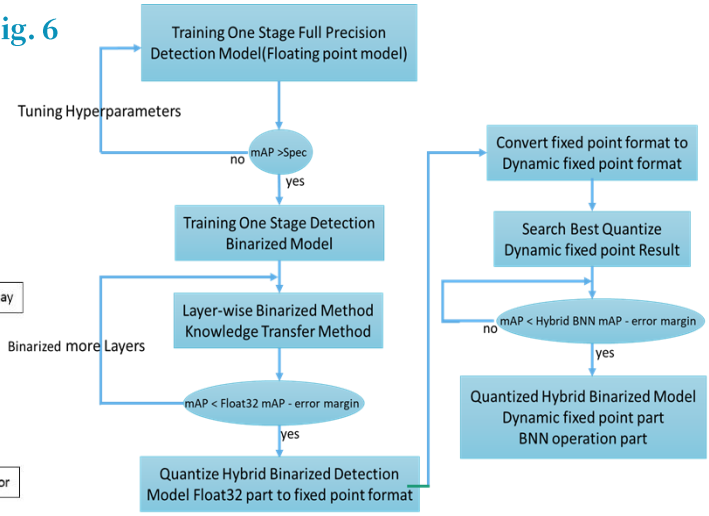


Fig. 6



https://sites.google.com/site/yc0405/

On Self-Maneuvered Patrolling Robots with Artificial Intelligence and Multi-Sensory Data Fusion Technology

Principal Investigator

Prof. Yu-Chee Tseng

PAIR

Summary

In this project, we focus on the development of AI and multi-sensory data fusion technology. Based on the developed approaches, we can improve machine (i.e., robot) perception of environment and increase the intelligence of robots. The core research topic is how to utilize AI to carry out a wide range of sensory data analysis, and then extract high-level information to enhance the efficiency of work as well as provide advanced robot-human services. The proposed system is able to recognize people even without capturing human biological features.

Keywords

Image Recognition, Depth Learning, Data Fusion, Wearable Computing, Robot, Drone, Person Identification (PID).

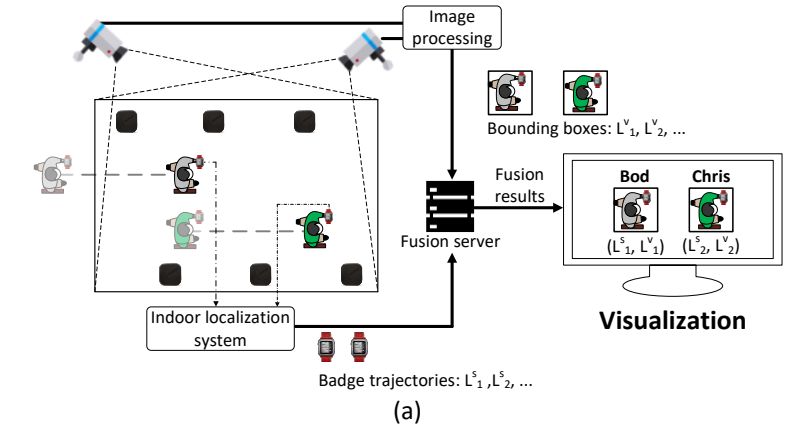
Innovations

- We develop several fusion systems to integrate AI and multi-sensory information. The fusion systems are able to recognize people and visualize their profiles. The systems could be applied on surveillance, robot, vehicle, drone, etc. The following are developed applications:
    - [Supervision] Fusion of surveillance camera (RGB) and location data: We design a pairing mechanism to couple human objects with their IDs by fusing trajectory data from cameras and the BLE beacon-based positioning system. And we develop a PID system to tag user profiles on video in real time. Fig. 1 (a) and (b) shows the architecture and PID results of Supervision.
    - [Third Eye] Fusion of depth camera (RGB-D) and inertial sensor data: We develop a PID system with a RGB-D camera and wearable devices. The system is implemented in a robot to track a specific person. Fig. 1 (c) shows the tracking robot we are working on with Prof. Wayne Wang (NTNU).
- [V2X Eye] Fusion of driving camera and V2X: We extend the fusion system on V2X environment. In order to improve vehicle perception, profiles of cars around you could be displayed on the driving window. Fig. 2 shows the enhanced driving window by V2X Eye. (2019 IEEE APNOMS Best Paper Award)
- [SkyEye] Fusion of drone camera and wearables: We use a drone to identify ground human objects, and develop a PID system to tag wearable IoT information on drone videos in real time. Fig. 3 shows our conceptual result after tagging IoT data. To the best of our knowledge, this is the first work integrating IoT data and computer vision from a drone camera.
- [Night vision] Fusion of non-RGB vision and wearables: We extend the PID system by using no RGB information in low-light conditions. Our system fuses the user motion-related features retrieved from wearable inertial sensors and LiDAR point clouds. It helps visualize personal profiles easily in LiDAR scans as shown in Fig. 4.

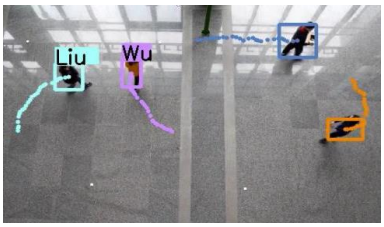
Benefits

We fuse AI technologies and multi-sensory computing to develop data fusion systems. Several applications have been developed, including: [Supervision] achieves PID on long-distance cameras for surveillance; [Third Eye] achieves PID and tracking robots; [V2X Eye] helps understand the driving environment; [SkyEye] realizes PID in the drone view; [Night vision] identifies persons without biological features in dark conditions. Other on-going Researches: Integrating commercial out-of-the-shelf devices into our data fusion system, including ROS robot, smart watch, etc. The system can improve machine perception in a smart environment and increase the intelligence of robots.

Fig. 1



(a)



(b)

(c)

Fig. 2

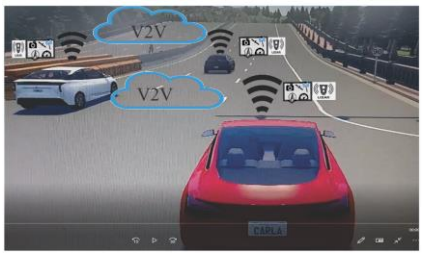


Fig. 3

Fig. 4

Fig. 1
(a) Supervision architecture, (b) Supervision result, and (c) Tracking robot.

Fig. 2
Enhanced driving window by V2X Eye.

Fig. 3
Tagging wearable IoT information in the drone view by SkyEye.

Fig. 4
Identifying person without RGB cameras by Night vision.

http://fox1.csie.ncu.edu.tw/Kcfan-website/indexcn.html

A Deep Learning-Based Gesture Interface and Value-Added Location Services

Principal Investigator

Prof. Kuo-Chin Fan

Summary

The contactless human computer interaction systems including stationary and mobile interfaces are proposed in this project. We develop three major components providing different styles of usage manners so that easy and joyful ways of interacting with various applications can be achieved. Users can operate the system without touching via hand gestures, virtual keyboard, or in-air handwriting. They can also quickly acquire location-related information by taking pictures of textual signs on streets. In addition, image categorization and social network forecasting schemes are devised, such as obtaining well-classified site-specific images automatically and predicting future social networks by taking the dynamic nature of it into account. All these systems are implemented using deep learning technologies.

Keywords

Contactless Human Computer Interaction, Location Semantics, Virtual Keyboard, In-air Input, Dynamic Social Network

Innovations

- We design contactless human computer interaction systems including a stationary interface, a mobile interface, and an image categorization and social network prediction system, as shown in Fig. 1.
- We train several deep neural networks for gesture recognition and apply them in a context-aware manner to enhance the recognition rate and provide a user-friendly interactive experience. Fig. 2 demonstrates some examples.
- An easy way for multi-language input is proposed via the virtual keyboard, which enables users to operate in air.
- Store-related information can be automatically displayed on mobile devices equipped with cameras by taking a picture on the texts of store signs. We make this scenario possible by text recognizing in complicated street view images. See Fig. 3 for an illustration.
- We propose an approach to classify unlabeled images related to a specific site and arrange them into pre-defined categories, providing users easy glances at well-classified photos and filtering unrelated photos of the site.
- To input characters with smart glasses, we design a robust finger detection mechanism (40 fps and 8.3 pixel error in average) and achieve good recognition performance via ego-view in-air handwriting. Fig. 4 shows some recognition results.

Benefits

- We design the gesture recognition modules by combining 3DCNN mechanism with LSTM, which extract both spatial and temporal features at the same time and obtain more satisfactory recognizing results.
- The text recognition technique for street view images developed in this project is robust (text detection performance achieves 87.7% in hmean) for store/traffic signs with both horizontal and vertical texts.
- We obtain great performance for finger detection by combining real and synthesized finger images in the training phase.
- In addition to developing unlabeled image categorizing tools, we also implement a Web interface. By entering one specific location URL of Google map into the Web interface, well-classified images of that particular location will be displayed.
- Through the auto-encoding scheme, we effectively compress the social network, reduce the computational cost of LSTM and thus do more accurate prediction.

Fig. 1

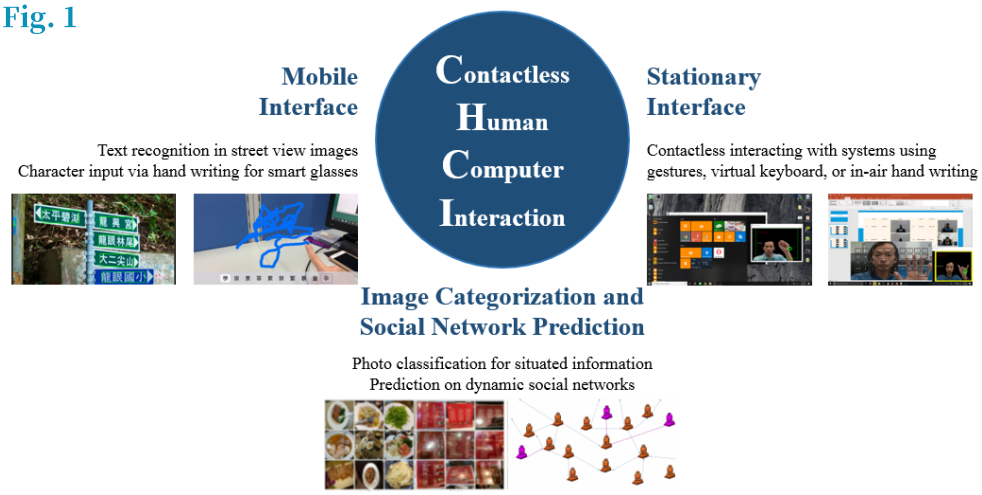


Fig. 2

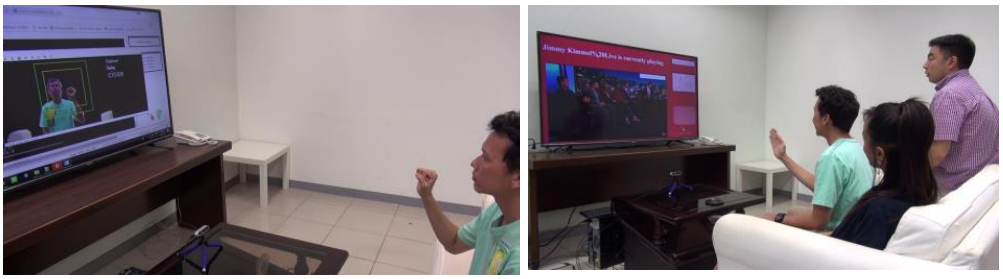


Fig. 3

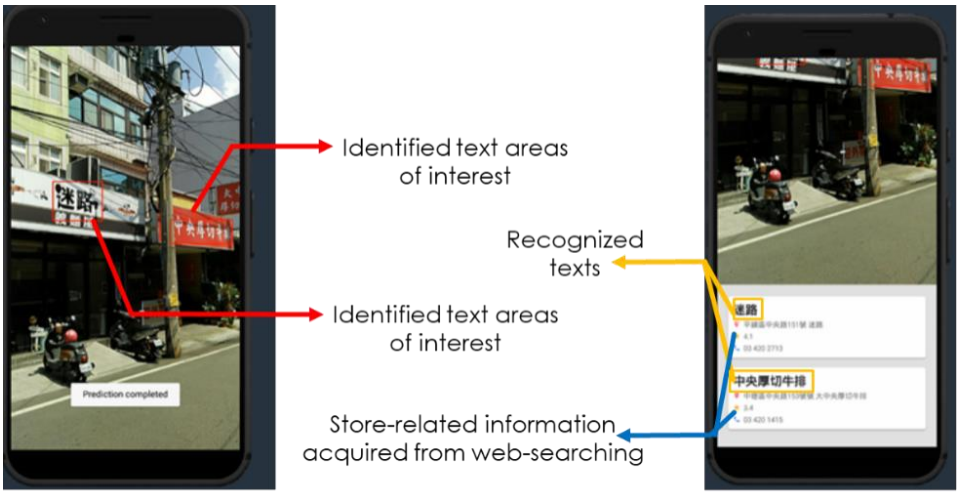


Fig. 4

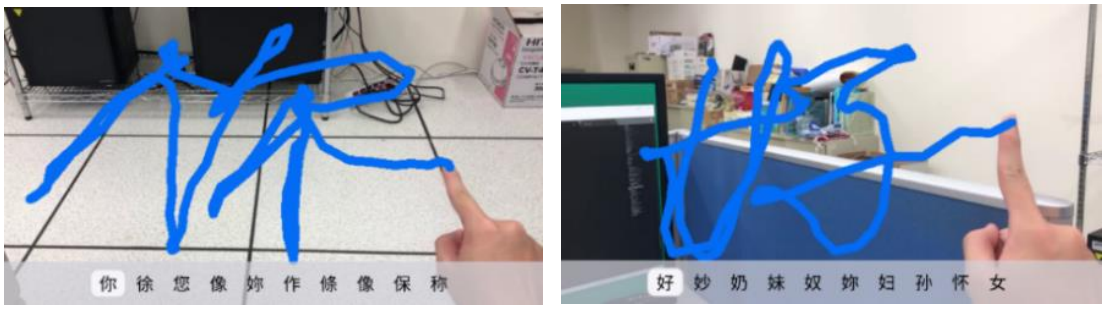


Fig. 1
The project develops three major components: a stationary interface, a mobile interface, and an image categorization and social network prediction system.

Fig. 2
Samples of training (left) and using gestures to interact with applications (right).

Fig. 3
Text recognition of street view images.

Fig. 4
Character input via in-air handwriting for smart glasses.

An Intelligent Partner for Visually Impaired People

Principal Investigator

Prof. Wen-June Wang

Summary

We are developing an integrated system to facilitate visually impaired people in daily activities. The system will ensure that users have both safe outdoor activities and convenient indoor actions. Using Deep Learning technologies, our team is designing and implementing a navigation device, whether it is wearable or carried via a robot, to help visually impaired persons to travel on the road, recognize around environment and communicate with the device.

Keywords

Deep learning, Machine Learning, Robots, Wearable Devices, Semantic Segmentation, Ground Truth Generator, Visually Impaired Supporting System, Human-Machine Dialog, Indoor Living Support

Innovations

- A guiding robot is implemented (see Fig.1) to assist visually impaired people for outdoor navigation.
- Deep learning technology is employed to find obstacles and their distances.
- Guiding robot can travel on the right side of the road, avoid obstacles and arrive the destination.
- Street Logos are recognized by the robot eyes. Suggestions based on recognition results further assists users to find new location and door of stores.
- A new semantic segmentation tool can help the recognition system to automatically collect a few millions of street photos for neural network training (Fig. 2).
- A cloud-based dialog navigation agent (CDNA) system is designed and will be integrated with the robot. Through deep learning-based natural language dialog, the robot can further assist the user. The developed system is illustrated in Fig. 3.
- Via the deep-learning-based object recognition modules, the indoor supporting system will issue an alarm signal whenever the setups at home are changed and assist the user to find personal belongings.
- Based on different deep learning modules, the system guides the visually impaired through the signs indoor (Fig. 4), and detects the dangerous signs as well as obstacles while they are walking (Fig. 5).
- A wearable device for helping visually impaired people (see the right one in Fig. 1).

Benefits

- The robot hardware and its neural network system are developed and tested on streets. Preliminary results show that the navigation system can guide blind people to travel in a university campus.
- The robot eye system can recognition 14 types of street logos. Preliminary result shows a 91.3% of accuracy with a minimal but sufficient set of training data.
- The ground truth generator had already collected up to 8 million street photos in 6 major cities in Taiwan. Our target is to collect 10 million around cities in Taiwan. The marked ground truth includes at least 100 thousands of street photos with 14 type of logos.
- The human-machine dialog system is developed and tested by using information near a university campus. The dialogue is effective in assisting the visually impaired people.
- Indoor objects and setup can be precisely identified and located. The design of the smart living supporting system is proven to be feasible.
- A wearable device prototype for helping blind people.

Fig. 1

Fig. 2

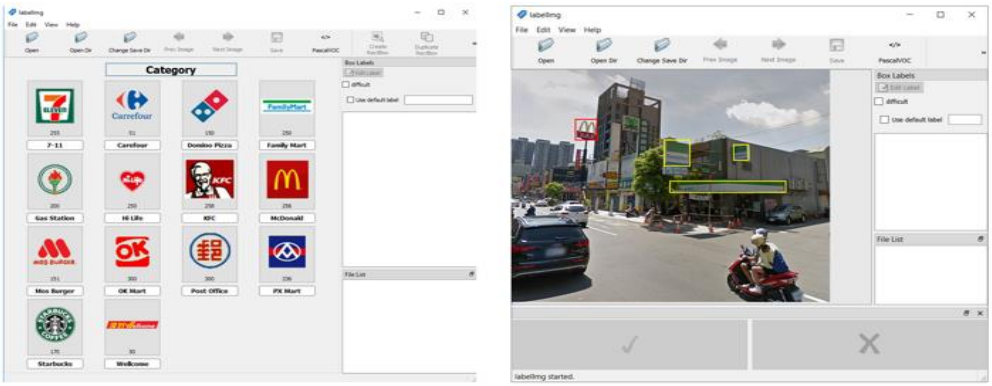


Fig. 3

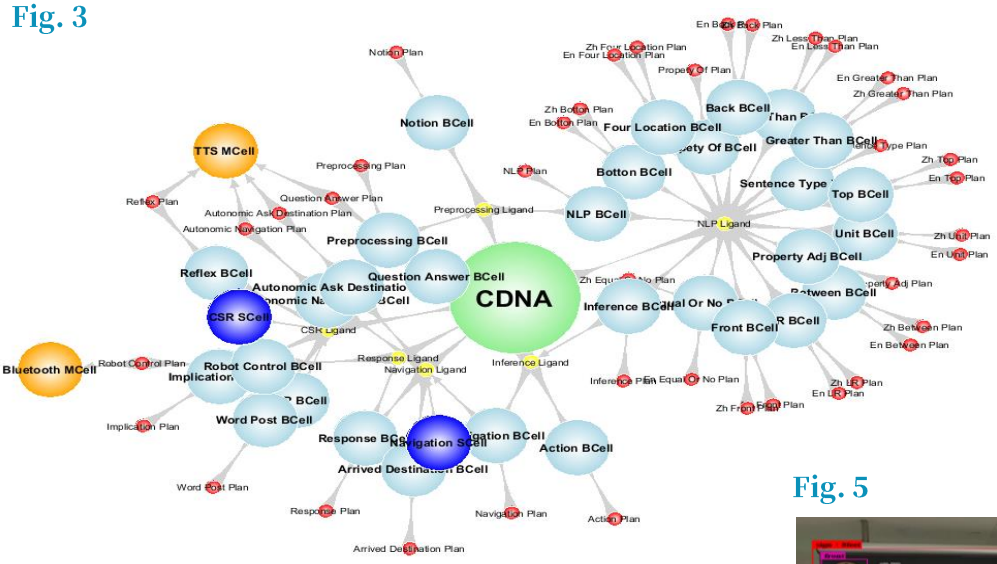


Fig. 4

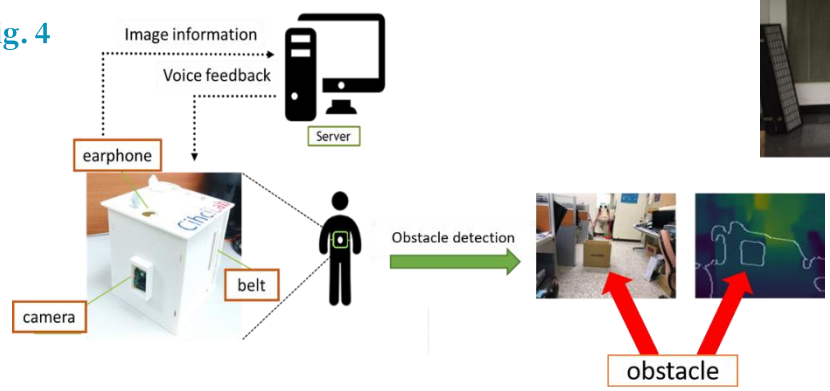


Fig. 5

Fig. 1
The Guiding Robot and Its Usage Scenario, Wearable Device.

Fig. 2
An Automatic Street Logo Recognition, Collection and Editing Tool.

Fig. 3
The Cloud-based Human-Machine Dialog System.

Fig. 4
Obstacle Detection Process.

Fig. 5
The detected indoor signs.

Development of Artificial Intelligence-Based Brain Computer Interface

Principal Investigator

Prof. Kuo-Kai Shyu

Summary

Human brain is the most complicated system in the world. The intricate interconnection of neuron groups enables the construction of various brain functions. In our project, we aim to build EEG databases and apply artificial intelligence (AI) to clinical, healthcare, and educational applications. With the development of wearable EEG system, EEG data under different circumstances can be collected. AI systems built on the cloud or edge computing are used to recognized users' intentions and produce feedbacks to achieve brain-machine interaction, called AI-based brain computer interface (BCI). Our current achievements can be summarized into three categories. For clinical applications, we have built a stroke EEG database and a post-stroke epilepsy EEG database. Convolutional neural network (CNN) is used to evaluate stroke patients' rehabilitation progress and their incidence rate to incur seizure attack. For homecare applications, an AI-based BCI system, an AI emotion detection system and vital-sign wearable devices are developed. The AI-based BCI is currently designed using CNN, for severely paralyzed patients to help them communicate with external environments. The AI emotion detection system is designed using three-layer Long-Short Term Memory (LSTM) network which is used to identify users' emotional states, such as anxiety, angry, happy, etc. The vital-sign wearable devices collect subjects' PPG and ECG data which enables the construction of big data for hypertension, diabetes, vascular sclerosis evaluations. For educational applications, an instant biofeedback system (IBS) is designed. Students' cognitive loads and emotional statues are detected based on instantaneous EEG analyses. Students learning states are feedback to tutors in classroom environment and guide teachers to help students achieve better learning performances.

Keywords

Artificial intelligence, Closed-loop brain computer interface, and electroencephalography (EEG)

Innovations

(I) Clinical Applications:

(A) Stroke Rehabilitation: (1). Subacute stroke patients with mild impairment. (2). Movement-related brain oscillations database (3). Event-related synchronization (ERS) at both ipsilesional and contralesional hemispheres negatively correlated with motor recovery 3 month of stroke ($r=-0.66\sim-0.78$, $p<0.005$) (as shown in Fig. 1).

(B) Post-stroke Epilepsy: (1). After-stroke EEG database construction (2). Develop and implement a feature-based fusion model using various machine learning approaches (3). 21-channel EEG measurement (4). ~75% accuracy for epilepsy detection (as shown in Fig. 2).

(II) Healthcare Applications:

(A) Brain Computer Interface (BCI) Prosthesis: (1). Motor imagery EEG database (2). Implementation of a Dry-Electrode EEG BCI system using Convolutional Neural Network (as shown in Fig. 3). (3). Eight-channel EEG system (4). ~80% Accuracy for imagery movements.

(B) Emotion Detection: (1). Emotion EEG database (2). Real-time emotion recognition in EEG data using Long-Short Term Memory network (as shown in Fig. 4a and 4b). (3). ~78% dominance detection, ~82% valence detection, ~81% arousal detection.

(III) Educational Application:

(1). Emotion EEG Detection- Real-time emotion recognition (2). Cognitive load Detection- Real-time cognitive load recognition (as shown in Fig. 5). (3). Self-efficacy Detection- Real-time self-efficacy recognition

Note: consistent data revealed that negative emotions may generate extraneous cognitive load—and then impact learning performance and self-efficacy.

(IV) Technical Core:

Hardware achievements: (1). High-sensitivity EEG dry electrode with high-performance impedance matching circuit (as shown in Fig. 6a). (2). Multi-channel wireless EEG with 4, 8, 20, and 40 channels. (3). Wearable

vital-sign monitoring watch for real-time monitoring of heart-rate, heart-rate variability, cuffless blood pressure, exercise ECG. (4). 3D posture recorder using eleven inertial measurement units (IMU) with 100 Hz digitization rate (as shown in Fig. 6b)

Benefits

(I) Clinical Applications: (A) Stroke Rehabilitation: Movement-related brain oscillations provide an objective biomarker as a potential predictor of longitudinal motor recovery and also therapeutic target of neuromodulation. (B) Post-stroke Epilepsy: Using AI techniques to develop a EEG-based epilepsy detection system for enhancing doctors' clinical diagnosis of patients after a stroke.

(II) Healthcare Applications: (A) Brain Computer Interface (BCI) Prosthesis: Using AI to improve the performance of EEG-based BCI will allow paralyzed patients (lock-in syndrome) to achieve communication with external environments. (B) Emotion Detection: Using AI to detect user's emotion statuses can help people to manage their mental stress and emotions.

(III) Educational Application: (A) To evaluate cognitive load/emotion/self-efficacy students, physicians, or pilots perceptive in real-time. (B) Teachers/Instructor make modifications in their instruction or adjustments to learner's work content based on information from EEG prediction.

Fig. 1
Stroke Rehabilitation.

Fig. 2
Post-stroke Epilepsy.

Fig. 3
Brain Computer Interface (BCI) Prosthesis.

Fig. 4
Emotion Detection.

Fig. 5
Education Application.

Fig. 6
Technical Core.

Fig. 1

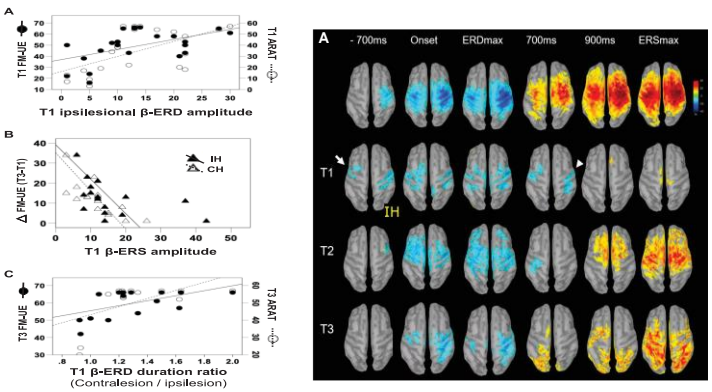


Fig. 2

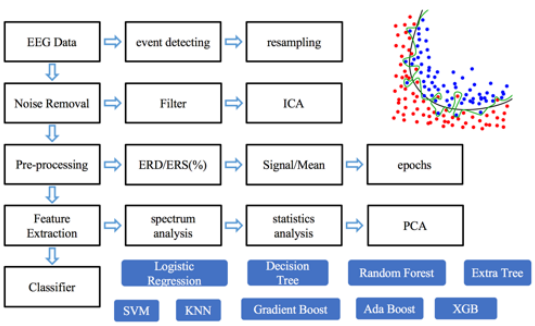


Fig. 3

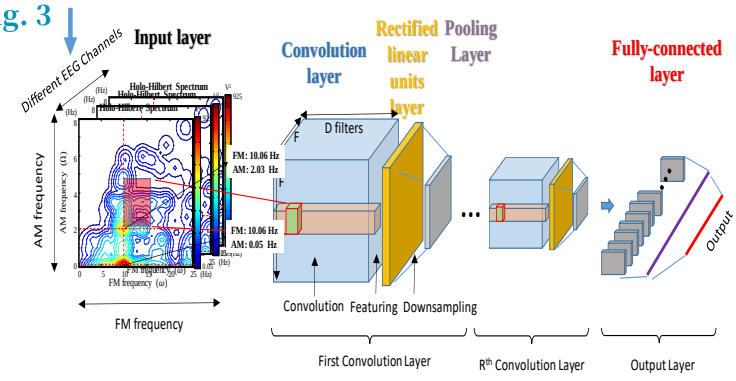


Fig. 4

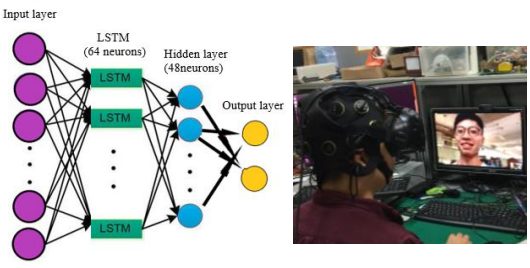


Fig. 5

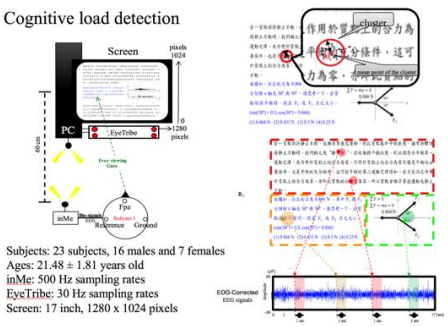


Fig. 6

Deep Intelligence Based Spoken Language Processing

Principal Investigator

Prof. Jia-Ching Wang

Summary

Speech is not only the most natural means of communications among people, but also the most effective means of human-computer interaction. Enabling a computer to process spoken language like a human is a great problem, which scholars have been trying to solve for decades. Deep learning puts this goal into practice. To such an end, this project develops a plan for processing spoken language that integrates speech processing, acoustic signal processing, natural language processing, and deep learning techniques. Key techniques such as intelligent multi-channel speech processing and speech separation, robust speech recognition, spoken language translation, speech emotion recognition, and open field dialogue will be developed. Native languages and dialects such as Minnan are particularly addressed.

Keywords

Spoken language processing, speech separation, speech recognition, spoken language translation, speech emotion recognition, dialogue system, deep learning.

Innovations

- In front-end processing, we propose a deep learning-based multi-channel speech enhancement algorithm, which integrates the beamforming technique and the deep neural networks (Fig. 1).
- For speech separation, we present a separation method based on complex GP latent variable model. The proposed method projects the spectrum into a low-dimensional subspace.
- We propose the hierarchical extreme learning machine (HELM) for audio-visual speech enhancement as an alternative model for the speech enhancement task (Fig. 2). Besides, an HELM-based bone conducted speech enhancement method is also proposed.
- To achieve robust speech recognition (Fig. 3), we propose a novel use of the graph-regularization based methods to enhance speech features by preserving the inherent manifold structures of the magnitude modulation spectra. We also propose a semi-supervised approach to construct the acoustic model so that lower resource is required.
- For native language speech recognition, we have developed a Taiwanese Minnan speech recognizer.
- In machine translation, we present a bi-directional translator between English and Chinese (Tab. 1). The translator can handle influent input sentences.
- We propose an acoustic-based speech emotion recognition system, which uses multiple feature extraction net-work based on deep learning, as well as a new recurrent neural network that we have developed. Besides, we al-so propose a semantic-based speech emotion recognition system.
- In order to understand the semantic of the dialogue, we develop language understanding techniques for dialogue systems (Fig. 6).

Benefits

- This project develops an English learning robot based on ASUS Zenbo. Through intelligent speech recognition technology, the English learning robot enables the user to practice pronunciation and articulation within non-English environments (Figs. 5 ,6).
- The robustness techniques alleviate the undesired impact caused by environmental distortions so as to make automatic speech recognition systems maintain an acceptable performance level.
- The spoken language processing techniques that are developed herein will be used on an intelligent interactive platform that can be applied to real world applications.

Fig. 1

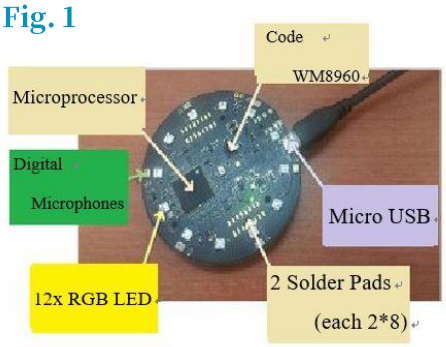


Fig. 2

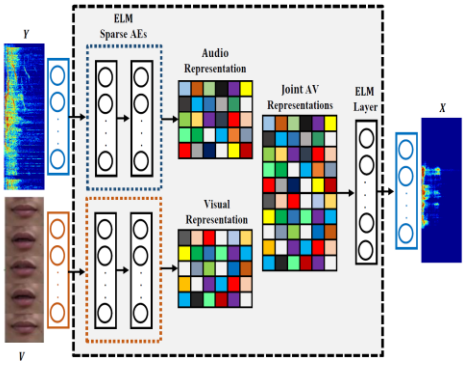
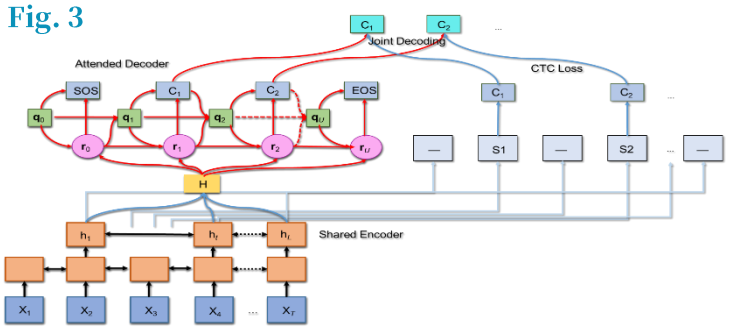


Fig. 3



Tab. 1

Source	Reference (Human)	Baseline	Proposal System
Wang Zhen was quite outstanding in the recent two races. He won a silver medal for the athletics world championship held in Beijing last year and was upgraded to the individual project champion in the world championship this May.	王鎮近兩個賽季狀態突出，去年北京田徑世錦賽收穫銀牌，今年 5 月團體競走是錦賽上又升級個人項目冠軍。	在最近的兩場比賽中，王振民獲得了去年北京奧運會競技錦標賽的金牌，今年 5 月他被提升為世界錦標賽的個人項目冠軍。	王振南在最近的兩次比賽中都取得了突出成績，去年他在北京舉辦的世界田徑錦標賽中獲得了一枚銀牌，今年 5 月他被提升為世界錦標賽中的個人項目冠軍。
So the high temperature allowance is not within the minimum wage, it must be issued in cash and not substituted by physical objects as drinks.,	因此高溫津貼不再最低工資範圍內，必須以現金形式發放，不能用飲料等實物代替。	因此，高溫津貼不屬於最低工資的範圍，必須以現金髮放，而不能以諸如飲料等物理物品代替。	因此，高溫補貼不屬於最低工資，必須以現金支付，而不能用飲料等實物取代。
People who sent a certain amount of mass fraud messages may be punished for crime of swindling even if they did not succeed in fraud.	羣發詐騙短信達到一定數量，即使為詐騙得手，也可以詐騙罪論處。	發出一定數量的詐欺信息的人可能會因為受詐騙罪而受到懲罰，即使他們在詐欺中沒有成功。	發出一定數量詐欺訊息的人，即使沒有詐欺行為，也可能因詐騙犯罪而受到懲罰。

Fig. 4

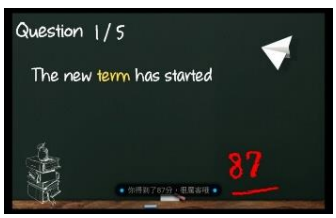


Fig. 5

Fig. 6

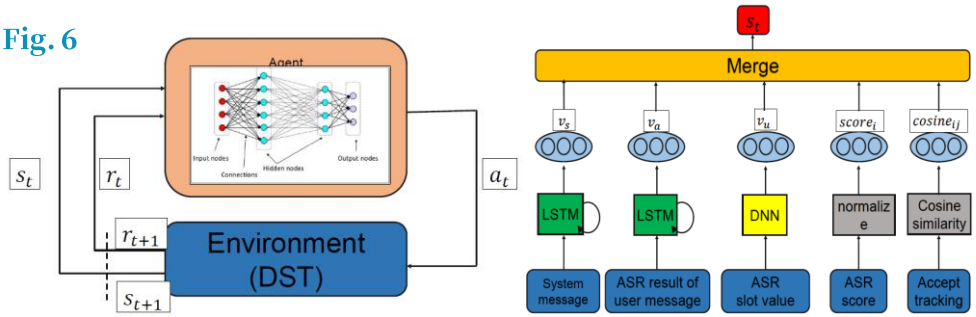


Fig. 1
The microphone array processing prototype.

Fig. 2
Audio-visual speech enhancement.

Fig. 3
The flowchart of speech recognition using connectionist-temporal-classification and attention model.

Tab. 1
Examples of English to Chinese Translation.

Fig. 4
Interface of the English learning robot.

Fig. 5
Photo of a user and the English learning robot.

Fig. 6
Deep-Q network based dialog state tracking.

http://fm.mgt.ncu.edu.tw/teacher/Jin_Huei_Yeh.htm

The Impact, Adaption and Response of Individual, Firm and Social-Economic System from the Induction and Adoption of AI

Principal Investigator

Prof. Jin-Huei Yeh

Summary

This 4-year project aims to examine the impacts from the prevailing state-of-the-art AI/digitalization and automation technology on human and organizational behavior, corporate governance, industrial supply chain and evolution, economic growth/development, social inequality and security. By borrowing the lens and implications from social materiality to comprehend the decision/procedure of an organization on its adoption of AI technology, we disentangle how organization behavior interact with a technology in practice, enact structures which shape their situated use of that technology and the changing face of jobs. Given the potential social consequences of reshaping job opportunities, the project particularly emphasizes on the social security design from the legitimate perspective. Our subproject also covers applications of AI/learning analytical tools in the fields management and social science to mitigate the social inequality that is caused by asymmetric information or unfair competition in various markets. By reviewing the global researches and learning from the experience across countries, we expect to shed new lights toward future tentative solutions or remedies for our domestic citizens, corporates and policy makers for future in a hope to mitigate the undesired downsides from a governance and management point of view.

Keywords

Artificial intelligence, organization behavior, technology adoption and diffusion, information asymmetry, industrial structure, inequality, social materiality, social security, social design

Innovations

- We conduct a comprehensive survey and study on how AI/digitalization lead to different strategic thinking; and the adoption and diffusion of AI turning into productivity growth of firms in the short run and long run.
- We disentangle how organization behavior interacts with AI technology in practice, enacts structures which shape their situated use of that technology, accommodates the field reality of human-automation collaboration. and adapts to the changing face of jobs and the social consequences of reshaping jobs.
- We emphasize the social security design from the legitimate perspective via considering the scheme of basic income/pension and how to strike a balance between the reform of system and the guaranty of social insurance.

Benefits

- We expect to provide a comprehensive and through understanding of the impacts, responses and adaptiveness of firms, industry and society as a whole from the adoption of AI/digitalization related technology.
- By surveying local and global researches, experience across countries, we expect the research outputs to shed new lights and insightful implications for future policy and education from social and management perspectives.
- Characterize the impacts of AI tech and digitalization on the evolution of organization/jobs, and employment relationships; and the types of adoption and diffusion of AI/digital technology among Taiwanese enterprises.
- Findings on US and Taiwanese industry level dynamic firm size distributions over the past 30 years suggest that (1) leading firms getting bigger in the most recent decade; (2) the adoption of AI technology enhances a firm's monopoly power that changes the competition environments and entry/exit dynamics of the industry.
- A preliminary open platform of anti-pump-and-dump stock price manipulation detection is built to mitigate the information asymmetry and protect the relatively uninformed publics from being manipulated in the capital market.

- Understanding, with cross-country comparisons, the impacts, threats and challenges of the social security system confronting to the reshape of future workforce given rise to the evolving adoption of AI/automation innovation and thinking for tentative proposal strategies/policies at institution/government level to mitigate the potential risks.

Fig. 1

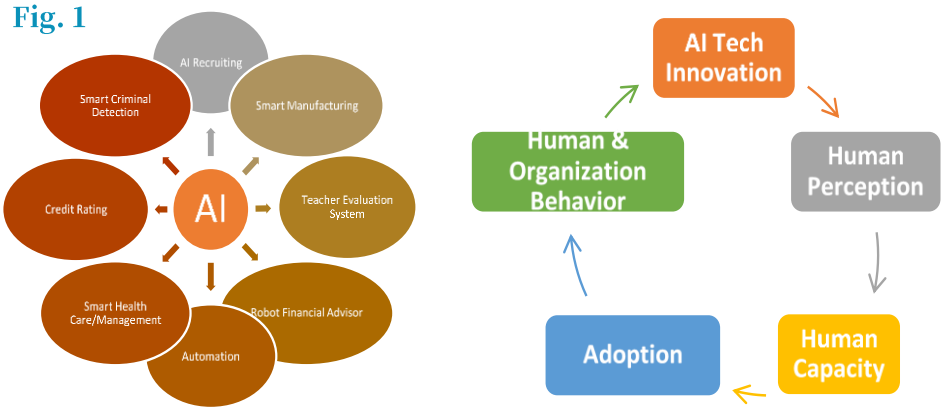


Fig. 2

Fig. 3

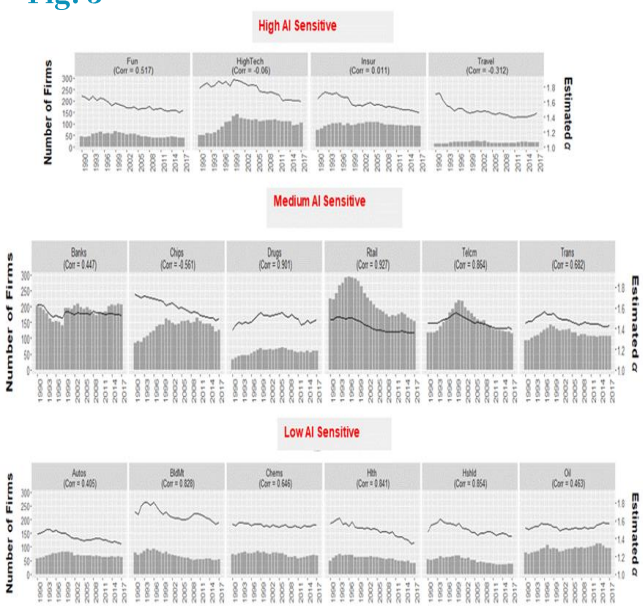


Fig. 4

Fig. 1
Sociomateriality and Affordance perspective toward Fields of AI technology application.

Fig. 2
The Framework and Implementation Ideas of the Integrated Project.

Fig. 3
Estimated Taill Index of firm size across industries over time with different AI/Digitalization sensitivity.

Fig. 4
Anti-Information Asymmetry Pump&Dump Fraud Detection Platform.

http://isdlab.ie.ntnu.edu.tw/isdwebsite/teacher-1.html

Development of Theory and Systems of Robot Learning from Human Demonstration (LfD)-Development of Learning from Human Demonstration Robot

Principal Investigator

Prof. Wei-Yen Wang

PAIR

Summary

This project proposes a learning from demonstration (LfD) system that allows robots to be not only taught by human via demonstration but also adjusted by themselves. Moreover, a multimodal perception approach is provided so that the robotic arm can perform flexible automation without complicated coding by professionals. In addition, the system is implemented on an edge AI chip to increase the overall performances and flexibility, expanding the potential of the system to be incorporated with other applications. In the future, we do not only focus on vision-based tasks, but also consider other scenarios where voice and touching are required to be learned in the LfD system.

Keywords

Learning from demonstration (LfD), 6 DoF dual arm robot, multimodal perception approach, artificial general intelligence (AGI), robotic calligraphy system, edge AI chip.

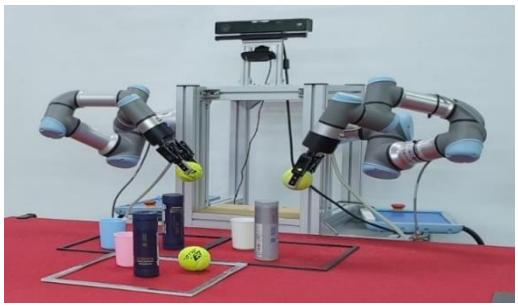
Innovations

- This project develops a multimodal perception approach to recognize face, gestures, and objects so that a dual robotic arm system is capable of performing complex tasks after human demonstration without complicated coding by professionals. This provides high feasibility for flexible automation.
- A calligraphy system learns how to write Chinese calligraphy according to a novel proposed hypothesis generation model based on cognitive psychology.
- We design a portable delta-like robot manipulator to write Chinese calligraphy. Moreover, a spring system is employed to reduce the mechanical vibration of machine.
- Our research team and a tech company jointly unveiled the world's smallest AI chip, with a dimension of 0.7 by 0.7 centimeters.

Benefits

- Using multimodal perception approach for face, gesture and object recognitions, the robotic arm can perform flexible automation stably without complicated coding.
- A dual robotic arm system is used for the proposed LfD system to learn from human demonstrations and complete tasks such as automatic pick-and-place and human-robot collaboration in both simulated and real world environment..
- A novel hypothesis generation model is proposed to emulate a human cognition system, solving tasks using previous learning experiences.
- This project implements the proposed intelligent algorithms on an edge AI chip to increase the overall performances and flexibilities, expanding the potential of the system to be incorporated with other applications.

Fig. 1



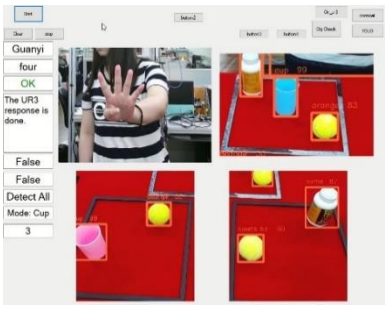
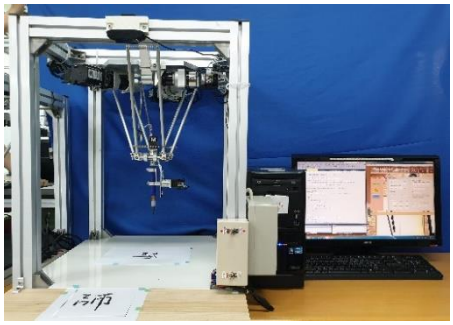


Fig. 2



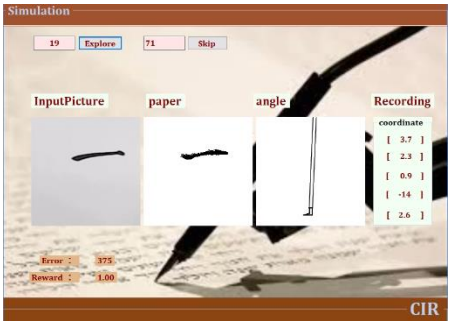


Fig. 3

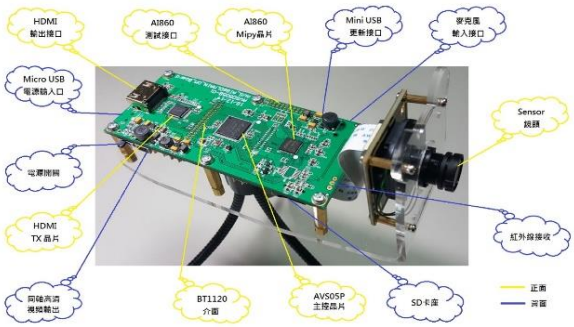


Fig. 4

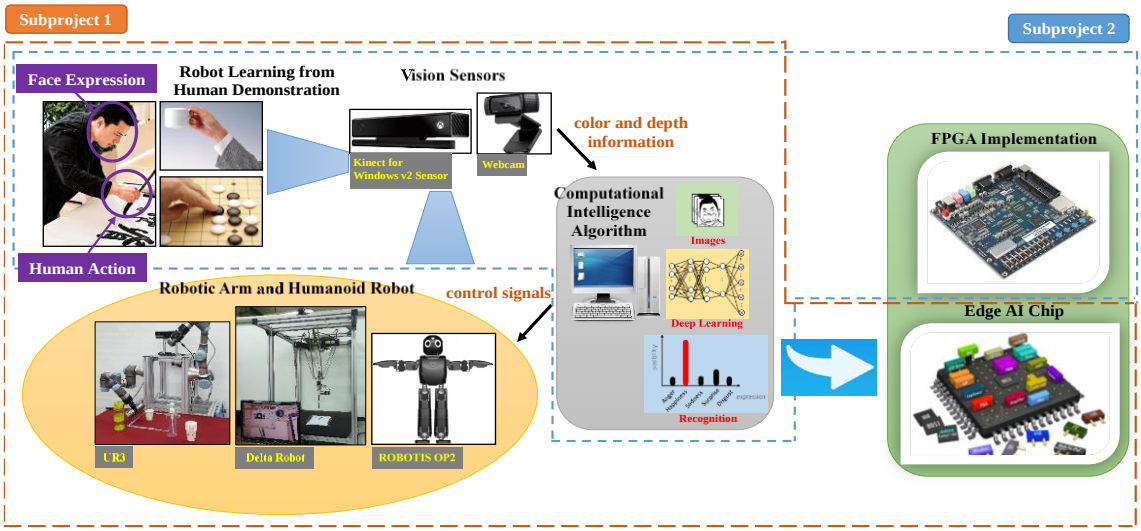


Fig. 1
Robot learning from human demonstration using multimodal perception approach.

Fig. 2
Robotic calligraphy system with self-error-correction.

Fig. 3
Development of the edge AI chip.

Fig. 4
System architecture of robot learning from human demonstration.

http://isdlab.ie.ntnu.edu.tw/isdwebsite/teacher-2.html

Development of Theory and Systems of Robot Learning from Human Demonstration (LfD)-Human Action and Face Expression Analysis System Based on 3-D Images

Principal Investigator

Prof. Chen-Chien Hsu

Summary

This project aims to develop a learning from demonstration (LfD) robot system with mobile capability for navigation in an indoor environment to provide home service/care duties. Based on the achievements obtained in the first year, a vision-based LfD system is developed this year to provide better understanding of the behaviors of a human demonstrator through robust action and object recognition, based on which a mimicking robot can reproduce the tasks demonstrated by the human under different circumstances. To validate the proposed approach, manipulation of a coffee maker is adopted as an example. To obtain a mobile LfD system, FPGA implementation of an improved V-SLAM is currently under development to provide a low cost and high computational efficiency realization, where a two-fold process constructed by front-end and back-end is employed. Currently, we have built an object tracking system on FPGA hardware for evaluating the performance of the feature detection and matching modules that we have designed. Before the end of this year, a visual odometry implemented on FPGA can be developed for the LfD robot system.

Keywords

learning from demonstration (LfD), deep learning, multi-action recognition, object recognition, mimic robot, object tracking, visual odometry, VSLAM, FPGA.

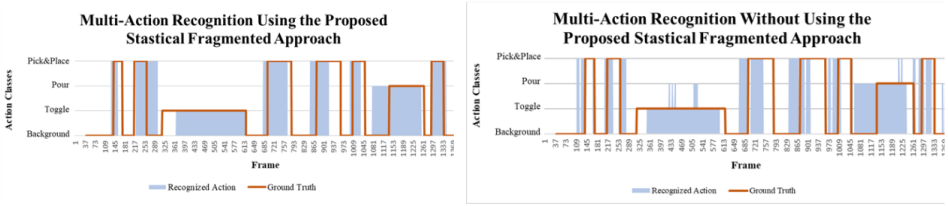
Innovations

- A multi-action recognition approach based on an I3D deep learning network followed by a statistical fragmentation of actions is developed to enhance the action recognition results.
- Object recognition using YOLO is incorporated with multi-action recognition to establish an action base corresponding to the behavior that users demonstrated. Based on the sequential order of actions in the action base, a mimicking robot can reproduce the overall demonstrated behavior via motion planning by inverse kinematics or pre-programming.
- Different from existing datasets where large outdoor actions are provided, a LfD dataset is established by collecting more than 9 classes of actions for home service/care usage.
- We develop a hardware-implemented object tracking system on a FPGA platform, where SIFT and matching algorithms are optimized and designed to improve their overall hardware efficiencies.
- The front-end of the V-SLAM algorithm is implemented on FPGA hardware to design a visual odometry with low cost and high computational efficiency.

Benefits

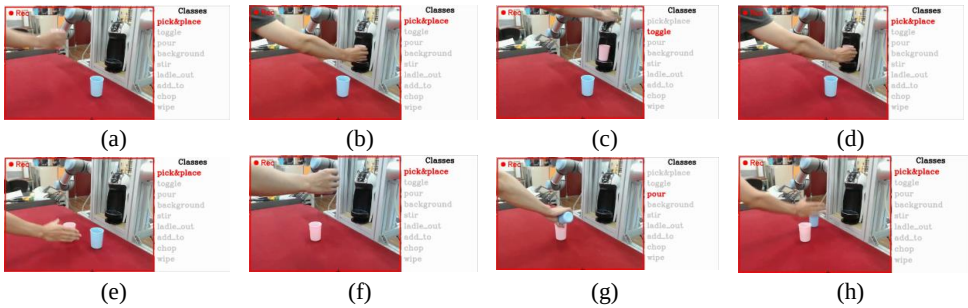
- Because of the proposed statistical fragmentation method, action recognition results are greatly improved as shown in Fig. 1, where noises in the recognition results have been significantly removed.
- Taking the manipulation of a coffee maker as an example, effectiveness of the proposed LfD system incorporating multi-action recognition and object recognition is confirmed. Fig. 2 shows the human demonstration and the recognized actions indicated in red on the right-hand side. Reproduction of the demonstrated actions by the mimicking robot is shown in Fig. 3, where an action base in the left side of each figure indicates the current status of the UR3 robot arm.
- A two-fold process of a V-SLAM constructed by front-end and back-end is shown in Fig. 4 (a). Using the KITTI dataset, the average translation error and rotation error of the 11 sequences are reduced by around 30% compared to ORB-SLAM2. The localization and mapping results of the first sequence in shown in Fig. 4 (b), and the architecture of the front-end implemented on a FPGA platform is shown in Fig. 4 (c). Because frame rate of feature extraction and matching modules can reach 50 fps, the FPGA-implemented front-end of the V-SLAM is expected to have a processing speed up to 40 fps, providing real-time operation for the LfD system.

Fig. 1



(a) (b)

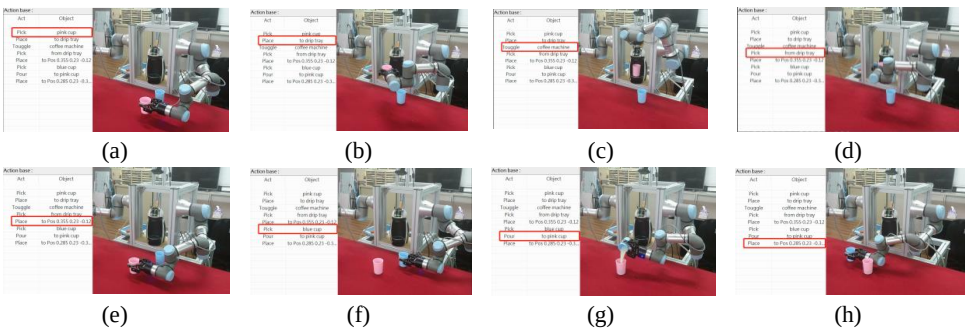
Fig. 2



(a) (b) (c) (d)

(e) (f) (g) (h)

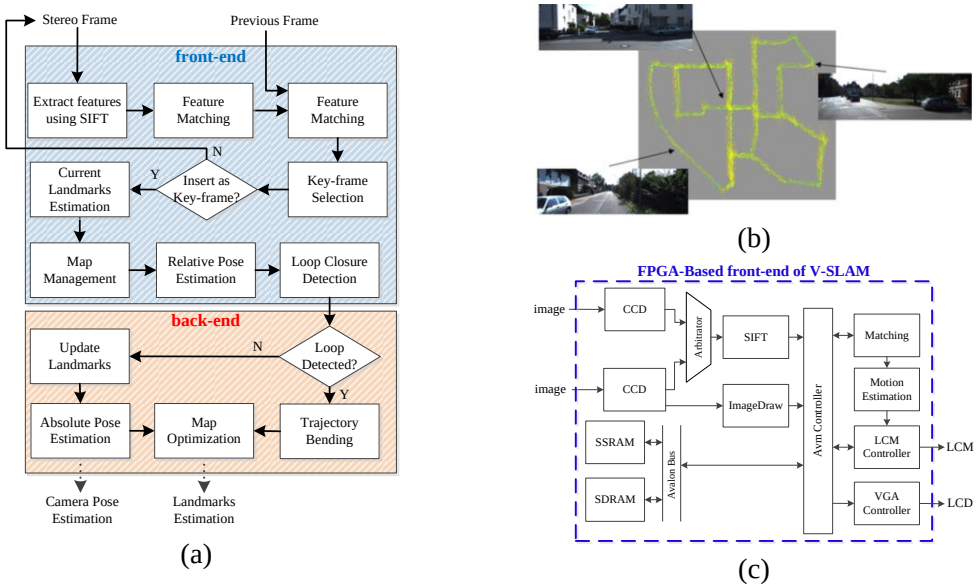
Fig. 3



(a) (b) (c) (d)

(e) (f) (g) (h)

Fig. 4



(a) (b) (c)

Fig. 1
Multi-action recognition results with and without the proposed statistical fragmentation method.

Fig. 2
Human demonstration and recognized actions in manipulating a coffee maker.

Fig. 3
LfD robot reproduces the demonstrated behavior according to the action base.

Fig. 4
(a) The architecture of the FPGA-based V-SLAM.
(b) Localization and mapping results of the V-SLAM using the first sequence of the KITTI dataset.
(c) The FPGA-based front end of the V-SLAM.

https://sites.google.com/site/jackybaltes/home?authuser=0

Safe Explainable AI via Behavior Decomposition

Principal Investigator

Prof. Jacky Baltes

Summary

The goal of this project is to develop novel algorithms that are able to transform neural networks and deep learning architectures learned into representations that are susceptible to analysis and verification. Artificial neural networks and especially deep learning approaches are popular at the moment, because of their outstanding performance on a variety of tasks. One important drawback of artificial neural networks is the fact that they act as a black box and that is impossible to extract the knowledge from the network. Therefore, a user can never be sure if the network learned the correct function or not. This may lead to poor and incorrect performance of the network or even biases against certain classes of users (e.g., misclassifying images of black people or Asians). Since more and more AI algorithms directly and substantially impact people's lives, there has been a recent push towards explainable AI, that is AI algorithms whose performance can be interpreted, analyzed, and understood by humans.

We propose behavior tree programs, an extension of Brook's subsumption architecture, as an intermediate representation suitable to model the important aspects of perception, motion planning, and goal reasoning of several important classes of robot systems. We use high accuracy mobile manipulation as an umbrella technology covering our applications, because of its importance to Taiwan industry. We investigate and evaluate our approach in three domains: (a) self-driving warehouse robots, (b) nuclear power plant operation robots, and (c) robot athletes. These robot systems pose unique and important challenges for AI. In the self-driving robot application, we investigate methods for converting and visualization the mapping from images (e.g., red, green, and yellow pixels) into perceptions (e.g, a traffic light that is currently green is 30m in front of the car). The goal is to evaluate the robustness of the perception against other images. In the nuclear power plant operation robot domain, we investigate the motion plans generated by a drone through reinforcement learning for decommissioning generator set (e.g., evaluate radiation space distribution for the planning of decommissioning strategy or separate waste in different radiation level). In the robot athlete domain, we transpile goal directed behavior into motion plans (i.e., action sequences). The goal is to develop humanoid robots that are able to compete against humans at an Olympic level. We plan to participate at the FISU World University Games in 2022.

Initially, we use a white box approach, that is, we use knowledge about the internal structure of the system in our conversion. For example, the output of deep learning network or the reinforcement learner will be used directly. In the last stage, we will treat the system as a black box and infer an approximate behavior tree program for a robot system without knowledge about its internal structure.

Keywords

Explainable AI, behavior tree decomposition, high accuracy mobile manipulation, nuclear plant operation robot, autonomous warehouse robots.

Innovations

- Novel behaviour tree based programming language will allow us to evaluate safety of opaque AI systems and and to explain the decisions and motion plans of a robot system
- Development of high accuracy mobile manipulation applications will bring important new capabilities to Taiwan industry.

Benefits

We investigate and evaluate our approach in three domains: (a) warehouse robots, (b) nuclear power plant operations robots, and (c) robot athletes.

- In our project, we focus on a programming language that allows us to easily express the control flow behavior of robotic applications.
- Based on modern AI Architectures using deep learning as shown in Fig. 1, the programming language should be abstract enough so that it allows us to reason about the beliefs and intentions of the robot, as this is important for explainable AI.
- Through the development of an integrated and applied robotics project, the students acquire knowledge in a practical and contextualized way. This systemic approach not only helps on the knowledge consolidation, but also, and more importantly, prepare the students in their social and interdisciplinary skills.
- Problem solving requires research, creativity, logical reasoning and action planning. Group work helps them develop their capacity of communicating ideas, reason and negotiate.
- Leadership skills are also developed in this way. All these social skills are as important as the technical ones, and maybe even more important in the long term -- given the fast pace of technological advances, tools, methods, technologies, they are likely to change, be updated and replaced in a short period of time. Social skills, like the ones mentioned above, help students be prepared for these changes.

Fig. 1
Modern AI Architecture Using Deep.

Fig. 2
High Accuracy Mobile.

Fig. 1

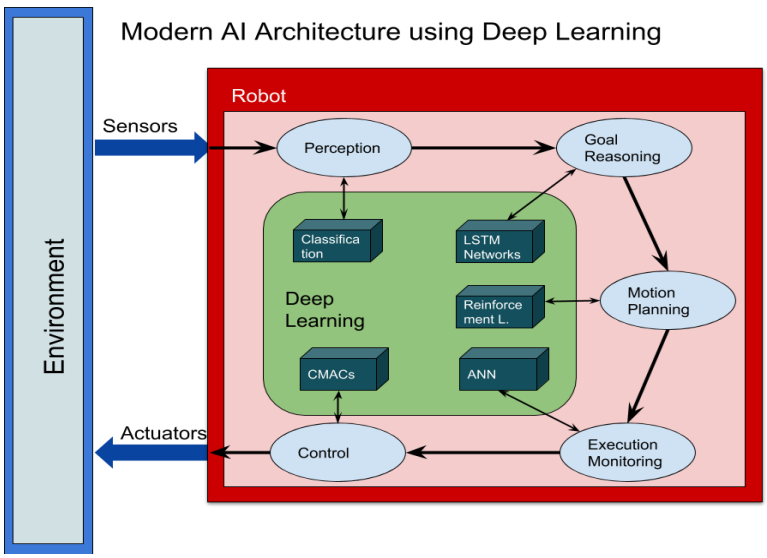
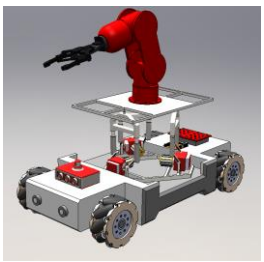


Fig. 2



Intelligent Conversational Robot with Deep Natural Language Understanding

Principal Investigator

Prof. Wen-Lian Hsu

Summary

Our goal of developing an intelligent conversational assistant for natural language problems is to build an example of explainable AI (XAI) with applications such as a finance-related chatbot. Our system solves the user queries and produces step-by-step explanations similar to a real human. If the assistant makes a mistake, we can easily unravel the error. When it solves a problem correctly, the explanation would allow the user to understand the logic behind it. We adopt the statistical principle-based approach (SPBA) that preserves the advantages of both statistical machine learning (SML) and rule-based (RB) methods while avoiding their pitfalls. Our robot learns to solve a problem in a way similar to human learning. Although the assistant for a specific domain, e.g. finance, involves limited domain knowledge and simpler sentential structure, major components of natural language understanding (NLU) such as anaphora resolution, ellipsis, fronting, entailment, parsing, ontology construction, and inference, are all necessary to be incorporated. Hence, after building a conversational assistant, we would have also acquired fundamental modules for many NLU applications including math problem solver and other task-oriented agents. Currently, we have achieved 80% of our goal, and are working rigorously toward semi-automatic ontology construction and natural language generation modules.

Keywords

Statistical Principle-based Approach (SPBA), Natural Language Processing, Chatbot, Virtual Assistant, Language Understanding, Intelligent Agents.

Innovations

Unlike most machine learning systems which exhibit an F-score of around mid-80 percent, pattern recognition capabilities of a conversational assistant normally require an F-score in the high-90's in order to lay down logical and legible natural language explanations. Furthermore, entailment and inference are crucial in constructing an all-in-one chatbot system. To achieve this, we require a large amount of common-sense knowledge. However, new techniques for extracting such knowledge and making use of it are yet to be developed. We have arranged such knowledge in the form of verb and noun frames, including ontologies, and various equivalent descriptions in our knowledge base and inference system, *InfoMap*. Below, we briefly describe our innovations.

- **Knowledge acquisition.** The main goal is to learn ontologies semi-automatically. Words in an ontology of another word are usually collocated with that word. Via deep learning, we attempt to learn noun-noun (NN), noun-verb (NV), adjective-noun (AN), and noun-classifier (NC) pairs. Currently, we have achieved 95% accuracy for NC pair, and 90% for the other pairs. NC pairs are crucial for solving problems and other pairs are the backbone for resolving ambiguity in natural language.
- **Anaphora and ellipsis resolution.** We resolve anaphora and ellipsis in natural language inquiries, and reconstruct topic-fronting using collocation pairs.
- **Natural language scripts.** Scripts are important for our agent. Instead of formal logic representations, we have adopted scripts to produce comprehensible answers for users.
- **Dialogue system.** After solving a query, our system can take further questions, such as those regarding different aspects of the problem, as well as questions that alter parameters of the problem. In the future, we shall work on questions raised by users who encounter specific difficulties.
- **Natural language generation.** Given a sentence, we can generate ontological instances. We attempt to generate natural sentences by inverting the above process through ontological instances. This step is crucial to question answering and response generation in a dialogue system. We will complete the Chinese agent system and extend it to other languages in the near future. (Fig. 1)

Benefits

- **Knowledge accumulation.** Unlike the parameters learned in SML systems and their isolated modules, our system learns an ontological knowledge, which can be categorized, accumulated, and reused similar to encyclopedia. It can even be adapted according to the level of intelligence of the user. Moreover, our core program can be easily extended to other dialogue systems to achieve deep understanding of domain knowledge provided that an ontology is available.
- **Narration.** Many existing chatbot systems employ logical forms for the computer to perform understanding. Traditionally, the inference process of such a system is very difficult for human to comprehend. Since we are designing a system for a general audience, it is essential to provide solutions similar to those by humans with many alternative descriptions. Besides, the inference in everyday conversation is not too complex, so simple entailment relations should suffice. Currently, we attempt to generate natural sentences as mentioned above. In the future, such a process will be further extended to generate paragraphs and stories.
- **NL agents.** Our agent is built on the SPBA, with its core, the InfoMap kernel. At the heart of InfoMap we employ three key components, i.e., entailment scripts, semantic frames, and instance maps. An overall flowchart of the SPBA framework is shown in Fig. 2. By means of this kernel, we can match natural language sentences via semantic frames to concepts in the knowledge base and build instance maps. Then, we follow scripts to utilize frames and concepts to simulate the thinking process of humans. In a step-by-step fashion, we capture key concepts and generate the correct answer. In the future, we also plan to build a general language education system. It will go beyond simple grammar checker and offer paraphrasing or writing suggestions, in order to substantially enhance students' writing ability.

Fig. 1
Understanding an inquiry regarding financial information in order to generate the solution.

Fig. 2
Flowchart of the proposed SPBA-based conversational system.

Fig. 1

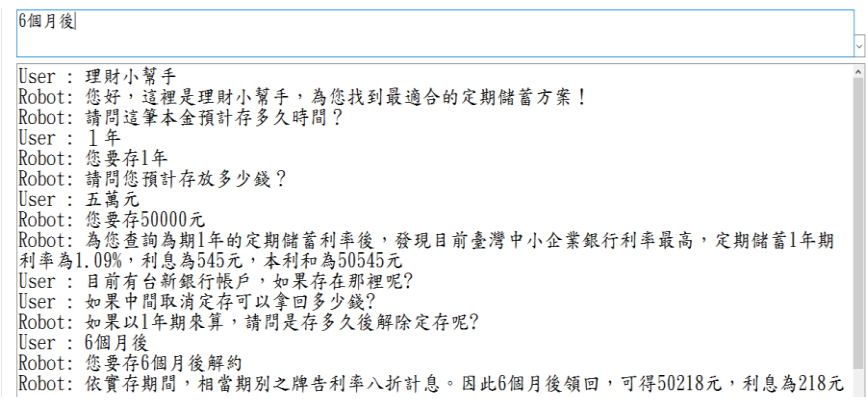
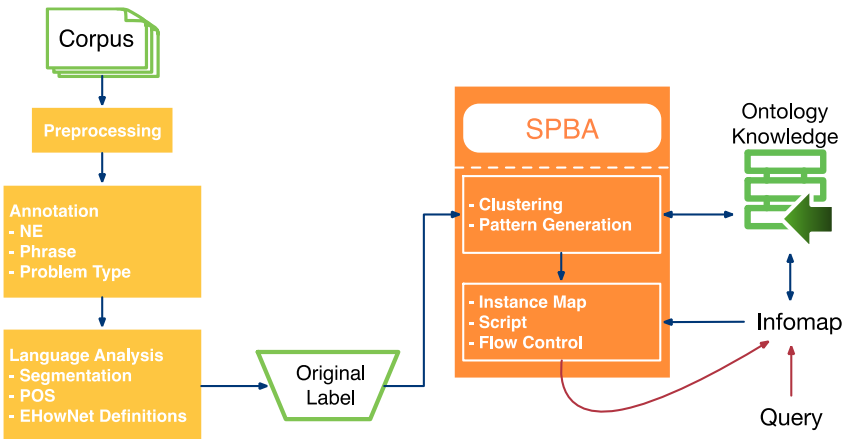


Fig. 2



http://www.cs.nccu.edu.tw/~jsliu/lab/c_1.html

Advancing Drone Intelligence: Core Technology Development and Contextual Applications

Principal Investigator

Prof. Jyi-Shane Liu

Summary

Our research aims to develop an autonomous intelligent control system for UAV and advance artificial intelligence components for the UAV industry. The technical agenda is to establish intelligent UAV control techniques by advancing from basic-level intelligent flight to more complex adaptive flight. This includes dynamic space modeling, sense-and-avoid and adaptation on task control. The project is composed of four functional modules. The cognition module is responsible for situation awareness and decision-making, such as using SLAM to identify the location of UAV relative to the inspected target as (Fig. 1). The motion module provides motion planning and camera path planning. The perception module develops sensory analysis and recognition algorithm. The interface module supports data visualization and human-machine interaction with task control user interface. All modules contribute to a common knowledge model that includes space, mission, and self-status. Through integration, our current autonomous intelligent control system can perform tasks such as building inspection, area scan and delivery.

Keywords

Autonomous Intelligent Drone, Intelligent Flight Control, Mission Intelligence, Drone Applications, Artificial Intelligence, Robotics, Computer Vision, Computer Graphics, Real-Time Intelligent System, Deep Learning

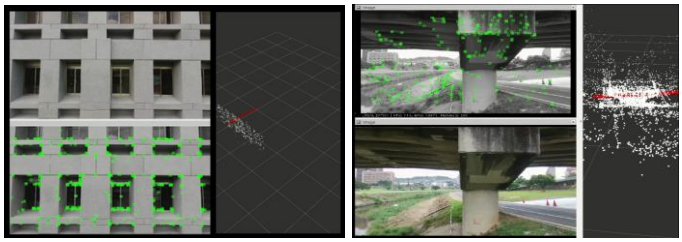
Innovations

- A path planning algorithm that can be used to calculate flight path (Fig. 2) and provides path correction when the UAV deviates from intended route due to various factor.
- By combining traditional computer vision techniques and neural networks approaches, we aim to develop a series of visual recognition models (Fig. 3 & Fig. 6(b)) and algorithms (Fig. 6(a)(c)) specifically for UAV to be able to execute missions autonomously in various designed scenarios.
- Mission-oriented ground control station (Fig. 4) to support researchers and users to monitor UAV behavior during mission execution.
- An autonomous intelligent control system framework aims to integrate sensory recognition, flight behavioral control, path planning and user interface. This framework, based on Behavior Tree (Fig. 5) serves as an integrated architecture to flexibly incorporate related technical advances according to customized needs.

Benefits

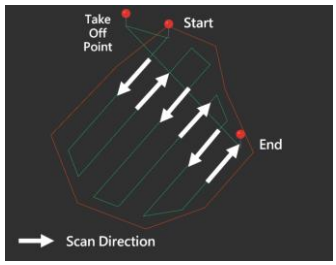
- Many research focused on advancing specific functionality of drone. However, technical validation with UAV require software/hardware integration and safety measures. The autonomous intelligent control system, contributed by this project, serves as an integrated robotic framework that can incorporate and replace individual functional modules. This benefit researchers who needs to test a developing algorithm but lacks other system components to validate its performance in the real world.
- Task control system based on Behavior Tree (Fig. 5) that can easily reorganize to form new task behavior.
- Field testes system is in real-world, providing insights for AIs to interact with real environment (Fig. 6).
- The ground control station accelerates research by visualizing salient information related to performance evaluation. It also aims to provide safety measures and alert when researchers fail to notice unexpected events that could cause accidents.
- Visual recognition capability specific for UAV is essential towards the autonomous intelligent system. Various technical methods for different use cases provide a much wider option for different types of mission.

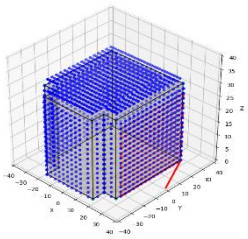
Fig. 1



(a) (b)

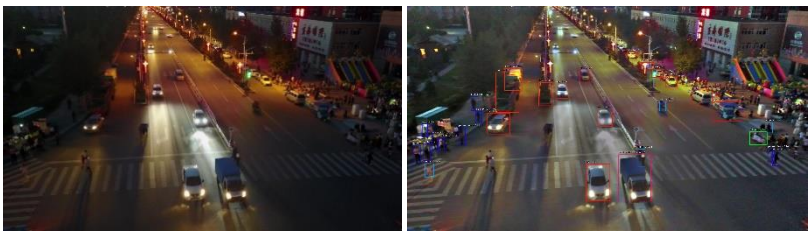
Fig. 2





(a) (b)

Fig. 3



(a) (b)

Fig. 4

Fig. 5

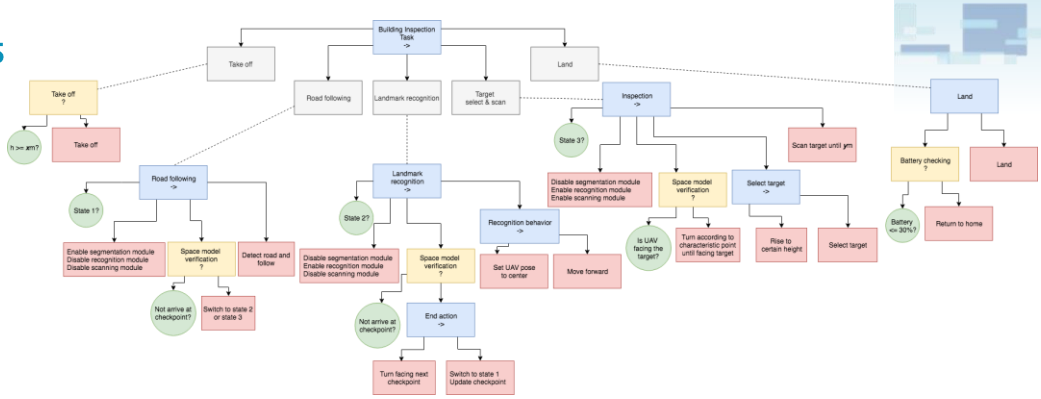


Fig. 6

(a) (b) (c)

Fig. 1
(a)(b) Building and bridge pier inspection based on SLAM.

Fig. 2
(a) 2D Area coverage scan path planning.
(b) 3D Building coverage scan path planning.

Fig. 3
(a) Original image.
(b) Object detection in degraded image.

Fig. 4
(a) Ground control station in manual control mode.
(b) UAV performance that allows observation from 3 sides given 3 images that compose by UAV swarms.

Fig. 5
Building inspection using Behavior Tree for task control.

Fig. 6
(a) Accurate line following on vertical surface.
(b) Road following along a riverside park.
(c) Building inspection.

http://www.ce.nchu.edu.tw/wb_teacher02.asp?url=42&cno=4&tno=6

Intelligent Agricultural Cultivation Support System Integrating UAV Surveillance

Principal Investigator

Prof. Ming-Der Yang

Summary

This research project aims to establish an intelligent agricultural cultivation support system, by integrating unmanned aerial vehicle (UAV) surveillance and artificial intelligent (AI) analytical techniques. The support system provides a great help in various current agricultural management challenges, such as sparse agricultural data and manpower limitation. Three major tasks are proposed in order to achieve the goal over the four years' project span, including forming a UAV multi-source image database, developing relevant AI image process technologies, and establishing a UAV image analysis cloud platform (Fig.1). Meanwhile, the UAV flight standard operation procedure (SOP) for rice cultivation surveillance at different stages is established. Rice, the major staple food crop in Taiwan, is selected and monitored for two whole crop life cycles. More than 12 TB UAV multi-source image data, including visible spectrum, multispectral, thermal, and hyperspectral images, were collected from three study sites, NCHU (National Chung Hsing University) Experimental field, NCHU Agricultural Experimental Station, and TARI (Taiwan Agriculture Research Institute) Experimental Field. We have developed various of models and applications, such as seedling positioning/counting, leaf color analysis, plant height analysis, green coverage ratio analysis, yield prediction, grain moisture content assessment, damage assessment, crop recognition using Convolutional Neural Networks (CNN) and Deep Neural Networks (DNN)(Fig.2, 3) with edge computing capabilities. We also developed a cloud platform of Aerial Agriculture Analysis for tasks, such as image mosaicking, image texture analysis, vegetation index analysis, and 3D model construction functions.

Keywords

Unmanned aerial vehicle (UAV), UAV multi-source image database, UAV image cloud platform, edge computing, Convolutional Neural Networks (CNN), Deep Neural Networks (DNN).

Innovations

- A UAV agricultural multi-source image database was established to cover the complete life cycle of rice growth stages. More than 12 TB UAV images, including visible spectrum, multispectral, thermal, and hyperspectral images, were collected (Fig.4).
- A variety of image analyses covering rice growth stages were implemented for seedling positioning/counting (Fig.5), leaf color analysis, plant height analysis, green coverage ratio analysis, yield prediction, grain moisture content assessment, damage assessment, and crop recognition.
- Convolutional Neural Networks (CNN) and Deep Neural Networks (DNN) were used with edge computing capability to develop a real-time crop recognition model for rice.
- A cloud platform of Aerial Agriculture Analysis was established to provide UAV image analysis and expert advice support system.

Benefits

- The project completed 727 UAV aerial imaging tasks and established a UAV agricultural multi-source image database consisting of more than 12 TB images, such as visible light, multi-spectral, and thermal image data.
- Various field surveys were also conducted to establish a link between processed UAV images and rice growth for further applications.
- Many applications have been developed, such as seedling positioning/counting, leaf color analysis, plant height analysis, green coverage ratio analysis, yield prediction, grain moisture content assessment, damage assessment (Fig.6), and crop recognition (Fig.7), which can be applied to other crops in the future.
- A real-time crop recognition model with edge computing capacity was developed based on CNN and DNN for real-time classification and object identification.
- The cloud platform of Aerial Agriculture Analysis (Beta) has been established proving functions of UAV image uploading, automatic orthophoto mosaicking, and Google Map overlapping with GPS information. (http://uav-fly.nchu.edu.tw/agri/)

Achievements (2018 - 2019.09)

- Future Tech. 2019: Breakthrough Award & Highlight Innovation
- G0V grant 2019: Best proposal award
- CTCI AI Creative competition: Honorable mention
- Ag Thon 2018: 2nd place
- Patent pending: 2
- Technology licensing: 1
- Industry-Academy cooperation: 9

Fig. 1

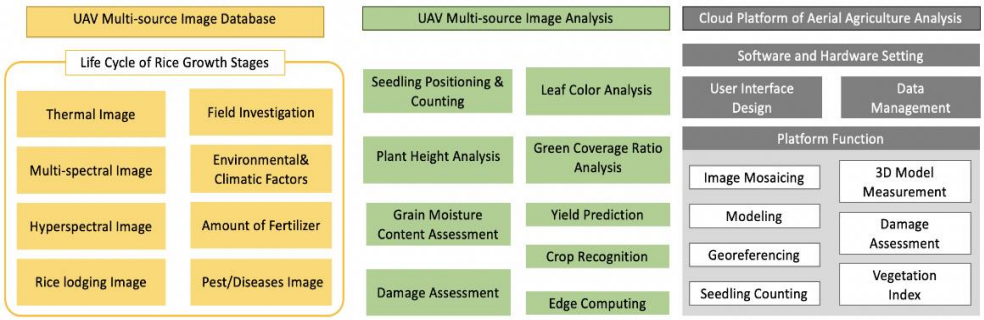


Fig. 2

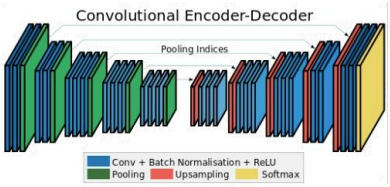


Fig. 3

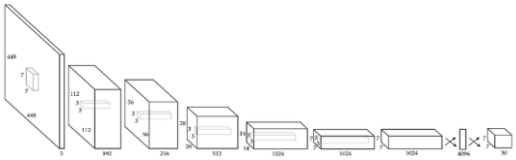


Fig. 4

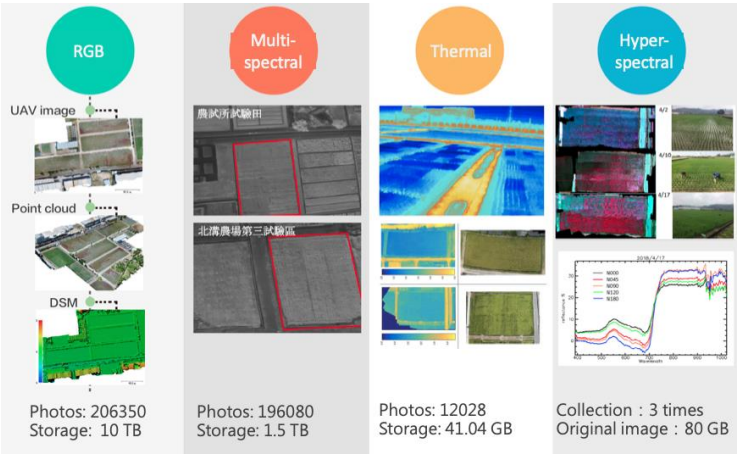


Fig. 5

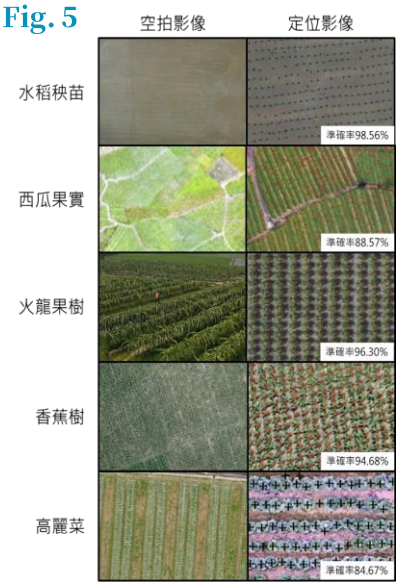


Fig. 6

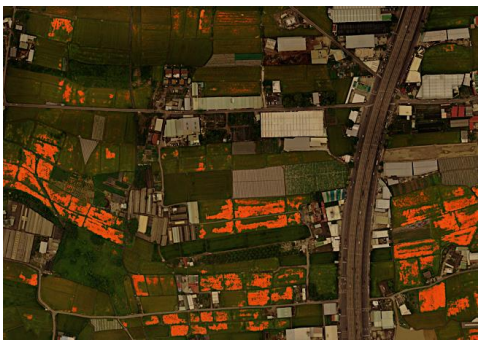


Fig. 7

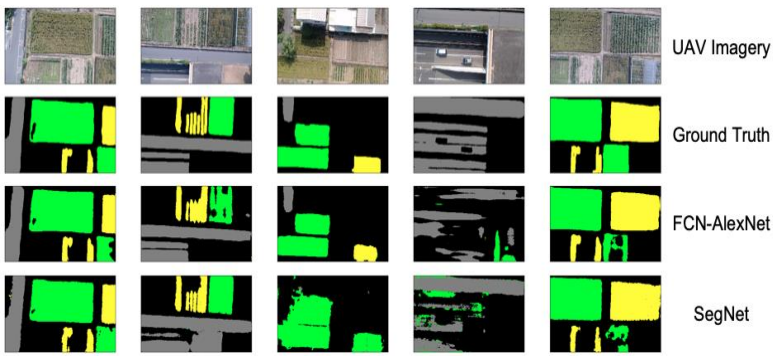


Fig. 1
Schematic chart of three major tasks in this project.

Fig. 2
SegNet Architecture.

Fig. 3
YOLO Architecture.

Fig. 4
Multi-source UAV images.

Fig. 5
Plant positing/counting.

Fig. 6
Rice lodging investigation.

Fig. 7
Crop recognition.

Artificial Intelligence on the Assistance of Learning Board Games for Human

Principal Investigator

Prof. Shi-Jim Yen

Summary

The goal of this project is to develop the AI techniques that assists human to learn board games techniques easily and efficiently. Board games training to the field of artificial intelligence is like trials over drosophila to the medical field. A series of AlphaGo and Alpha Zero programs create a history for AI. We believe that computer games will also play the role of the drosophila of computer-assisted learning. Because this project will investigate many kinds of games, it is expected that innovative Deep Learning(DL)/Reinforcement Learning(RL)/Deep Reinforcement Learning(DRL) techniques will be developed. In terms of the development of society, with the rise of AI, humans need better logical decision-making ability and critical thinking skills more than ever. Board games are great tools for exercising logical thinking. For children, playing board games helps them develop logical thinking abilities, and cultivates their resilience and perseverance. It is shown by a Gerontology research institute in Japan that Go could improve cognitive functions. As such, this project has positive effects on the development of society.

Keywords

Board Game, Alpha Go, Alpha Zero, Deep Learning, Reinforcement Learning, Deep Reinforcement Learning, Residual Network, Generative Adversarial Network, FeUdal networks, Seed Optimization.

Innovations

- We use DRL to train the models for personalized ladder and visual learning based on Scaffolding Learning theory. It could improve the quality of our AI Go services. The services are running on an online AI-assisted board games learning platform for teaching, learning and playing Go. It will be the largest AI Go service with one billion annual AI service count.
- We cooperate with Facebook ELF team on PolyGames. PolyGames is a framework developed by deep reinforcement learning and Alphazero techniques.
- We propose a novel approach using deep learning to obtain an encoding scheme so that the compression ratio is suitable for Chinese dark chess endgame databases, it could increase the storage efficiency of Chinese dark chess programs.
- We introduce a deep learning competition framework on Othello for education.
- We use Seed optimization framework on board games, including Draughts, Othello, and Block Go. The framework could increase the strength of those game programs.

Benefits

- Computer board game is considered to be the drosophila of artificial intelligence. In the field of assisting human learning, we believe computer board games can also play the role of the drosophila. By exploring various types of games in this project, we could spark numerous new DL/RL/DRL technologies (Fig. 1).
- We collaborate with the Nihon Ki-in (Japan Go Association), and with professional player Ming-wan Wang, and deploy an ample number of AI Go bots (Fig. 2) on their UGEN(Yugen no Ma) Go playing website (Fig. 3). Currently, these bots are very popular among Japanese players, with over 250 million cumulative service count.
- We develop a life-wide Go learning system (Fig. 4) and build a web based APP for the system (Fig. 5).
- PolyGames is a general reinforcement learning game framework which can be applied on many board games. For Othello, Hex and other games, PolyGames may be the strongest game program in the world.
- We win more than 10 gold medals in Computer Olympian computer game tournaments and other international computer game competitions.

Fig. 1
The bot engine framework in this project.

Fig. 2
Our AI service for the UGEN Go website.

Fig. 3
In the UGEN(幽玄之間) website, our project provided over 259 million cumulative AI Go service.

Fig. 4
The architecture for the life-wide system .

Fig. 5
The life-wide Go learning system.

Fig. 1

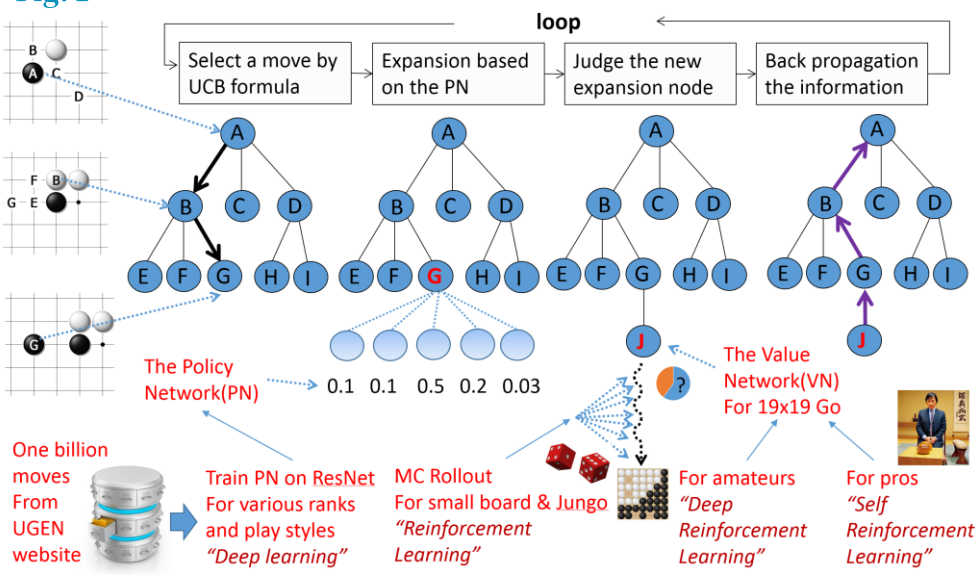


Fig. 2

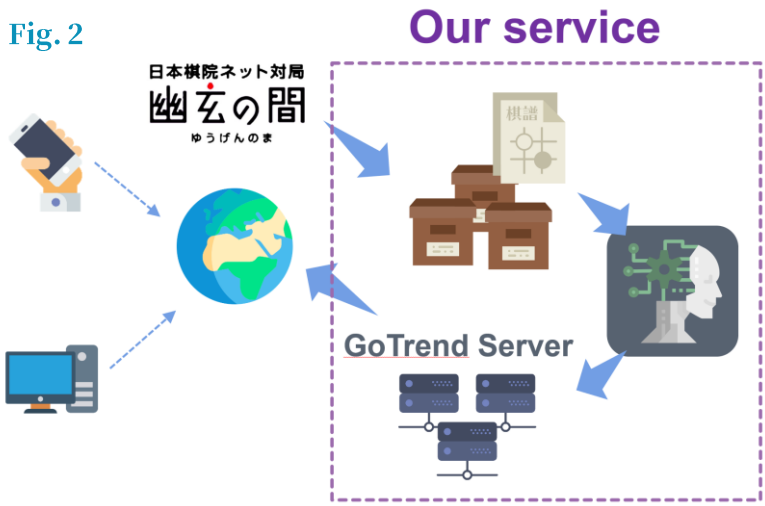


Fig. 3

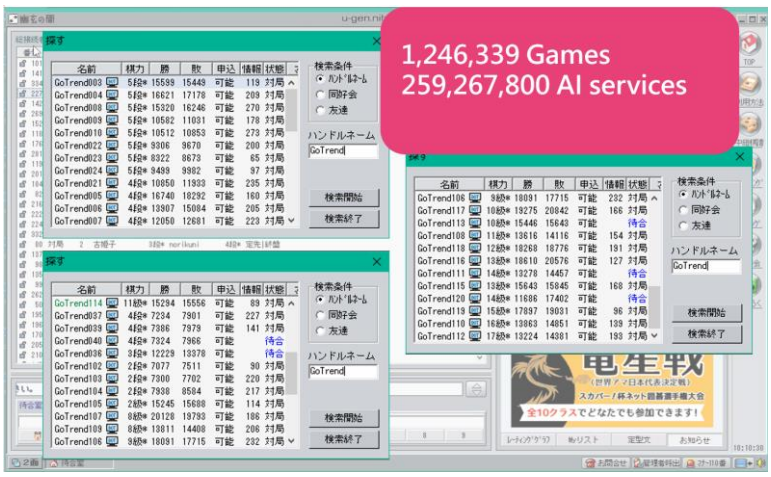


Fig. 4

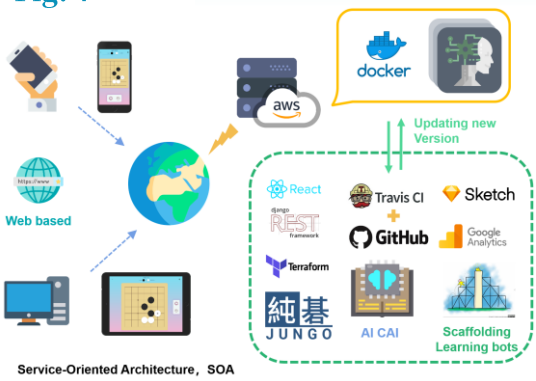
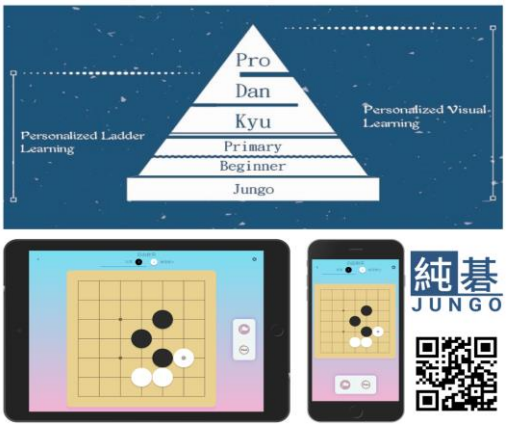


Fig. 5



Threat, Challenge, and Turning Point of AI for News Communication

Principal Investigator

Prof. Chi-Ying Chen

Summary

An interdisciplinary effort is needed for solving the fake news crisis, because the solutions depend not only on AI, but also on social and political input. It's easy to create fake news (for example, Deepfake) through AI technology, but using AI technology to fight fake news is far from enough since the problem comes from the misuse of social media and its high speed spreading into the crowd. We propose a good solution through a collaborative project that develops a trusting news platform by AI and blockchain technology and effectively utilizes the common supervisory power of the society. This project contains three sub-projects covering different research aspects. Sub-project 1 will investigate these technical components: (1) AI blockchain based crowd sourcing trusting news ranking mechanisms as well as ecosystem models; (2) AI detects the tampering of multimedia materials; (3) AI uses evidence to evaluate the truthfulness of a news; and (4) AI and blockchain integration. Sub-project 2 and 3 discuss the social and legal aspect respectively.

Keywords

Fake news, Trustfulness of news, Artificial intelligence, Blockchain, Science and technology law, Deep learning.

Innovations

Sub-project 1–Technology Perspective:

- We propose truthfulness news ranking mechanisms based on AI blockchain crowd sourcing.
- We propose multimedia identification technologies based on deep learning technology to effectively identify forged multimedia materials.
- We propose evidence based AI mechanisms to evaluate the truthfulness of a news by finding related evidences.
- We propose mechanisms and system design research to integrate AI and blockchain for trusting news platform.

Sub-Project 2: Humanity Perspective:

- We propose to use the ELM model to investigate the public's information processing and evaluation of fake news, and evaluate the trusting news platform by its agenda setting function.
- We propose to explore the AI impact on journalists and practitioners as well as study fact-checking mechanism.

Sub-Project 3: Law Perspective:

- We propose to set up an internet news reliability identification system that conforms with our national legal system to further establish the law and morality rules of internet news propagation.

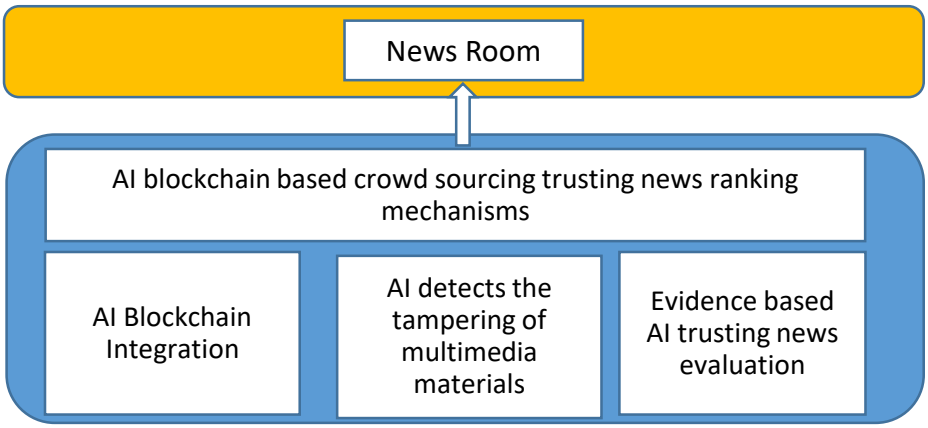
Benefits

- By combining the AI/blockchain/social/humanity/law aspects and technologies in building an AI blockchain platform which can leverage AI and blockchain accountable and traceable crowd sourcing mechanism to rank the trusting news.
- Mechanisms and system design research to integrate AI and blockchain for trusting news platform. The social and legal aspect respectively so as to guide the construction of this platform with trusting news ecosystem in line with the humanity and social needs.

System architecture

The system architecture is shown in Fig 1. We define the trusting news as the verifiable news. The system leverages the integration of AI and Blockchain technology to evaluate the truthfulness of a news by the weighting factors from crowd sourcing, domain experts, and evidence based AI. Blockchain provides a transparent news trace and message owner provenance for crowd and experts to express their views of a given news statement in a blockchain news room. Once a given news statement is created in a news room, the AI component will work to detect the tampering of the multimedia material, and find evidences to evaluate the truthfulness of a news.

Fig. 1

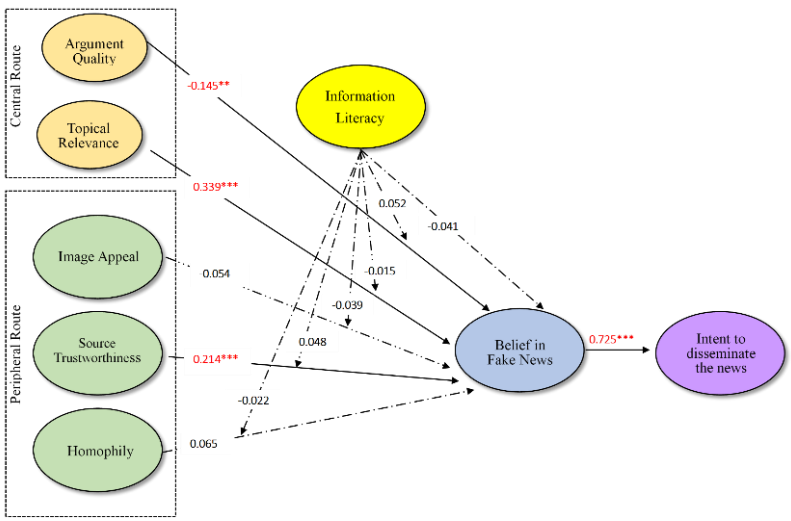


Humanity Perspective

Factors associated with belief or disbelief in false news:

- users depend on the argument quality to recognize the falsehood of the messages.
- Topical Relevance and Source Trustworthiness make user vulnerable to believing in false news.
- Information Literacy does not moderate the relationship.

Fig. 2



Law Perspective

- Fake News has become so intertwined with our daily lives that it generates grand issues such as privacy endangerment, social stability breakdown, democracy recession, and even sovereignty loss in Taiwan.
- The filter bubble problem and Middle Media problem remains very challenging. The first being people trusting what they want to trust hence they would see news as genuine if the news confirms to their beliefs. The second being groups trying to stop correct and reliable news to spread out, stopping the public to receive rightful information.
- Under certain situation and circumstances, all speeches, whether truthful or false, should have some degree of protection.
- New regulations on online political advertising should be in compliance with existing rules and they must be further suited for the new media environments.

Fig. 1
System Architecture.

Fig. 2
Information Processing Model of False News.

Artificial Intelligence in Asset Allocation, Derivatives Pricing and Risk Management

Principal Investigator

Prof. Min-Teh Yu

PAIR

Summary

This project was proposed in response to the Ministry of Science and Technology (MOST) for promoting the development of various AI innovation research centers. The project applies various AI technologies in asset allocation, derivatives pricing and risk management. The first focus of this project is on how to form a high-yield portfolio and control the portfolio risks using AI. This project applies GPU and financial theory to provide a more advanced model for capturing the risk profile of the financial assets, markets, and institutions. A stable, secure, and low-cost asset allocation AI Robo-advisor will be designed. The project also uses AI to develop fast and accurate derivatives pricing models for financial institutions, which can correctly evaluate the price and hedge ratio of derivatives. We also explore how Monte Carlo simulation with GPU can enhance the pricing and risk management for derivatives.

Keywords

Artificial Intelligence; Machine Learning; Portfolio Management; Trading Strategy; Stock Picking; Market Timing; Smart-AI Index ETF; Stochastic Volatility; Derivatives; Options; Market Calibration; Graphics Processing Unit; Compute Unified Device Architecture; Economic Scenario Generator; Risk Management; Monte-Carlo Simulation; Value at Risk; Expected Shortfall.

Innovations

- Evaluate the stock picking and market timing strategies using machine learning techniques with corporate and market data. It then applies the machine learning methods to predict future states of the economy and stock market using macroeconomic variables to develop a smart-AI index ETF.
- Propose a spherical Monte Carlo estimator to improve the efficiency of Monte Carlo integration in calculating risk measures under Basel Accords and prices and Greeks letters of financial derivatives. Applications in portfolio management and credit scoring models using machine learning are also investigated.
- Develop a multi-agent deep reinforcement learning framework with a two-level nested agent structure to learn effective portfolio management. The proposed RL system consists of a specially designed reward function and a novel policy network structure. It trained the automatic trading agents to achieve better investment performance with lower risk under realistic transaction cost settings.
- The project integrates the financial engineering and behavior finance theories to construct a three-stage AI-Robo advisor model, with the first phase of asset selection, the second phase of optimizing portfolio, and the third phase of automatic rebalancing. The current asset selection phase has integrated the AI neural network algorithm for picking assets. This study verified the AI-Robo advisor model with the Franklin funds market and Taiwan's top 150 stocks market as investment targets. The empirical results confirm that the AI-Robo advisor model applied in two different markets can significantly beat the different benchmark indexes.
- Apply deep learning techniques to resolve the computational complexity of Monte Carlo simulation to produce timely price and risk measures for a large portfolio of assets and liabilities of financial institutions. Unlike existing pricing models based on neural networks, deep learning techniques are used to calibrate the parameters in theoretical financial engineering models. Hence, the proposed approach can save lots of computation time in comparision with conventional pricing models.
- Apply the CUDA (Compute Unified Device Architecture) of NVIDIA to provide timely measures of value at risk, expected shortfall, BIS Ratio for banks, and Solvency II ratio for insurance companies using AI-assisted derivatives pricing models with GPU parallel computing.

Benefits

This project aims at developing and applying AI to three areas in finance: asset allocation, derivatives pricing, and risk management. The research problems in the six sub-projects are interacted and highly related to AI. Also, the six sub-project investigators have successful experiences both in academia and industry. This project is to solve timely issues confronted in the current financial industry.

- In 'AI' and 'Derivatives Pricing,' the project is to provide fast and accurate derivatives pricing methods and hedge ratio of derivatives. It enables financial institutions to issue derivatives at low risk with the proposed dynamic hedging principle and allows investors to invest more with the hedging tools.
- In 'AI' and 'Risk Management,' with the GPU, Monte Carlo techniques, and financial theory, the project aims at providing more advanced methods and suitable models to capture risk profile of the financial products to facilitate practitioners engaging in trading activities in pursuit of higher trading volume and exposure with a lower risk. The goal of proposed multi-agent reinforcement learning is to model investment profits and risk measurement simultaneously in asset allocation and portfolio management problems.
- In 'AI' and 'Asset Allocation,' the project aims at providing a more stable, secure, and low-cost asset allocation. AI Robo-Adviser is expected to provide the general public with high-quality financial services that were only accessible for high-income customers.

Fig. 1

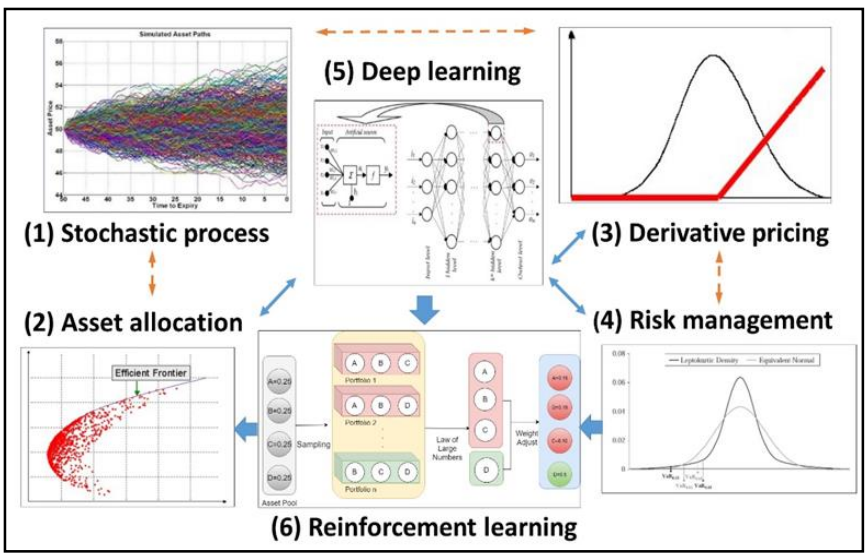


Fig. 1

This project combines

(1) Asset stochastic process (Geometric Brownian Motion, Heston model, Libor Market Model, Monte Carlo Simulation),

(2) Asset allocation (Markowitz MV optimization model、Black-Litterman model, Behavior Finance, Robo-advisor),

(3) Derivatives pricing and hedging (FX, G2++, Libor Market Model, GPU),

(4) Risk management (VaR with Skewness, Kurtosis),

(5) Traditional deep learning techniques in financial applications (Convolutional neural network, Deep neural network, Recurrent neural network),

and (6) Reinforcement learning with risk management for asset allocation (portfolio management, policy network, multi-agent).

A Home-Based Bio-Neuro-Feedback Intervention Programs Integrated System Using Artificial Intelligence Affective Computing and Multimodal Physiological Signals Monitoring for Patients with High-Risk of Cardiovascular Disorders

Principal Investigator

Prof. I-Mei Lin

PAIR

Summary

The aim of this project is to develop an artificial intelligence (AI) platform with home-based psychological intervention programs for patients with high-risk of cardiovascular disorder. This AI platform includes the affective computing and multimodal physiological signals detection system using TSMC 28nm chip. The first subproject aims to induce neutral, anger, happiness, and sadness by scripts and measure physiological responses including electrocardiography (ECG), photoplethysmography (PPG), and electroencephalography (EEG). The second and third subprojects aim to develop an integrated system of artificial intelligence affective computing and multimodal physiological signal detections. The physiological signal parameters of patients will be analyzed under different emotional states through affective computing algorithm, further providing bio-neuro-feedback treatment modules as well as evaluating the efficacy of the treatment. There are three core technologies in this project: (1) The AI affective computing and multimodal physiological signals was be administered. (2) The AI-based platform of affective computing and multimodal physiological signals detection will be developed. (3) The AI and home-based bio-neuro-feedbacks will be applied to patients with high-risk of cardiovascular disorder. This project anticipates that high-risk patients will demonstrate better autonomic nervous system activity and more relax brain waves than controls after AI home-based bio-neuro-feedback intervention.

Keywords

Artificial intelligence, affective computing, multimodal physiological signals, home-based training, bio-neuro-feedback, cardiovascular disorder, electrocardiography, photoplethysmography, electroencephalography

Innovations

- With numerous studies published by the Lab of Kaohsiung Medical University (KMU), bio-neuro-feedbacks were confirmed as an evidence-based clinical intervention. The risks of morbidity and adverse prognosis of cardiovascular diseases are expected to decrease. The hospital-based treatment will be transferred to home-based training by wearable devices with AI-platform in the future.
- The Lab of National Chung Cheng University (CCU) applied AI-based affective computing to monitor ECG, PPG and EEG signals and achieved high (> 90%) accuracy in identifying different emotional states.
- AI-edge devices and chip designs have been successfully developed by the Lab of National Chiao Tung University (NCTU) with optimized performances for wearable systems. This project can significantly impact and improve the quality and portability of medical grade home-based devices and treatments.

Benefits

- To develop an AI-based platform with affective computing processes and multimodal physiological signals monitoring.
- To develop AI-based treatment modules of bio-neuro-feedbacks, which featured innovation and international competitiveness.
- To facilitate the prevention of cardiovascular diseases through home-based treatment modules.
- To promote the diagnostic efficiency of medical professionals, provide the feedback of treatment efficacy, achieve maximum clinical benefits, improve self-management of patients, and reduce professional medical burden and costs.
- To develop prospective technology and cultivate interdisciplinary talents in the fields of psychology, electrical engineering, and electronics engineering.

Fig. 1

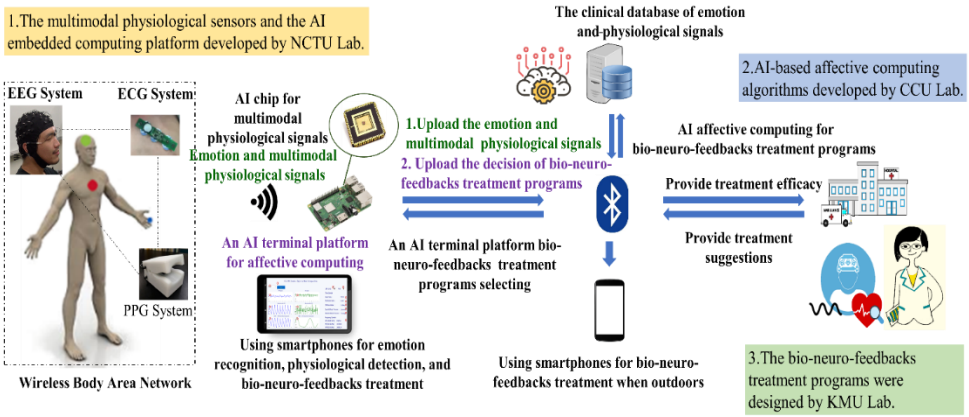


Fig. 2

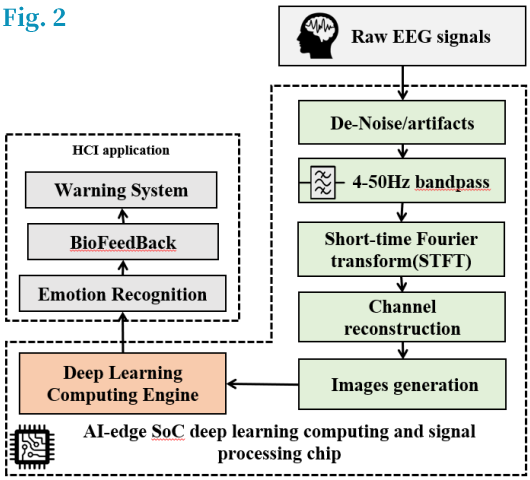


Fig. 3

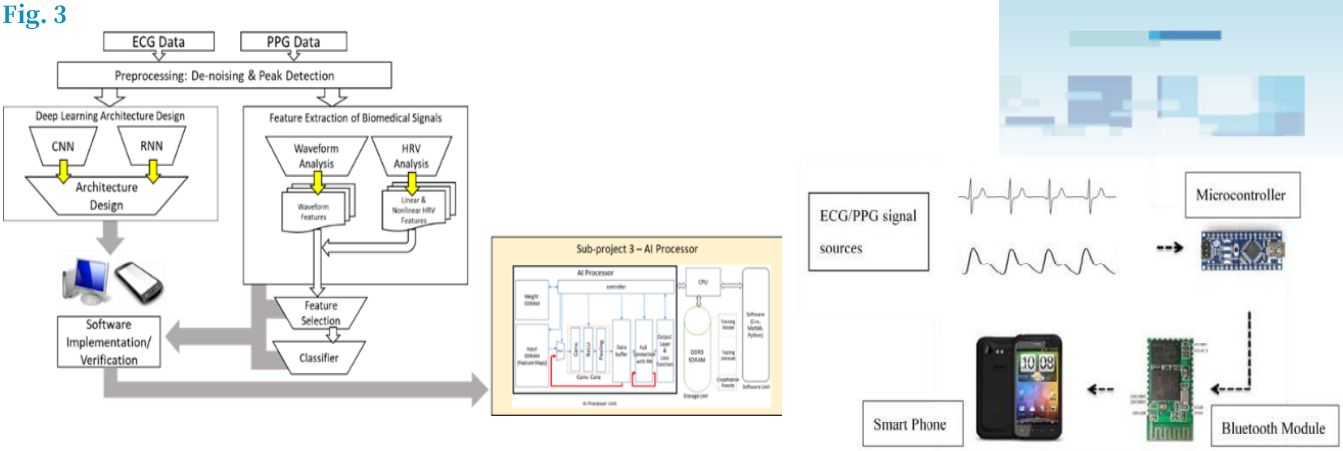


Fig. 4

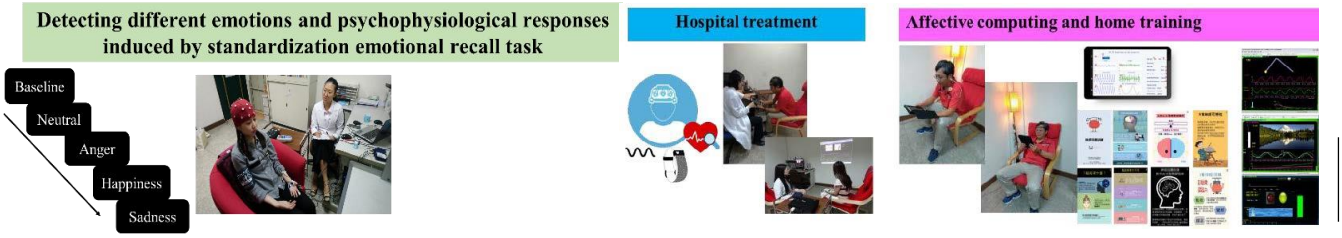


Fig. 1
The AI-based platform operating affective computing algorithms and physiological signals processing and providing bio-neuro-feedback treatment modules.

Fig. 2
Design flow of the EEG based emotion recognition system.

Fig. 3
Research design and system verification process of the ECG/PPG based emotion recognition system (left), and test platform of the real-time ECG/PPG identification algorithm based on smart phone (right).

Fig. 4
The affective computing and multimodal physiological signals monitoring for patients with high-risk of cardiovascular disorders, and the home-based bio-neuro-feedback treatment modules.

Pervasive Artificial Intelligence Research Labs 2019 Annual Report

Funding Agency Ministry of Science and Technology, Taiwan

Editor in Chief Yu-Chee Tseng

Editors
Ted Kuo, Po-Chi Hu, Li-Chun Wang, Hsueh-Ming Hang, I-Chen Wu, Jen-Hui Chuang, Jiun-in Guo, Kuo-Chin Fan, Wen-June Wang, Kuo-Kai Shyu, Jia-Ching Wang, Jin-Huei Yeh, Wei-Yen Wang, Chen-Chien Hsu, Jacky Baltes, Wen-Lian Hsu, Jyi-Shane Liu, Ming-Der Yang, Shi-Jim Yen, Chi-Ying Chen, Min-Teh Yu, I-Mei Lin

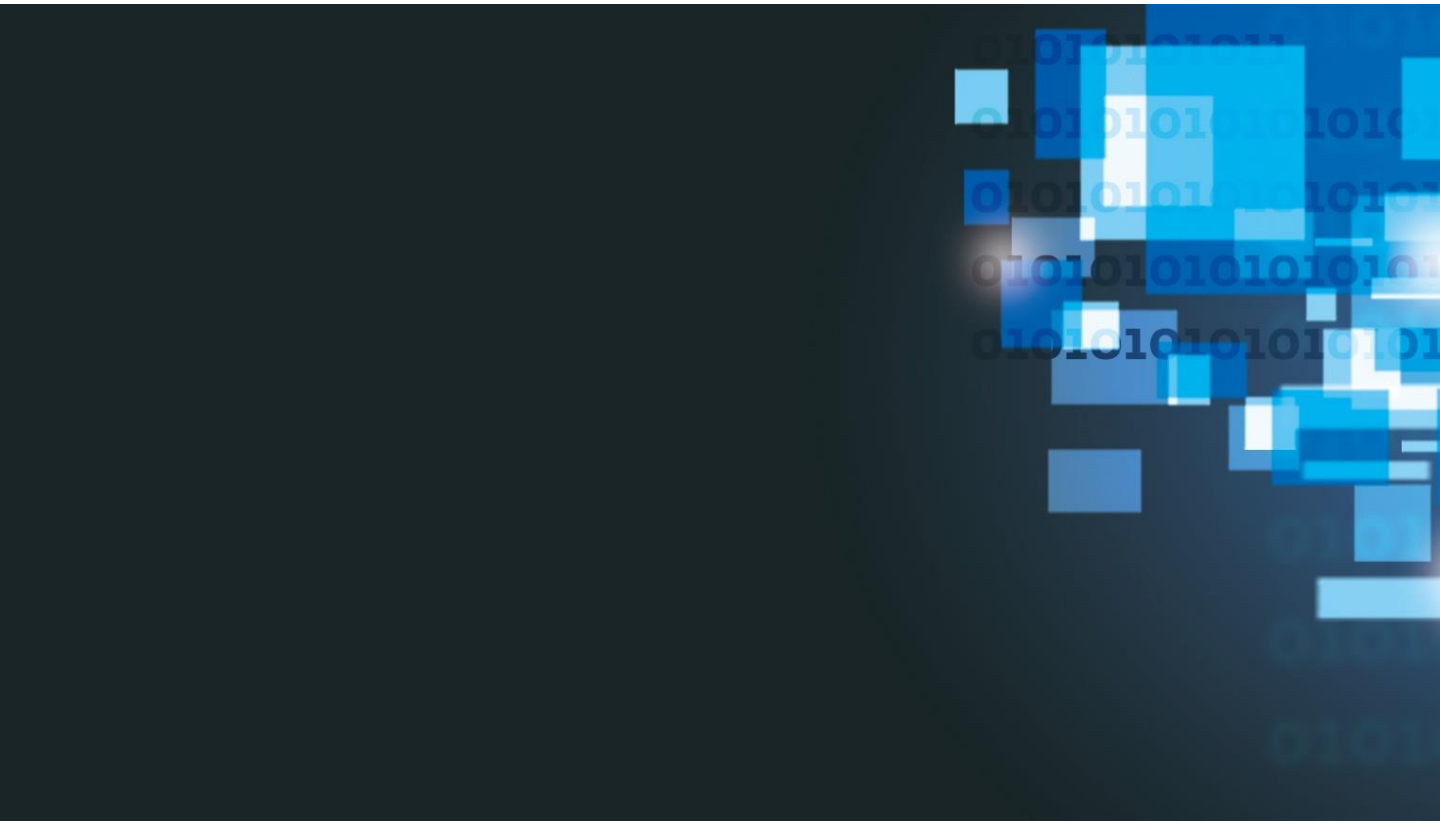
Associate Editors Li-Jen Wang, Shu-I Lin, Ya-Chiao Chuang, Yi-Ling Kuo, Shu-Hsin Ho, Ya-Wen Lu, Chao-Ho Lin, Fen-Yu Chang

Art Designer Ming-Lun Hsieh, Shu-Hsin Ho

Contact info@pairlabs.ai

Publication Date November 11, 2019

Publisher
Pervasive Artificial Intelligence Research (PAIR) Labs
National Chiao Tung University
1001 Da Hsueh Road, Hsinchu, Taiwan 30010
+886-3-571-2121 ext. 52957
www.pairlabs.ai/en



PAIR Labs, 1001 Da Hsueh Rd., Hsinchu, Taiwan 30010
https://pairlabs.ai/en