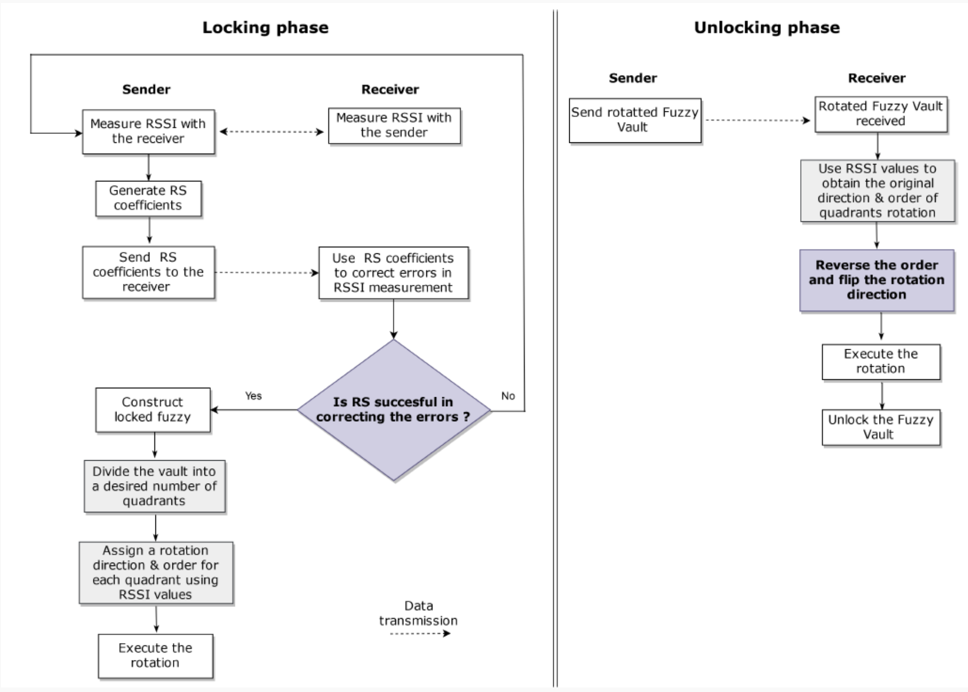
IEEE CONSUMER TECHNOLOGY SOCIETY (CTSOC)

ISSUE: MAY 2022

CTSOC-NCT NEWS ON CONSUMER TECHNOLOGY



Rotation Assisted Fuzzy Vault (RAFV): Method to Protect Fuzzy Vault Construction

TABLE OF CONTENTS

EDITOR'S NOTE

On behalf of the Editorial Board of IEEE CTSoc News on Consumer Technology (NCT) editor-in-chief Wen-Huang Cheng, and editors, Yafei Hou, Luca Romeo and Jianlong Fu, I am pleased to introduce the May issue of the News on Consumer Technology (NCT).

This issue starts with a cover story that introduces a scheme to secure fuzzy vault enabled authentication in body area networks-based smart healthcare, published in IEEE Consumer Electronics Magazine. This work proposed the Rotational Assisted Fuzzy Vaults (RAFV) scheme that strengthens the security of any authentication solution using the fuzzy vault construction approach with improved communication overhead while not compromising security.

This is followed by an interview with Prof. Yong Man Ro from Korea Advanced Institute of Science and Technology (KAIST), South Korea where he shared his perspectives, experiences, and advice on conducting research in the topics of human interpretable deep leaning, and human-oriented multi-modal applications

Finally, this issue ends with a feature article brought by Dr. Pierre Bonnet from The French Agricultural Research Centre for International Development (CIRAD), Dr. Alexis Joly from the National Institute for Research in Digital Science and Technology (Inria) in France, and Dr. Sue Han Lee from Swinburne University of Technology. The article introduces the Pl@ntNet project, an award-winning interdisciplinary research project that supports biodiversity through a collaborative platform based on deep learning that engages the general public as active contributors. The project has been ongoing for more than 10 years to-date with active web and mobile applications for automated image-based plant species classification that anyone can use and contribute as citizen scientists.

Have a nice read!

Yuen Peng Loh
Editor of NCT

Yafei Hou

ARTICLE TITLE
Securing Fuzzy Vault Enabled Authentication in Body Area Networks-Based Smart Healthcare

AUTHOR(S)
Jack Hodgkiss; Soufiene Djahel

JOURNAL TITLE
IEEE Consumer Electronics Magazine

JOURNAL VOLUME AND ISSUE
Volume: 11, Issue: 1, Jan. 1 2022

DATE OF THE ARTICLE
07 May 2020

PAGE NUMBERS FOR THE ARTICLE
6 - 16

The emergence of body area networks (BANs) has largely used for real-time sensing of human biometrics and for remote control of smart wireless devices, which in turn are beginning to revolutionize healthcare industry. However, BANs are vulnerable to a myriad of security attacks. Due to their limited power and computational capabilities, the processor-intensive cryptographic techniques need to be avoided. This article aims to the authentication service for BAN sensors and proposes a scheme named "RAFV: Rotational Assisted Fuzzy Vaults" to harden the security of any authentication solution using the fuzzy vault construction approach. The evaluation results have shown it can successfully conceal the secret of the vault even if the locking elements are known to the adversary. Also, it may improve upon communication overhead without compromising its security and less additional computational overhead.

INTERVIEW WITH RESEARCH TEAM OF HUMAN MULTIMODAL LEARNING IN KAIST, SOUTH KOREA

Professor Yong Man Ro
School of Electrical Engineering,
Korea Advanced Institute of Science and Technology (KAIST), Korea

Yong Man Ro is a Professor at School of Electrical Engineering, KAIST (Korea Advanced Institute in Science and Technology), South Korea. Prof. Ro established Image and video systems (IVY) Laboratory and the center for applied research in AI (CARAI) at KAIST in 1997 and 2019, respectively. Among the years, his research team has been conducting research in a wide spectrum of image and computer vision research topics. Recent research interests are deep learning, machine learning in computer vision and image processing, multimodal learning, explainable AI and Robust AI. He and his team has produced over 140 SCI indexed papers and 329 international conference papers.

Human Multimodal Research Team: (Left to right) Minsu Kim (PhD candidate), Jeung Hun Yeo (PhD candidate), Joanna Hong (PhD candidate), Prof. Yong Man Ro, Jeongsoo Choi (PhD candidate), Chae Won Kim (MS candidate), Hyun Jun Kim (PhD candidate), Bella Godiva(MS candidate), Se Jin Park (PhD candidate)

What are your team's main objectives and research topics?

In my recent research team in KAIST, main objectives are to design and develop creative and efficient learning methods and neural architectures to provide useful deep representations of multi-modal data for diverse applications. In particular, we are conducting research to develop advanced neural learning methods such as self-supervised and unsupervised, to put the technologies into more practical situations beyond the supervised learning which requires large human resources. Also, we try to provide a creative method to utilize rich multi-modal information even if some modalities are not available during inference, in order to give the consumers more powerful performances.

The main research topics of my team are developing human interpretable deep learning models, human-oriented multi-modal applications such as audio-visual speech recognition, talking face generation, speech synthesis, and multimodal dialogue systems.

In your opinion, why are those research topics interesting for you compared to other AI topics?

Research topics covered by my team are conducted with the hope of producing useful impacts on the community. My team works along with industry and government to keep being updated with the current demand of the real world. For me, research would be more interesting if it is not only kept on a paper but also be able to be applied in the community and contribute to the improvement of people's life quality. We believe that AI would become more meaningful when it can work along with humans. Therefore, we need to equip AI with senses that human has like audio, vision, and language to enable them to communicate with humans. This goal is hoped to be achieved through human-oriented multimodal research.

What are the biggest obstacles you're facing with your current research topics, and how do you deal with them?

Over the years, machine learning technology has evolved, and the size of the model has become wider and deeper. In the multimodal field, in particular, the amount of data required is even greater because the model is more complex and numerous than the single model. At the same time, there is a condition that data of various domains must be well-aligned for multimodal learning. These data are hard to obtain naturally in the real world, and there are many difficulties in collecting them.
The simplest and obvious way is to collect and label more real-world data. It is very time-consuming and expensive work, but it is something that all researchers must do consistently. Therefore, we need to cooperate with the various institution and research teams to do it efficiently. Another way is data generation. It is difficult to generate data, but once set, it is much cheaper and easier to collect data in massive quantities. In addition, we can generate data that can never be obtained naturally, so it is more helpful for various data augmentation. Finally, we adopt few-shot learning techniques to overcome data sparsity. Nowadays, many learning techniques for few-shot learning such as meta-learning and transfer learning have developed a lot, so if these methods are properly applied, effective learning is possible even on limited data.

Due to the heterogeneous nature of data, combining modalities in multimodal learning is problematic. Do you have any strategy for combining them effectively?

Multimodal learning is difficult due to the heterogeneous nature of data. Reducing heterogeneity gap, induced by inconsistent distribution of different modalities, is still considered as a challenging problem in various task (e.g. metric learning and knowledge distillation in multimodal).
To avoid heterogeneity gap issue in multimodal learning, we utilize multi-modality associative bridging through memory. The memory network maps each multimodal feature to each latent space. To use multimodal data without gap, the features are saved in each space. And, we can load the other modal data using one of the multimodal data. Therefore, learning one-to-one mapping through memory, can combine multimodal data effectively.

Human multimodal learning is very useful for many applications. For practical, a high-performance and efficient model is highly recommended. Could you tell us a bit about a robust multimodal learning system?

Human-oriented robust multimodal system contains human-AI dialog system which utilizes multimodal information: text, video, sound, etc. Using multimodal information, various robust techniques have been developed: speech enhancement and recognition system, speech reconstruction system, facial video generation system, automatic human-AI dialog system, and so on.
A robust multimodal learning system may require training with huge and diverse datasets. Since multimodal dataset has complex characteristics due to its diverse domain condition, combining all the modalities in the same representation space would provide the model's robustness. To do so, we also need multi-task and transfer learning among various dataset with various modalities.

Currently, Covid-19 has a huge impact on human life change. If you have an unlimited research budget from any source, what kinds of AI applications or systems will you develop to help people stay safe from Covid-19?

Because of Covid-19, Many people are reducing human contact and increasing their time at home. We will develop an audiovisual analysis system to help facilitate communication with others remotely within the home. The system contains audiovisual speech recognition in the wild using deep learning and this can help communication in the way of reducing the influence of noisy environment.
Also, demand for video conferencing has grown rapidly since advent of the advent of Covid-19. We will be able to develop technologies about generating talking face from speech information in real time using deep learning for both keeping privacy of participants and providing video conferencing service stably by reducing the amount of information transmitted.

In the last 10 years, technology has rapidly advanced, including Artificial Intelligence (AI). From your perspective, what will AI and deep learning be in the next 10 years?

AI evolution is approaching faster than we think. At the moment, AI is on the front lines of the industry. It's becoming rare to find workplaces that do not apply AI in their

work. Various lines of AI research are pouring out every year, and the industry is rapidly turning the research into useful technologies. I'm very excited about what would come in the next 10 years and how much more helpful it will turn out to be. It will become a part of not just an organization, but our daily lives in a wider scope and at a lower cost easily accessible for everyone.
As it integrates into most aspects of our lives and becomes pervasive, all of us should prepare for such transformations so that no one is marginalized from the educated. For AI applications to become reality, a user-friendly interface will have to be developed along with the AI research. Also, we should utilize AI as an assistant, more like a tool to augment human skills. Thus, we would have to make choices on restrictions of the use that benefit us so that it does not yield privacy and cyber-related crimes.

Do you have any encouragement messages for young scientists that are interested in your research field?

The field of AI has shown remarkable progress and results in single modality realms. However, research not only remains in single modalities, but is expanding to using various modalities, combining language, vision and audio. In parallel to the development of multimodal learning, datasets are becoming more large-scale and high resolution to support multiple modalities. By learning joint representations and aligning these modalities, we are able to capture more complex and meaningful information compared to using single modalities. Nevertheless, the real-world contains even more data beyond textual, audio and visual senses, and multimodal learning will become an important stepping stone to representing the real-world in the future. Researching multimodal learning can become burdensome and tedious, especially if you are in academia and have limited means, as training requires increasingly more resources. However, if you have the passion and patience to discover and solve problems, you will be naturally guided during the process to make an impact on the real-world. It is important to keep up with the rapidly changing field of research, but also be patient on working with projects. Take it slow, but pursue with passion!

PL@NTNET: HARNESSING THE POWER OF ARTIFICIAL INTELLIGENCE FOR BIODIVERSITY CONSERVATION

Pierre Bonnet, Ph.D.
pierre.bonnet@cirad.fr
Permanent researcher (researcher in Biodiversity informatics)
Cirad-Bios, Amap mixt research unit, Montpellier, France

Alexis Joly, Ph.D.
alexis.joly@inria.fr
INRIA Research Director (DR 2), INRIA, ZENITH team, France

Sue Han Lee, Ph.D
shlee@swinburne.edu.my
Lecturer, Faculty of Engineering, Computing and Science, Swinburne University of Technology

In the age of advancing technology, artificial intelligence systems have become increasingly ubiquitous to many aspects of society. Not only were such technology researched and developed for the benefit and convenience of regular consumers, various initiatives have also been put in motion for the benefit of various societal challenges such as the environment. Pl@ntNet[1] is one such initiative that was launched as early as 2009 as a theoretical research project to experiment image recognition techniques to monitor biodiversity. It was established through an interdisciplinary collaboration between several French research centres in science and technology and biodiversity, namely The French Agricultural Research Centre for International Development (CIRAD), the French National Institute for Research in Digital Science and Technology (INRIA), the French Research Institute for Agriculture, Food and the Environment (INRAE), and the French Research Institute for Development (IRD).

The History

The project started out with the focus on developing image recognition algorithms for plant species identification and in particular with the search for similarities and more recently supported by deep learning convolutional neural networks technology with a focus on user engagement and collaboration. Between the years 2011 and 2013, Pl@ntNet officially launched what is termed as "Citizen Science Platforms" with the release of the first web application (Goëau et al., 2011) followed by the mobile application (Goëau et al., 2013) respectively. In 2015, with the rise deep learning technology, the researchers' work has been accelerated and the Pl@ntNet team was among the first one to develop a plant identification system with a scale as large as 8,000 species. The citizen science aspect is in reference to the availability of the platforms to not only botanists, but also free to use by the general public as active users and contributors. The free app

[1] https://plantnet.org/en/

that can be downloaded into mobile devices enables anyone to take a photo of a plant for identification, explore from a list of potential species that might match as well as "vote" or validate the seemingly correct species.

Subsequently in 2018, Pl@ntNet successfully conducted the first ecological study using the data contributed by citizens using the app (Botella et al., 2018). In the study, species distribution modelling of invasive alien plants based on citizen collected data were compared to data inventories made by experts with reasonable predictive effectiveness for some species. As of 2020, Pl@ntNet is also widely used for teaching and raising awareness of biodiversity conservation and agroecology (Bonnet et al., 2020), for instance early prediction of invasive species through species distribution monitoring. Moreover, Pl@ntNet data has been integrated into the Global Biodiversity Information Facility (GBIF) database. In 2021, after more than 10 years from the official launch of the citizen science platform, Pl@ntNet received the Inria – French Academy of Sciences – Dassault Systems Innovation Award which honours research teams that have been active in issues of transfer and innovation in the field of computer science and mathematics. This is in recognition of the research that has supported biodiversity with the collaborative platform based on deep learning for plant identification.

The Applications

The core focus of the project at the beginning has been the development of image recognition algorithms for plant species identification and in particular with the search for similarities and more recently supported by deep learning Convolutional Neural Network (CNN) technology. This is then integrated into application platforms for ease of use. The approach of the platform is to not only provide the name of the plant to be identified, but also to provide a list of most similar plant observations to the ones submitted by the end users in order to have more engagement with them. The users are therefore involved in the process because they will be presented with choices to select the correct plant if necessary. This in turn supports a continuous improvement of the backend algorithm.

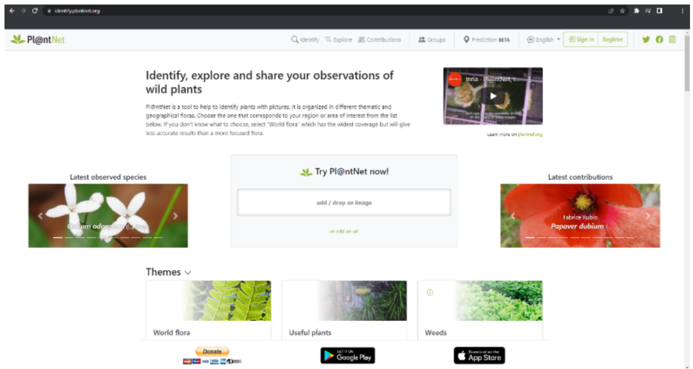


Fig. 1. Pl@ntNet web app interface for plant species recognition from an image.

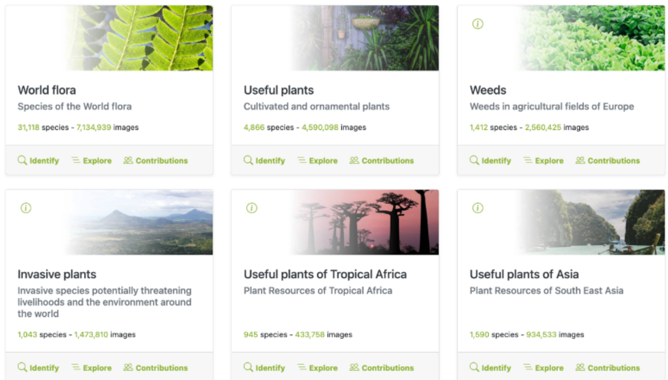


Figure 2. Graphic interface of Pl@ntNet web app showing the thematic floras covered by the citizen science platform.

Figures 1 and 2 shows the web application interface of the platform for the recognition as well as exploration. The mobile application as shown in Figure 3, is relatively more convenient on the field whereby users are able to easily conduct identification from their mobile devices on the fly. The recognition can be performed using images of different organs of the plant, such as the flower, leaf, fruit, as well as the bark. The application then engages the user to validate the recognition results among a list similar plants, if possible. On the other hand, other users are also able to validate the contributions of each other in a collaborative manner either in a personal capacity to verify the correctness of the images in the database through exploration, or to participate in groups of similar interests. This collaborative platform fosters the spirit of community among users from different walks of life to exchange knowledge of a common interest in biodiversity. The application has been fast growing, where a few years ago there are up to 20 million users of the app with about 500,000 uses per day during spring peaks, and still increasing.

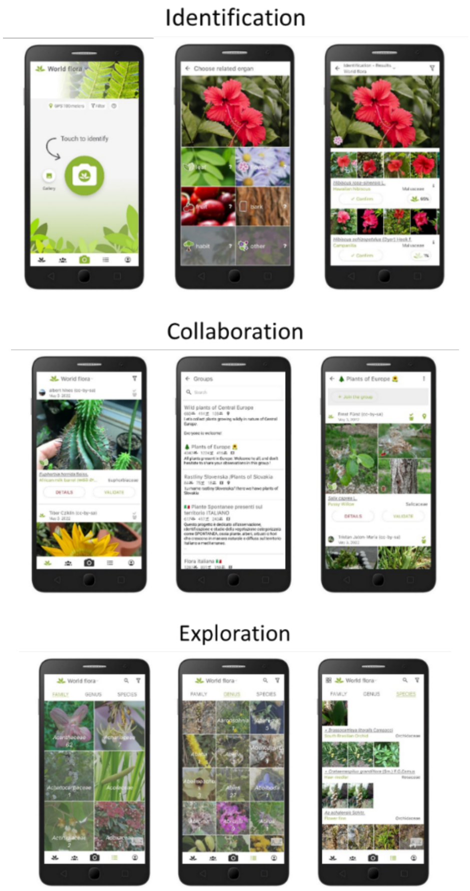


Figure 3. Mobile application of Pl@ntNet provides an engaging interface for users to perform identification of plant species, explore plants in the database, and collaborate with other users in the community.

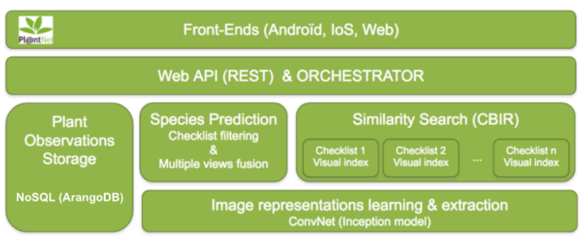


Figure 4. Pl@ntNet system architecture at the beginning of 2017 (Affouard et al., 2017).

The Backend

The Pl@ntNet system architecture can be divided into five main modules as shown in Figure 4 (Affouard et al., 2017). The first is the Plant Observation Storage that contains plant (images of flowers, leaf, etc., species) and provenance (device, user, etc.) data as well as geo-locations of observations, all store within Arango (NoSQL document storage). The Image representation learning and extraction engine is a CNN that is periodically trained using validated observations that were collected from the app users for continuous model improvement. Next is the Species Prediction module that uses the aforementioned CNN for prediction when observations (images) are supplied by the end user through the apps. Other than a single image-based prediction, a weighted average approach has been incorporated for conditions where the user is able to supply images of various organs of a single plant in order to “fuse” the predictions for optimal results of the most probably species. Following that, the feature vector from the CNN is also extracted to support the Similar Search module, specifically Content-Based Image Retrieval (CBIR) operation that uses the compressed binary code of the feature vector from unsupervised hashing for nearest neighbours search. Figure 5 illustrates the concept of the CNN for species classification, and the feature vector extracted to perform CBIR. Finally, the Representational State Transfer (REST)-ful web API is used to manage the data exchanges between the server and front-ends. Users are able to access the rich information such as images available in the database of the plants and species, or descriptive web pages to provide further information.

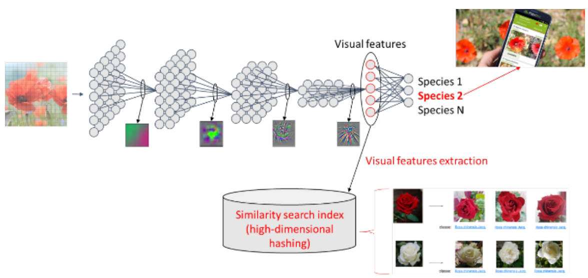


Figure 5. Conceptual illustration of plant species image recognition technology. Convolutional Neural Network is trained for species classification and visual features from the CNN is extracted for similarity search.

Now and the Future

Currently, Pl@ntNet has been adapted to multiple regions including the Americas, Africa, Oceania and Asia, a vast expansion from the initial coverage of Western European flora at the start of the project, and now includes up to 38,045 plant species. To broaden the coverage of the application to remote areas such as dense evergreen forests or high mountain environments, the Pl@ntNet team is currently experimenting with dedicated technologies for offline use of the identification search engine, as people living in remote areas have less access to new technologies (such as recent smartphones, broadband internet, etc.). This is important to mitigate geographical bias in order to ensure a more homogeneous geographical coverage for biodiversity monitoring.

As part of the Cos4cloud[2] European research project, Pl@ntNet is also investing in making its plant species identification service available in the form of an API[3]. This new service has already enabled nearly 5 thousand researchers and app developers to test it and/or integrate it into their own data workflow. These new uses of this service could very soon be available through a wide range of connected sensors or autonomous robots.

REFERENCES

[1] Goëau, H., Joly, A., Selmi, S., Bonnet, P., Mouysset, E., Joyeux, L., Molino, J.F., Birnbaum, P., Bathelemy, D. and Boujemaa, N., 2011, November. Visual-based plant species identification from crowdsourced data. In Proceedings of the 19th ACM international conference on Multimedia (pp. 813-814).

[2] Goëau, H., Bonnet, P., Joly, A., Bakić, V., Barbe, J., Yahiaoui, I., Selmi, S., Carré, J., Barthélémy, D., Boujemaa, N. and Molino, J.F., 2013, October. Pl@ntnet mobile app. In Proceedings of the 21st ACM international conference on Multimedia (pp. 423-424).

[3] Affouard, A., Goëau, H., Bonnet, P., Lombardo, J.C. and Joly, A., 2017, April. Pl@ntnet app in the era of deep learning. In ICLR: International Conference on Learning Representations. Toulon, France.

[4] Botella, C., Joly, A., Bonnet, P., Monestiez, P. and Munoz, F., 2018. Species distribution modeling based on the automated identification of citizen observations. Applications in Plant Sciences, 6(2), p.e1029.

[5] Bonnet, P., Joly, A., Faton, J.M., Brown, S., Kimiti, D., Deneu, B., Servajean, M., Affouard, A., Lombardo, J.C., Mary, L. and Vignau, C., 2020. How citizen scientists contribute to monitor protected areas thanks to automatic plant identification tools. Ecological Solutions and Evidence, 1(2), p.e12023.

[2] https://cos4cloud-eosc.eu/

[3] https://my.plantnet.org/