

Singapore Management University

Institutional Knowledge at Singapore Management University

Research Collection School Of Computing and
Information Systems

School of Computing and Information Systems

10-2021

AlxFood'21: 3rd workshop on AlxFood

Ricardo GUERRERO

Michael SPRANGER

Shuqiang JIANG

Chong-wah NGO

Singapore Management University, cwngo@smu.edu.sg

Follow this and additional works at: https://ink.library.smu.edu.sg/sis_research



Part of the [Databases and Information Systems Commons](#), and the [Graphics and Human Computer Interfaces Commons](#)

Citation

GUERRERO, Ricardo; SPRANGER, Michael; JIANG, Shuqiang; and NGO, Chong-wah. AlxFood'21: 3rd workshop on AlxFood. (2021). *Proceedings of the 29th ACM International Conference on Multimedia, MM 2021, Virtual Conference, October 20-24*. 5688-5689.

Available at: https://ink.library.smu.edu.sg/sis_research/6629

This Conference Proceeding Article is brought to you for free and open access by the School of Computing and Information Systems at Institutional Knowledge at Singapore Management University. It has been accepted for inclusion in Research Collection School Of Computing and Information Systems by an authorized administrator of Institutional Knowledge at Singapore Management University. For more information, please email cherylids@smu.edu.sg.

AlxFood'21: 3rd Workshop on AlxFood

Ricardo Guerrero

Samsung AI Center

Cambridge, UK

r.guerrero@samsung.com

Shuqiang Jiang

Chinese Academy of Sciences

Beijing, China

sqjiang@ict.ac.cn

Michael Spranger

Sony Corporation

Tokyo, Japan

Michael.Spranger@sony.com

Chong-Wah Ngo

Singapore Management University

Singapore

cwnngo@smu.edu.sg

ABSTRACT

Food and cooking analysis present exciting research and application challenges for modern AI systems, particularly in the context of multimodal data such as images or video. A meal that appears in a food image is a product of a complex progression of cooking stages, often described in the accompanying textual recipe form. In the cooking process, individual ingredients change their physical properties, become combined with other food components, all to produce a final, yet highly variable, appearance of the meal. Recognizing food items or meals on a plate from images or videos, their physical properties such as the amount, nutritional content such as the caloric value, food attributes such as the flavor, elucidating the cooking process behind it, or creating robotic assistants that help users complete that cooking process, is of essential scientific and technological value yet technically extremely challenging. The 3rd AlxFood workshop was held as a half-day workshop in conjunction with the 29th ACM International Conference on Multimedia (ACM MM 2021), in Chengdu, China and virtually.

CCS CONCEPTS

• **Computing methodologies** → **Machine learning algorithms**;
• **Mathematics of computing** → *Probability and statistics*; • **Information systems** → *Information retrieval*.

KEYWORDS

Computational food analysis, multimodal, food

ACM Reference Format:

Ricardo Guerrero, Michael Spranger, Shuqiang Jiang, and Chong-Wah Ngo. 2021. AlxFood'21: 3rd Workshop on AlxFood. In *Proceedings of the 29th ACM International Conference on Multimedia (MM '21), October 20–24, 2021, Virtual Event, China*. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3474085.3478573>

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

MM '21, October 20–24, 2021, Virtual Event, China

© 2021 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-8651-7/21/10.

<https://doi.org/10.1145/3474085.3478573>

1 INTRODUCTION

Food and cooking are an ubiquitous and fundamental part of human life, that not only provides sustenance, but it is also associated with well-being and culture. Food analysis presents exciting research and application challenges for modern AI systems, particularly in the context of multimodal data such as images or video, in an attempt to automate food understanding.

Many fundamental AI and machine learning problems, such as fine-grained object category recognition, multi-label structured prediction, image segmentation, cross-modal analysis, retrieval, and captioning, activity understanding and recognition, natural language understanding, linguistic and visual-linguistic reference resolution, multimodal grounding, task representation, inference and planning, multimodal knowledge representation and extraction, human-robot interaction, and design of novel user interfaces & user experiences, are inherently present in the context of food AI.

The AlxFood Workshop aims to include both academic and industry perspectives, and targets researchers and practitioners from any field of AI with an interest in food and cooking. Any analytic methods developed in this context have the potential to advance the state-of-the-art in AI as well as revolutionize how people interact with food, including improving health, wellness, and the food preparation itself. A growing health crisis in the US related to diet is further exacerbated by our aging population and sedentary lifestyles. Six of the ten leading causes of death in the United States, including cancer, diabetes, and heart disease, can be directly linked to diet. Therefore, AlxFood Workshop is a timely event as it aims to bring together domain experts in nutrition, as invited speakers, and computational scientists. The AlxFood 2021 Workshop is the third such workshop after AlxFood 2019/2020 were held in conjunction with IJCAI.

2 TOPIC COVERED

The AlxFood workshop series is opened to anyone who works on the intersection between AI and food. Therefore, a non-exhaustive list of covered topics and in no particular order is as follows:

- Multimodal food recognition, meal to ingredient level to recipe
- Multimodal estimation of food & ingredient amounts
- Multimodal estimation of meal attributes, including flavor
- Cross-modal food and recipe retrieval
- Planning in the cooking domain for recipe and dish creation

- AI and ML in novel recipe generation (computational creativity)
- Food domain knowledge representation: Recipes, cooking procedure, ingredient interaction, flavour, nutrition, and health facts
- Knowledge extraction for food domain knowledge: ontologies, recipes
- Multimodal grounding and reference resolution
- Cooking process analysis from video and images
- Interfaces for food related AI systems. UX and Interactive agents for multimodal food logging
- Food manipulation and cooking skill learning for robots (robot learning)

3 RELATIONSHIP TO PREVIOUS WORKSHOPS

The first and second AIXFood workshops were successfully organized in IJCAI'2019 and IJCAI'2020. Dr. Spranger, who is on the organizing committee of this workshop, were also workshop organizers of AIXFood 2019 and 2020.

Other recent related workshops were held at ICPR2020, ACM-MM2019 and ECAI2018, notably, the 6th and 5th International Workshop on Multimedia Assisted Dietary Management, and the 10th Workshop on Multimedia for Cooking and Eating Activities.

4 WORKSHOP ORGANIZERS

Ricardo Guerrero. is a Senior Researcher at Samsung AI Center based in Cambridge, UK. There he undertakes research on computational food analysis, such as cross-modal food understanding, synthesis and retrieval. His main research interests include computational food analysis, machine learning, computer vision, medical image processing and multimodal data analysis, with applications to fundamental problems of visual representation and recognition, tracking, action recognition, and human-computer interaction.

Michael Spranger. is a Deputy General Manager and Senior Researcher at Sony Corporation based in Tokyo, Japan. There, he is leading R&D for cooking robots, in particular, knowledge extraction from images and text, as well as skill acquisition for cooking manipulation. Michael has co-organized various Workshops and tutorials relevant for food such as a Workshop on Shared Autonomy (IJCAI/ECAI-18) as well as tutorials on natural language interaction for robots (IJCAI-15) and neural-symbolic reasoning and learning (IJCAI/ICML-18).

Shuqiang Jiang. is a Professor at Chinese Academy of Sciences. His research interests include multimedia content analysis and retrieval, image/video understanding, and multimodal intelligence. He has authored or coauthored more than 150 papers. He was supported by NSFC Excellent Young Scientists Fund in 2013, Young top-notch talent of Ten Thousand Talent Program in 2014. He is the senior member of IEEE and CCF, member of ACM, Associate Editor of IEEE Multimedia, Multimedia Tools and Applications. He is the vice chair of IEEE CASS Beijing Chapter and vice chair of ACM SIGMM China chapter. He is actively involved in research of food and cooking analysis.

Chong-Wah Ngo. is a Professor at Singapore Management University. His research interests are in multimedia search and computing. He has been serving the technical program committees of numerous multimedia and information retrieval conferences including ACM Multimedia (MM), ACM SIGIR, International Conf. on Multimedia Retrieval (ICMR) and International Conf. on Multimedia and Expo (ICME). In addition, he was an Associate Editor of IEEE Trans. on Multimedia (2011-14), and is on the editorial board of IET Computer Vision and ITE Trans. on Media Technology and Applications. He is on the steering committee of TRECVID and ICMR (Int. Conf. on Multimedia Retrieval); conference co-chair of PCM 2018, ICMR 2015; program co-chairs of ACM MM 2019, MMM 2018, ICMR 2012, MMM 2012 and PCM 2013. He is founding leader of VIREO (Video REtrieval grOUp). He also served as the chairman of ACM (Hong Kong Chapter) during 2008 and 2009. He is actively involved in research of food and cooking analysis.

5 PROGRAM COMMITTEE

- Mikhail Fain, independent.
- Uchenna Akujuobi, Sony Corporation.
- Paolo Napoletano, University of Milano - Bicocca, Italy.
- Fangda Han, Rutgers University, USA.
- Kiyoharu Aizawa, The University of Tokyo, Japan.
- Donghyeon Park, Korea University, South Korea.
- Fengqing Zhu, Purdue University.
- Keiji Yanai, UEC Tokyo, Japan.
- Hai X. Pham, Samsung AI Center - Cambridge, UK.
- Minyoung Kim, Samsung AI Center - Cambridge, UK.

6 KEYNOTE SPEAKERS

Vladimir Pavlovic. is a Professor of Computer Science at Rutgers University in New Jersey, USA, and a Principal Scientist and the Director of Future Interactions at the Samsung AI Center in Cambridge, UK. Vladimir's research interests include probabilistic machine learning, multimodal representation learning, and next generation human sensing. Over the past twenty years Vladimir has published extensively in the domains of computer vision and human-computer interaction, including his seminal works on human gesture modeling and recognition, human motion analysis, and non-verbal human affect understanding. At Rutgers, he co-leads the Center for Accelerated Real Time Analytics (CARTA), <https://carta.rutgers.edu>, is a member of the Executive Committee of Rutgers Center for Cognitive Science (RUCCS), <https://ruccs.rutgers.edu>, and an associate member of the Center for Quantitative Biology (CQB), <https://cqb.rutgers.edu>, and Management Science and Information Systems at Rutgers Business School. Vladimir received his Ph.D. in Electrical and Computer Engineering from the University of Illinois at Urbana-Champaign.

Alex Wiltchko. is a senior research scientist at Google Brain. He obtained his PhD in neurobiology from Harvard Medical School, where he built new methods for understanding and parsing behavior and body language. He then co-founded Whetlab, an ML-as-a-service company, which was acquired by Twitter in 2015. At Google Brain, Alex works on new tools for machine learning, as well as applying ML to problems in biology and chemistry, e.g. bringing the sense of smell to computers.