3rd International Conference on
Activity and Behavior Computing (ABC)

Oct. 22nd - 23rd, 2021, Bangkok, Thailand
(**Changed to Online**)

https://abc-research.github.io/

abc@sozolab.jp

Final Program – at a glance: Times are in JST (+9hr GMT)

Detailed Program with ABSTRACTs from Page 12 ~

Oct. 22nd (Fri.), 2021

OPENING, 13:00-13:10

SESSION 1: REGULAR PAPERS, 13:10-14:10

SESSION 2: CHALLENGE + AWARDS, 14:20-15:50

SESSION 3: REGULAR PAPERS, 16:00-17:00

KEYNOTE 1, 17:00-18:00

NETWORKING, 18:00-18:20

Oct. 23rd (Sat.), 2021

SESSION 4: REGULAR PAPERS, 13:00-14:00

SESSION 5: WORK IN PROGRESS, 14:10-15:00

SESSION 6: PANEL: THE FUTURE OF ACTIVITY AND BEHAVIOR COMPUTING, 15:10-17:10

KEYNOTE 2, 17:10-18:10

AWARDS + CLOSING, 18:10-18:30

NETWORKING, 18:30-18:50

About

Focusing on the vision-based and sensor-based recognition and analysis of human activity and behavior, this book gathers selected papers presented at the 3rd International Conference on Activity and Behavior Computing (ABC 2021), https://abc-research.github.io/ - on October 22-23, 2021. The respective chapters cover action recognition, action understanding, behavior analysis, and related areas. Human-robot interaction is continuously evolving in industrial context for the betterment of services as well as ease and safety of human lives. Human-robot collaboration (HRC) is an important area for the future. Hence, with a goal of recognizing human activities in HRC context, a new challenge is thrown in this event called "Bento Packaging Activity Recognition Challenge", where a dataset collected in a HRC context is introduced. For the dataset collection, an environment of Lunch-box (widely known as 'bento' in Japan) packaging on a moving conveyor belt performed by human subjects is created. The challenge participants were asked to predict the activities of these human subjects during the packaging. Selected challenge papers are included in this book. The book addresses various challenges and aspects of human activity recognition in both the sensor-based and vision-based domains, making it a unique guide to the field.

Message

Focusing on the vision-based and sensor-based recognition and analysis of human activity and behavior, this book gathers selected papers presented at the 3rd International Conference on Activity and Behavior Computing (ABC 2021), https://abc-research.github.io/ - on October 22-23, 2021. The respective chapters cover action recognition, action understanding, behavior analysis, and related areas. The 3rd Activity and Behavior Computing conference aims to be a venue to discuss some of the facets of computing systems, which are able to sense, recognize, and eventually understand human activities, behaviors, and the context within which they occur. This in turn enables a wide range of applications in a large variety of disciplines.

We received 21 submissions. After rigorous review by related top experts (number of reviewers per paper: min. 2 to maximum 5), and review rebuttal process, 9 papers were accepted and included in this book. The process is summarized as follows: "The program chairs along with the general chairs had a rigorous meeting to decide the fate of the papers at the first round. After the rebuttals, another meeting was arranged and the final decisions are made. Even after the final decision, the program committee did a few stages of minor editing for some papers to enrich the quality of the chapters. The resulting selection of papers reflects the broad interests of this nascent community and is important to take stock of the state of the art and continued research challenges in the field." A reviewer had to answer several questions and hence the review process was rigorous; e.g., Contribution to ABC, Decision, What category best fits this paper?, Originality, Revision Requirements, and Your level of expertise on this subject.

Human-robot interaction is continuously evolving in industrial context for the betterment of services as well as ease and safety of human lives. Now-a-days we can see increasing collaborative works between humans and robots in almost every industry, from automobile manufacturing to food packaging. Despite the fact that the collaboration is making the overall industrial processes optimized and flexible, activity recognition during these processes is becoming more complex. Furthermore, to improve the interactions between human and robot, it is necessary for the robots to have a better understanding of human behavior and activities.

In recent years, there have been many works on identifying complex human activities and more research focusing on identifying human activities in Human-robot collaboration (HRC) scenario. Research in this area can result in building a better system. So, with a goal of recognizing human activities in HRC context, in Bento Packaging Activity Recognition Challenge we shared a dataset collected in a HRC context. For the dataset collection, we created an environment of Lunch-box (widely known as 'bento' in Japan) packaging on a moving conveyor belt performed by human subjects. The challenge participants were asked to predict the activities of these human subjects during the packaging. Apart from the regular 9 papers, 7 selected papers for this challenge are included along with the summary of the challenge.

We gladly thank the authors for their contributions, the reviewers for their time to review rigorously. We are thankful to other members and advisors of this conference. There are two keynote speakers who delivered the talks. We are thankful to them:

- Björn W. Schuller, Fellow IEEE (Imperial College London/UK and University of Augsburg/Germany) – for his talk on "Multimodal Sentiment Analysis: Explore the No Pain, Big Gain Shortcut"; and
- Koichi Kise (Osaka Prefecture University/Japan) – for his talk on "Reading of Reading for Actuating: Augmenting Human Learning by Experiential Supplements".

We are thankful to the 11 Panel Speakers on the "The Future of Activity and Behavior Computing", and the Session Chairs for their valuable time. We thank the volunteers and the participants for their efforts to make it a wonderful event.

Finally, this book presents some of the latest researches in the field of activity and behavior computing. We hope that it will serve as a reference for researchers and practitioners in academia and industry related to human activity and behavior computing. We look forward to meet you all in the upcoming International Conference on Activity and Behavior Computing (ABC) in the coming years! Stay safe and stay strong.

Best regards,

Osaka, Japan	Md Atiqur Rahman Ahad
Fukuoka, Japan	Sozo Inoue
Brighton, UK	Daniel Roggen
Tokyo, Japan	Kaori Fujinami

Committee

GENERAL CHAIRS
Sozo Inoue, Kyushu Institute of Technology / AUTOCARE LLC, Japan
Md Atiqur Rahman Ahad, University of Dhaka, Bangladesh / Osaka University, Japan

PROGRAM CHAIRS
Daniel Roggen, University of Sussex, UK
Kaori Fujinami, Tokyo University of Agriculture and Technology, Japan

ADVISORY BOARD
Paul Lukowicz, DFKI, Germany
Kristof Van Laerhoven, University of Siegen, Germany
Daniel Roggen, University of Sussex, UK
Michael Beigl, Karlsruhe Institute of Technology, Germany
Jeffry Cohn, University of Pittsburgh, USA
Anind Dey, University of Washinton, USA
Nobuo Kawaguchi, Nagoya University, Japan

PUBLICITY CHAIRS
Paula Lago, Universidad Nacional Abierta y a Distancia, Colombia
Anja Exler, Adastra, Germany
Hiroki Watanabe, Hokkaido University, Japan

PUBLICATION CHAIR
Tahera Hossain, Aoyama Gakuin University, Japan

LOCAL ORGANIZATION CHAIR
Phond Phunchongharn, KMUTT, Thailand

PROGRAM COMMITTEE
Ren Ohmura, Toyohashi University of Technology, Japan
Kazuya Murao, Ritsumeikan University, Japan
Kristof Van Laerhoven, University of Siegen, Germany
Susanna Pirttikangas, University of Oulu, Finland
Philipp M. Scholl, University of Freiburg, Germany
Hristijan Gjoreski, Ss. Cyril and Methodius University, Macedonia
Katsuhiko Kaji, Aichi Institute of Technology, Japan
Tsuyoshi Okita, Kyushu Institute of Technology / Riken, Japan
Paula Lago, Universidad Nacional Abierta y a Distancia, Colombia
Stephan Sigg, Aalto University, Finland

Hiroki Watanabe, Hokkaido University, Japan
Pascal Hirmer, University of Stuttgart, Germany
Shigeyuki Miyagi, The University of Shiga Prefecture, Japan
Yu Enokibori, Nagoya University, Japan
Anja Exler, Adastra, Germany
Kristina Yordanova, University of Rostock, Germany
Trung Thanh NGO, Osaka University, Japan
Upal Mahbub, University of Meryland, College Park, USA
Atsushi Shimada, Kyushu University, Japan
Guillaume Lopez, Aoyama Gakuin University, Japan
Atsushi Nakazawa, Kyoto University, Japan
Liming Chen, Ulster University, UK
Christopher Nugent, Ulster University, UK
François Charpillet, INRIA, France
Maria Del Pilar Villamil Giraldo, University of Los Andes, Colombia
Takuro Yonezawa, Nagoya University, Japan
Anna Yokokubo, Aoyama Gakuin University, Japan
Yuuki Nishimura, University of Tokyo, Japan

SECRETARIES
Nattaya Mairittha, Kyushu Institute of Technology, Japan
Tittaya Mairittha, Kyushu Institute of Technology, Japan

Highlighting Points

- Presents some of the latest researches in the field of activity and behavior computing
- Gathers extended versions of selected papers presented at 3rd ABC 2021, 22-23 Oct. 2021
- Serves as a reference for researchers and practitioners in academia and industry

Keynote Speech 1

BJÖRN W. SCHULLER, FELLOW IEEE

Imperial College London, UK

University of Augsburg, Germany

MULTIMODAL SENTIMENT ANALYSIS: EXPLORE THE NO PAIN, BIG GAIN SHORTCUT

Multimodal Sentiment Analysis usually comes at massive cravings for labelled data. And, this data is best digested by deep nets well prepared by expert chefs who know their job all too well. For each modality flavour such as audio, text, video, or even physiology, a different recipe is best suited, and this is best topped by some fancy well-fitting multimodal fusion layers or techniques. This makes one wonder if there was any short-cut to perfect Multimodal Sentiment Analysis ideally saturating with little portions of data and preparable even by the domain layman. In other words: Can we solve sentiment analysis with minimal labelling effort and some black-box AI that takes it all from there, even if the data is largely heterogenous in terms of involved modalities. This talk invites to explore such an avenue with the steps of self-learning representations, coupling analysis and synthesis of sentiments for data augmentation, and autonomously learning reinforced, cross-modally, and self-supervised at scale. It will be garnished by insights into recent challenges organised by the presenter including MuSe and Interspeech ComParE. If all goes well, we shall arrive soon at instant multimodal sentiment analysis that fully satisfies.

Short Biography: Björn W. Schuller received his diploma, doctoral degree, habilitation, and Adjunct Teaching Professor in Machine Intelligence and Signal Processing all in EE/IT from TUM in Munich/Germany. He is Full Professor of Artificial Intelligence and the Head of GLAM - the Group on Language, Audio, & Music - at Imperial College London/UK, Full Professor and Chair of Embedded Intelligence for Health Care and Wellbeing at the University of Augsburg/Germany, co-founding CEO and current CSO of audEERING – an Audio Intelligence company based near Munich and in Berlin/Germany, Guest Professor at Southeast University in Nanjing/China and permanent Visiting Professor at HIT/China amongst other Professorships and Affiliations. Previous stays include Full Professor at the University of Passau/Germany, Key Researcher at Joanneum Research in Graz/Austria, and the CNRS-LIMSI in Orsay/France. He is a Fellow of the IEEE and Golden Core Awardee of the IEEE Computer Society, Fellow of the BCS, Fellow of the ISCA, President-Emeritus of the AAAC, and Senior Member of the ACM. He (co-)authored 1,000+ publications (40k+ citations, h-index=91), is Field Chief Editor of Frontiers in Digital Health and was Editor in Chief of the IEEE Transactions on Affective Computing amongst manifold further commitments and service to the community. His 30+ awards include having been honoured as one of 40 extraordinary scientists under the age of 40 by the WEF in 2015. First-in-the-field of Affective Computing and Sentiment analysis challenges such as AVEC, Interspeech ComParE, or MuSe have been initiated and by now organised overall more than 25 times by him. He is an ERC Starting and DFG Reinhart-Koselleck Grantee, and consultant of companies such as Barclays, GN, Huawei, Informetis, or Samsung.

Keynote Speech 2

KOICHI KISE

Osaka Prefecture University, Japan

READING OF READING FOR ACTUATING: AUGMENTING HUMAN LEARNING BY EXPERIENTIAL SUPPLEMENTS

Reading is a fundamental activity for learning languages. By "reading" a reading behavior of a learner, we can know the level of the learner's knowledge and mental states. The result of "reading" can be used to improve learners' behavior by using various actuators, i.e., methods of giving feedback. In general, effective actuators depend on learners and thus can be found by analyzing the experience of learning. In my talk, I introduce recent results of the project called "experiential supplements," focusing on application to language learning. By analyzing human experiences, we obtain pieces of information called experiential supplements, which make it easier to follow other learners' successful learning experiences. As a result, learners can learn a language more efficiently and effectively.

Short Biography: Koichi Kise received B.E., M.E., and Ph.D. degrees in communication engineering from Osaka University, Osaka, Japan, in 1986, 1988 and 1991, respectively. From 2000 to 2001, he was a visiting researcher at German Research Center for Artificial Intelligence (DFKI), Germany. He is now a professor of the Department of Computer Science and Intelligent Systems, Osaka Prefecture University, Japan. With Prof. Andreas Dengel, DFKI, he founded in 2008 the Institute of Document Analysis and Knowledge Science (IDAKS), Osaka Prefecture University, and now works as the director. He has received awards including best paper awards of three major international conferences in the field of document analysis, i.e., ICDAR (international conf. on document analysis and recognition, in 2007 and 2013), DAS (document analysis systems, in 2010) and ICFHR (international conf. on frontiers in handwriting recognition, in 2010). He was the chair of IAPR TC11 (reading systems, 2012-2016), and a member of IAPR conferences and meetings committee. He has been an Editor-in-Chief of International Journal of Document Analysis and Recognition. He also worked for international conferences including as the general chair of ICDAR2017, a track chair of the document analysis track of ICPRs (2012, 2018), and a program co-chair of ICDAR2013, 2015 and ACPR2013, 2015. His research interests are in the areas of document analysis, human behavior analysis and learning augmentation.

Table of Contents

Final Program with ABSTRACTs: Times are in JST (+9hr GMT)

Oct. 22nd (Fri.), 2021

OPENING, 13:00-13:10

SESSION 1: REGULAR PAPERS, 13:10-14:10

Chair: Kazuya Murao, Ritsumeikan University, Japan

Chapter#6: BoxerSense: Punch Detection and Classification Using IMUs

*Yoshinori Hanada (Aoyama Gakuin University)**

Physical exercise is essential for living a healthy life since it has substantial physical and mental health benefits. For this purpose, wearable equipment and sensing devices have exploded in popularity in recent years for monitoring physical activity, whether for well-being, sports monitoring, or medical rehabilitation. In this regard, this paper focuses on introduces sensor-based punch detection and classification methods towards boxing supporting system which is popular not only as a competitive sport but also as a fitness standard for people who wish to keep fit and healthy. The proposed method is evaluated on 10 participants where we achieved 98.8% detection accuracy, 98.9% classification accuracy with SVM in-person-dependent (PD) cases, and 91.1% classification accuracy with SVM in person-independent (PI) cases. In addition, we conducted a preliminary experiment for classifying 6 different types of punches performed from both hands for two different sensor positions (right wrist and upper back). The result suggested that using an IMU on the upper back is more suited for classifying both hand punches than an IMU on the right wrist. To provide feedback in real-time, we estimated the real-time performance of each classification method and found out all our methods could classify a single punch in less than 0.1 seconds. The paper also discussed some points of improvement towards a practical boxing supporting system.

Chapter#3: Using LUPI to Improve Complex Activity Recognition

Kohei Adachi (Kyushu Institute of Technology); Paula Lago (Universidad Nacional Abierta y a Distancia); Yuichi Hattori (Kyushu Institute of Technology); Sozo Inoue (Kyushu Institute of Technology)*

Sensor-based activity recognition can recognize simple activities such as walking and running with high accuracy, but it is difficult to recognize complex activities such as nursing care activities and cooking activities. One solution is to use multiple sensors, which is unrealistic in real life. Recently, LUPI (Learning using Privileged Information) has been proposed, which enables training using additional information only in the training phase. In this paper, we used LUPI for improving the accuracy of complex activity recognition. In short, training is performed with multiple sensors during the training phase, and a single sensor is used during testing. We used four published datasets for evaluating our proposed method. As a result, our proposed method improves up to 16% in F1-score compared to the baseline method when we used random-split cross-validation of each subject.

Chapter#8: A data-driven approach for online pre-impact fall detection with wearable devices
*Takuto Yoshida (Nagoya University)**

The implementation of wearable airbags to prevent fall injuries depends on accurate pre-impact fall detection and a clear distinction between activities of daily living (ADL) and them. We propose a novel pre-impact fall detection algorithm that is robust against ambiguous falling activities. We present a data-driven approach to estimate the fall risk from acceleration and angular velocity features, and use thresholding techniques to robustly detect a fall before impact. In the experiment, we collect simulated fall data from subjects wearing an inertial sensor on their waist. As a result, we succeeded in significantly improving the accuracy of fall detection from 50.00% to 96.88%, the recall from 18.75% to 93.75%, and the specificity 81.25% to 100.00% over the baseline method.

SESSION 2: CHALLENGE + AWARDS, 14:20-15:50
Chair: Sozo Inoue, Kyushu Institute of Technology

Chapter#12: Lunch-Box Preparation Activity Understanding from MoCap Data Using Handcrafted Features
Yeasin Arafat Pritom (University of Dhaka); Md. Sohanur Rahman (University of Dhaka); Hasib Ryan Rahman (University of Dhaka); M. Ashikuzzaman Kowshik (University of Dhaka); Md Atiqur Rahman Ahad (University of Dhaka)*

Japanese packed-meal or Bento preparation or packaging automated by a robot is a recent challenge. It is one of the latest cementations in human activity recognition systems. It shows contingency on physical gesture recognition techniques using some sensor-based datasets, especially some motion capture data or skeleton data. To get a firm grip on the recognition, we (Team Boson Kona) implement a method dealing with some handcrafted features gained from the filtered data. We employ several preprocessing steps to handle any inconsistency of the dataset: after cleaning the data, we use filtering and windowing techniques to process the raw data. Then we figure out multiple statistical and geometrical features from joints based handcrafted features. After considering the most significant features for model training, we try out Support Vector Machine classifier, Random Forest classifier and Extra Tree classifier. After applying all these methodologies, we obtain maximum accuracy of 64.6% using the Extra Tree classifier.

Chapter#11: Identification of Food Packaging Activity Using MoCap Sensor Data
Adrita Anwar (University of Dhaka); Malisha Islam Tapotee (University of Dhaka); Purnata Saha (University of Dhaka); Md Atiqur Rahman Ahad (University of Dhaka)*

The automation system has brought a revolutionary change in our lives. Food packaging activity recognition can add a new dimension to industrial automation systems. However, it is challenging to identify the packaging activities using only skeleton data of the upper body due to the similarities between the activities and subject-dependent results. Bento Packaging Activity

Recognition Challenge 2021 provides us with a dataset of 10 different activities performed during Bento box packaging in a lab using MoCap (motion capture) sensors. Bento box is a single-serving packed meal that is very popular in Japanese cuisine. In this paper, we develop methods using the classical machine learning approach, as the given dataset is small compared to other skeleton datasets. After preprocessing, we extract different hand-crafted features and train different models like Extremely Randomized Trees, Random Forest, and XGBoost classifiers, and select the best model based on cross-validation score. Then we explore different combinations of features and use the best combination of features for prediction. By applying our methodology, we achieve 64% accuracy and 53.66% average accuracy in 10-fold cross-validation and leave-one-subject-out cross-validation, respectively.

Chapter#10: Can Ensemble of Classifiers Provide Better Recognition Results in Packaging Activity?

Promit Basak (University of Dhaka); A.H.M Nazmus Sakib (University of Dhaka); Shahamat Mustavi Tasin (University of Dhaka); Syed Doha Uddin (University of Dhaka); Md Atiqur Rahman Ahad (University of Dhaka)*

Skeleton-based Motion Capture (MoCap) systems have been widely used in the game and film industry for mimicking complex human actions for a long time. MoCap data has also proved its effectiveness in human activity recognition tasks. However, it is a quite challenging task for smaller datasets. The lack of such data for industrial activities further adds to the difficulties. In this work, we have proposed an ensemble-based machine learning methodology that is targeted to work better on MoCap datasets. The experiments have been performed on the MoCap data given in the Bento Packaging Activity Recognition Challenge 2021. Bento is a Japanese word that resembles lunch-box. Upon processing the raw MoCap data at first, we have achieved an astonishing accuracy of 98% on 10-fold Cross-Validation and 82% on Leave-One-Out-Cross-Validation by using the proposed ensemble model.

Chapter#17: Summary of the Bento Packaging Activity Recognition Challenge

Kohei Adachi (Kyushu Institute of Technology); Shamma Alia (Kyutech); Nazmun Nahid (Kyushu Institute of Technology); Haru Kaneko (Kyushu Institute of Technology); Paula Lago (Universidad Nacional Abierta y a Distancia); Sozo Inoue (Kyushu Institute of Technology)*

With a goal of making the interactions between humans and robots more straightforward, we organized Bento Packaging Activity Recognition Challenge as a part of The 3rd International Conference on Activity and Behavior Computing. In this work, we analyze and summarize the approaches of submission of the challenge. In a dataset consisting of Bento packing activities, the subjects perform 10 activities in total with a conveyor belt. The challenge started on 1st June, 2021 and continued until 25th August, 2021. The participant teams used the given dataset to predict the10 activities and they were evaluated using accuracy. The winning team achieved about 64% accuracy on testing data.

Chapter#13: Bento Packaging Activity Recognition Based on Statistical Features

Faizul Rakib Sayem (University of Dhaka); Md. Mamun Sheikh (University of Dhaka); Md Atiqur Rahman Ahad (University of Dhaka)*

Due to the fast advancements of low-cost micro-embedded sensors and MoCap sensors, human action recognition has become an essential study topic and is garnering a lot of interest in different sectors. Recently, it is drawing a lot of attention in human-robot collaboration to assist human to preform regular tasks step-wise because it is difficult to obtain human labor at a lower wage to monitor industrial works. In this work, we have presented a straightforward machine learning paradigm to recognize 10 different Bento (lunch-box) packaging activities in real-time world. Unlike other skeleton-based human activity recognition domain, it is a very challenging task due to the absence of lower-body marker information. Under these circumstances, we have provided an in-depth statistical analysis of different Bento packaging activities. After feature extraction process, we have used several machine algorithms and obtained best results in Random Forest Classifier using hyperparameter tuning. We have achieved 64.9% validation accuracy using leave-one-out method.

Chapter#14: Using k-Nearest-Neighbours Feature Selection for Activity Recognition

Björn Friedrich (Carl von Ossietzky Universität); Tetchi Ange-Michel Orsot (unaffiliated); Andreas Hein (Carl von Ossietzky Universität)*

As team 2A&B Core we contributed a k-Nearest-Neighbour approach for classifying 10 different activities for the Bento Packing Activity Recognition Challenge. We used hand-engineered features from motion capture data. Before classifying the data we scaled it and applied a Principal Component Analysis. We used the k-Nearest-Neighbour classifier to find the best feature and we found that the best feature was one that was related to a fixed position. Our approach achieved an accuracy of 38.78% on our validation subject and 42.00% on our test subject.

Chapter#15: Bento Packaging Activity Recognition from Motion Capture Data

Jahir Ibna Rafiq (University of Asia Pacific); Shamaun Nabi (University of Asia Pacific); Al Amin (University of Asia Pacific); Shahera Hossain (University of Asia Pacific)*

Human activity recognition (HAR) has been an important research field for more than a decade due to its versatile applications in different area. It has gained significant attention in the healthcare domain. Although it has similarity with other form of activity recognition, it offers a unique set of challenges. Body movements in a food preparation environment are considerably less than many other activities of interest in real world. In this paper, a comprehensive solution has been demonstrated for the bento box packaging challenge activity recognition. In this case, we present a well-planned approach to recognize activities during packaging tasks from Motion Capture data. We use dataset obtained from motion capture system where subjects have 13 markers on their upper body area, and by special use of cameras and body suit. We obtain around 50,000 sample for each of the activity. We reduce the data dimensionality and make the data suitable for the classification purpose by extracting reliable and efficient features. After feature extraction process, three different classifiers e.g., Random Forest classifier, ExtraTrees classifier

and Gradient Boosting classifier are compared to check the result. We conclude that this challenging dataset has been observed to work most efficiently for Random Forest classifier using hyperparameter tuning.

Chapter#16: Bento Packaging Activity Recognition with Convolutional LSTM using Autocorrelation Function and Majority Vote

Atsuhiro Fujii (Ritsumeikan University); Kazuki Yoshida (Ritsumeikan University); Kiichi Shirai (Ritsumeikan University); Kazuya Murao (Ritsumeikan University)*

This paper reports Bento Packaging Activity Recognition Challenge by team "RitsBen" held in the International Conference on Activity and Behavior Computing (ABC 2021). Our approach used an autocorrelation function in the preprocessing to isolate the data since the dataset was given with repetitive activity. We then use a model that implements convolutional layers, and LSTM. The final decision is made by majority vote using sigmoid predictions output from all body parts. The loss is calculated using BCEWithLogitsLoss for each body part. The evaluation results showed that average accuracy of 0.123 were achieved among subjects 1, 2, and 3 in leave-one-subject-out manner. However, we did not achieve high accuracy as the possibility that the extraction of repetitive actions was not correct.

SESSION 3: REGULAR PAPERS, 16:00-17:00

Chair: Michael Beigl, Karlsruhe Institute of Technology, Germany

Chapter#5: Using Human Body Capacitance Sensing to Monitor Leg Motion Dominated Activities with a Wrist Worn Device

Sizhen Bian (DFKI); Siyu Yuan (TU Kaiserslautern); Vitor Fortes Rey (DFKI); Paul Lukowicz (DFKI GmbH, Germany)*

Inertial Measurement Unit (IMU) is currently the dominant sensing modality in sensor-based wearable human activity recognition. In this work, we explored an alternative wearable motion-sensing approach: inferring motion information of various body parts from the human body capacitance (HBC). While being less robust in tracking the body motions, HBC has a property that makes it complementary to IMU: it does not require the sensor to be placed directly on the moving part of the body of which the motion needs to be tracked. To demonstrate the value of HBC, we performed exercise recognition and counting of seven machine-free leg-alone exercises. The HBC sensing shows significant advantages over the IMU signals in both classification (0.89 vs 0.78 in F-score) and counting.

Chapter#2: Open-Source Data Collection for Activity Studies at Scale

Alexander Hoelzemann (University of Siegen); Kristof Van Laerhoven (University of Siegen); Jana Sabrina Pithan (University Of Vechta)*

Activity studies range from detecting key indicators such as steps, active minutes, or sedentary bouts, to the recognition of physical activities such as specific fitness exercises. Such types of activity recognition rely on large amounts of data from multiple persons, especially with deep learning. However, current benchmark datasets rarely have more than a dozen participants.

Once wearable devices are phased out, closed algorithms that operate on the sensor data are hard to reproduce and devices supply raw data. We present an open-source and cost-effective framework that is able to capture daily activities and routines, and which uses publicly available algorithms, while avoiding any device-specific implementations. In a feasibility study, we were able to test our system in production mode. For this purpose, we distributed the Bangle.js smartwatch as well as our app to 12 study participants, who started the watches at a time of individual choice every day. The collected data was then transferred to the server at the end of each day.

Chapter#1: Toward the Analysis of Office Workers' Mental Indicators Based on Wearable, Work Activity, and Weather Data

Yusuke Nishimura (Kyushu Institute of Technology); Tahera Hossain (Kyushu Institute of Technology); Akane Sano (Rice University); Shota Isomura (NTT Data Institute of Management Consulting); Yutaka Arakawa (Kyushu University); Sozo Inoue (Kyushu Institute of Technology)*

In recent years, many organizations have prioritized efforts to detect and treat mental health issues. In particular, office workers are affected by many stressors, and physical and mental exhaustion, which is also a social problem. To improve the psychological situation in the workplace, we need to clarify the cause. In this paper, we conducted a 14-day experiment to collect wristband sensor data as well as behavioral and psychological questionnaire data from about 100 office workers. We developed machine learning models to predict psychological indexes using the data. In addition, we analyzed the correlation between behavior (work content and work environment) and the psychological state of office workers to reveal the relationship between their work content, work environment, and behavior. As a result, we showed that multiple psychological indicators of office workers can be predicted with more than 80\% accuracy using wearable sensors, behavioral data, and weather data. Furthermore, we found that in the working environment, the time spent in `web conferencing', `working at home (living room)', and `break time (work time)' had a significant effect on the psychological state of office workers.

KEYNOTE 1, 17:00-18:00

Chair: Anton Nijholt, University of Twente, The Netherlands

Björn W. Schuller, Fellow IEEE (Imperial College London/UK and University of Augsburg/Germany)

Title: Multimodal Sentiment Analysis: Explore the No Pain, Big Gain Shortcut

NETWORKING, 18:00-18:20

Oct. 23rd (Sat.), 2021

SESSION 4: REGULAR PAPERS, 13:00-14:00

Chair: Takuro Yonezawa, Nagoya University, Japan

Chapter#7: FootbSense: Soccer Moves Identification Using a Single IMU

Yuki Kondo (Aoyama Gakuin University); Shun Ishii (Aoyama Gakuin University); Hikari Aoyagi (Aoyama Gakuin University); Tahera Hossain (Aoyama Gakuin University); Anna Yokokubo (Aoyama Gakuin University); Guillaume Lopez (Aoyama Gakuin University)*

Although wearable technologies are commonly used for sports at elite levels, these systems are expensive and it is still difficult to recognize detailed player movements. We introduce a soccer movements recognition system using a single wearable sensor to aid the skill improvement for amateur players. We collected 3-axis acceleration data of 6 soccer movements and validated the proposing system. We also compared 3 sensor locations to find the best accurate location. With Ensemble Bagged Trees classification method, we achieved 78.7% classification accuracy of 6 basic soccer movements from the inside ankle sensor. Moreover, our results show that it is possible to distinguish between running and dribbling, passing and shooting, even though they are similar movements in soccer. Besides, the second highest accuracy was achieved from a sensor placed on the upper part of the back, which is a safer wearing position compared to other locations. These results suggest that our approach enables a new category of wearable recognition system for amateur soccer.

Chapter#9: Modeling Reminder System for Dementia by Reinforcement Learning

Muhammad Fikry (Kyushu Institute of Technology); Nattaya Mairittha (Kyushu Institute of Technology); Sozo Inoue (Kyushu Institute of Technology)*

Prospective memory refers to preparing, remembering, and recalling plans that have been conceived in an intended manner. Various busyness and distractions can make people forget the activities that must be done the next time, especially for people with cognitive memory problems such as dementia. In this paper, we propose a reminder system with the idea of taking time and response into consideration to assist in remembering activities. Using the reinforcement learning method, this idea predicts the right time to remind users through notifications on smartphones. The notification delivery time will be adjusted to the user's response history, which becomes feedback at any available time. Thus, users will get notifications based on the ideal time for each individual either, either with repetition or without repetition, so as not to miss the planned activity. By evaluating the dataset, the results show that our proposed modeling is able to optimize the time to send notifications. The eight alternative times to send notifications can be optimized to get the best time to notify the user with dementia. This implies that our algorithm propose can adjust to individual personality characteristics, which might be a stumbling block in dementia patient care, and solve multi-routine plan problems. Our propose can be useful for users with dementia because we can remind very well that the execution time of notifications is right on target, so it can prevent users with dementia from stressing out over a lot of notifications,

but those who miss notifications can receive them back at a later time step, with the result that information on activities to be completed is still available.

Chapter#4: Attempts toward Behavior Recognition of the Asian Black Bears using an Accelerometer

Kaori Fujinami (Tokyo University of Agriculture and Technology); Tomoko Naganuma (Tokyo University of Agriculture and Technology); Yushin Shinoda (Tokyo University of Agriculture and Technology); Koji Yamazaki (Tokyo University of Agriculture); Shinsuke Koike (Tokyo University of Agriculture and Technology)*

In recent years, many organizations have prioritized efforts to detect and treat mental health issues. In particular, office workers are affected by many stressors, and physical and mental exhaustion, which is also a social problem. To improve the psychological situation in the workplace, we need to clarify the cause. In this paper, we conducted a 14-day experiment to collect wristband sensor data as well as behavioral and psychological questionnaire data from about 100 office workers. We developed machine learning models to predict psychological indexes using the data. In addition, we analyzed the correlation between behavior (work content and work environment) and the psychological state of office workers to reveal the relationship between their work content, work environment, and behavior. As a result, we showed that multiple psychological indicators of office workers can be predicted with more than 80\% accuracy using wearable sensors, behavioral data, and weather data. Furthermore, we found that in the working environment, the time spent in `web conferencing', `working at home (living room)', and `break time (work time)' had a significant effect on the psychological state of office workers.

SESSION 5: WORK IN PROGRESS, 14:10-15:00

Chairs:
Guillaume Lopez, Aoyama Gakuin University, Japan
Kaori Fujinami, Tokyo University of Agriculture and Technology, Japan

WIP#11: A Method for Action Quality Analysis from Skeleton Data of Rehabilitation Exercises

Faizul Rakib Sayem (University of Dhaka); Md. Mamun Sheikh (University of Dhaka); Md Atiqur Rahman Ahad (University of Dhaka)*

Researches in healthcare applications such as fitness tracking, fall detection, remote monitoring of elderly patients' daily activities are gaining much attention nowadays. To monitor and improve patients' states of physical and mental health, rehabilitation exercises bring a new dimension to the healthcare system. In this paper, we represent a feature-based machine learning approach, which can almost correctly evaluate the quality of rehabilitation exercises from a challenging benchmark dataset, called KIMORE dataset. We have provided a simple yet efficient machine learning paradigm for automatically assessing action quality from skeleton data of rehabilitation cases. Our method can compute a subject's quality of the performed actions, from skeletal joint position data only. In our proposed method, we apply a feature selection method after extracting features in order to avoid overfitting and get best results. Then, we exploit

different machine learning algorithms. Gradient Boosting combines the predictions from multiple decision trees to generate the final predictions. We randomly split the dataset into training and validation group for 30 times by exploiting stratify function. Using Gradient Boosting Regressor, we have found the lowest Mean Absolute Error (MAE) on the validation data. Our proposed method has achieved average MAE of about 0.0359 and 0.0348 for Exercise-1 and 5 of the KIMORE dataset, respectively. Based on our study, it is the lowest MAE on the actions.

WIP#12: Examination of analogy method of degree of interest in news articles by gaze information

Takuya Ogawa (Ehime University); Keiichi Endo (Ehime University); Hisayasu Kuroda (Ehime University); Shinya Kobayashi (Ehime University)*

There is a huge amount of news information on the Internet. This gives users the advantage of having access to a large amount of information. However, users are not always interested in all the news that is distributed. Therefore, when users read the news on the Internet, they need to select the information they are interested in. We have developed NEAR, a news application for smartphones, with the aim of solving the problem of information overload. However, in order for the app to learn the user's interest trends, it needed to have the user sort through a lot of news, which was a burden for the user. Therefore, we propose a method that focuses on the user's gaze information as a method for learning interests while reducing the burden of article selection by the user. By focusing on gaze, it is thought that it will be possible to analogize the degree of interest on a word-by-word basis rather than on an article-by-article basis, and it is expected that the number of article selection required for learning will be reduced.

WIP#15: A Comparative Analysis on Joint's Importance to Achieve Better Performance in Behavior Analysis in Human-Robot Collaborative Workspace

Nazmun Nahid (Kyushu Institute of Technology); Shotaro Yoshinaga (Kyushu Institute of Technology); Sozo Inoue (Kyushu Institute of Technology)*

Manufacturing systems capable of performing multiple tasks require different types of resources. The fully automated system using robots has high speed, precision, tirelessness, and strength, but it is expensive whereas, human workers are smart, creative, flexible, and able to use different tools in different situations but do not have the same production ability as robots. Human-Robot Collaboration (HRC) thus becoming the new frontier in the research area as it allows us to combine the advantages of robots along with the adaptability and cognitive aptitudes of human workers. Though it has tremendous potential we failed to utilize it to the fullest due to safety concerns. In this work, we have presented a brief literature review regarding these issues. From our analysis, a simplified human motion analysis system is required especially where human and robot works within an intimate distance. So, we have done an elaborate comparative performance analysis on different body markers indicating different joints of the human body. We have found that in the case of close distance HRC the highest accuracy of 99 % is achieved from the Head-Wrist and the Shoulder-Wrist combination. Also, Head sensor data has a great impact in this type of scenario.

WIP#23: Human Activity Recognition Based on Wavelet-Based Features along with Feature Prioritization

*Mahmudul Hasan Abid (Khulna University); Abdullah-Al Nahid (Khulna University, Khulna-9208, Bangladesh)**

Activity recognition from human action data is quite a challenging task in the biomedical data science community. The main challenge in dealing with human activity recognition (HAR) datasets is their high cardinality. So reducing cardinality is a cardinal area of research in the HAR field. In this research, we have tried to reduce the data dimensionality by utilizing future selection methods. This research work has extracted features using wavelet packet transform (WPT) and the cardinality of the feature set has been reduced by using the Genetic Algorithm (GA) technique. We also have ranked the selected features according to their importance based on their SHAP values. In the venture, we have got an interesting inspection that in HAR datasets, signal values lay into lower frequency regions mostly. The highest accuracy and f1-score which we have got are 94.74%, 94.73%, and 89.98%, 89.67% for the feature extracted and feature selected dataset respectively.

WIP#25: Recognizing Human Actions Utilizing Relative Velocity of Skeleton Joints through Statistical Features

*Anindya Das Antar (University of Dhaka); Masud Ahmed (University of Dhaka); Md Atiqur Rahman Ahad (University of Dhaka)**

Human activity or action recognition (HAR) becomes immensely essential for many applications in healthcare, security, surveillance, etc. in the field of computer vision and ubiquitous sensors. In the field of computer vision, RGB frames have been widely explored. In the last decade, depth maps and skeleton data become well-explored arenas for HAR. In this paper, we have explored velocity-based statistical feature (VSF) that encode motion information of 3D skeleton joints across video frames. We have kept track of changes in linear joint velocity (LJV) and angular joint velocity (AJV) of skeleton joints by extracting statistical features (e.g., mean, median, maximum, minimum, std), which we named as VSF. Most of the skeleton-based methods are explored on 2 or 3 datasets usually. However, we have explored 7 benchmark datasets to demonstrate the performance of our method. The explored datasets are Kinect Activity Recognition Dataset (KARD), UTKinect-Action3D dataset, MSR 3D Action Pairs Dataset, Office Activity Dataset (OAD), Florence 3D Dataset, CAD-60 Dataset, and MSR Daily Activity 3D Dataset. We employed leave-one-subject-out cross-validation ("New Person") along with dataset splitting framework for the experiment. Support vector machine classifier with linear kernel have achieved superior performance. The proposed scheme has demonstrated superior or comparable results in most of the cases. These results demonstrate the importance of the extracted features, which can be implemented in other domains, especially in healthcare applications based on skeleton data.

SESSION 6: PANEL: THE FUTURE OF ACTIVITY AND BEHAVIOR COMPUTING, 15:10-17:10

Chair: Md Atiqur Rahman Ahad, University of Dhaka, Bangladesh

Panelists:

- Stephan Sigg, Aalto University, Finland
- Takuro Yonezawa, Nagoya University, Japan
- Paula Lago, Universidad Nacional Abierta y a Distancia, Colombia
- Yu Enokibori, Nagoya University, Japan
- Kristof Van Laerhoven, University of Siegen, Germany
- Atsushi Nakazawa, Kyoto University, Japan
- Phond Phunchongharn, KMUTT, Thailand
- Christopher Nugent, Ulster University
- Guillaume Lopez, Aoyama Gakuin University, Japan
- Philipp Scholl, University of Freiburg, Germany
- Hristijan Gjoreski, Ss. Cyril and Methodius University, Macedonia

KEYNOTE 2, 17:10-18:10

Chair: Kaori Fujinami, Tokyo University of Agriculture and Technology

Koichi Kise (Osaka Prefecture University/Japan)

Title: Reading of Reading for Actuating: Augmenting Human Learning by Experiential Supplements

AWARDS + CLOSING, 18:10-18:30

NETWORKING, 18:30-18:50