Advanced systems for head scratch detection

Žygimantas Jočys

Submitted for the degree of MPhil

University of Sussex

June 2022

Declaration

I hereby declare that this thesis has not been and will not be submitted in whole or in part to another University for the award of any other degree.

Signature:

Žygimantas Jočys

Acknowledgements

First of all, I want to thank my supervisor Prof. Daniel Roggen at the Wearable Computing Lab. I thank him for his support and for providing me with excellent guidance, mentoring and research facilities.

I also would like to thank my wife for the ultimate support and for keeping up with all the ups and downs during this period.

Contents

List of Tables

List of Figures

Nomenclature

IMU - Inertial Measurment Unit

EPS - Electric Potential Sensor

ML - Machine Learning

DL - Deep Learning

UwB - Ultra Wide Band

TWR - Two-Way-Ranging

TDOA - Time Difference of arrival

AuNW - Ultrathin Nanowire

BLE - Bluetooth Low Energy

MF - Magnetic Field

ADC - Analog to Digital Converter

RF - Random Forest

k -NN - k - Nearest Neighbour

RTC - Real Time Clock

SPI - Serial Peripheral Interface

DAC - Digital to Analog Converter

UART - Universal Asynchronous Receiver/Transmitter

FFT- Fast Fourier Transforms

Abstract

Itching is a condition that affects a substantial group of people. This condition may be caused by different conditions, such as scabies, atopic dermatitis, or kidney failure; it can also be a symptom of a malignant condition, such as lymphoma. So far, a scratch was being detected by manually counting the occurances or using a bone-conducting microphone, which is uncomfortably set up. Thus, there is a need for a next-generation system that allows detecting scratches on multiple people simultaneously without invading patients' lives. Wearable sensors allow the ability to directly collect the data asynchronously from many people and detect activities by applying machine learning algorithms.

In this thesis, we propose using multimodal wearable sensors and combining the data from Inertial Measurement Units (IMU), Electric Potential Sensor (EPS) and a microphone using machine learning-based fusion for scalp scratch detection. In this thesis, we describe the results on three problems: (1) the impact of fusing EPS and IMU for scratch detection, (2) the ambient microphone's ability to detect scratch, (3) the future direction for next-generation scratch detection system.

We evaluated the fusion of EPS and IMU on a constrained dataset that mimics an office worker's daily activities, which we collected in the Wearable Technologies Lab at the University of Sussex. We showed that multimodal fusion is superior to using a wrist-worn IMU solely. For the (2) objective, we collected a small dataset from four people showing that an ambient microphone can be a powerful modality for scratch detection.

Finally, we propose a clear direction for future research that involves a wide-scale dataset collection, novel hardware, and powerful Deep Learning algorithms to power the next generation scratch detection system.

Chapter 1

Introduction

In this chapter, we are presenting the motivation for this thesis. Moreover, we detail the contributions that have been made, that include a dataset collections, analyses and the future direction for scratch detection. The work resulted in multiple publications detailed below. We also give a thesis overview.

1.1 Motivation

We are interested to monitor scratching, because itching is a condition which affects a substantial group of people. This condition may be caused by scabies, atopic dermatitis, or kidney failure; it can also be a symptom of a malignant condition, such as lymphoma [27]. Thus, we need an accurate and scalable monitoring system that is able to detect scratches accurately and not bounding the subjects ability to live a daily life.

In this thesis, we are exploring wearable sensors, multi-modal fusion and the future direction to create a next generation scratch detection system, that would allow for scalable and accurate scratch recognition. Three main components are required to recognise human activities using machine learning based methods:

- **Sensors** to collect the signals from the real world.
- **Datasets** to train and evaluate the scratch detection system.
- **Software** to detect scratching from the observed signals.

1.2 Research contributions

In this thesis we present five main contributions:

Dataset "Scratch in office environment" is a next generation dataset to evaluate machine learning algorithms for fusion for scratch detection. It was collected in the premises of the University of Sussex from 10 individuals performing daily activities. They wore four IMU sensors and one EPS sensor. This dataset is first of the kind that allows to explore the importance of different sensors and how the fusion can impact the accuracy of daily activities and scratching events.

Analysis of the multimodal fusion on the dataset "Scratch in office environment" showed that we can achieve higher accuracies by combinining IMU, EPS and position of the hand. Moreover, we showed that hand coordinate relative to the torso is the most informative feature for human activity recognition.

Proof of Concept Dataset "Sound of Scratch" was collected from four individuals to show the effectiveness of sound modality for scratch detection from a wrist-worn microphone.

Analysis of sound importance for scratch recognition has been performed to show that sound is a modality that can be used to detect scratch.

Future Direction for Scratch detection is proposed for the next generation scratch detection system.

1.3 Publications

- Zygimantas Jocys, Arash Pouryazdan, and Daniel Roggen. **Multimodal fusion of IMUs and EPS body-worn sensors for scratch recognition.** In Proceedings of the 14th EAI International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth '20). Žygimantas Jočys did all the dataset collections, analyses and the writing of the paper, Arash Pouryazdan provided guidance with respect to the EPS, Daniel Roggen provided guidance and reviewed the text in the paper.

- Zygimantas Jocys and Daniel Roggen. **Scalp scratch detection using a wrist-worn microphone.** *Not published yet.* Žygimantas Jočys did all the dataset collections, analyses and the writing of the paper, Daniel Roggen provided guidance and reviewed the text in the paper.

- Lugoda, P., Costa, J. C., Garcia-Garcia, L. A., Pouryazdan, A., Jocys, Z., Spina, F., Salvage, J., Roggen, D., Münzenrieder, **N.Coco Stretch: Strain Sensors Based on Natural Coconut Oil and Carbon Black Filled Elastomers**. Žygimantas

Jočys collected the dataset and did the analysis related to the sensor's applications for human activity recognition.

1.4 Outline of the thesis

- **Chapter 2: Literature review for Sensors, Human Activity Recognition, Scratch detection.** We review previously used techniques to detect scratches and their limitations. Moreover, we evaluate different sensor modalities to potentially fit on a wrist and be part of a wrist-worn scratch detection system. As a summary in this section we provide a breakdown of current challenges for scratch detection system, that are improved upon in this thesis.

- **Chapter 3: Multimodal fusion of IMUs and EPS body-worn sensors for scratch recognition.** We present an investigation of sensors and algorithms to realise a wearable scratch detection device. We collected a dataset, where each user wore 4 inertial measurement unit (IMU) sensors and one electric potential sensor (EPS). The dataset contained 813 scratching instances and 5 h 15 min of recorded data. We investigated the trade-offs between the number of devices worn (comfort) and accuracy.

- **Chapter 4: Scalp scratch detection using a wrist-worn microphone.** We evaluate how sound, collected by a wrist-worn microphone, can be used to recognise and distinguish scratch occurrences. We collected a dataset from 4 users, where each user conducted different scratches in environments with different specifications. Subsequently we performed different analysis to see how well scratch can be detected. These results indicate that an ambient microphone may be one modality to include for a future multimodal scratch detection device.

- **Chapter 5: Future direction, large scale dataset collection, EPS based localisation, next generation deep learning methods** We propose next steps for next generation hardware, dataset and software. First of all, we propose that the future sensor should consists of a wrist-worn IMU,EPS and ambient microphone and a neck-worn signal generator. Second, we provide guidelines for a real life dataset collection, that could be a gold standard dataset to evaluate the efficacy of the scratch detection system. Finally, we propose the general direction for deep learning algorithms for more accurate human activity recognition.

Chapter 2

Scratch detection

In this section, we review previously used techniques to detect scratches and their limitations. Moreover, we evaluate different sensor modalities to potentially fit on a wrist and be part of a wrist-worn scratch detection system. As a summary in this section we provide a breakdown of current challenges for scratch detection system, that have improved upon in this thesis.

2.1 Review of scratch detection methodologies

In previous works, an itch [5] was described as an irritating feeling that causes the desire to scratch. A chronic itch can be a reaction of skin diseases, kidney failure, cancers, and neurological disorders. Given that the disease might need urgent medical attention, scratches should be monitored as they can lead to a faster disease identification.

Clinical and consumer goods trials using questionnaires and manually annotated observed behaviour, are not effective methods of evaluating the efficacy of treatment for a large numbers of users. For example, to evaluate if treatment improves an itching condition, we would need to measure how much a participant scratches before and after the treatment. In the early days, scratching was observed manually, without any technological help. In research by [11], scratching was monitored in 40 hospital patients by the staff. During this research project, the amount of scratching before and after treatment was noted. Observing the changes in scratching occurences, the treatment can be evaluated as effective, or not. However, this system is not suitable to be deployed on a large group of people (hundreds). This test, if conducted with many people, would show the efficacy of the treatment, if significantly reduced occurrences of scratching was witnessed after the treatment.

Another approach is to video record nocturnal scratches by using an infrared camera. In one experiment, which was conducted with seven users [10], the infrared camera was used, so that the user would not be interrupted during the night and the scratches were annotated later.

Human activities have been commonly used to define human behavioural patterns. It allows the quantitative evaluation of how much an activity occurs and to understand the underlying relationships between activities. For effective human activity recognition, a sensor system is needed to effectively perceive the environment and get enough information to distinguish between activities. Human activity recognition is comprised of three different parts: sampling from the sensor, feature extraction, and classification. In this work, we will analyse and survey current sensor technologies and decide which ones are the most suitable.

Currently, the most common approach to classify activities is to use accelerometers [6]. It allows for the monitoring of activities without invading user privacy. Not only does it allow for systems to be deployed on many users simultaneously, but accelerometers also enable real-time and real-world monitoring without the need of an observer to invade the subjects' privacy. Triaxial accelerometers [20] on the wrist for scratch deployment were noted to be successful and achieved high sensitivity (0.96) and specificity (0.92). In this previous research, 12 people participated; they were monitored in a controlled environment. The duration of the dataset collection was 140 s per user. Each person scratched the back of his/her head, leg, and/or elbow and later walked and rolled around. A similar work was done by detecting scratching by using mobile systems [14]; this method achieved 90% accuracy.

However, even though a single modality has shown good results, there are many new sensors, which can potentially be incorporated in such a system and increase the accuracy of scratch detection.

For instance, acoustic sensors [17] have been used for scratch detection. The sound sensor uses body conduction to record the sound from the wrist. The dataset was collected on four volunteers over 6 h during the night. As the room was silent, it was easier to distinguish between a scratch and no scratch, and a 'determination ratio' of 0.98 was achieved.

Electric potential sensors (EPSs) were explored for scratch sensing [21]. This work only reported a visual observation that EPS signals are correlated to a scratch, but the authors did not demonstrate any automated activity recognition based on this modality.

Machine learning models vary in complexity, computational time, and their ability to generalise. Logistic regressions [20] and recurrent neural networks [16] have been used thus far. In the research [16], 24 atopic dermatitis patients spent 2–5 nights in the hospital and wore two accelerometer devices (GeneActiv, Activinsights Ltd.) on each wrist. Only two classes were labelled: scratch or null class, and as the data needed to be split for each class equally, a considerable amount of recording data was lost for the training of the model. There is still room for new machine learning models and considerable progress needs to be made with the current techniques.

In the previous works, all the currently collected datasets for scratch detection were either very short (150 s) or very simplistic (two classes: scratch or null class), which does not allow one to evaluate how the system would perform in a more realistic scenario. Moreover, given that in each research, there was one sensor modality per dataset, there is a need to explore different fusion options between different sensors.

2.2 Sensor technology overview

In the table below we review different sensor modalities and their application for scratch detection.

Sensor	Description
Acceleration	Accelerometers [20] on the wrist for scratch detection have been successful and achieved high sensitivity (0.96) and specificity (0.92). Accelerometers measure the change in velocity. The force caused by vibration or a change in motion (acceleration) causes the mass to squeeze the piezoelectric material, which produces an electrical charge that is proportional to the force exerted upon it.

Body-conduction microphone	A body-conduction microphone is shaped like a wristwatch and adheres to the skin of the dorsal portion of the forearm [17]. The key device embedded in this unit is the thin piezoelectric element. The fingertips and nails make specific sounds and imperceptible vibrations when scratching the body surface. This detection system detects imperceptible vibration from fingertips at the wrists with a high sensitivity. To prevent invasion of the electrical and mechanical noise from outside the measuring surface, the piezoelectric element is sealed in a concave aluminum cap (1-mm thick) with polyurethane gel and silicone sealant. Additionally, the piezoelectric element adheres to the skin surface with viscous polyurethane gel. Therefore, the body-conducted sound energy is efficiently transmitted to the piezoelectric element. The analysis program detects scratching behaviour based on the acquired sound signal information.
Bone-conduction microphone	A bone-conduction microphone [18] measures the mechanical vibration between the nail and the skin that is scratched. Scratching sounds, induced by the vibration, change according to scratching intensity or skin condition. Therefore, the information about the scratching sounds can help us to evaluate the characteristics of scratching. In this study, a nail-mounted compact microphone sensor for measuring human scratching sounds was designed and fabricated. The power spectrum showed that it can also be used to distinguish the intensity of the scratch.
Microphone	A microphone converts sound into a small electrical current. Sound waves hit a diaphragm that vibrates, moving a magnet near a coil. In some designs, the coil moves within a magnet. Other microphones, such as condenser microphones, work on the principle of capacitance. A normal ambient microphone could be integrated into an IMU and be used for scratch detection.

Ultra wide band (UwB)-based localisation two-way-ranging (TWR)	In TWR, the distance from a tag to the anchor is obtained by sending a packet back and forth. By measuring how long it took for the packet to return, the tag can estimate the distance to the anchor. For positioning, the tag initiates communication with the anchors and ranges with the anchors, one by one. Once the tag has ranged with at least three, ideally four anchors, it can compute its position by trilateration. The platform is commercially available at [2].
UwB-based localisation time difference of arrival (TDOA)	In TDOA mode, tags will periodically transmit a UWB blink, with or without any scheduling or regard for other tags or anchors. Because the tag must only send a single UWB transmission, the positioning can happen fast and at low power. With the commonly used Aloha protocol, tags only send and never receive resulting in incredibly low power consumption, with battery lifetimes of several years. However, in order to limit the interference with other tags, the update rate is somewhat limited. Alternatively, the tags can be scheduled, as in TWR+, to achieve an even higher update rate. The platform is commercially available at [2].
IMU-based localisation	Four IMUs can be used together to localise each joint. The IMUs are located on the torso, upper arm, lower arm and hand. It uses quaternions to compute relative orientations and it uses some predetermined limb lengths to compute the coordinates of the hand related to the torso.
Wireless technology (Wi-fi)	It uses multiple wi-fi broadbands to detect the changes in a room, which can be used for scratch detection. It has shown, nocturnal scratching can be detected with high accuracy. However, the main limitation is that if the position of the objects inside the room changes, it cannot recognise the scratch. Wireless technologies have been applied to human activity recognition in [29].

Pressure sensor on fingertip	Flexible pressure sensors [12] presented a sandwiched ultrathin gold nanowire (AuNW)-impregnated tissue paper between two thin polydimethylsiloxane sheets. The sensing mechanism is due to pressing force-dependent contact between AuNWs and interdigitated electrode arrays. Unlike a bulk rigid planar metal, soft tissue paper has porous and rough surfaces with hairy AuNWs. The number of AuNWs bridging finger electrode pairs depends on the external forces applied. On applying an external pressure, a small compressive deformation of tissue paper enables more AuNWs in contact with finger electrodes, leading to more conductive pathways. This caused an increase in current when a fixed voltage of 1.5 V was applied.
Pressure sensor in clothes to measure angles	The same sensor [12] can be used to sense the pressure created in clothes (elbows, shoulders, knees). As scratching involves the movements of joints, this sensor would allow the recognition of changes related to scratching.
Bluetooth	Bluetooth Low Energy (BLE) technology can be a very good alternative for localisation. With the increasing distance from the transmitter, the received signal strength decreases and the travel time from the transmitter to the receiver increases. Two basic approaches are used—triangulation and fingerprinting. Triangulation uses estimation of the distance from several transmitters based on signal attenuation and time characteristics of the signal propagation. Fingerprinting is a localisation method comprising of two phases. In the first phase—learning—vectors are collected, which are composed of received signal strength indicator (RSSI) values and optional extra features measured by a measuring device in the known locations. Indoor localisation has been shown effective [13].

Electric potential sensor (EPS)	An EPS [21] is classified as a capacitive sensor with the ability to measure small variations in electric potential or an electric field remotely and accurately. When a person scratches their scalp, an electric charge is generated on the hair and the sensor can detect this disturbance of the electric field.
EPS-based localisation	EPS-based localisation [24] can be used to sense the phase of the ambient electric field from a body-worn sensor with respect to a reference. A signal generator could be worn around the neck and used as a reference point.
Magnetic field-based localisation	Magnetic field has been used for localisation in Finexus [8]. Magnetic field (MF) sensing is one clear approach for continuous, accurate and occlusion-free finger tracking. To localise electromagnets, Finexus leverages techniques similar to those used by the Global Positioning System (GPS). Intuitively, the system first calculates the distance between the electromagnet and four magnetic sensors, and then uses trilateration to identify the electromagnet's 3D position.
Camera	Cameras have been used for scratch detection for a while, where they manually annotate if the user scratches or not. In one experiment [10] with seven participants, the infrared camera was used to film them during the night, so that they would not be interrupted during the sleep and the scratches were annotated later.

Table 2.1: Sensor modalities, that could potentially be used for scratch detection with the principles of how they work.

Following the sensor modality descriptions in Table 2.1, the next section compares them to evaluate their suitability for human activity recognition and more specifically, for scratch detection.

2.2.1 Suitability of sensors for naturalistic scratch detection

In this work, we want to evaluate what sensor modalities can be worn to collect data throughout an entire day. To do this, the sensor device must be attached to the body without hindering the person's daily activities, but it must still be able to capture important data for scratch detection.

For a thorough assessment of the sensor modalities suitability, we need to assess different properties regarding comfort, integration to the main platform, limitations, etc. First of all, we want the sensor system to be comfortable. This means having the smallest number of devices on the body. Ideally, they should be located on the wrist, as people are used to wear watches and they should not bind the user to a location (the person should not be put in a room but should be allowed to live their normal life). From the technical side, the sensor modality need to have the ability to be easily integrated to the Bluesense platform or be commercially available.

Sensor	Number of devices for both wrists	Comfortability 1 (uncomfortable)- 5 (comfortable)	Previous work on scratch detection	Located on the wrist	Sensitivity to ambient signals 1-5	Commercially available	Integration with Bluesense 1 (complicated) - 3 (easy)	Location independent
IMU	2	5	✓	✓	1	✓	3	✓
Body-conduction microphone	2	3	✓		2		1	✓
Bone-conduction microphone	2	1	✓		2		1	✓
Microphone	2	5		✓	5	✓	3	✓
UWB-based localisation	4	3			1	✓	1	✓
IMU-based localisation	7	1			1		1	✓
Wireless technology (Wi-fi)	0	5			5	✓		
Pressure sensor on fingertip	2	2			5		3	✓
Pressure sensor in clothes to measure the angles	2	3			3		1	✓
bluetooth localisation	3	4		✓	1	✓	1	✓
EPS)	2	5	✓	✓	4	✓	1	✓
EPS-based localisation	3	3			4		1	✓
Magnetic field-based localisation	4	1			1		1	✓
Camera	0	5	✓	✓	1		1	

Table 2.2: Pros and cons of each sensor modality, where the modalities in bold were deemed suitable and are or will be tested with additional experiments. We graded based on personal opinion the comfortability, sensitivity and integration of the signal.

In table 2.2, we scored all the sensors based on comfortability, sensitivity and integration to Bluesense. Comfortability was scored based on the following: If it can be worn on the only on the wrist like a watch it gets a score of 5, if it requires extra sensors or a set up that is so uncomfortable that the user would not be able to pursue their daily activities it gets a score of 1, everything in between gets scored based on personal opinion. Sensitivity was evaluated based on experience of dealing with the sensor and evaluating the signal when ambient noise is present or not, when considering the head scratching signal of the particular sensor modality. The integration with Bluesense was graded based on the development required for each modality.

Table 2.2 shows, that based on comfort, the possibility to integrate it to the Bluesense platform and the possibility to fit the sensors in one wrist-worn device, we chose to investigate the following sensor modalities:

- IMU
- Microphone
- EPS
- EPS-based localisation

All of these sensors could fit onto the wrist; however, additional signal generators are needed to include the hand location information in our system. We propose to attach the signal generator to a collar using a badge-like case. Furthermore, IMU-based localisation was tested as well, to evaluate the importance of hand coordinated relative to the torso for scratch detection task.

2.3 Limitations of current scratch detection systems

To sum up the limitations of previously work that we are addressing in this thesis are the following:

- No comprehensive dataset to train and test and evaluate scratch detection accuracy in a real world scenario.
- No use of multiple modalities fused together to realise a sophisticated scratch detection system.
- No evaluation of ambient microphone for scratch detection.

Chapter 3

Multi-modal fusion using EPS and IMU

In this chapter we present an investigation of sensors and algorithms to realise a wearable scratch detection device. We collected a dataset, where each user wore 4 inertial measurement unit (IMU) sensors and one electric potential sensor (EPS). Data was collected from nine users, where each user followed a 40-min protocol, which involved scratching different parts of head, shoulder, and leg, as well as other activities such as walking, drinking water, brushing teeth, and typing to a computer. The dataset contained 813 scratching instances and 5 h 15 min of recorded data. We investigated the trade-offs between the number of devices worn (comfort) and accuracy. We trained the k-NN and random forest algorithms by using between 1 and 5 features per channel. We concluded that a scratch could be detected with 80.7% accuracy by using the random forest algorithm on hand coordinates, which required four devices. However, an f1 score of 70% could be achieved with k-NN with IMU and EPS data, which only required one device. Moreover, the fusion of IMU data with EPS data improved the accuracy and reduced the deviation between the folds.

The novelty of this work is as follows:

- New dataset using four IMUs, 1 EPS, and hand coordinates with ten activities and nine users with 813 scratches and a total of 5 h 15 min of recorded data. The novelty of this dataset lies in the large number of sensor modalities, to push the machine learning models to their limits and evaluate how well slightly different scratches can be differentiated from other activities. The new dataset includes more scratch locations as well as a more realistic NULL class.

- Observation that EPS data combined with IMU data reduce the error variance between the folds and increase the accuracy of human activity recognition.

- Comparison of the k-nearest neighbour and random forest algorithms by using different numbers of features scored by mutual information. Previous work has not explored feature selection for scratch detection from different modalities.

- Exploration of trade-offs between accuracy, comfort, and the number of devices.

3.1 Multimodal scratch dataset

A new dataset was built using data collected from nine users. It contained two classes and 10 subclasses, which are shown in Table 3.1. These activities were chosen because of their similarity to scratching in the frequency domain and because they occur in everyday life. Each user wore four IMU sensors and one EPS.

Class	Subclass
Scratch	Top of the head
	Back of the head
	Side of the head
	Shoulder
	Leg
Null	Null
	Drink water
	Brush teeth
	Wash hands
	Walk
	Computer

Table 3.1: Ten classes of the dataset with a null class

As can be seen from Table 3.2, in all, we recorded 40 min of scratching and 5 h 15 min of data. The dataset had ten classes, and some activities were performed simultaneously in order to find the limits of the machine learning models and their ability to learn the subtle differences between very similar activities. In our case, the subtle differences were scratching the top of the head, back of the head, and side of the head.

	Dataset parameters
Age	22–40 years
Males	6
Females	3
Scratching instances	813
Total scratching time	40 min
Collection time per user	35 min
Total dataset time	5 h15 min

Table 3.2: Parameters of the dataset

The participants had different hair lengths, and there were six males and three females. The diversity of the users' biological features was needed to evaluate whether the model could generalise enough to detect the same activities on unseen users. There were no overweight or underweight participants; all of the participants had average body shapes. It was also important to have a long duration and a diverse null class, so that the model could learn the difference required to distinguish between scratching and other activities.

3.1.1 Sensor setup

The IMU is based on the BlueSense technology [23]. BlueSense gives the raw acceleration, rate of turn, and magnetic field data and can provide the quaternion data, which encode the orientation of the device. Moreover, the quaternion data from the four IMUs could be used to obtain the approximate hand coordinates (explained in Section 3.1.1). Furthermore, BlueSense can be extended with expansion boards containing additional sensors. For this dataset collection we used an expansion board with an EPS sensor [21].

EPS and IMUs

The sensor configuration for each user is shown in Figure 3.1 (c). This configuration was chosen in order to obtain the hand coordinates in the Cartesian system, the charge of the electric field, and the IMU data.

(a) IMU

(b) Electric potential sensor

(c) Sensor configuration during data collection. Each person wore four IMUs (on torso, upper arm, lower arm, and hand) and 1 EPS device extended on the IMU on the hand.

Figure 3.1: Imu (a) and EPS (b) were used for the data collection. They were attached to the body by using four straps.

For this work, a device was a single object, IMU with EPS was one device, a sensor was a sensor modality (IMU and EPS were two sensors). We referred to a *channel* as an output of a sensor. Sensors can have one or more outputs. For example, in this situation, an IMU had 13 channels. They were as follows: three channels of acceleration along the x, y, and z axes; three channels of the rate of turn along the x, y, and z axes; three channels of the magnetic field along the x, y, and z axes; and four channels indicating the device orientation in quaternions. EPS had only one output channel, which represented the electric potential.

The EPS, shown in Figure 3.1 (b), streamed the voltage data using the ADC channel to Bluesense [23]. Once the electric field was disrupted [21], it could be observed in the signal. Moreover, EPS had the ability to detect 50-Hz grid voltage when the computer was connected to the power source, which allowed us to accurately detect typing to a computer.

Hand coordinates

Hand coordinates were computed after the data collection. The quaternion data captured the orientation of each sensor. The sensor positioning is shown in Figure 3.1 (c). By getting the orientation of each sensor and by using a vector for each limb, we could sum up all of the Cartesian coordinates of each joint to obtain the hand coordinates.

3.1.2 Data collection

The dataset collection was approved by University of Sussex Ethical Committee, application n. ER/ZJ70/1. The participants were recruited in the Engineering and Informatics building for the participation in the data collection. During the data collection, the participants needed to follow the defined protocol.

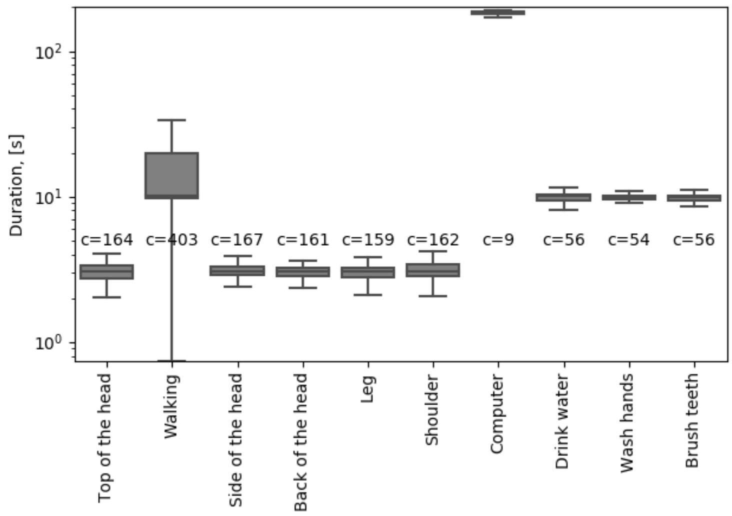


Figure 3.2: Duration of each activity in the collected dataset. The value of c denotes the number of occurrences of each activity across the entire dataset.

We defined a protocol, displayed in Figure 3.3, which allowed us to collect a wide variety of scratches. The data collection took place in the lab, and we attempted to obtain the most naturalistic dataset with a varied null class. In particular, we collected a dataset consisting of three different scratch locations on the head (top, side, and back) and scratch locations on other parts of the body (leg and shoulder). The scratches were performed with two different intensities (intense and moderate). In order to include a realistic null class in the dataset, we asked the participants to walk for 20 s in the office in between groups of scratches. In addition, we asked the participants to perform a number of other hand gestures, including simulating washing hands and brushing teeth, drinking water, and typing on a computer, as a way of including a more realistic set of activities to evaluate how well scratch could be distinguished from the other activities of daily living.

The data collection was a tedious process and included many activities. An application was used to show the current and the next activity on an IPad, using the application 'Seconds' [1]. Undesired deviation from the given protocol occurred during the collection. However, the labels of activities were adjusted after the data collection to adjust to the

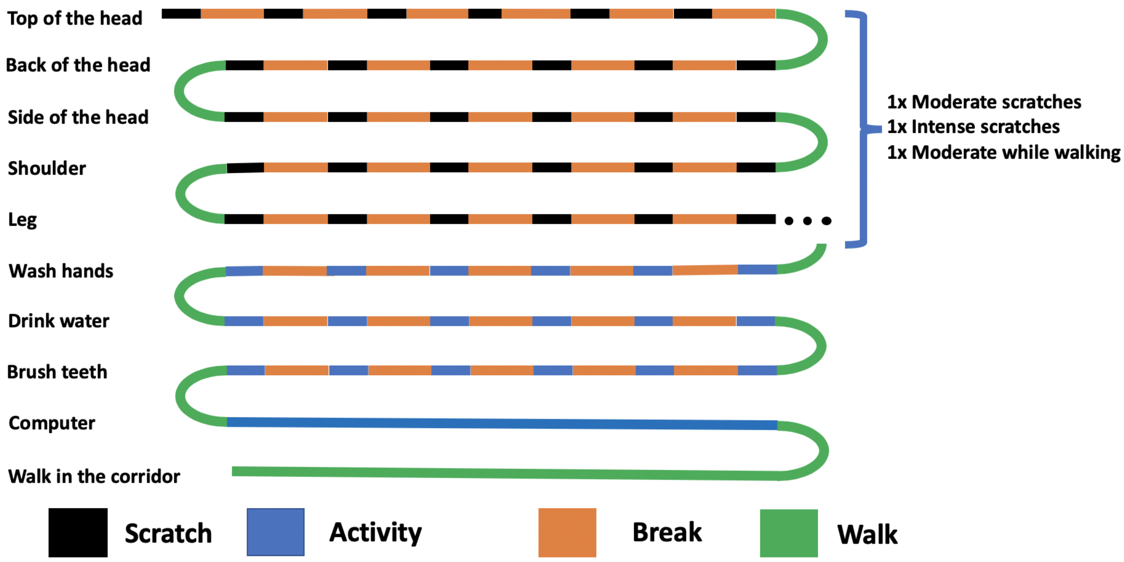


Figure 3.3: Protocol of data collection. First, the scratches were performed at three different intensities and all of the other activities for the null class were performed thereafter. A scratch occurred for 3 s, and then, there was a 10-s break. It was repeated six times, and before the next activity, a 20-s walk was performed.

deviations.

3.2 Human activity recognition pipeline

The human activity recognition process is a process that requires a specialised pipeline for each case. The pipeline that we used for the human activity recognition in this study is shown in Figure 3.4.

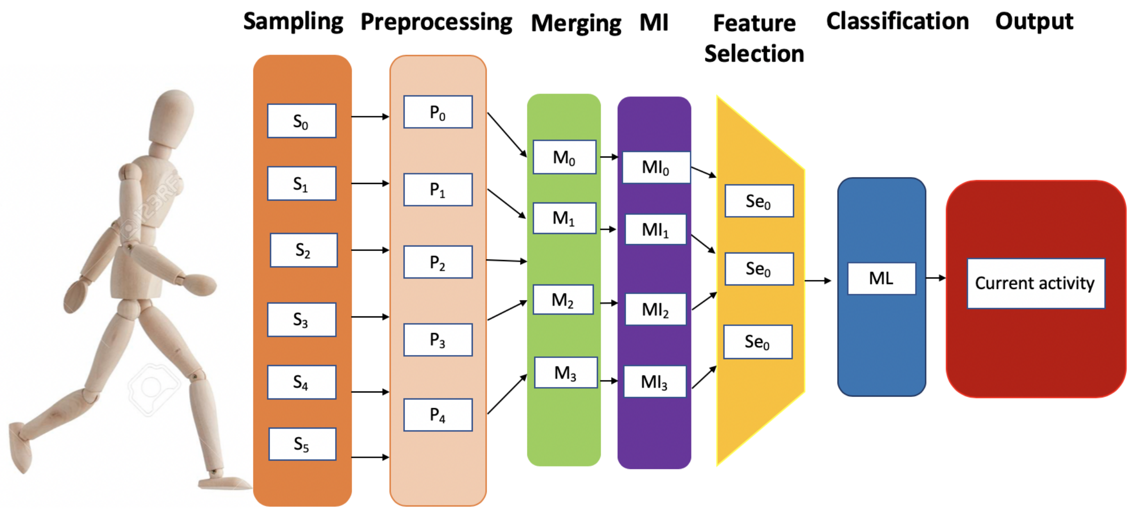


Figure 3.4: Human activity recognition pipeline for scratch detection. The data were sampled from the sensors and then preprocessed. The features with the highest MI score were used for the classification task.

3.2.1 Data Cleaning and Pre-processing

The data were collected using five devices (four IMUs and one EPS). However, with the additional development of the firmware, the IMUs and the EPS could be deployed using only one device. Thus, in the later stages, the number of devices that required EPS and IMUs was 1.

The sampling rates of the IMUs and the EPS were 100 Hz and 1 kHz, respectively. After sampling, the data were stored on the local SD cards. All the devices were synchronised; however, they did not log the data at exactly the same time. Because of the different time stamps, an interpolation technique was considered. As the quaternion data had a value of an angle, linear interpolation was not possible. Thus, an ASOF function was used. ASOF merged the data according to the nearest timestamps rather than the equal timestamps. A time delta equal to 10 ms was chosen: if the nearest timestamp was further than 10 ms, then the function did not choose the nearest value and assigned NaN.

The EPS was very sensitive, and hence, the collected data had a considerable amount of fluctuation. Thus, a low-pass Butterworth filter was applied to smooth the signal, and then the signal was resampled to match the 100-Hz IMUs' frequency.

3.2.2 Channels

Additional data, apart from the sampled data, were computed, in order to obtain more information:

- These additional data included the hand coordinates, which are described in Section 3.1.1.

- In the formula, m is any modality with the x, y, z projections and m_{xyz} is the magnitude, which was calculated to determine whether the magnitude could enable the device to achieve a relatively high performance.

$$m_{xyz} = \sqrt{m_x^2 + m_y^2 + m_z^2} \tag{3.1}$$

The magnitude was computed using formula 3.1 for the acceleration, rotation, and hand coordinates.

Acc_x	$Quat_0$	Gyr_x	$Hand_x$	EPS
Acc_y	$Quat_1$	Gyr_y	$Hand_y$	-
Acc_z	$Quat_2$	Gyr_z	$Hand_z$	-
Acc_{xyz}	$Quat_3$	Gyr_{xyz}	$Hand_{xyz}$	-

Table 3.3: Channels. All the channels used in activity recognition are displayed including acceleration, rotation, orientation, hand coordinates, and the EPS.

Thereafter, sliding windows of the time series were generated with a window length of 0.4 s. This time was chosen on the basis of the fact that a scratch is an activity which occurs for a short duration of time.

3.2.3 Feature Extraction

From the features shown in Table 3.3, we had to extract features from the sliding windows. The features chosen for this case are shown in Table 3.4. The mean and the variance enabled us to defines the distribution. The percentiles allowed us to detect the key points in the distribution and avoid the outliers, contradictory to the minimum and the maximum functions. The mean crossing rate and the zero crossing rate were used to evaluate how periodic the signal was.

3.2.4 Feature Selection

Two of the most common methods used to select the most important features are the filter and wrapper methods. As we had 17 channels, there were 170 unique features. The wrapper method would take an unreasonable amount of time to find the best combination of features. Therefore, the filter method that used the mutual information [9] algorithm was used to select the features carrying the highest amount of information.

Domain	Features
Statistical	Mean
	Variance
	Percentile 25%
	Percentile 50%
	Percentile 75%
	Percentile 90%
	Mean crossing rate
	Zero crossing rate
Frequency	Energy

Table 3.4: List of features

$$\mathrm{I}(X;Y) = \sum_{y\in\mathcal{Y}}\sum_{x\in\mathcal{X}} p_{(X,Y)}(x,y)\log\left(\frac{p_{(X,Y)}(x,y)}{p_X(x)\,p_Y(y)}\right) \tag{3.2}$$

where p (X,Y) is the joint probability mass function of X, and $p_{(X)}$ and $p_{(Y)}$ are the marginal probability mass functions of X and Y, respectively, where X is the bin of the feature and Y is the class. For each feature, we created 100 equal-size bins and performed a small correction by adding 10^{-10} to the division so that when there were empty bins, division by zero was possible.

3.2.5 Machine Learning

As the range of the data varied considerably, all of the extracted features were normalised as shown in equation 3.3.

$$n_{normalized} = \frac{n_{feature} - \mu}{\sigma} \tag{3.3}$$

In the equation, µ is the mean of the feature in the training set, n_{raw} is the feature, and σ is the standard deviation of the feature in the training set.

k-Nearest Neighbour

Next, k-nearest neighbour is a model which is fast to train and has a proven record of successful applications in many areas. However, it is slow to compute predictions and is very susceptible to outliers. Therefore, choosing too many correlated features or features that did not bring valuable information toward the decision making, would only diminish

ML model	Modalities	Number of devices	Number of features				
			1	2	3	4	5
k-NN	xyz	4	38.01±0.7%	59.3±2.5%	61.2±1.9%	**62.7±1.9%**	62.1±2%
	IMU	1	41.4±6.7%	45.4±8.3%	46.3±7.8%	46.4±8%	**47±7.3%**
	IMU+EPS	1	43.2±2.1%	50.3±2.4%	52±2.1%	52.7±2.2%	**53.4±2.4%**
	IMU+xyz	4	43.8±6.2%	52.5±0.9%	55.9±3.2%	55.9±2.8%	**56.3±28%**
	IMU+xyz+EPS	4	47.9±1%	52.2±1.7%	54±2%	54.7±0.2%	**55.6±2%**
Random forest	xyz	1	55.1±2.5%	66.6±3.5%	66.7±3.2%	66.8±3.2%	**66.9±3.2%**
	IMU	1	45.6±8.2%	47.7±8.4%	48.1±8.5%	48.4±8.6%	**49±8.5%**
	IMU+EPS	1	45.2±4.2%	49.6±5.4%	50.6±5.7%	51.8±6%	**52.1±5.8%**
	IMU+xyz	4	50.6±3.3%	52.6±4%	51.96±4.5%	52±4.6%	**53.8±4.6%**
	IMU+xyz+EPS	4	56.3±2.2%	57.4±3%	57.8±3.1%	57.8±3.1%	**57.8±3%**

Table 3.5: Results table of the ten-class classification with the displayed Macro f1 score. Note that overall, a better performance was achieved with the random forest model. Further, xyz is a simplified notation for the hand coordinates data.

the model's ability to recognise activities. Therefore, the value of k was set as 100, and the sklearn implementation was used.

Random Forest

Random forest is ensemble-based learning method. This model has seen considerable success in regression and classification tasks. Given that the decision trees can distinguish important features, a larger number of features will lead to more accurate predictions. For the random forest model, we chose to use 100 trees. In this case, we used the random forest algorithm from the sklearn Python library.

3.2.6 Performance measurement

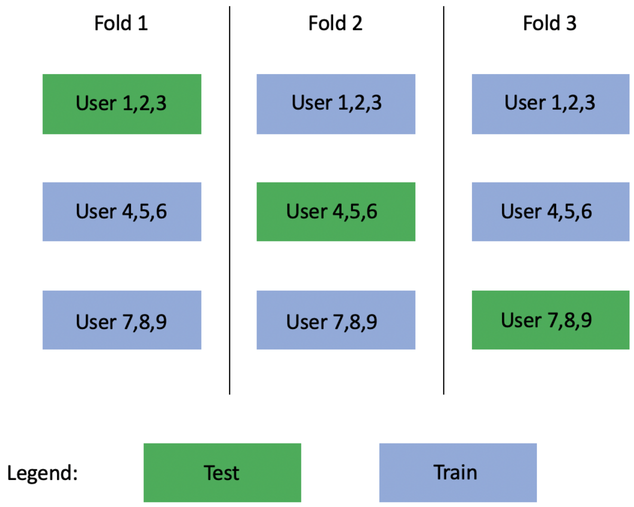


Figure 3.5: K-fold cross-validation where the data from three unseen users were left out for testing.

The goal of this project was to have a universal system for all the users. To achieve this goal, we used a three-fold cross-validation. During the cross-validation, the users were grouped into three groups of three users, and during each validation, each test was performed on the groups of three.

In this research project, a confusion matrix and the macro f1 score were chosen to evaluate the performance of the model and its ability to generalise. We chose to use the macro score in order to see how well the model recognised each class.

$$F_{1c} = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \tag{3.4}$$

$$F_{1Macro} = \sum_{c=1}^{n} F_{1c} \tag{3.5}$$

The macro f1 score was calculated using equation 3.5. The score for each class affected the overall results. Because of the class imbalance in the dataset (40 min from 5 h 15 min was scratches, and the rest was null), using the macro f1 score, we achieved a high performance when a majority of the data belonged to the null class.

Fine activity recognition

First, the models were tested to determine how well they could classify the ten classes. As there were very similar classes and multiple activities took place at the same time, this dataset was created to push the limits of the machine learning models.

Coarse scratch/non-scratch activity recognition

The 10-class classification task was meant to push the machine learning model to distinguish the subtle difference between the activities. By evaluating how well a scratch was distinguishable from a complex null class, we could determine how well a scratch could be detected in realistic scenarios. For this part of the experiment, the model was not retrained, but the labels were changed to either scratch or null. This implied that if ‘scratch top of the head’ was confused with ‘scratch back of the head’, the f1 score would not be reduced. The classes of the scratch and the null class are shown in Table 3.1.

3.3 Results

3.3.1 Most important features

With the use of mutual information, a heatmap of the most important features was produced; it is shown in Figure 3.6. For each channel, five features (which carried the highest amount of information) were selected and are displayed in Table 3.6.

Channel	f1	f2	f3	f4	f5
$\mathrm{Acc_x}$	var	90 %	75%	mean	median
$\mathrm{Acc_y}$	mean	75 %	90%	median	25%
$\mathrm{Acc_z}$	25%	mean	median	var	25%
$\mathrm{Gyr_x}$	var	25 %	90%	75%	median
$\mathrm{Gyr_y}$	var	90 %	25%	75%	median
$\mathrm{Gyr_z}$	var	90 %	75%	25%	median
$\mathrm{Quat_0}$	var	90 %	75%	25%	median
$\mathrm{Quat_1}$	75%	90 %	median	mean	25%
$\mathrm{Quat_2}$	90	75%	median	25%	mean
$\mathrm{Quat_3}$	var	90 %	75%	median	mean
$\mathrm{Acc_{xyz}}$	mean	median	energy	25%	75 % %
$\mathrm{Gyr_{xyz}}$	var	90 %	75%	mean	median
$\mathrm{Hand_x}$	var	90 %	25%	75%	median
$\mathrm{Hand_y}$	25%	median	75%	90 %	mean
$\mathrm{Hand_z}$	25%	mean	median	75%	90 %
$\mathrm{Hand_{xyz}}$	75 %	90 %	median	25 %	mean
Eps	mx0	var	90 %	mean	75 %

Table 3.6: Selected features for each channel

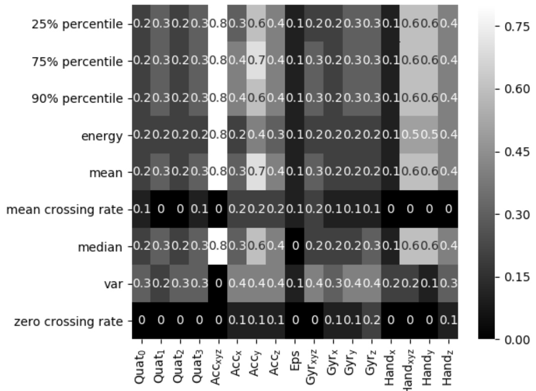


Figure 3.6: Feature heat map based on mutual information. The lighter shade indicates that the feature of the channel carried more information than the darker shade.

As can be seen, the variance carried the largest amount of information for most of the channels, and the percentiles showed good performance as well. The results displayed in Figure 3.6 revealed that the EPS carried the smallest amount of information as a channel.

The mean crossing rate and the zero crossing rate were not informative features in most of the cases.

3.3.2 Fine activity recognition results

In Table 3.5, the results are presented for the ten-class classification task using the k-NN and the random forest models trained on a different number of features from 1 to 5 per channel. The accuracy is shown as the mean and the standard deviation between the folds.

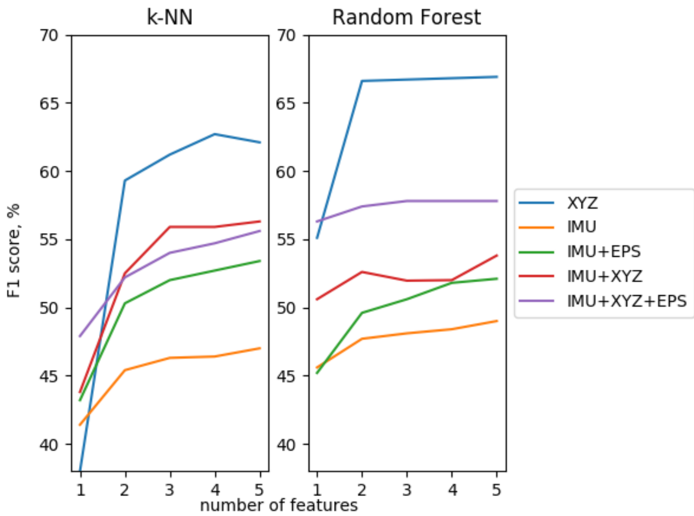


Figure 3.7: F1 score vs. number of features for ten-class classification

k-NN

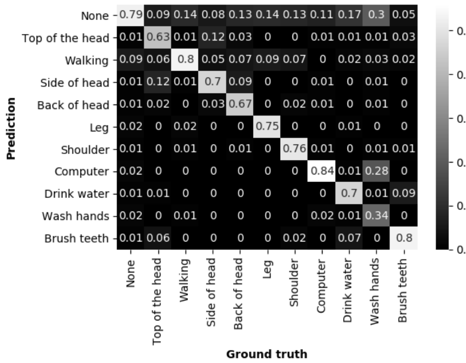
The best results were observed when only the hand coordinates with four features per channel were used. The f1 score reached 62.7%. To achieve this result, four devices were required. In contrast, with only the IMUs and the EPS, an f1 score of 52.3% was achieved and only one device was required.

Combining the hand coordinates with the IMU or IMU+EPS data only worsened the results.

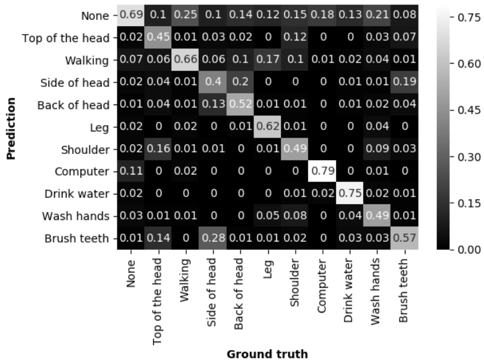
An f1 score of 53.4% was achieved with one device (IMUs and EPS), while with only IMUs, the best achieved result was 47%.

Random forest

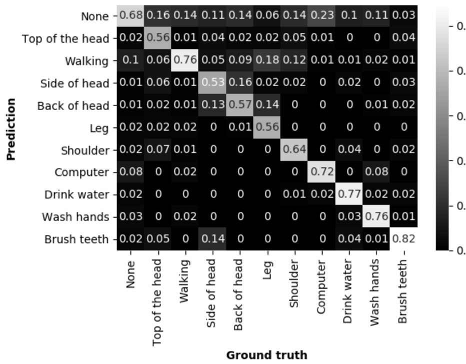
The random forest model showed an overall better performance. The best performance was observed when using hand coordinates with four features per channel, where an f1



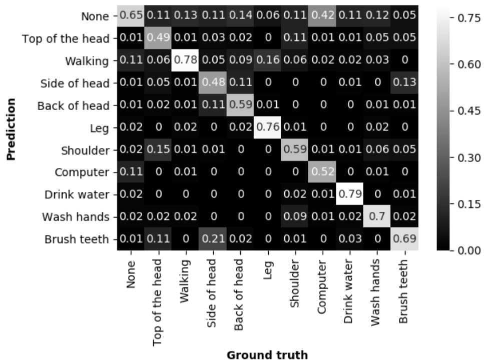
(a) Hand coordinates



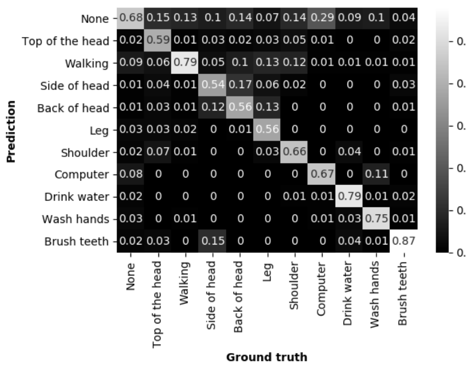
(b) IMU channel



(c) IMU and hand coordinates



(d) IMU and EPS channel



(e) Hand coordinates, IMU, and EPS

Figure 3.8: Confusion matrices when the inputs were two features per channel and the model was random forest.

score of 66.9% was achieved for the ten-class classification. However, this task required the attachment of four devices on the body. We observed that the performance improved with an increase in the number of features introduced.

An f1 score of 52.9% was achieved with one device (IMUs and EPS). Nonetheless, with IMUs only, the best achieved result was 49%.

In Figure 4.1, the confusion matrix shows which classes were mostly confused when tested on the RF model using two features per channel. In all the confusion matrices, note that 'Scratch the top of the head', 'Scratch the side of the head' and 'Scratch the back of the head' were considerably confused. In Figure (b), it can be seen that 'Brushing teeth' was confused with 'Scratch the side of the head' and vice versa, and in Figure (a), 'Washing hands' was confused with 'Typing on a computer'. When both the modalities were combined, as seen in Figure (c), the performance on these two classes drastically improved. When EPS was introduced, in Figure (d), it can be seen that the number of errors decreased in the scratch classes, as compared to when only IMUs were used.

3.3.3 Coarse scratch/non-scratch activity recognition results

It was very challenging to classify ten classes. However, to evaluate how well a scratch was recognisable, it was sufficient to distinguish a scratch from any other activity. In Table 3.7, the results are presented for the binary classification task.

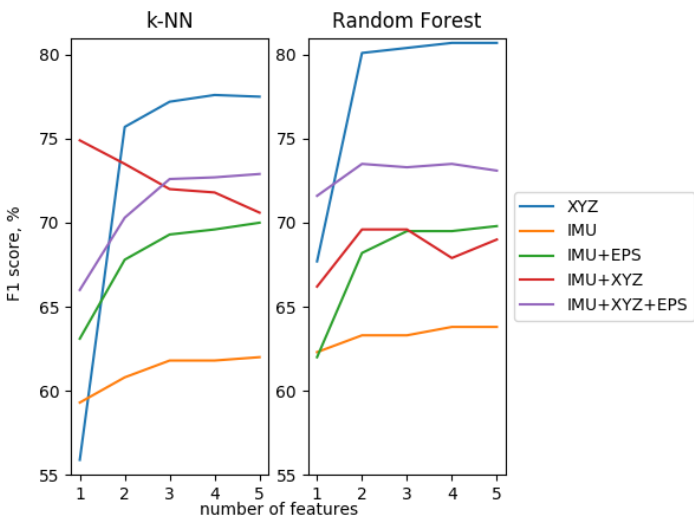


Figure 3.9: F1 score vs. number of features for binary classification

k-NN

The best result was achieved again by using the hand coordinates data. We achieved an f1 score of 77.6%. With a single device (IMUs and EPS), we achieved an f1 score of 70% with a standard deviation of 2.9%. By using only the IMUs, we achieved an f1 score of 62% with a standard deviation of 8.6%.

The EPS allowed us to reduce the standard deviation between the folds as compared to the IMU results. However, four devices were required with the extracted hand coordinates

ML model	**Modalities**	**Number of devices**	**Number of features**				
			1	2	3	4	5
k-NN	xyz	4	55.89±2.7%	75.7±3%	77.2±2.5%	**77.6±2.7%**	77.5±2.4%
	IMU	1	59.3±5.5%	60.8±8.5%	61.8±8.4%	61.8±8.4%	**62±8.6**%
	IMU+EPS	1	63.1±2.7%	67.8±2.9%	69.3±2.9%	69.6±2.8%	**70±2.9%**
	IMU+xyz	4	**74.9±1.4**%	73.5±78.3%	72±3.6%	71.8±3.8%	70.6±5.11%
	IMU+xyz+EPS	4	66±4.4%	70.3±4.5%	72.6±4.5%	72.7±4.5%	**72.9±4.7%**
Random forest	xyz	4	67.7±2%	80.1±2.7%	80.4±2.3%	80.7±2.7%	**80.7±2.6%**
	IMU	1	62.3±6.3%	63.3±8.7%	63.3±9.2%	063.8±10%	**63.8 ±10%**
	IMU+EPS	1	62.09±1.7%	68.2±6.4%	69.5±6.8%	69.5±7.6%	**69.8±7.2%**
	IMU+xyz	4	66.2±3.9%	69.6±2.7%	**69.6±2.5%**	67.9±3.3%	69±3.3%
	IMU+xyz+EPS	4	71.6±3.3%	73.5±3.4%	**73.7±3.3%**	73.5±3.1%	73.1±3.6%

Table 3.7: Results table of binary classification with the displayed macro f1 score. Note that overall, a better performance was achieved with the random forest model. Further, xyz is a simplified notation for the hand coordinates data.

to achieve the best results.

Random forest

The best score was achieved by using the random forest model with five features per channel. It reached an f1 score of 80.7% for scratch detection. With only the IMUs and the EPS, an accuracy of 69.8% was achieved, and the standard deviation between the folds was only 7.2%. The use of only IMUs resulted in an f1 score of 63.8% with a standard deviation of 10%.

The same behaviour as observed as that with the k-NN model. The EPS allowed us to reduce the standard deviation between the folds as compared to the IMU results. However, four devices were needed with the extracted hand coordinates to achieve the best results.

3.4 Discussion

3.4.1 Baseline Results

Currently, the standard approach toward human activity recognition is predicting an activity on the basis of the IMU data. In this work, the baseline results were achieved by using the extracted features from the IMU data. Moreover, these results only required the deployment of one device.

The baseline results of the new dataset were as follows:

- For the ten-class classification with k-NN, the best result was 47% ± 7.3%.
- For the ten-class classification with RF, the best result was 49% ± 8.5%.
- For the binary classification with k-NN, the best result was 62% ± 8.7%.
- For the binary classification with RF, the best result was 63.8% ± 10%.

3.4.2 Multimodal Fusion

The fusion of data between IMUs, EPS, and the extracted hand coordinates did not always result in a better performance.

IMUs and EPS

IMUs and EPS required one device on the wrist. Fusing the data and classification from the IMU and EPS data provided a slightly better performance than only using IMUs with the random forest and the k-NN models. Using the binary classification task with IMUs and EPS resulted in an accuracy of 70% when using k-NN and 69.8% when using RF. Moreover, with k-NN, the standard deviation decreased between the folds from 8.6% to 2.9%; it decreased from 10% to 7.2% in the case of the random forest model. Compared with the baseline results, there was a small improvement in the f1 score. In both the cases, fusing the IMU and EPS data decreased the standard deviation and improved the results obtained using both the models, but not significantly. The reduction in the standard deviation of the f1 score between the folds showed that the model's ability to recognise the activities was good for a diverse group of people. Moreover, as the EPS could be deployed on the IMU located on the hand, it did not create additional discomfort.

IMU and hand coordinates

For the ten-class and binary classifications, combining these modalities yielded better results than using only IMUs, but worse than when using only the hand coordinates. Using RF and binary classification, IMU achieved an accuracy of 63.8% $\pm$ 10%; with only the hand coordinates, the result was 80.7% $\pm$ 2.6%, and when combined, it achieved 69.6% $\pm$ 2.5%. The same behaviour was observed with k-NN and the ten-class classification task.

Each activity was associated more with a certain location relative to the torso than with specific movements. Thus, additional information could be redundant and decrease the accuracy. In this case, fusing the hand coordinates data with the IMU data did not provide with any gains, as compared to the results achieved using only the hand coordinates data.

IMU, hand coordinates, and EPS

Combining all the modalities outperformed the use of only the hand coordinates data. For the binary classification task using RF and binary classification, IMU achieved an accuracy of 63.8% $\pm$ 10%; with the use of only the hand coordinates data, the result was 80.7% $\pm$ 2.6%, and with the use of IMU with the hand coordinates data, the model achieved an accuracy of 69.6% $\pm$ 2.5%. Moreover, with the use of IMU, hand coordinates, and EPS data, the model was able to achieve an f1 score of 73% $\pm$ 2.6%.

The dataset was biased to the location, implying that each activity in the dataset was associated with a certain location relative to the torso and this particular feature worked exceptionally well with this dataset. The fusion of the dataset did not bring a drastic improvement in the accuracy but significantly decreased the standard deviation between the folds, as compared to the baseline results. However, it still did not outperform the achieved results when only the hand coordinates were used.

3.4.3 Number of Features

As can be seen from Figures 3.7 and 3.9, there was an increase in the accuracy, but later, the results plateaued. This might be caused by the mutual information algorithm, which did not take into account the fact that certain features were correlated and that the redundant information did not improve the performance. This was observed for both the classifiers, with the exception of k-NN with IMU and hand coordinates.

3.4.4 Comfort and Accuracy Trade-off

To deploy a scratch detection system for clinical trials, the smallest number of devices and the least invasive device must be chosen. Sensors set up with four devices (shown in Figure 3.1) cannot be used in any medical or clinical study, as this set up will make daily activities uncomfortable and there will be a higher risk of failure because of the large number of devices in use.

For the data collection in the present study, four devices were used. In general, more information gives better results. In Figures 3.5 and 3.7, it is shown that increasing the number of devices produced better results with the exception of k-NN and hand coordinates. However, discomfort is a major drawback for the deployment of a human activity recognition system. If comfort is the priority: one device with IMUs and EPS on the wrist might be sufficient with a 70% accuracy, and if accuracy is the priority, then with four devices (as shown in Figure 3.1) should be used. Note that an f1 score of 80.7% could be achieved by using only the hand coordinates.

3.4.5 Future Work

In Section 5, note that hand coordinates relative to the torso were needed to achieve the best performance for this dataset. To build a comfortable system, new localisation techniques should be explored, so that the sensors can fit on one wrist. We suggest to explore localisation techniques, such as ultra wide band. For example, PosXYZ [7] needs only two devices (thw slave and the master) to be deployed in such a system. The slave device will be attached to the wrist, and the master device will be the reference point. With the use of the ultra-wide-band technology, the location of the wrist as compared to the torso can be computed and can be used for human activity recognition. Thus, the hand coordinates will enable one to achieve higher accuracy, without needing four straps on the torso, upper arm, lower arm, and wrist, respectively.

Exploring different feature selection techniques, such as MRMR [26], can lead to higher performance. MRMR is a minimum redundancy feature selection algorithm, which also takes into consideration how redundant a feature is as compared to the other selected features. Moreover, to determine what accuracy can be achieved for this dataset, an investigation of deep learning models, such as DeepConvLSTM [19], can produce substantially higher accuracy results.

Finally, new hardware can be developed for head scratch detection, such as EPS-based glasses, which can work as a proximity sensor to detect when the hand is nearby, thus increasing the recognition of head scratches.

Even though in this study, we explored feature-based fusion, decision-based fusion has shown good performance as well [4]. Thus, exploring hierarchical classifiers with decision-based fusion should be able to yield even better scratch detection results.

The four participants in the study [25] were smokers, and they needed to tap the sensor to flag when the smoking happened. This gives the insight that in a similar study for scratch detection, active learning should be incorporated so that a user will be queried if he scratches and the scratch will be flagged.

3.5 Conclusions

During this experiment, we explored how the fusion of different sensor modalities contributed toward accurate scratch detection by using different numbers of features per channel and common machine learning models, such as k-NN and random forest. For this task, a dataset was built with the data collected for ten different activities to investigate the limitations of each model and explore the trade-off between the number of sensor modalities, number of features, and machine learning models.

The key results were as follows:

- The best baseline result for detecting a scratch with a simple IMU was an f1 score of 63.8% obtained using RF, which required only one device.
- The best result obtained using one device was an f1 score of 70% for scratch detection. It was achieved by using k-NN with IMU and EPS data.
- The best overall result was an f1 score of 80.7%. It was achieved for the binary scratch detection using the hand coordinates data and the RF model, which required four devices.

- Fusing EPS data with IMU data consistently increased the accuracy and reduced the deviation between the folds, as compared to using only the IMU data.

We found that hand coordinates alone enabled us to achieve the highest accuracy to detect all the activities. However, this dataset was biased to perform well on these data as each position was associated with a certain activity. However, with the current technology, such a model would require four IMUs on the torso, upper and lower arms, and the hand. Therefore, it is not convenient to use this setup on a large number of people.

For the best performance on this dataset, the hand coordinates data with five features should be used to achieve the highest accuracy of 80.7% for detecting scratches. However, if a comfortable system is a priority and accuracy can be sacrificed, then 70% accuracy can be achieved with a single device using IMUs and EPS.

Chapter 4

Scalp scratch detection using a wrist-worn microphone

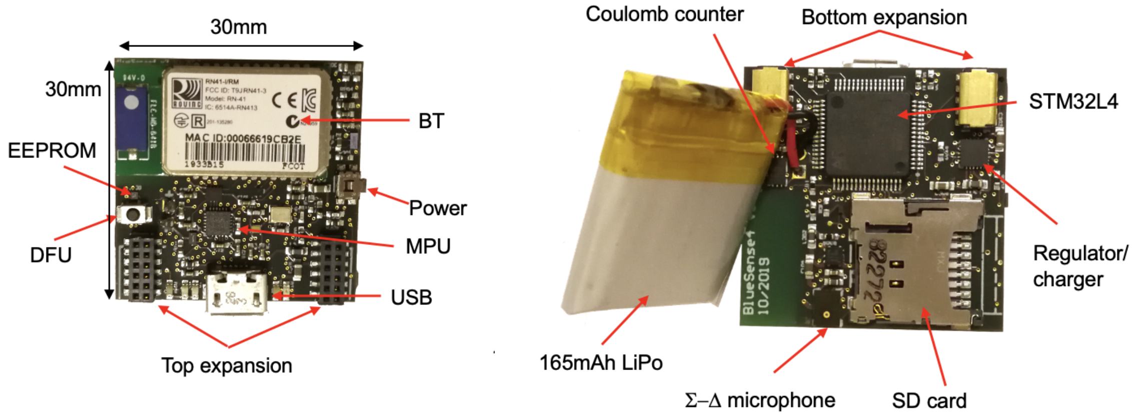
In this chapter, we evaluate how sound, collected by a wrist-worn microphone, can be used to recognise and distinguish scratch occurrences. We collected a dataset from 4 users, where each user conducted 20 scalp scratches, 10 shoulder scratches and 18 scalp scratches with and without TV noise being played in the background. The classification was done by training a random forest model with engineered features and was tested using leave-one-user-out cross-validation. In a silent environment we achieved sensitivity of 83.75% ± 8.8% and specificity 78.5% ± 4% and in an environment with TV noise present sensitivity decreased to 61%±20.45% and specificity 78%±9.6%. These results indicate that sound may be one modality to include for a future multimodal scratch detection device.

Previous work used uncomfortable and hard to install body or bone conduction based microphones (see sec 2). In this work, we explore to which extent a wrist-worn microphone can be used for scratch detection. The novelties are:

- An evaluation of more comfortable wearable wrist-worn microphone's ability to detect scratch.
- An anotated dataset for scratch detection, which incorporates also different levels of ambient noise.
- A frequency analysis of the scratch from the wrist worn microphone data.
- Analysis and visualisation of how the information of the extracted features change when background noise is present and when it is not. We used Mutual Information - based feature selection.

- An evaluation of scratch detection algorithms.

4.1 Wearable Sensing platform

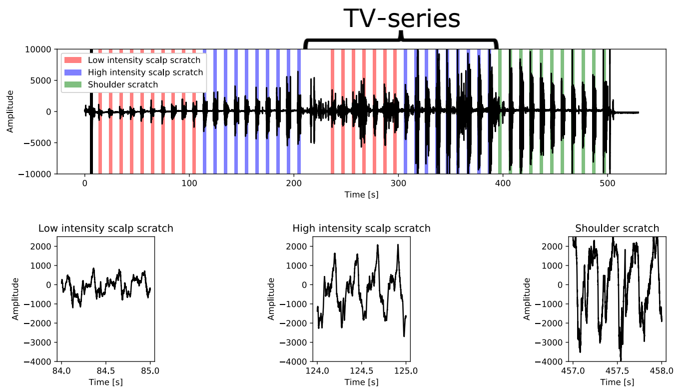


(a) (b)

Figure 4.1: The data collection platform is 30x30mm in size with connectors for additional expansions (a). The microphone is on the skin-facing side when worn on the body (b).

The platform [22] used for data collection is an in-house wearable sensing research platform. Its primary function is to be an Inertial Measurement Unit (IMU) and a digital microphone, which can be expanded using expansion connectors for sensor research purposes. The device is 30x30mm. It is based on an ARM Cortex M4 processor (STM32L496 from ST), which runs at 20MHz with our default firmware. The platform comprises a 9-axis inertial measurement unit (TDK Invensense ICM-20948), an digital MEMS microphone (ST MP34DT05-A), a micro-SD card, Bluetooth 2 and USB interfaces, a fuel gauge for built-in power measurements (LTC2942), an EEPROM to store configuration (M24128). The processor built-in real-time clock (RTC) is operated from a dedicated 32KHz quartz (10ppm frequency tolerance). The platform operates at 3V from a lithium polymer battery (165mAh) with a LTC3553 voltage regulator. The expansion connectors provide I2C, Serial Peripheral Interface (SPI), universal asynchronous receiver/transmitter (UART), analog-to-digital converter (ADC) inputs and digital to analog converter (DAC) output for expansion boards.

The device firmware has been designed for ease of use. It allows without any programming to acquire the data from the built-in sensors or external ADC inputs through a command line interface. Data can be streamed over Bluetooth or USB, or stored in the



(a) Data collection

(b) The sensor was attached on the wrist using a band.

Figure 4.2: The protocol contains 10 low-intensity, 10 high-intensity scalp scratches and 10 shoulder scratches which was recorded without any noise in the background. There were 18 recordings with noise in the background.

SD card. The current firmware allows to acquire in isolation IMU data, sound data, or analog inputs data from the expansion connector. It can also acquire multimodal data: i.e. it can simultaneously acquire IMU, sound and analog inputs in a multiplexed streaming and storage format. This is particularly relevant to acquire data for activity recognition based on a combination of multimodal sensors. The sample rate of IMU, microphone and ADC is fully configurable. The IMU data is also processed by the firmware to obtain the device orientation in quaternions. All data is time-stamped using the internal RTC.

The microphone is omnidirectional. It has a 64dB signal-to-noise ratio and a -26 dbFS sensitivity. It is on the bottom-side of the device (i.e. facing towards the skin if worn on the wrist), but it is not in direct contact with the skin due to the case. The microphone is clocked at 2MHz and provides a 1-bit digital pulse density modulation output, which comes straight from its sigma-delta analog to digital converter. The data is converted to an audio signal using a 3rd order Sinc filter, each with a decimation ratio of 82 which yields a 16-bit 8KHz audio signal with a dynamic range has been experimentally tuned for typical ambient sounds and speech ($\leq dB$).

4.2 Evaluation of scratch detection

The aim of this work is to evaluate the effectiveness of scratch detection from sound collected by a wrist-worn microphone using three steps: data collection, analysis and classification.

4.2.1 Dataset collection

The aim of this protocol (Table 4.2) is to evaluate:

1. if scratching is detectable with the microphone in a silent environment (activities 1, 2 and 5)

2. if scratching is detectable when there is background noise (activities 3,4)

3. if scalp scratching (activities 1, 2) can be distinguished from shoulder scratching (activity 5).

The dataset was collected from 4 users, as seen in table 4.1. To have a representative evaluation we collected a dataset from 2 males with different hair lengths, as well as 2 females: one with loose hair and one with a tight pony tail.

User	Gender	Hair type	Hair length
1	Female	Straight (tight pony tail)	40cm
2	Male	Black, straight	10cm
3	Female	Straight loose	50cm
4	Male	Straight loose	20 cm

Table 4.1: Characteristics of the participants in the dataset.

For the data collection each user wore a sensor (Figure 4.1) on their main hand (Figure 4.2 (b)). As it is common to wear watches and bracelets on the wrist it is also the most adequate place for the device's location. For large scale data collections the device is not more invasive than a smart watch.

The protocol was shown on an Ipad using "Seconds Pro" [1]. This was a constrained recording. It was done to show that scratch can be recognised using wrist-worn microphone and if it can be recognised with ambient noise present. The audio was recorded at 8kHz and was stored to a SD card.

Acitivity ID	Activity	Repetition	Duration	Wait time
1	Low intensity scalp scratch	10	3 s	7 s
2	High intensity scalp scratch	10	3 s	7 s
3	Low intensity scalp scratch while watching TV-series (distance=50 cm, 1m, 2m)	3, 3, 3	3 s	7 s
4	High intensity scalp scratch while watching TV-series (distance=50 cm, 1m, 2m)	3, 3, 3	3 s	7 s
5	Shoulder scratch	10	3 s	7 s

Table 4.2: Protocol of data collection used to collect the dataset. Each participant did 30 scalp scratches and 10 decoy scratches while scratching the shoulder rather than the head. The wait time is the between scratches, where a person does nothing in order to visually assess from the signal when the scratch occurs.

4.2.2 Signal Visualisation

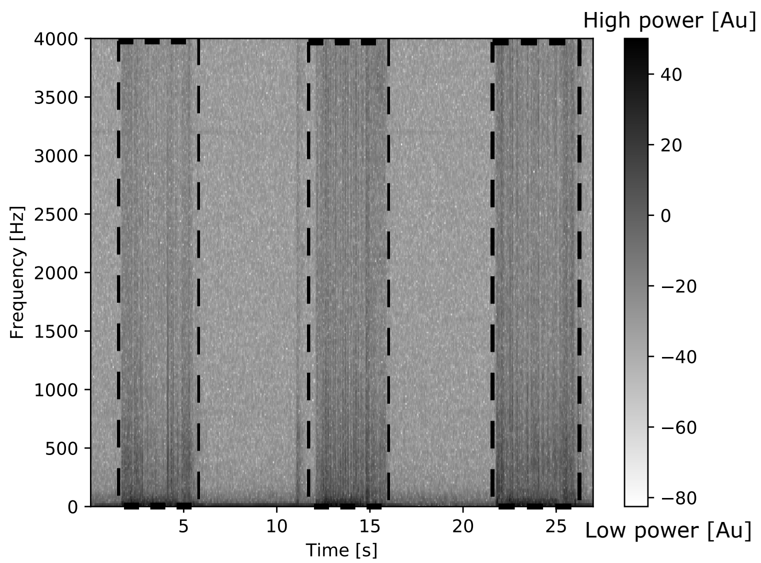


Figure 4.3: Spectogram of an excerpt of the recording (user 2). Three scratches (dashed boxes) and the null class are visibile.

The spectogram was created by computing Fast fourier transforms of a sliding window of 1000 samples (0.125 seconds) with an overlap of 50%. The higher size of the window was chosen as the noise originating from body movements is captured by the microphone, as seen in the high intensity scalp scratch in figure 4.2(a). From the figure 5.4, it can be seen that there is no observable clusters of dominant frequencies.

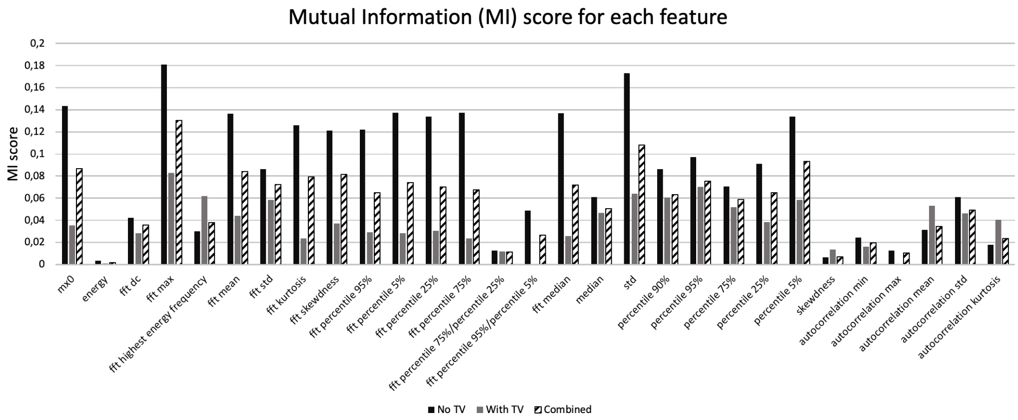


Figure 4.4: The Mutual Information score is shown for each feature for data, that was collected under three different conditions: without background noise (no TV), with background noise (with TV) and when these two datasets are combined. It can be seen, that when the TV-series are on, the MI score drops compared to when the scratches occur in a silent environment.

4.2.3 Feature Analysis

Two of the most common methods used to select the most important features are the filter and wrapper methods. The wrapper method would take an unreasonable amount of time to find the best combination of features from every permutation of the 85 features. Therefore, the filter method that used the mutual information [9] algorithm was used to select the features carrying the highest amount of information.

$$\mathrm{I}(X;Y) = \sum_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} p_{(X,Y)}(x,y) \log \left(\frac{p_{(X,Y)}(x,y)}{p_X(x)\, p_Y(y)} \right) \tag{4.1}$$

where p (X,Y) is the joint probability mass function of X, and $p_{(X)}$ and $p_{(Y)}$ are the marginal probability mass functions of X and Y, respectively, where X is the bin of the feature and Y is the class. For each feature, we created 100 equal-size bins and performed a small correction by adding 10^{-10} to the division so that when there were empty bins, division by zero was not occurring.

We show in figure 4.4, that the features extracted from the frequency domain (Fourier transform) carries the largest amount of information in the silent background. However, when TV noise are being played in the background (noisy), the amount of information that each feature carries for scratch detection is smaller.

In the table 4.3 we show, that the top 5 features in the silent and noisy backgrounds. It can be seen, that the MI scores are much lower from the data, which was collected with TV-series in the background, compared to the data collected in a silent background. The

	Silent background		Noisy background	
Order	Feature	MI score	Feature	MI score
1	FFT max	0.181	FFT max	0.083
2	stdD	0.173	75%	0.07
3	mx0	0.143	std	0.064
4	FFT 5%	0.137	fft max f	0.0602
5	FFT 75%	0.137	90 %	0.06

Table 4.3: The top features are shown with the associated Mutual Information score. FFT maximum amplitude had the largest MI score in both cases. However, the top features when noise is present have a much smaller MI score compared to the features extracted from the data, which was collected in silence.

top feature FFT maximum amplitude carried the largest amount of information. A key difference in both cases is that when TV-series are not present, top 5 features are from the frequency domain, while when TV-series are not present two (75% and 90% of the window) of the top 5 features are frequency invariant.

4.2.4 Classification

We trained a classification task using random forest (100 estimators with 5 max depth) model with Gini impurity and leave-one-user-out cross-validation. To evaluate if a user-independent scratch recognition model can be trained to work on unseen users, we train on User 1, 2, 3 and test on User 4; and then we repeat this with different test users. For the classification task we extracted the features (Table 4.4) from a sliding window. The sliding window size is 1000 samples (0.125 seconds) with an overlap of 50%. The null class contains everything except scratch. The labels were split into overlapping windows and the dominant label in the window was chosen. We calculated sensitivity or specificity by defining scratch as positives (P) and null class as negatives (N). To evaluate the performance we used sensitivity and specificity (Equation 1), with the $F1_{weighted}$ (Equation 2), where n_i is the class size and N is the total number of samples.

$$sensitivity = \frac{P_{True}}{P},\ specificity = \frac{N_{True}}{N} \tag{4.2}$$

$$F1_{weighted} = \sum_i 2w_i \frac{precision_i\ recall_i}{precision_i + recall_i},\ w_i = \frac{n_i}{N} \tag{4.3}$$

Type	Features	Dimension
$\mathcal{F}$	DC component of FFT	1
	Highest FFT value and frequency	2
	Max, Min, Mean, std of the FFT values	2
	Kurtosis, skewness of FFT	2
	Percentiles 5, 25, 50, 75, 90 of FFT	5
$\mathcal{T}$	Mean, std	1
	Percentiles 5, 25, 50, 75, 90	5
	Auto-correlation (min,max,mean,std)	4
	Mean-crossing rate	1
	Kurtosis, Skewness	2
$\mathcal{E}$	Energy of 50 Hz bands	60
	Total number of features	**87**

Table 4.4: We used 85 extracted features from the sliding windows of size 1000 (0.125 s) and 50% overlap. $\mathcal{T}$ are the features extracted from the sliding window, $\mathcal{F}$ - from Fourier transforms, $\mathcal{E}$ - energy of the frequency bands.

For the 1st aim (ref section 4.1) we trained and tested only on scratches (scalp and shoulder scratches (Activities 1, 2, 5) are one class) with no TV-series present. The model achieved sensitivity of 83.75%±8.8% and specificity of 78.5%±4%. $F1_{weighted}$ score was 82%. The worst performance is obtained with User 1, where the sensitivity is 72%.

For the 2nd aim, where we trained on all the data, but tested only on the section where TV-series (Activities 3, 4) are present. This yielded sensitivity of 61%±20.45%, specificity of 78%±9.6% and a $F1_{weighted}$ score of 73.75%.

For the 3rd aim we wanted to evaluate if shoulder scratch can be distinguished from scalp scratches. We performed a classification with 3 classes: scalp scratch, shoulder scratch and null class (Activities 1, 2, 5). The model was trained and tested only on the data where TV-series are not present in the background. We achieved that on average scalp scratch was recognised with an accuracy of 43%±15.8%, shoulder scratch with an accuracy of 39.25%±14.13%, null class 68%±14.54% and the $F1_{weighted}$ score was 65.4%.

Test number	Environment	Classes	User 1	User 2	User 3	User 4	Accuracy	$F1_{weighted}$
1	Silent	Accuracy of scratch	72 %	97 %	83 %	83 %	83.75%±8.8%	82%±2.3%
		Accuracy of null	83 %	75 %	74 %	82 %	78.5%±4%	
2	TV-Series	Accuracy of scratch	42 %	87%	40 %	75 %	61% ±20.45%	73.75% ±3.87%
		Accuracy of null	81 %	62%	88%	81%	78%±9.6%	
3	Silent	Accuracy of scalp scratch	51 %	40%	20%	63%	43±15.8%	65.4%±9.23%
		Accuracy of shoulder scratch	30 %	21 %	54%	52%	39.25%±14.13%	
		Accuracy of null	79 %	43 %	76 %	74 %	68% ±14.54%	

Table 4.5: The results of the classification show that in a silent environment scratch can be detected with 83.7% accuracy, while the null class with 78%. With TV-series in the background scratch with 61% and null with 78%. Shoulder scratches cannot be distinguished from scalp scratches. The accuracy is rate at which the model predicts the class correctly.

4.3 Discussion

The sensor platform is available to use with multiple sensor modalities. We show that sound is a strong predictor and can be used for sensor fusion. The fusion should improve the overall scratch detection. The dataset was collected from 4 users with different hair styles. We showed that by doing leave-one-user-out cross-validation sensitivity can be 83.75% and specificity 78.5% in a silent environment. Compared to the body-conduction microphone in [17] our performance is lower as our microphone captures the ambient sound whereas the body-conduction microphone is isolated using a gel and needs to be tightly fixed to the hand. Thus, there is a trade-off of comfort and performance.

Moreover, we observed that the MI scores of the features extracted from the data collected with background noise are smaller than from the data collected in silence for scratch detection. It means, if we were to use features selected in the silent environment and were to use them for classification when there is ambient noise, the performance is likely to degrade.

Furthermore, sound made by scratching is dependent on the type of the hair. For example, the lowest performance was achieved on User 1, who had hair tied in a tight pony tail. Another limitation is scratch detection when ambient noise is present. In this case we achieved a sensitivity of 61%±20.45%.

We observed that scratches do not seem to have a particular frequency signature, and thus filtering the ambient sound does not seem practical. When we listened to the recordings, the scratch was audible despite the TV-series playing in the background. It means that the sound of scratching is captured and more sophisticated algorithms, such as Deep Learning could be tried. Microphone beamforming could help to enhance the directional sensitivity of our device.

4.4 Conclusion

In this chapter, we show that sound can be used as a sensor modality to distinguish scratch. In a silent environment the sensitivity and specificity are 83.75% and 78.5%. On the other hand, in a noisy environment the sensitivity and specificity are 59.25% and 77.25%. Moreover, we came to the same conclusion by computing Mutual Information scores to see the importance of the features. The features extracted from the data that was collected with noise in the background had significantly smaller score compared with the data that was collected in a silent environment. With the ambient noise the model with the chosen features is not reliable enough to detect scratches with a high accuracy. It enforces the idea, that additional sensor modalities should be used together to distinguish the scalp scratches, as it will allow to compensate when ambient noise is present.

Ambient noise is a challenge and appropriate techniques should be used to diminish the impact of the noise on the predictions. A microphone would work extremely well for scratches, where there is a limited amount of noise, but for real daily activities it is not accurate enough to be deployed alone. However, it could significantly increase the scratch detection accuracy if it would be fused with other modalities. Our research platform allows us to jointly acquire sound and IMU [22] data and this will be explored in the future.

Chapter 5

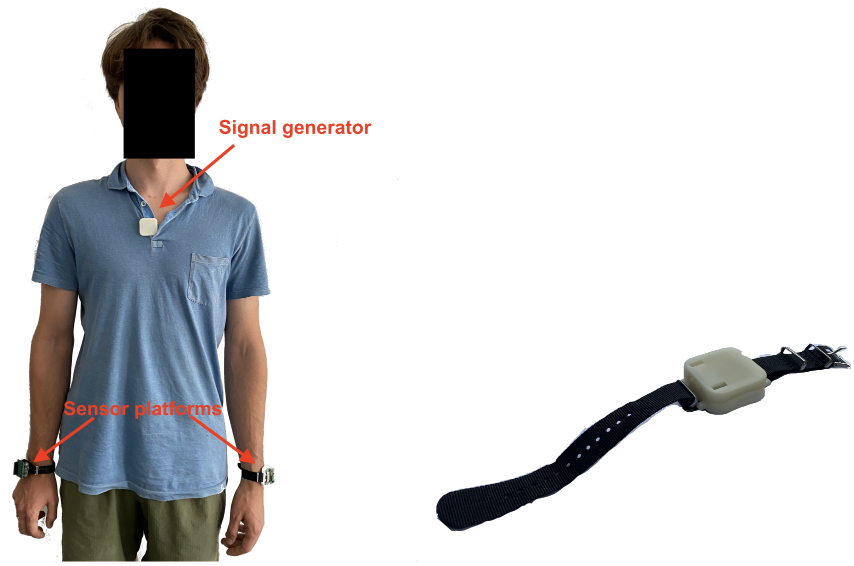
Future work

In this chapter, we propose next steps for next generation hardware, dataset and software. First of all, we propose that the future sensor should consist of a wrist-worn IMU,EPS and ambient microphone and a neck-worn signal generator. Second, we provide guidelines for a real life dataset collection, that could be a gold standard dataset to evaluate the efficacy of the scratch detection system. Finally, we propose the general direction for deep learning algorithms for more accurate human activity recognition.

5.1 Next Generation Hardware

5.1.1 Proposed set-up

The case is made from plastic and the wristband is made from nylon, to reduce the risk of an allergic reaction.



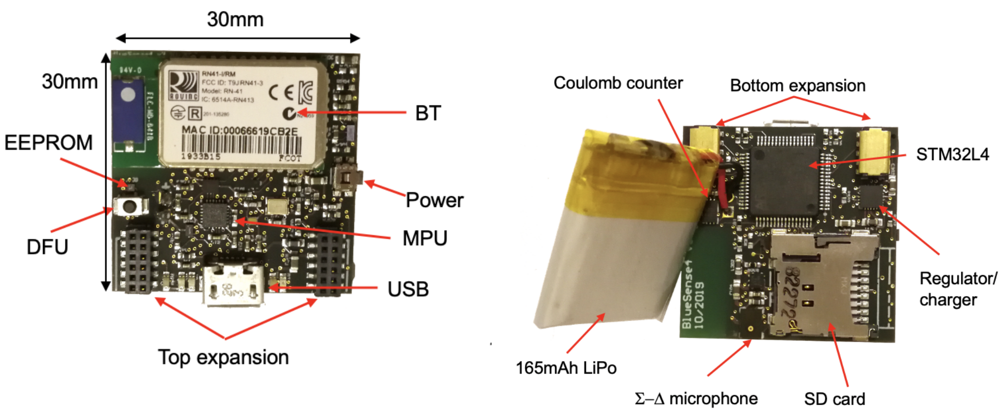
(a) Sensor set up on the body (b) Sensor platform in the case

Figure 5.1: In figure (a), the participant wears the sensors on the wrists. In figure (b), the case with the wrist band is shown.

5.1.2 Bluesense

The platform used for data collection is an in-house wearable sensing research platform. Its primary function is to be an IMU and a digital microphone, which can be expanded using expansion connectors for sensor research purposes. The device is 30x30 mm. It is based on an ARM Cortex M4 processor (STM32L496 from ST), which runs at 20 MHz with our default firmware. The platform comprises a 9-axis inertial measurement unit (TDK Invensense ICM-20948), a digital MEMS microphone (ST MP34DT05-A), a micro-SD card, Bluetooth 2 and USB interfaces, a fuel gauge for built-in power measurements (LTC2942), an EEPROM to store configuration (M24128). The processor built-in real-time clock (RTC) is operated from a dedicated 32KHz quartz (10ppm frequency tolerance). The platform operates at 3V from a lithium polymer battery (165mAh) with a LTC3553 voltage regulator. The expansion connectors provide I2C, SPI, UART, ADC inputs and DAC output for expansion boards.

The device firmware has been designed for ease of use. It allows without any programming to acquire the data from the built-in sensors or external ADC inputs through a command line interface. Data can be streamed over Bluetooth or USB, or stored in the SD card. The current firmware allows to acquire in isolation IMU data, sound data, or analog inputs data from the expansion connector. It can also acquire multimodal data: i.e. it can simultaneously acquire IMU, sound and analog inputs in a multiplexed streaming and storage format. This is particularly relevant to acquire data for activity recognition based on a combination of multimodal sensors. The sample rate of IMU, microphone and ADC is fully configurable. The IMU data is also processed by the firmware to obtain the device orientation in quaternions. All data is time-stamped using the internal RTC.



(a) IMU

(b) Electric Potential Sensor

Figure 5.2: The data collection platform is 30x30mm with connectors for additional expansions (a). The microphone is on the skin-facing side when worn on the body (b). The values of the resistors and capacitors will be determined by the wanted frequency of the signal.

IMU

The ICM-20948, also known as inertial measurment unit, is a multi-chip module (MCM) consisting of two dies integrated into a single QFN package. One die houses a 3-axis gyroscope, a 3-axis accelerometer, and a Digital Motion Processor^{TM} (DMP). The other die houses the AK09916 3-axis magnetometer from Asahi Kasei Microdevices Corporation. The ICM-20948 is a 9-axis MotionTracking device all in a small 3x3x1 mm QFN package. This chip is integrated to the Bluesense sensor platform and the data can be recorded to the SD card.

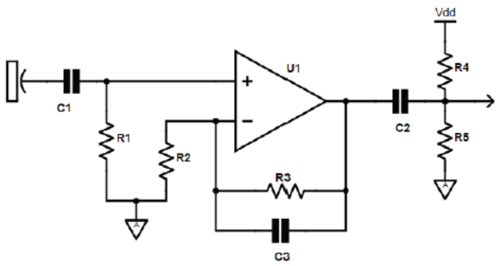
Microphone

The microphone is omnidirectional. It has a 64 dB signal-to-noise ratio and a -26 dbFS sensitivity. It is on the bottom-side of the device (i.e. facing towards the skin if worn on the wrist), but it is not in direct contact with the skin due to the case. The microphone is clocked at 2 MHz and provides a 1-bit digital pulse density modulation output, which comes straight from its sigma-delta analog to digital converter. The data is converted to an audio signal using a 3rd order Sinc filter, each with a decimation ratio of 82 which yields a 16-bit 8 KHz audio signal with a dynamic range has been experimentally tuned for typical ambient sounds and speech (¡90dB).

5.1.3 EPS

EPS technology

EPS is design to measure electrophysiological signals, such as ECG and EMG with a flat frequency response from 0.2 Hz to 20 kHz. The sensor has 20 Gohm of input resistance and input capacitance of 15 pF. This allows remote measurement of small variations in voltage signals and the ambient electric field non-invasively. Capacitive sensors have been used in the environment for human-computer interaction with limited work on body-worn sensor, except where the authors carried out a similar experiment using another type of body-worn capacitive sensor to detect walking and jogging. However, their design topology suffered from saturation due to the static charge built up while EPS was DC-stable, which means it recovers from saturation without the need to switch circuitry. It has previously been shown that EPS can also measure the amount of static charge build-up.



(a) EPS circuit

(b) EPS on Bluesense

Figure 5.3: EPS circuit (a) and the EPS on bluesense (b) are used.

EPS-based localisation

The same EPS will be used for localisation with a signal generator. A CMOS 555 timer IC produces a 50% duty cycle square wave. Its output is sent to a low-pass RC filter that filters out the harmonics, leaving only the fundamental sine wave. Some distortion is common as it's difficult to completely eliminate the harmonics. A more selective LC filter can be used to improve sine wave quality.

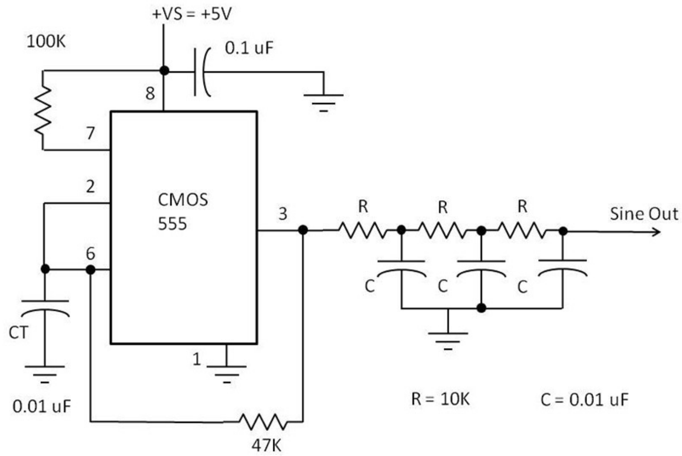
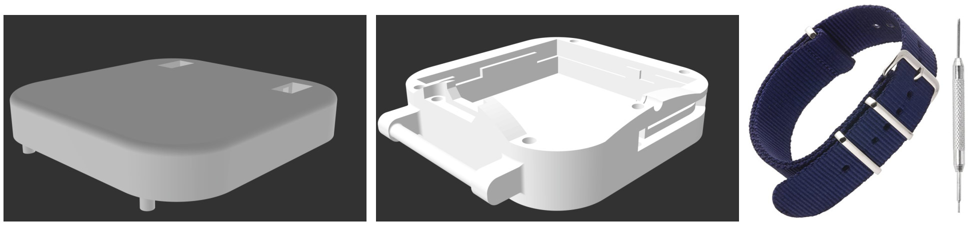


Figure 5.4: The CMOS 555 timer IC produces a 50% duty cycle square wave and the RC circuit is used as a filter.

5.1.4 Device case

The initial case design is shown in Figure 5.5. Case parts (a) and (b) will be adapted after finalising the EPS design and when the exact height of the Blusense is known, gaps in the case sides will be added to allow resistless air flow (needed for the microphone).



(a) Top case (b) Bottom case (c) Wrist band

Figure 5.5: The (a) and (b) parts of the case will be built using the 3D printer and the wristband (c) will be bought from Amazon [3].

5.1.5 Set-up instructions

1. Make sure that the battery is charged and there is a compatible SD card.
2. Turn on Bluesense.
3. Calibrate it.
4. Format the SD card.
5. Attach the EPS extension board.
6. Start logging.

7. Collect the dataset.

8. Quit logging.

9. Export the data.

5.2 The immediate tasks for finalising the sensor platform

Some of the tasks are not finalised and will need to be done in the future:

- EPS-based localisation and detailed characterisation of the needed frequencies and voltages.

- Exploration of making the sound anonymous by removing certain frequencies but keeping the ones needed for scratch detection.

- Testing the full device for the dataset collection.

5.2.1 Sensor production

Each part of the proposed device will need to be manufactured and assembled together. The Bluesense platform is already available with the integrated IMUs and microphone; a custom EPS will be built for the specified frequencies. Moreover, given that the required characterisations will be achieved for EPS-based localisation, the circuit can be built in the lab. The goal should be to build three pairs of devices for this stage.

5.2.2 EPS-based localisation

The EPS-based localisation technique has been shown to be feasible for this task in recent work [24]. In this work, we want to be able to use a signal generator to see, if it can be used as a proximity sensor around the neck and if it could work as a localisation technique in the scalp scratch recognition task. To evaluate this idea's suitability the following steps:

1. Characterise the voltages needed for proximity detection using EPS.

2. Characterise the frequencies needed so no interference occur with the noise signals from the main grid.

3. Evaluate the usability of this sensor modality for scratch detection by placing the signal generator on the neck and using the EPS on the wrist.

5.2.3 Testing

When the sensor platform is ready, the data collection should be tested and Bluesense configured for the different sample rates of the different sensor modalities. The testing procedure protocol will include 3x3 s scratches with a 5 s rest in between, 10 s of walking and 3x3 s scratches. Then, each individual modality should be checked if scratching and walking can be distinguished solely from the signal. Moreover, it should be tested how long the sensor platfrom can record at the chosen sampling rates and indentify the limitations of the battery and the SD card, as it will be important to be aware of these limitations during the data collections.

5.3 Comprehensive real world dataset collection

Currently, there are no gold standard datasets that would allow to evaluate scratch detection systems in a real world scenario. We need to have such a dataset to have unbiased performance evaluation. Such a system, would need to consist of a varied and long null class and imperfect and varied " natural scratches" done by different types of people with varied types of hair, skin, bodytypes. Moreover, we would need to record a high variaty of scratches, which can differ in frequency, speed and type of movement.

Thus, we propose to collect a scratch detection dataset at least from 20 people and of two parts: constrained dataset collection (30 min), real-life dataset collection (2h). The constrained dataset collection will be used to evaluate the properties of scratch and identify the limitations of each sensor for the application of scratch detection. The real world dataset collection would be collected in a real world scenario, where the participant will have a high level objectives, that the person would need to accomplish and during these two hours the person will be recorded and his scratch will be annotated afterwards.

5.3.1 P1: Constrained dataset collection

During the constrained dataset collection we want to collect scratches in environments, that would allow to assess the limitations of the sensor as well as identifying different scratching modes. We propose the following scenario for the constrained scratch detection to identify the limitations of the sensors:

Electric Potential Sensor

To identify the limitations of EPS for scratch detection, we need to see how well scratch can be recognised with different levels of electric field noises.

Thus the scenario will be the following to identify the limitations:

- 5 sets of scratch while walking on the carpet
- 5 sets of scratch while being next to the computer with 50 Hz main grid.
- 10 sets at different distance from a white noise generator.
- 10 sets at different distance from a pink noise generator.

This scenario will allow to characterise the limitations of the EPS sensor for scratch detection and to identify the applicability domain of EPS for localisation and scratch detection.

Microphone

To make a comprehensive review of sensor's ability to detect scratch, we need to identify all the limitations of an ambient wrist-worn microphone for scratch detection. We will use the following scenario to identify the limitations:

- 5 sets of scratch while walking on the carpet
- 5 sets of scratch while scratching with music in the background
- 10 sets at different distance from a white noise generator using a local sound source with decibels measured at each position.
- 10 sets at different distance from a pink noise generator using a local sound source with decibels measured at each position.

Inertial Measurement Unit

To identify the limitations of a wrist-worn microphone for scratch detection, we need to measure how well scratch can be recognised from other similar activities.

We will use the following to identify the limitations:

- 5 sets of scalp scratch
- 5 sets of hair comb.

- 5 sets of drinking tea.
- 5 sets putting earphones inside the ears.
- 5 sets of putting glasses on.

5.3.2 P2: Real-life dataset collection

The expected duration for this session is 2 hours. Participants would be asked to perform higher order tasks on the university campus which would also be video recorded. Participants would be followed by an annotator, who would use wearable glasses to film everything to later annotate the scratch and non-scratch activity ground truths. Examples of higher order tasks include: “Find a book on Deep Learning in the library” or “Go to the canteen and drink a cup of tea”.

During all these auxiliary activities we are going to monitor the amount of natural scratch that is occurring. The data collection will be filmed with smart glasses in order to annotate afterwards. The annotation happens from the recorded video, where the annotator specifies what activity happens on the video in that particular time. Even though no personal information would be collected during the data collection, several videos would be recorded. Given that the dataset (anonymised sensor data without the video) would be made public, the users would be asked if some screenshots and video extractions can be used for presentations during conferences. If they do not agree with the above the video would be deleted after annotation. From Inertial Measurement Unit Electric Potential data the user cannot be identified. We would remove certain frequencies so that from sound (collected by microphone data) people would not be identifiable.

5.4 Next generation machine learning models for multi modal fusion

In this thesis, we only explored classical machine learning methods and we did not have the opportunity to explore deep learning models for scratch detection. It has been shown, that deep learning is a powerful tool for human activity recognition, but it has not been adapted for multimodal fusion. We propose a conceptual framework for next generation deep learning models for human activity recognition that allows for easier transfer learning between modalities and would not require to retrain the model from scratch when a new modality is introduced.

A key problem with the current state-of-the-art Deep Learning models is that if you add an additional modality it requires to retrain the model completely, simply because your input shapes do not match. When in reality you need to be able to model the temporal relationships between different modalities. It means, that when you add a new modality you would need to fine-tune a part of it to let the model learn, how the signal from the new modality is related to the signal from the existing modalities.

This is a crucial feature needed for life-long learning as such a system needs to have the ability to effortlessly add and substract different sensors without needing to retrain the whole system. Thus, we propose to investigate the combination of graph neural networks and transformers.

5.4.1 Deep Convolutional Transformers

The goal of this work is to evaluate if the adaptation of the transformer architecture can allow achieve better accuracy for human activity recognition. Moreover, the transformer has been used only as a sequence-to-sequence model, thus as it is learning how to return the sequence and not a scalar classification is not ideal and means how to reduce the dimensionality are needed for an effective learning.

In this experiment we would experiment with different dimensionality reduction techniques to identify the best way of aggregating different signals.

For the moment state-of-the art human activity recognition. can perform as well as an LSTM for human activity recognition and to evaluate the trade-offs for this particular application. The current state-of-the-art model for human activity recognition uses DeepConvLstm from [19] and with slight variations better performances have been a achieved with similar architectures. As Transformers seem to be much more effective for sequence-to-sequence task we want to adapt the sequence-to-sequence model for human activity recognition in the Transformer paper [28], they presented that the transformer excelled current state-of-the art models by the following components:

1. Training times.
2. Memory requirements.
3. Performance trade-offs.

In the figure 5.6 we propose an By increasing the width of the window we will observe how the performance changes when we use the transformer encoder instead of the LSTM. We are going to test if the self-attention is able to remove the vanishing gradient problem

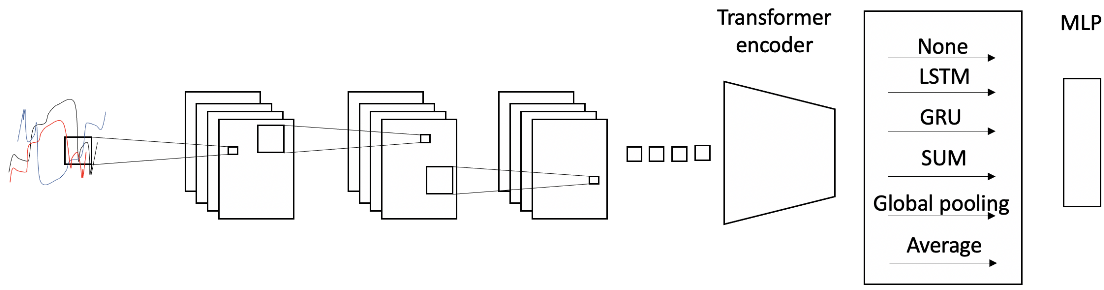


Figure 5.6: We will explore architecture based on DeepConvLSTM and we will use a transformer followed by an LSTM,GRU, Sum, Pooling or average functions to see how the performance differs.

for the LSTMs and if the LSTM can be changed to a simpler dimensionality reduction function.

The main evaluation will be done on the Opportunity dataset, but to collaborate the tests will also be performed on Skoda. The Novelty of this work will be the the evaluation of the novel architecture (with Transformers) for human activity recognition.

- **The goal** is to evaluate the dimension reduction techniques, as shown in the figure 5.6, from the computed new embeddings for time-series classification. We will evaluate how each different dimensionality reduction technique affects the performance of the classifier.

- **The datasets** that are going to be used are: PAMAP, OPPORTUNITY and Scratch in office environment.

- **The benchmark** It will be compared against different dimensionality reduction techniques and known benchmark results such as DeepConvLSTM.

- **The problem** is that transformers are sequence-to-sequence models and it has not been explored how well different techniques can reduce the dimentionality.

- **Our approach** is to use the convolutional layers from the DeepConvLSTM and add a transformer encoder, which will compute the representation of the change in with respect to other changes. Different dimensionality reduction techniques will be explored on the output from the transformer encoder to understand the trade-offs between performance of the model and computational costs.

- **Contributions:**

  – Evaluate dimensionality reduction algorithms for transformer decoder.

– See if vanishing gradients can be avoided with self-attention as the input would be a sum of all the timepoints. So during training and forward pass during timestep t the model will have information about timestep t-N.

– Laying the foundations for 4.2.3 section, where we will explore how we can have multimodal fusion and effective dimensionality reduction techniques will need to be used.

5.4.2 Investigation of Transformers

The goal of this experiment is to improve and adapt the transformers architecture for human activity recognition. The main limitations are that the universal transformer [28] is not extremely effective: does not learn the local dependencies and the memory requirements expands quadraticaly with the length of the windows, we need to explore different architectures to adapt the functionality for ies, as in the paper [15]. We will develop a custom querying algorithm with the local CNNs that will allow use less training time and learn short term dependencies. The goal is to optimise the Transformers performance for Human activity recognition from wearable sensor data. This will be evaluated on Opportunity, Skoda and the Scratch dataset . The novelty of this work, will be a novel variant of transformer adapted for human activity recognition.

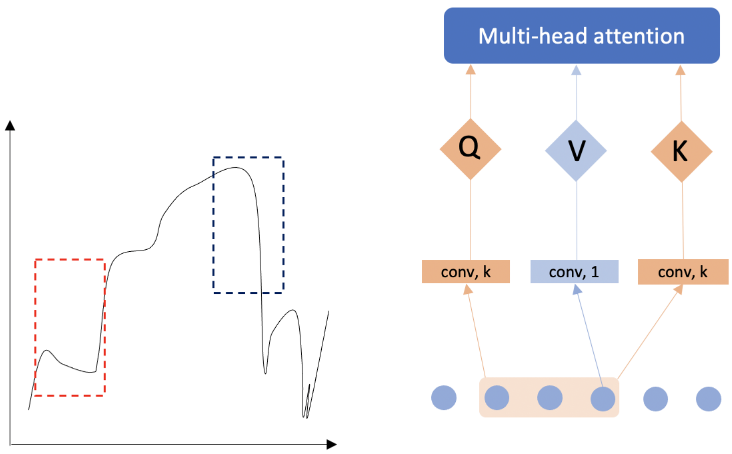


Figure 5.7: We will explore different ways how to effectively compute the keys and queries for optimal human activity recognition.

Causal convolutions [15], shown in Figure 5.7, in the multi-head attention are used to learn the local context of the signal where the attention was applied on using causal convolutions. Rather than using convolution of kernel size 1 with stride 1 (matrix multiplication), we employ causal convolution of kernel size k with stride 1 to transform inputs (with proper paddings) into queries and keys. Note that causal convolutions ensure that the current position never has access to future information. By employing causal convolution, generated queries and keys can be more aware of local context and hence, compute their similarities by their local context information, e.g. local shapes, instead of point-wise values, which can be helpful for learning accurate relations.

In this, part we would like to explore if we can split long series into non-overlapping windows and apply causal convolutional neural networks with a layer of perceptron to learn the deep the representations of past series, these representations would to fit longer time-series and incorporate more information without consuming unreasonable amounts of memory.

For this part we will compare to the results achieved with a transformer without local convolutions and to the state-of-the-art models.

- **The goal** is to evaluate different ways how to compute the keys, queries and values for 1D signal and to evaluate how it affects the performance.

- **The datasets** that are going to be used are: PAMAP, OPPORTUNITY and Scratch in office environment.

- **The benchmark** Different ways how to compute keys, querries and values will be compared.

- **The problem** is that the transformer model is not able to learn local relations, thus we will investigate different approaches of how we can investigate different metrics.

- **Our approach** is to explore the following ways how to compute Q, K, V for better classification accuracy and computational performance:

  - Traditional way - the self-attention is computed on raw points.
  - Convolutions- self-attention is computed the 1D CNN output.
  - Causal convolutions - self-attention is computed on the point and convolutions related to the previuos points.
  - Attentive convolutions- the self-attention is computed on points that are computed from the product of the 1D CNN output and the points.

  Moreover, we will explore how it affects the prediction, by using it as a sequence-to-sequence model and predicting the future series.

- **Contributions:**

– Evaluate how different ways of compute keys affect the performance and the computational requirements.

– Evaluate how it different Q, V, K affects the prediction of time-series.

5.4.3 Transformers for Fusion

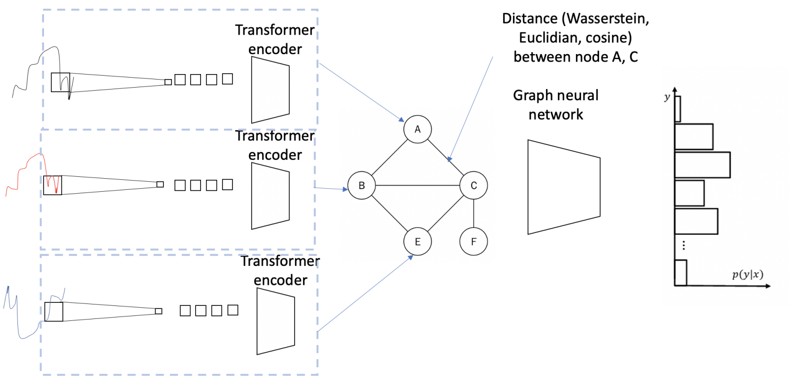


Figure 5.8: We show a new multi-tasl architecture that leverages a contrastive loss will allow to effectively retrain models as the same encoder will be used. Moreover, we propose to use a GNN where the edge will be distance between representations. It will allow for the model to effectively learn the needed features in ensemble.

The goal will be to investigate different fusion techniques for data which is sampled at different rates from different sensor modalities. More over, as the information lies in different domains it is important that each model learns the needed features in each domain. Fusion of different sensors that has different sample rates has not been explored widely. We want to investigate different techniques for effective and transparent fusion of different sensor modalities. For this part we will investigate fusion techniques on *the Scratch dataset*. The following steps will be explored:

1. **Decision fusion**. Majority vote of models from different sensor modalities.

2. **Before final layer concatenation.** We will concatenate the representations of different sensor modalities and pass them together through the final MLP layer.

3. **Before final layer sum.** We will sum the representations of different sensor modalities and pass them together through the final MLP layer.

4. **Dilated and strided convolutions to match the dimensions in the transformer.** As the sampling rates are different dilated and strided convolutions should enable to match the dimensions and feed together to the transformer.

5. **Graph neural network**. Once we are able to compute the representations we can use a graph neural neural network to learn what makes the activity an activity. The graph consists of nodes and edges, where the node is the representation that is being learned for the signal and the and edge will be a distance between the representations. We will explore different distance to see which one allows to learn the relations the best. We will test Euclidian, Wasserstein, Cosine distances.

The steps 1, 2, 3 will be used as a benchmark to see if the model is able to recognise the activities when we are learning the modalities together, these are the most common approach in feeding data from different sensor modalities. We will do a more in depth analysis in how 2 and 3 are different as concatenations allows for a better recall for the model, while summing allows to learn more complex relations between the sensor modalities.

While the sampling rates are different of different sensor modalities. For example motion is captured at 250Hz, while Audio is captured at 16 kHz and EPS at 3 kHz. It means that different size windows should be chosen and different features are learned. Thus, with step 2 we will manually adjust the strided and dilated convolutions for an optimal performance to compute the queries and keys before feeding to the transformer.

However, as HAR is a complicated problem for transfer learning, because the information lies in different domains for each sensor modality. We will investigate if we can learn a more general approach using contrastive loss.

- **The goal** is to evaluate the new architecture presented in figure 5.8.
- **The datasets** that are going to be used are: PAMAP, OPPORTUNITY and Scratch in office environment.
- **The benchmark** will be used DeepConvLSTM and the work done in the previuos section. Moreover, we will compare the proposed fusion technique with the different fusion techniques mentioned earlier.
- **The problem** is that fusion techniques use nowadays do not enforce learning the relations between different modalities.
- **Our approach** will comprise of multiple steps and the model will be developed and tested incrementally.
    1. The first step is too evaluate what performance can be achieved once we use this network with similar sensor modalities or same modalities at different locations. For this task we will only use the acceleration and gyroscope data.

2. The second step is too evaluate if the same transformer and encoder can benefit from additional sensor modalities, where the information is in the different domains. For this task we will use Scratch in office environment dataset and we will add sound and EPS data to see if the same encoder can learn how to relate those.

3. If the same encoder does not work, we will use a subset of the data to evaluate on which set of features the information lies and we will use different encoders to learn in the frequency domain and in the "normal" domain.

Chapter 6

Conclusions

In this chapter, we discuss the contributions made and their impact to human activity recognition, scratch detection, multi-modal fusion. Finally, we discuss the limitations of the work.

6.1 Summary of achievements

During the past decade we had many technological advances. Yet, previously scratch detection relied on human annotators, then they relied on cameras until later on wearable sensors were introduced. Wearable sensors allow to monitor the amount of scratch that occur in daily uninterrupted human life during a prolonged period of time.

Moreover, each sensor modality has it's own applicability domain and naturally there will be false predictions. Using multimodal fusion we increase the accuracy of scratch detection. During this thesis we focused on sound, electric field, motion and hands position relative to the torso, which are the modalities that can be used individually to detect scratching.

To this day, in the field, data is a key limitation to create more sophisticated human activity recognition pipelines. Data recording is expensive. Data annotation requires to recruit participants, video record their activities and later on annotated.

In this thesis we made the following contributions:

- In the 2nd chapter we did a comprehensive review of methods for scratch detection and of sensor modalities that can be used for human activity recognition and in particular scratch detection. Moreover, we systematically reviewed the pros and cons of each sensor modality and it's potential to be used on a wrist-worn device. Finally, we summarised the main limitations of the previuos work, which we are adressing in this thesis.

- In the 3rd chapter, we investigated sensors and algorithms to realise a wearable scratch detection device. The work consisted of two parts: a dataset collection and analysis of different algorithms for the multimodal fusion. We collected a dataset from nine users containing 813 scratching instances and 5 h 15 min of recorded data. We concluded that a scratch could be detected with 80.7% accuracy by using the random forest algorithm on hand coordinates, which required four devices. However, an f1 score of 70% could be achieved with k-NN with IMU and EPS data, which only required one device. Moreover, the fusion of IMU data with EPS data improved the accuracy and reduced the deviation between the folds.

- In the 4th chapter, we evaluated how sound, collected by a wrist-worn microphone, can be used to recognise and distinguish scratch occurrences. We collected a dataset from 4 users, where each user performed different scratches. The classification was done by training a random forest model with engineered features and was tested using leave-one-user-out cross-validation. In a silent environment we achieved sensitivity of 83.75% $\pm$ 8.8% and specificity 78.5% $\pm$ 4% and in an environment with TV noise present sensitivity decreased to 61%$\pm$20.45% and specificity 78%$\pm$9.6%. These results indicate that sound may be one modality to include for a future multimodal scratch detection device.

- In the 5th chapter, we proposed that the future sensor should consists of a wrist-worn IMU,EPS and ambient microphone and a neck-worn signal generator. Second, we provided guidelines for a real life dataset collection, that could be a gold standard dataset to evaluate the efficacy of the scratch detection system. Finally, we proposed the general direction for deep learning algorithms for more accurate and flexible human activity recognition.

6.2 Limitations

Each of our contributions presents limitations, either in terms of assumptions made, or in terms of results obtained. We discuss these limitations in the following subsections and present related perspectives to address these issues in chapter 5 where we present the future direction.

6.2.1 Limitations of the collected datasets

We have collected two datasets and we are highlighting the limitations of each dataset.

IMU and EPS

This dataset was collected in a constrained environment from a limited amount of people. The participant was instructed to perform scratches, which did not allow to collect natural scratches. Moreover, it was collected in one isolated environment and the sensor modalities were not pushed to the limits. Thus, in Chapter 5 we wrote the guidelines for gold standard dataset, which would be collected from a large number of people in a natural setting.

Microphone

This dataset was collected in a constrained environment from only four people with very few scratching occurrences. It was done to evaluate the feasibility of the sensor modality, however we would need more experiments to evaluate the limitations of scratch detection.

6.2.2 Limitations of the analysis

For both methods we only used traditional machine learning methods to assess the feasibility of using and combining sensor modalities. In Chapter 5, we propose next generation deep learning methods for more accurate human activity recognition using state-of-the-art techniques. These new methods should allow more easily retrain models when new sensor modalities are introduced.

Bibliography

[1] Seconds interval timer, 2010-2019.

[2] Pozyx, 2015.

[3] Amazon nylon nato watch straps by sniper bay™: Military style divers bands: 18mm 20mm 22mm 24mm (18mm, navy blue): Amazon.co.uk: Watches, 2020.

[4] Oresti Banos, Miguel Damas, Hector Pomares, Fernando Rojas, Blanca Delgado-Marquez, and Olga Valenzuela. Human activity recognition based on a sensor weighting hierarchical classifier. *Soft Comput.*, 17(2):333–343, February 2013.

[5] Diana Bautista, Sarah Wilson, and Mark Hoon. Why we scratch an itch: The molecules, cells and circuits of itch. *Nature neuroscience*, 17:175–82, 02 2014.

[6] Akram Bayat, Marc Pomplun, and Duc A. Tran. A study on human activity recognition using accelerometer data from smartphones. *Procedia Computer Science*, 34:450 – 457, 2014. The 9th International Conference on Future Networks and Communications (FNC'14)/The 11th International Conference on Mobile Systems and Pervasive Computing (MobiSPC'14)/Affiliated Workshops.

[7] F. Bonnin-Pascual and A. Ortiz. An uwb-based system for localization inside merchant vessels. In *2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, pages 1559–1562, Sep. 2019.

[8] Ke-Yu Chen, Shwetak N. Patel, and Sean Keller. Finexus: Tracking precise motions of multiple fingertips using magnetic sensing. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, page 1504–1514, New York, NY, USA, 2016. Association for Computing Machinery.

[9] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing).* Wiley-Interscience, USA, 2006.

[10] Toshiya Ebata, Hiroshi Aizawa, and Ryoichi Kamide. An infrared video camera system to observe nocturnal scratching in atopic dermatitis patients. *The Journal of Dermatology*, 23(3):153–155, 1996.

[11] Kaoru Endo, Hozumi Sano, Takayuki Fukuzumi, Jun Adachi, and Toshiyuki Aoki. Objective scratch monitor evaluation of the effect of an antihistamine on nocturnal scratching in atopic dermatitis. *Journal of Dermatological Science*, 22(1):54 – 61, 1999.

[12] Shu Gong, Willem Schwalb, Yongwei Wang, Yi Chen, Yue Tang, Jye Si, Bijan Shirinzadeh, and Wenlong Cheng. A wearable and highly sensitive pressure sensor with ultrathin gold nanowires. *Nature Communications*, 5(1):3132, 2014.

[13] Pavel Kriz, Filip Maly, and Tomas Kozel. Improving Indoor Localization Using Bluetooth Low Energy Beacons. *Mobile Information Systems*, 2016:2083094, 2016.

[14] Jongin Lee, Dae-ki Cho, Seokwoo Song, SeungHo Kim, Eunji Im, and John Kim. Mobile system design for scratch recognition. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, page 1567–1572, New York, NY, USA, 2015. Association for Computing Machinery.

[15] Shiyang Li, Xiaoyong Jin, Yao Xuan, Xiyou Zhou, Wenhu Chen, Yu-Xiang Wang, and Xifeng Yan. Enhancing the locality and breaking the memory bottleneck of transformer on time series forecasting. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems 32*, pages 5243–5253. Curran Associates, Inc., 2019.

[16] A. Moreau, P. Anderer, M. Ross, A. Cerny, T. H. Almazan, and B. Peterson. Detection of nocturnal scratching movements in patients with atopic dermatitis using accelerometers and recurrent neural networks. *IEEE Journal of Biomedical and Health Informatics*, 22(4):1011–1018, July 2018.

[17] Yuichi Noro, Youichi Omoto, Koji Umeda, Futa Tanaka, Yousuke Shiratsuka, Tomomi Yamada, Kenichi Isoda, Kimiko Matsubara, Keiichi Yamanaka, Esteban C. Gabazza, Masakatsu Nishikawa, and Hitoshi Mizutani. Novel acoustic evaluation system for scratching behavior in itching dermatitis: Rapid and accurate analysis for nocturnal scratching of atopic dermatitis patients. *The Journal of Dermatology*, 41(3):233–238, 2014.

[18] Takeshi Okuyama, Kazuki Hatakeyama, and Mami Tanaka. Measurement of human scratch behavior using compact microphone. volume 45, pages 731–737, 01 2014.

[19] Francisco Javier Ordóñez and Daniel Roggen. Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition. *Sensors*, 16(1), 2016.

[20] J. Petersen, D. Austin, R. Sack, and T. L. Hayes. Actigraphy-based scratch detection using logistic regression. *IEEE Journal of Biomedical and Health Informatics*, 17(2):277–283, March 2013.

[21] A. Pouryazdan, R. J. Prance, H. Prance, and D. Roggen. Wearable electric potential sensing: A new modality sensing hair touch and restless leg movement. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, UbiComp '16, pages 846–850, New York, NY, USA, 2016. ACM.

[22] Daniel Roggen. Arm cortex m4-based extensible multimodal wearable platform for sensor research and context sensing from motion sound. *Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers*, 45:284–289, 2020.

[23] Daniel Roggen, Arash Pouryazdan, and Mathias Ciliberto. Poster: Bluesense - designing an extensible platform for wearable motion sensing, sensor research and iot applications. In *Proceedings of the 2018 International Conference on Embedded Wireless Systems and Networks*, EWSN ’18, pages 177–178, USA, 2018. Junction Publishing.

[24] Daniel Roggen, Arash Pour Yazdan, Francisco Javier Ordóñez Morales, Robert J. Prance, and Helen Prance. Electric field phase sensing for wearable orientation and localisation applications. In *Proceedings of the 2016 ACM International Symposium on Wearable Computers*, ISWC '16, page 52–53, New York, NY, USA, 2016. Association for Computing Machinery.

[25] P. M. Scholl and K. van Laerhoven. A feasibility study of wrist-worn accelerometer based detection of smoking habits. In *2012 Sixth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*, pages 886–891, July 2012.

[26] A. Torralba and A. Oliva. Depth estimation from image structure. *IEEE Transactions on Pattern Analysis Machine Intelligence*, 27(09):1226–1238, sep 2002.

[27] R. Twycross, M.W. Greaves, H. Handwerker, E.A. Jones, S.E. Libretto, J.C. Szepietowski, and Z. Zylicz. Itch: scratching more than the surface. *QJM: An International Journal of Medicine*, 96(1):7–26, 01 2003.

[28] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Ł ukasz Kaiser, and Illia Polosukhin. Attention is all you need. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems 30*, pages 5998–6008. Curran Associates, Inc., 2017.

[29] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu. Device-free human activity recognition using commercial wifi devices. *IEEE Journal on Selected Areas in Communications*, 35(5):1118–1131, 2017.

Appendix A

Multimodal fusion of IMUs and EPS body-worn sensors for scratch recognition

Multimodal fusion of IMUs and EPS body-worn sensors for scratch recognition

Zygimantas Jocys
University of Sussex
Falmer, UK
z.jocys@sussex.ac.uk

Arash Pouryazdan
University of Sussex
Falmer, UK
a.pouryazdan@sussex.ac.uk

Daniel Roggen
University of Sussex
Falmer, UK
d.roggen@sussex.ac.uk

ABSTRACT

In order to develop and evaluate the extent to which itching affects a person's daily life, it is useful to develop automated means to recognise the action of scratching. We present an investigation of sensors and algorithms to realise a wearable scratch detection device. We collected a dataset, where each user wore 4 inertial measurement unit (IMU) sensors and one electric potential sensor (EPS). Data were collected from nine users, where each user followed a 40-min protocol, which involved scratching different parts of head, shoulder, and leg, as well as other activities such as walking, drinking water, brushing teeth, and typing to a computer. The dataset contained 813 scratching instances and 5 h 15 min of recorded data. We investigated the trade-offs between the number of devices worn (comfort) and accuracy. We trained the k-NN and random forest algorithms by using between 1 and 5 features per channel. We concluded that a scratch could be detected with 80.7% accuracy by using the random forest algorithm on hand coordinates, which required four devices. However, an f1 score of 70% could be achieved with k-NN with IMU and EPS data, which only required one device. Moreover, the fusion of IMU data with EPS data improved the accuracy and reduced the deviation between the folds. This expanded the state-of-the-art method by opening up new trade-offs between accuracy and comfort for future research.

KEYWORDS

Scratch detection, feature selection, machine learning, activity recognition, sensor technologies.

ACM Reference Format:

Zygimantas Jocys, Arash Pouryazdan, and Daniel Roggen. 2020. Multimodal fusion of IMUs and EPS body-worn sensors for scratch recognition. In *14th EAI International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth '20), May 18–20, 2020, Atlanta, GA, USA.* , 11 pages. https://doi.org/10.1145/3421937.3421987

1 INTRODUCTION

Itching is a condition which affects a substantial group of people. This condition may be caused by scabies, atopic dermatitis, or kidney failure; it can also be a symptom of a malignant condition, such as lymphoma [19]. Scratching occurs as a result of itching, but the itching is increased while scratching. Scratching also produces wounds on the skin, which create even more discomfort. Thus, itching and scratching are correlated [8].

Furthermore, atopic dermatitis causes a high intensity of itching followed by a period of respite. The amount of scratching which occurs over night and the impact that it has on the skin cannot be evaluated by the user himself as he is not consciously aware of it. We need a system which can, during such flares, warn the user to apply the treatment on the affected area. Such a system will improve the user's sleep quality by reducing the intensity of the cycle: itching, scratching, more intense itching.

Moreover, automatic monitoring for atopic dermatitis clinical trials will enable a more accurate measurement of how the frequency and the intensity of scratching change before and after the user starts treatment.

In previous research, acceleration sensors with simple models such as logistic regression [14] and more complex deep learning models [11] showed good success in detecting a scratch. In this paper, we present a new systematic approach toward the evaluation of the required sensor modalities and the complexity of machine learning models to successfully detect scratches. The novelty of this work is as follows:

- Survey of previous scratch detection work to evaluate the dataset required to represent the daily lives of humans.
- New dataset using four IMUs, 1 EPS, and hand coordinates with ten activities and nine users with 813 scratches and a total of 5 h 15 min of recorded data. The novelty of this dataset lies in the large number of sensor modalities, to push the machine learning models to their limits and evaluate how well slightly different scratches can be differentiated from other activities. The new dataset includes more scratch locations as well as a more realistic NULL class.
- Observation that EPS data combined with IMU data reduce the error variance between the folds and increase the accuracy of human activity recognition.
- Comparison of the k-nearest neighbour and random forest algorithms by using different numbers of features scored by mutual information. Previous work has not explored feature selection for scratch detection from different modalities.
- Exploration of trade-offs between accuracy, comfort, and the number of devices.

2 RELATED WORK

In previous works, an itch [3] was described as an irritating feeling that causes the desire to scratch. A chronic itch can be a reaction

PervasiveHealth '20, May 18–20, 2020, Atlanta, GA, USA

ACM ISBN 978-1-4503-7532-0/20/05…$15.00
https://doi.org/10.1145/3421937.3421987

of skin diseases, kidney failure, cancers, and neurological disorders. Given that the disease might need urgent medical attention, scratches should be monitored as they can lead to a faster disease identification.

In the early days, scratching was observed manually, without any technological help. In the research [9], scratching was monitored on 40 patients in the hospital by the staff. During this project, the amount of scratching occurring before and after the treatment was noted. Observing the change in the scratch quantity before and after the treatment allows one to evaluate the extent to which the treatment reduces the occurrences of scratches. This system is not suitable for deployment on a large group (hundreds) of people.

Another approach is to video record nocturnal scratches by using an infrared camera. In one experiment, which was conducted with seven users [7], the infrared camera was used, so that the user would not be interrupted during the night and the scratches were annotated later.

Currently, the most common approach to classify activities is to use accelerometers [4]. It allows for the monitoring of activities without invading user privacy. Not only does it allow for systems to be deployed on many users simultaneously, but accelerometers also enable real-time and real-world monitoring without the need of an observer to invade the subjects' privacy. Triaxial accelerometers [14] on the wrist for scratch deployment were noted to be successful and achieved high sensitivity (0.96) and specificity (0.92). In this previous research, 12 people participated; they were monitored in a controlled environment. The duration of the dataset collection was 140 s per user. Each person scratched the back of his/her head, leg, and/or elbow and later walked and rolled around. A similar work was done by detecting scratching by using mobile systems [10]; this method achieved 90% accuracy.

However, even though a single modality has shown good results, there are many new sensors, which can potentially be incorporated in such a system and increase the accuracy of scratch detection.

For instance, acoustic sensors [12] have been used for scratch detection. The sound sensor uses body conduction to record the sound from the wrist. The dataset was collected on four volunteers over 6 h during the night. As the room was silent, it was easier to distinguish between a scratch and no scratch, and a 'determination ratio' of 0.98 was achieved.

Electric potential sensors (EPSs) were explored for scratch sensing [15]. This work only reported a visual observation that EPS signals are correlated to a scratch, but the authors did not demonstrate any automated activity recognition based on this modality.

Machine learning models vary in complexity, computational time, and their ability to generalise. Logistic regressions [14] and recurrent neural networks [11] have been used thus far. In the research [11], 24 atopic dermatitis patients spent 2–5 nights in the hospital and wore two accelerometer devices (GeneActiv, Activinsights Ltd.) on each wrist. Only two classes were labelled: scratch or null class, and as the data needed to be split for each class equally, a considerable amount of recording data was lost for the training of the model. There is still room for new machine learning models and considerable progress needs to be made with the current techniques.

In the previous works, all the currently collected datasets for scratch detection were either very short (150 s) or very simplistic (two classes: scratch or null class), which does not allow one to evaluate how the system would perform in a more realistic scenario. Moreover, given that in each research, there was one sensor modality per dataset, there is a need to explore different fusion options between different sensors.

We incorporated four IMUs, extracted the hand coordinates, and used EPS for scratch detection during the dataset collection. Moreover, there is a need for a challenging dataset, which has multiple classes and represents a more realistic world. A more complex dataset will allow to evaluate how different sensor modalities contribute toward an efficient and effective human activity recognition.

3 MULTIMODAL SCRATCH DATASET

A new dataset was built using data collected from nine users. It contained two classes and 10 subclasses, which are shown in Table 1. These activities were chosen because of their similarity to scratching in the frequency domain and because they occur in everyday life. Each user wore four IMU sensors and one EPS.

Class	Subclass
Scratch	Top of the head
	Back of the head
	Side of the head
	Shoulder
	Leg
Null	Null
	Drink water
	Brush teeth
	Wash hands
	Walk

Table 1: Ten classes of the dataset

As can be seen from Table 2, in all, we recorded 40 min of scratching and 5 h 15 min of data. The dataset had ten classes, and some activities were performed simultaneously in order to find the limits of the machine learning models and their ability to learn the subtle differences between very similar activities. In our case, the subtle differences were scratching the top of the head, back of the head, and side of the head.

	Dataset parameters
Age	22–40 years
Males	6
Females	3
Scratching instances	813
Total scratching time	40 min
Collection time per user	35 min
Total dataset time	5 h15 min

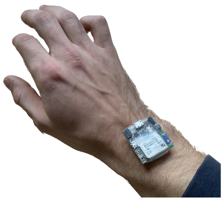
Table 2: Parameters of the dataset

The participants had different hair lengths, and there were six males and three females. The diversity of the users' biological features was needed to evaluate whether the model could generalise enough to detect the same activities on unseen users. There were no overweight or underweight participants; all of the participants had average body shapes. It was also important to have a long duration and a diverse null class, so that the model could learn the difference required to distinguish between scratching and other activities.

3.1 Sensor setup

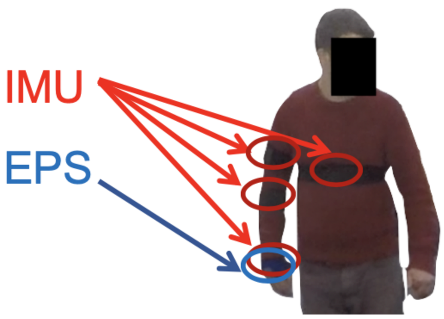
The IMU is based on the BlueSense technology [16]. BlueSense gives the raw acceleration, rate of turn, and magnetic field data and can provide the quaternion data, which encode the orientation of the device. Moreover, the quaternion data from the four IMUs could be used to obtain the approximate hand coordinates (explained in Section 3.1.2). Furthermore, BlueSense can be extended with expansion boards containing additional sensors. In this work, we used an expansion board with an EPS sensor [15].

3.1.1 EPS and IMUs. The sensor configuration for each user is shown in Figure 1 (c). This configuration was chosen in order to obtain the hand coordinates in the Cartesian system, the charge of the electric field, and the IMU data.



(a) IMU

(b) Electric potential sensor



(c) Sensor configuration during data collection. Each person wore four IMUs (on torso, upper arm, lower arm, and hand) and 1 EPS device extended on the IMU on the hand.

Figure 1: Imu (a) and EPS (b) were used for the data collection. They were attached to the body by using four straps.

In this work, a device was a single object, IMU with EPS was one device, a sensor was a sensor modality (IMU and EPS were two sensors). We referred to a *channel* as an output of a sensor. Sensors can have one or more outputs. For example, in this work, an IMU had 13 channels. They were as follows: three channels of acceleration along the x, y, and z axes; three channels of the rate of turn along the x, y, and z axes; three channels of the magnetic field along the x, y, and z axes; and four channels indicating the device orientation in quaternions. EPS had only one output channel, which represented the electric potential.

The EPS, shown in Figure 1 (b), streamed the voltage data using the ADC channel to Bluesense [16]. Once the electric field was disrupted [15], it could be observed in the signal. Moreover, EPS had the ability to detect 50-Hz grid voltage when the computer was connected to the power source, which allowed us to accurately detect typing to a computer.

3.1.2 Hand coordinates. Hand coordinates were computed after the data collection. The quaternion data captured the orientation of each sensor. The sensor positioning is shown in Figure 1 (c). By getting the orientation of each sensor and by using a vector for each limb, we could sum up all of the Cartesian coordinates of each joint to obtain the hand coordinates.

3.2 Data collection

This work was approved by University of Sussex Ethical Committee, application n. ER/ZJ70/1. The participants were recruited in the Engineering and Informatics building for the participation in the data collection. During the data collection, the participants needed to follow the defined protocol.

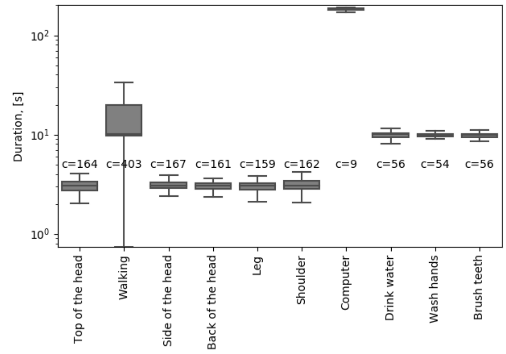


Figure 2: Duration of each activity in the collected dataset. The value of c denotes the number of occurrences of each activity across the entire dataset.

We defined a protocol, displayed in Figure 3, which allowed us to collect a wide variety of scratches. The data collection took place in the lab, and we attempted to obtain the most naturalistic dataset with a varied null class. In particular, we collected a dataset consisting of three different scratch locations on the head (top, side, and back) and scratch locations on other parts of the body (leg and shoulder). The scratches were performed with two different intensities (intense and moderate). In order to include a realistic null class in the dataset, we asked the participants to walk for 20 s in the office in between groups of scratches. In addition, we asked the participants to perform a number of other hand gestures, including simulating washing hands and brushing teeth, drinking water, and typing on a computer, as a way of including a more realistic set of activities to evaluate how well scratch could be distinguished from the other activities of daily living.

The data collection was a tedious process and included many activities. An application was used to show the current and the next activity on an IPad, using the application 'Seconds' [1]. Undesired

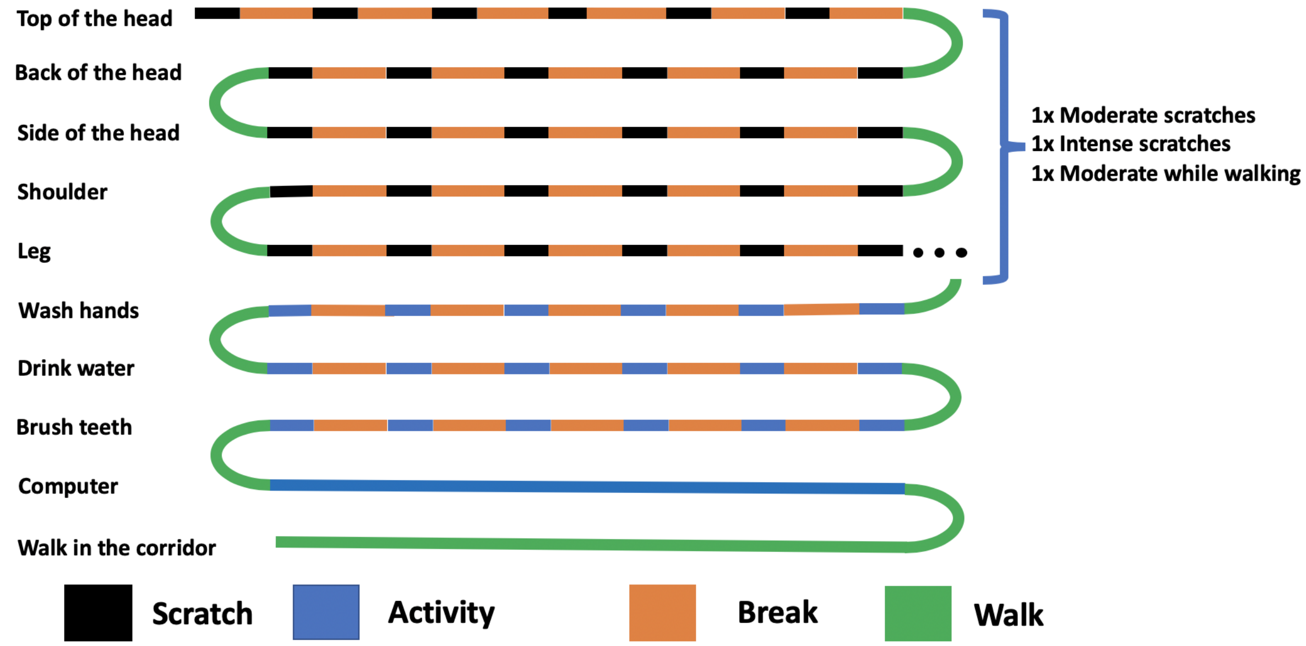


Figure 3: Protocol of data collection. First, the scratches were performed at three different intensities and all of the other activities for the null class were performed thereafter. A scratch occurred for 3 s, and then, there was a 10-s break. It was repeated six times, and before the next activity, a 20-s walk was performed.

deviation from the given protocol occurred during the collection. However, the labels of activities were adjusted after the data collection to adjust to the deviations.

4 HUMAN ACTIVITY RECOGNITION PIPELINE

The human activity recognition process is a process that requires a specialised pipeline for each case. The pipeline that we used for the human activity recognition in this study is shown in Figure 4.

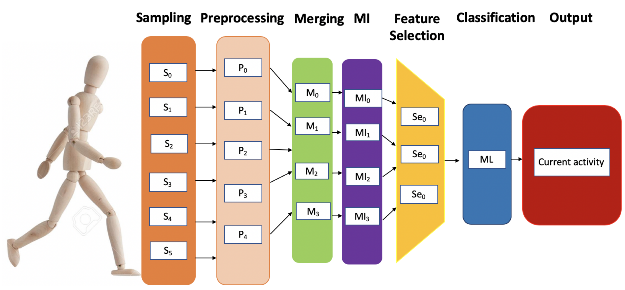


Figure 4: Human activity recognition pipeline for scratch detection. The data were sampled from the sensors and then preprocessed. The features with the highest MI score were used for the classification task.

4.1 Data Cleaning and Pre-processing

The data were collected using five devices (four IMUs and one EPS). However, with the additional development of the firmware, the IMUs and the EPS could be deployed using only one device. Thus, in the later stages, the number of devices that required EPS and IMUs was 1.

The sampling rates of the IMUs and the EPS were 100 Hz and 1 kHz, respectively. After sampling, the data were stored on the local SD cards. All the devices were synchronised; however, they did not log the data at exactly the same time. Because of the different time stamps, an interpolation technique was considered. As the quaternion data had a value of an angle, linear interpolation was not possible. Thus, an ASOF function was used. ASOF merged the data according to the nearest timestamps rather than the equal timestamps. A time delta equal to 10 ms was chosen: if the nearest timestamp was further than 10 ms, then the function did not choose the nearest value and assigned NaN.

The EPS was very sensitive, and hence, the collected data had a considerable amount of fluctuation. Thus, a low-pass Butterworth filter was applied to smooth the signal, and then the signal was resampled to match the 100-Hz IMUs' frequency.

4.2 Channels

Additional data, apart from the sampled data, were computed, in order to obtain more information:

- These additional data included the hand coordinates, which are described in Section 3.1.2.

- In the formula, m is any modality with the x, y, z projections and m_{xyz} is the magnitude, which was calculated to determine whether the magnitude could enable the device to achieve a relatively high performance.

$$m_{xyz} = \sqrt{m_x^2 + m_y^2 + m_z^2} \quad (1)$$

The magnitude was computed using formula 1 for the acceleration, rotation, and hand coordinates.

Acc_x	$Quat_0$	Gyr_x	$Hand_x$	EPS
Acc_y	$Quat_1$	Gyr_y	$Hand_y$	-
Acc_z	$Quat_2$	Gyr_z	$Hand_z$	-
Acc_{xyz}	$Quat_3$	Gyr_{xyz}	$Hand_{xyz}$	-

Table 3: Channels. All the channels used in activity recognition are displayed including acceleration, rotation, orientation, hand coordinates, and the EPS.

Thereafter, sliding windows of the time series were generated with a window length of 0.4 s. This time was chosen on the basis of the fact that a scratch is an activity which occurs for a short duration of time.

4.3 Feature Extraction

From the features shown in Table 3, we had to extract features from the sliding windows. The features chosen for this case are shown in Table 4. The mean and the variance enabled us to defines the distribution. The percentiles allowed us to detect the key points in the distribution and avoid the outliers, contradictory to the minimum and the maximum functions. The mean crossing rate and the zero crossing rate were used to evaluate how periodic the signal was.

Domain	Features
Statistical	Mean
	Variance
	Percentile 25%
	Percentile 50%
	Percentile 75%
	Percentile 90%
	Mean crossing rate
	Zero crossing rate
Frequency	Energy

Table 4: List of features

4.4 Feature Selection

Two of the most common methods used to select the most important features are the filter and wrapper methods. As we had 17 channels, there were 170 unique features. The wrapper method would take an unreasonable amount of time to find the best combination of features. Therefore, the filter method that used the mutual information [6] algorithm was used to select the features carrying the highest amount of information.

$$\mathrm{I}(X;Y) = \sum_{y\in\mathcal{Y}} \sum_{x\in\mathcal{X}} p_{(X,Y)}(x,y) \log\left(\frac{p_{(X,Y)}(x,y)}{p_X(x)\,p_Y(y)}\right) \quad (2)$$

where p (X,Y) is the joint probability mass function of X, and $p_{(X)}$ and $p_{(Y)}$ are the marginal probability mass functions of X and Y, respectively, where X is the bin of the feature and Y is the class. For each feature, we created 100 equal-size bins and performed a small correction by adding 10^{-10} to the division so that when there were empty bins, division by zero was possible.

4.5 Machine Learning

As the range of the data varied considerably, all of the extracted features were normalised as shown in equation 3.

$$n_{normalized} = \frac{n_{feature} - \mu}{\sigma} \quad (3)$$

In the equation, μ is the mean of the feature in the training set, n_{raw} is the feature, and σ is the standard deviation of the feature in the training set.

4.5.1 *k-Nearest Neighbour.*

Next, k-nearest neighbour is a model which is fast to train and has a proven record of successful applications in many areas. However, it is slow to compute predictions and is very susceptible to outliers. Therefore, choosing too many correlated features or features that did not bring valuable information toward the decision making, would only diminish the model's ability to recognise activities. Therefore, the value of k was set as 100, and the sklearn implementation was used.

4.5.2 *Random Forest.*

Random forest is ensemble-based learning method. This model has seen considerable success in regression and classification tasks. Given that the decision trees can distinguish important features, a larger number of features will lead to more accurate predictions. For the random forest model, we chose to use 100 trees. In this case, we used the random forest algorithm from the sklearn Python library.

4.6 Performance measurement

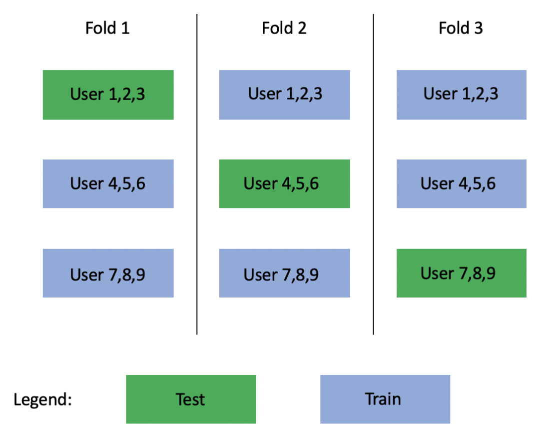


Figure 5: K-fold cross-validation where the data from three unseen users were left out for testing.

The goal of this project was to have a universal system for all the users. To achieve this goal, we used a three-fold cross-validation.

ML model	Modalities	Number of devices	Number of features				
			1	2	3	4	5
k-NN	xyz	4	38.01±0.7%	59.3±2.5%	61.2±1.9%	**62.7±1.9%**	62.1±2%
	IMU	1	41.4±6.7%	45.4±8.3%	46.3±7.8%	46.4±8%	**47±7.3%**
	IMU+EPS	1	43.2±2.1%	50.3±2.4%	52±2.1%	52.7±2.2%	**53.4±2.4%**
	IMU+xyz	4	43.8±6.2%	52.5±0.9%	55.9±3.2%	55.9±2.8%	**56.3±28%**
	IMU+xyz+EPS	4	47.9±1%	52.2±1.7%	54±2%	54.7±0.2%	**55.6±2%**
Random forest	xyz	1	55.1±2.5%	66.6±3.5%	66.7±3.2%	66.8±3.2%	**66.9±3.2%**
	IMU	1	45.6±8.2%	47.7±8.4%	48.1±8.5%	48.4±8.6%	**49±8.5%**
	IMU+EPS	1	45.2±4.2%	49.6±5.4%	50.6±5.7%	51.8±6%	**52.1±5.8%**
	IMU+xyz	4	50.6±3.3%	52.6±4%	51.96±4.5%	52±4.6%	**53.8±4.6%**
	IMU+xyz+EPS	4	56.3±2.2%	57.4±3%	57.8±3.1%	57.8±3.1%	**57.8±3%**

Table 5: Results table of the ten-class classification with the displayed Macro f1 score. Note that overall, a better performance was achieved with the random forest model. Further, xyz is a simplified notation for the hand coordinates data.

During the cross-validation, the users were grouped into three groups of three users, and during each validation, each test was performed on the groups of three.

In this research project, a confusion matrix and the macro f1 score were chosen to evaluate the performance of the model and its ability to generalise. We chose to use the macro score in order to see how well the model recognised each class.

$$F_{1c} = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \tag{4}$$

$$F_{1Macro} = \sum_{c=1}^{n} F_{1c} \tag{5}$$

The macro f1 score was calculated using equation 5. The score for each class affected the overall results. Because of the class imbalance in the dataset (40 min from 5 h 15 min was scratches, and the rest was null), using the macro f1 score, we achieved a high performance when a majority of the data belonged to the null class.

4.6.1 Fine activity recognition. First, the models were tested to determine how well they could classify the ten classes. As there were very similar classes and multiple activities took place at the same time, this dataset was created to push the limits of the machine learning models.

4.6.2 Coarse scratch/non-scratch activity recognition. The 10-class classification task was meant to push the machine learning model to distinguish the subtle difference between the activities. By evaluating how well a scratch was distinguishable from a complex null class, we could determine how well a scratch could be detected in realistic scenarios. For this part of the experiment, the model was not retrained, but the labels were changed to either scratch or null. This implied that if 'scratch top of the head' was confused with 'scratch back of the head', the f1 score would not be reduced. The classes of the scratch and the null class are shown in Table 1.

5 RESULTS

5.1 Most important features

With the use of mutual information, a heatmap of the most important features was produced; it is shown in Figure 6. For each channel, five features (which carried the highest amount of information) were selected and are displayed in Table 6.

Channel	f1	f2	f3	f4	f5
Acc_x	var	90 %	75%	mean	median
Acc_y	mean	75 %	90%	median	25%
Acc_z	25%	mean	median	var	25%
Gyr_x	var	25 %	90%	75%	median
Gyr_y	var	90 %	25%	75%	median
Gyr_z	var	90 %	75%	25%	median
$Quat_0$	var	90 %	75%	25%	median
$Quat_1$	75%	90 %	median	mean	25%
$Quat_2$	90	75%	median	25%	mean
$Quat_3$	var	90 %	75%	median	mean
Acc_{xyz}	mean	median	energy	25%	75 % %
Gyr_{xyz}	var	90 %	75%	mean	median
$Hand_x$	var	90 %	25%	75%	median
$Hand_y$	25%	median	75%	90 %	mean
$Hand_z$	25%	mean	median	75%	90 %
$Hand_{xyz}$	75 %	90 %	median	25 %	mean
Eps	mx0	var	90 %	mean	75 %

Table 6: Selected features for each channel

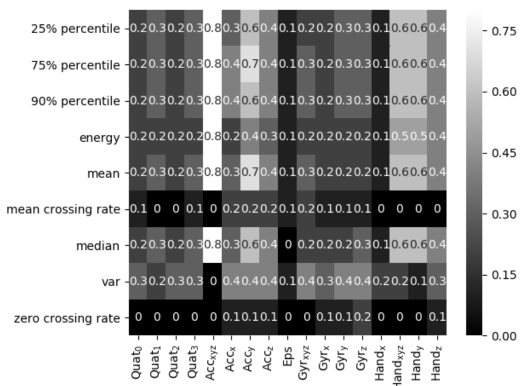


Figure 6: Feature heat map based on mutual information. The lighter shade indicates that the feature of the channel carried more information than the darker shade.

As can be seen, the variance carried the largest amount of information for most of the channels, and the percentiles showed good performance as well. The results displayed in Figure 6 revealed that the EPS carried the smallest amount of information as a channel. The mean crossing rate and the zero crossing rate were not informative features in most of the cases.

5.2 Fine activity recognition results

In Table 5, the results are presented for the ten-class classification task using the k-NN and the random forest models trained on a different number of features from 1 to 5 per channel. The accuracy is shown as the mean and the standard deviation between the folds.

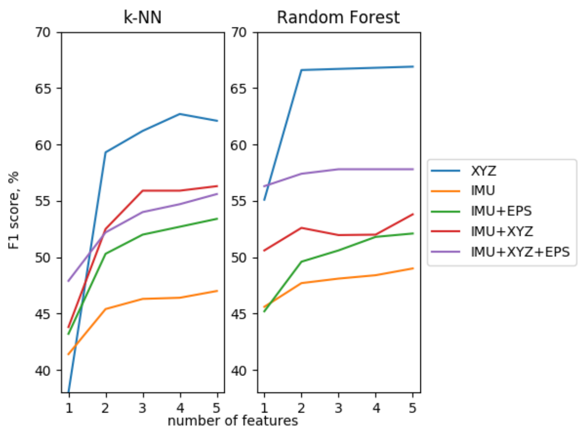


Figure 7: F1 score vs. number of features for ten-class classification

5.2.1 k-NN. The best results were observed when only the hand coordinates with four features per channel were used. The f1 score reached 62.7%. To achieve this result, four devices were required. In contrast, with only the IMUs and the EPS, an f1 score of 52.3% was achieved and only one device was required.

Combining the hand coordinates with the IMU or IMU+EPS data only worsened the results.

An f1 score of 53.4% was achieved with one device (IMUs and EPS), while with only IMUs, the best achieved result was 47%.

5.2.2 Random forest. The random forest model showed an overall better performance. The best performance was observed when using hand coordinates with four features per channel, where an f1 score of 66.9% was achieved for the ten-class classification. However, this task required the attachment of four devices on the body. We observed that the performance improved with an increase in the number of features introduced.

An f1 score of 52.9% was achieved with one device (IMUs and EPS). Nonetheless, with IMUs only, the best achieved result was 49%.

In Figure 9, the confusion matrix shows which classes were mostly confused when tested on the RF model using two features per channel. In all the confusion matrices, note that 'Scratch the top of the head', 'Scratch the side of the head' and 'Scratch the back of the head' were considerably confused. In Figure (b), it can be seen that 'Brushing teeth' was confused with 'Scratch the side of the head' and vice versa, and in Figure (a), 'Washing hands' was confused with 'Typing on a computer'. When both the modalities were combined, as seen in Figure (c), the performance on these two classes drastically improved. When EPS was introduced, in Figure (d), it can be seen that the number of errors decreased in the scratch classes, as compared to when only IMUs were used.

5.3 Coarse scratch/non-scratch activity recognition results

It was very challenging to classify ten classes. However, to evaluate how well a scratch was recognisable, it was sufficient to distinguish a scratch from any other activity. In Table 7, the results are presented for the binary classification task.

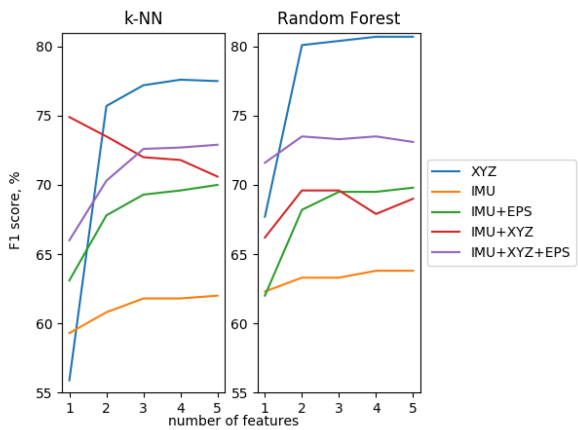


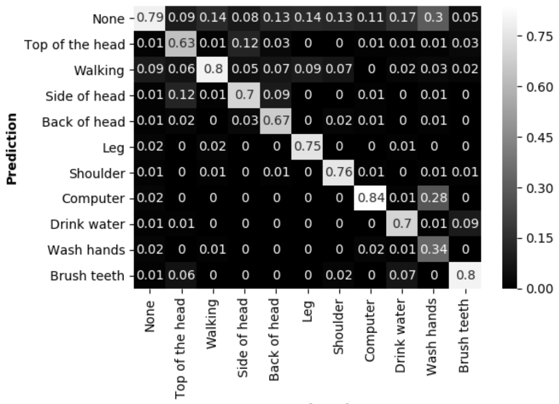
Figure 8: F1 score vs. number of features for binary classification

5.3.1 k-NN. The best result was achieved again by using the hand coordinates data. We achieved an f1 score of 77.6%. With a single device (IMUs and EPS), we achieved an f1 score of 70% with a standard deviation of 2.9%. By using only the IMUs, we achieved an f1 score of 62% with a standard deviation of 8.6%.

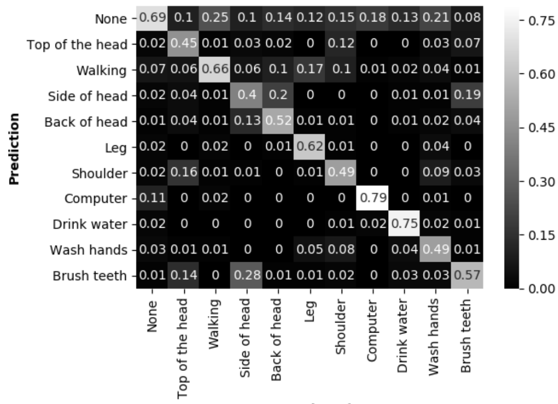
The EPS allowed us to reduce the standard deviation between the folds as compared to the IMU results. However, four devices were required with the extracted hand coordinates to achieve the best results.

5.3.2 Random forest. The best score was achieved by using the random forest model with five features per channel. It reached an f1 score of 80.7% for scratch detection. With only the IMUs and the EPS, an accuracy of 69.8% was achieved, and the standard deviation between the folds was only 7.2%. The use of only IMUs resulted in an f1 score of 63.8% with a standard deviation of 10%.

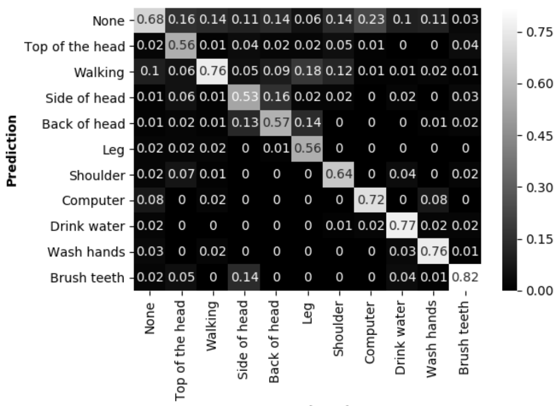
The same behaviour as observed as that with the k-NN model. The EPS allowed us to reduce the standard deviation between the folds as compared to the IMU results. However, four devices were



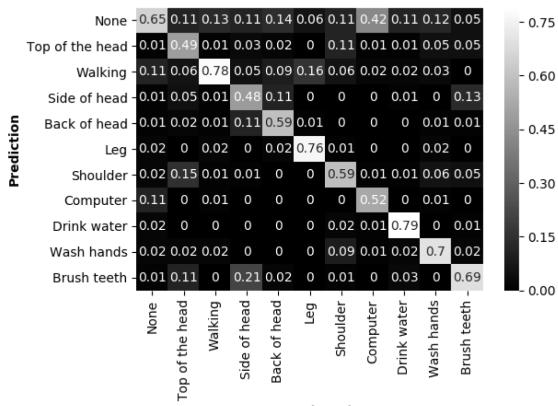
(a) Hand coordinates



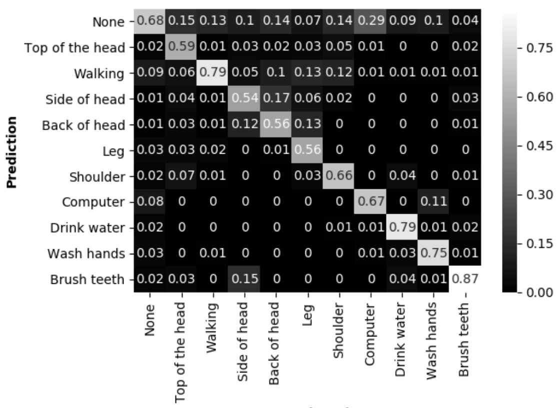
(b) IMU channel



(c) IMU and hand coordinates



(d) IMU and EPS channel



(e) Hand coordinates, IMU, and EPS

Figure 9: Confusion matrices when the inputs were two features per channel and the model was random forest.

ML model	Modalities	Number of devices	Number of features				
			1	2	3	4	5
k-NN	xyz	4	55.89±2.7%	75.7±3%	77.2±2.5%	**77.6±2.7%**	77.5±2.4%
	IMU	1	59.3±5.5%	60.8±8.5%	61.8±8.4%	61.8±8.4%	**62±8.6%**
	IMU+EPS	1	63.1±2.7%	67.8±2.9%	69.3±2.9%	69.6±2.8%	**70±2.9%**
	IMU+xyz	4	**74.9±1.4%**	73.5±78.3%	72±3.6%	71.8±3.8%	70.6±5.11%
	IMU+xyz+EPS	4	66±4.4%	70.3±4.5%	72.6±4.5%	72.7±4.5%	**72.9±4.7%**
Random forest	xyz	4	67.7±2%	80.1±2.7%	80.4±2.3%	80.7±2.7%	**80.7±2.6%**
	IMU	1	62.3±6.3%	63.3±8.7%	63.3±9.2%	063.8±10%	**63.8 ±10%**
	IMU+EPS	1	62.09±1.7%	68.2±6.4%	69.5±6.8%	69.5±7.6%	**69.8±7.2%**
	IMU+xyz	4	66.2±3.9%	69.6±2.7%	**69.6±2.5%**	67.9±3.3%	69±3.3%
	IMU+xyz+EPS	4	71.6±3.3%	73.5±3.4%	**73.7±3.3%**	73.5±3.1%	73.1±3.6%

Table 7: Results table of binary classification with the displayed macro f1 score. Note that overall, a better performance was achieved with the random forest model. Further, xyz is a simplified notation for the hand coordinates data.

needed with the extracted hand coordinates to achieve the best results.

6 DISCUSSION

6.1 Baseline Results

Currently, the standard approach toward human activity recognition is predicting an activity on the basis of the IMU data. In this work, the baseline results were achieved by using the extracted features from the IMU data. Moreover, these results only required the deployment of one device.

The baseline results of the new dataset were as follows:

- For the ten-class classification with k-NN, the best result was 47% ± 7.3%.
- For the ten-class classification with RF, the best result was 49% ± 8.5%.
- For the binary classification with k-NN, the best result was 62% ± 8.7%.
- For the binary classification with RF, the best result was 63.8% ± 10%.

6.2 Multimodal Fusion

The fusion of data between IMUs, EPS, and the extracted hand coordinates did not always result in a better performance.

6.2.1 IMUs and EPS. IMUs and EPS required one device on the wrist. Fusing the data and classification from the IMU and EPS data provided a slightly better performance than only using IMUs with the random forest and the k-NN models. Using the binary classification task with IMUs and EPS resulted in an accuracy of 70% when using k-NN and 69.8% when using RF. Moreover, with k-NN, the standard deviation decreased between the folds from 8.6% to 2.9%; it decreased from 10% to 7.2% in the case of the random forest model. Compared with the baseline results, there was a small improvement in the f1 score. In both the cases, fusing the IMU and EPS data decreased the standard deviation and improved the results obtained using both the models, but not significantly. The reduction in the standard deviation of the f1 score between the folds showed that the model's ability to recognise the activities was good for a diverse group of people. Moreover, as the EPS could be deployed on the IMU located on the hand, it did not create additional discomfort.

6.2.2 IMU and hand coordinates. For the ten-class and binary classifications, combining these modalities yielded better results than using only IMUs, but worse than when using only the hand coordinates. Using RF and binary classification, IMU achieved an accuracy of 63.8% ± 10%; with only the hand coordinates, the result was 80.7% ± 2.6%, and when combined, it achieved 69.6% ± 2.5%. The same behaviour was observed with k-NN and the ten-class classification task.

Each activity was associated more with a certain location relative to the torso than with specific movements. Thus, additional information could be redundant and decrease the accuracy. In this case, fusing the hand coordinates data with the IMU data did not provide with any gains, as compared to the results achieved using only the hand coordinates data.

6.2.3 IMU, hand coordinates, and EPS. Combining all the modalities outperformed the use of only the hand coordinates data. For the binary classification task using RF and binary classification, IMU achieved an accuracy of 63.8% ± 10%; with the use of only the hand coordinates data, the result was 80.7% ± 2.6%, and with the use of IMU with the hand coordinates data, the model achieved an accuracy of 69.6% ± 2.5%. Moreover, with the use of IMU, hand coordinates, and EPS data, the model was able to achieve an f1 score of 73% ± 2.6%.

The dataset was biased to the location, implying that each activity in the dataset was associated with a certain location relative to the torso and this particular feature worked exceptionally well with this dataset. The fusion of the dataset did not bring a drastic improvement in the accuracy but significantly decreased the standard deviation between the folds, as compared to the baseline results. However, it still did not outperform the achieved results when only the hand coordinates were used.

6.3 Number of Features

As can be seen from Figures 7 and 8, there was an increase in the accuracy, but later, the results plateaued. This might be caused by the mutual information algorithm, which did not take into account the fact that certain features were correlated and that the redundant information did not improve the performance. This was observed for both the classifiers, with the exception of k-NN with IMU and hand coordinates.

6.4 Comfort and Accuracy Trade-off

To deploy a scratch detection system for clinical trials, the smallest number of devices and the least invasive device must be chosen. Sensors set up with four devices (shown in Figure 1) cannot be used in any medical or clinical study, as this set up will make daily activities uncomfortable and there will be a higher risk of failure because of the large number of devices in use.

For the data collection in the present study, four devices were used. In general, more information gives better results. In Figures 5 and 7, it is shown that increasing the number of devices produced better results with the exception of k-NN and hand coordinates. However, discomfort is a major drawback for the deployment of a human activity recognition system. If comfort is the priority: one device with IMUs and EPS on the wrist might be sufficient with a 70% accuracy, and if accuracy is the priority, then with four devices (as shown in Figure 1) should be used. Note that an f1 score of 80.7% could be achieved by using only the hand coordinates.

6.5 Future Work

In Section 5, note that hand coordinates relative to the torso were needed to achieve the best performance for this dataset. To build a comfortable system, new localisation techniques should be explored, so that the sensors can fit on one wrist. We suggest to explore localisation techniques, such as ultra wide band. For example, PosXYZ [5] needs only two devices (thw slave and the master) to be deployed in such a system. The slave device will be attached to the wrist, and the master device will be the reference point. With the use of the ultra-wide-band technology, the location of the wrist as compared to the torso can be computed and can be used for human activity recognition. Thus, the hand coordinates will enable one to achieve higher accuracy, without needing four straps on the torso, upper arm, lower arm, and wrist, respectively.

Exploring different feature selection techniques, such as MRMR [18], can lead to higher performance. MRMR is a minimum redundancy feature selection algorithm, which also takes into consideration how redundant a feature is as compared to the other selected features. Moreover, to determine what accuracy can be achieved for this dataset, an investigation of deep learning models, such as DeepConvLSTM [13], can produce substantially higher accuracy results.

Finally, new hardware can be developed for head scratch detection, such as EPS-based glasses, which can work as a proximity sensor to detect when the hand is nearby, thus increasing the recognition of head scratches.

Even though in this study, we explored feature-based fusion, decision-based fusion has shown good performance as well [2]. Thus, exploring hierarchical classifiers with decision-based fusion should be able to yield even better scratch detection results.

The four participants in the study [17] were smokers, and they needed to tap the sensor to flag when the smoking happened. This gives the insight that in a similar study for scratch detection, active learning should be incorporated so that a user will be queried if he scratches and the scratch will be flagged.

7 CONCLUSIONS

During this experiment, we explored how the fusion of different sensor modalities contributed toward accurate scratch detection by using different numbers of features per channel and common machine learning models, such as k-NN and random forest. For this task, a dataset was built with the data collected for ten different activities to investigate the limitations of each model and explore the trade-off between the number of sensor modalities, number of features, and machine learning models.

The key results were as follows:

- The best baseline result for detecting a scratch with a simple IMU was an f1 score of 63.8% obtained using RF, which required only one device.
- The best result obtained using one device was an f1 score of 70% for scratch detection. It was achieved by using k-NN with IMU and EPS data.
- The best overall result was an f1 score of 80.7%. It was achieved for the binary scratch detection using the hand coordinates data and the RF model, which required four devices.
- Fusing EPS data with IMU data consistently increased the accuracy and reduced the deviation between the folds, as compared to using only the IMU data.

We found that hand coordinates alone enabled us to achieve the highest accuracy to detect all the activities. However, this dataset was biased to perform well on these data as each position was associated with a certain activity. However, with the current technology, such a model would require four IMUs on the torso, upper and lower arms, and the hand. Therefore, it is not convenient to use this setup on a large number of people.

For the best performance on this dataset, the hand coordinates data with five features should be used to achieve the highest accuracy of 80.7% for detecting scratches. However, if a comfortable system is a priority and accuracy can be sacrificed, then 70% accuracy can be achieved with a single device using IMUs and EPS.

8 ACKNOWLEDGEMENTS

This work was partly funded by EPSRC grant EP/S513921/1.

REFERENCES

[1] [n.d.]. Seconds Interval Timer. https://www.runloop.com/seconds-interval-timer-for-iphone
[2] Oresti Banos, Miguel Damas, Hector Pomares, Fernando Rojas, Blanca Delgado-Marquez, and Olga Valenzuela. 2013. Human Activity Recognition Based on a Sensor Weighting Hierarchical Classifier. *Soft Comput.* 17, 2 (Feb. 2013), 333–343. https://doi.org/10.1007/s00500-012-0896-3
[3] Diana Bautista, Sarah Wilson, and Mark Hoon. 2014. Why we scratch an itch: The molecules, cells and circuits of itch. *Nature neuroscience* 17 (02 2014), 175–82.
[4] Akram Bayat, Marc Pomplun, and Duc A. Tran. 2014. A Study on Human Activity Recognition Using Accelerometer Data from Smartphones. *Procedia Computer Science* 34 (2014), 450 – 457. http://www.sciencedirect.com/science/article/pii/S1877050914008643 The 9th International Conference on Future Networks and Communications (FNC'14)/The 11th International Conference on Mobile Systems and Pervasive Computing (MobiSPC'14)/Affiliated Workshops.
[5] F. Bonnin-Pascual and A. Ortiz. 2019. An UWB-based System for Localization inside Merchant Vessels. In *2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*. 1559–1562.
[6] Thomas M. Cover and Joy A. Thomas. 2006. *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, USA.
[7] Toshiya Ebata, Hiroshi Aizawa, and Ryoichi Kamide. 1996. An Infrared Video Camera System to Observe Nocturnal Scratching in Atopic Dermatitis Patients.

The Journal of Dermatology 23, 3 (1996), 153–155. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1346-8138.1996.tb03990.x

[8] T Ebata, S Iwasaki, R Kamide, and M Niimura. 2001. Use of a wrist activity monitor for the measurement of nocturnal scratching in patients with atopic dermatitis. *The British journal of dermatology* 144 (02 2001), 305–9.

[9] Kaoru Endo, Hozumi Sano, Takayuki Fukuzumi, Jun Adachi, and Toshiyuki Aoki. 1999. Objective scratch monitor evaluation of the effect of an antihistamine on nocturnal scratching in atopic dermatitis. *Journal of Dermatological Science* 22, 1 (1999), 54 – 61. http://www.sciencedirect.com/science/article/pii/S0923181199000481

[10] Jongin Lee, Dae-ki Cho, Seokwoo Song, SeungHo Kim, Eunji Im, and John Kim. 2015. Mobile System Design for Scratch Recognition. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1567–1572. https://doi.org/10.1145/2702613.2732820

[11] A. Moreau, P. Anderer, M. Ross, A. Cerny, T. H. Almazan, and B. Peterson. 2018. Detection of Nocturnal Scratching Movements in Patients with Atopic Dermatitis Using Accelerometers and Recurrent Neural Networks. *IEEE Journal of Biomedical and Health Informatics* 22, 4 (July 2018), 1011–1018.

[12] Yuichi Noro, Youichi Omoto, Koji Umeda, Futa Tanaka, Yousuke Shiratsuka, Tomomi Yamada, Kenichi Isoda, Kimiko Matsubara, Keiichi Yamanaka, Esteban C. Gabazza, Masakatsu Nishikawa, and Hitoshi Mizutani. 2014. Novel acoustic evaluation system for scratching behavior in itching dermatitis: Rapid and accurate analysis for nocturnal scratching of atopic dermatitis patients. *The Journal of Dermatology* 41, 3 (2014), 233–238. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/1346-8138.12405 https://onlinelibrary.wiley.com/doi/abs/10.1111/1346-8138.12405

[13] Francisco Javier Ordóñez and Daniel Roggen. 2016. Deep Convolutional and LSTM Recurrent Neural Networks for Multimodal Wearable Activity Recognition. *Sensors* 16, 1 (2016). https://www.mdpi.com/1424-8220/16/1/115

[14] J. Petersen, D. Austin, R. Sack, and T. L. Hayes. 2013. Actigraphy-Based Scratch Detection Using Logistic Regression. *IEEE Journal of Biomedical and Health Informatics* 17, 2 (March 2013), 277–283.

[15] A. Pouryazdan, R. J. Prance, H. Prance, and D. Roggen. 2016. Wearable Electric Potential Sensing: A New Modality Sensing Hair Touch and Restless Leg Movement. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct (UbiComp '16)*. ACM, New York, NY, USA, 846–850. http://doi.acm.org/10.1145/2968219.2968286

[16] Daniel Roggen, Arash Pouryazdan, and Mathias Ciliberto. 2018. Poster: BlueSense - Designing an Extensible Platform for Wearable Motion Sensing, Sensor Research and IoT Applications. In *Proceedings of the 2018 International Conference on Embedded Wireless Systems and Networks (EWSN ’18)*. Junction Publishing, USA, 177–178. http://dl.acm.org/citation.cfm?id=3234847.3234874

[17] P. M. Scholl and K. van Laerhoven. 2012. A Feasibility Study of Wrist-Worn Accelerometer Based Detection of Smoking Habits. In *2012 Sixth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*. 886–891. https://doi.org/10.1109/IMIS.2012.96

[18] A. Torralba and A. Oliva. 2002. Depth Estimation from Image Structure. *IEEE Transactions on Pattern Analysis Machine Intelligence* 27, 09 (sep 2002), 1226–1238.

[19] R. Twycross, M.W. Greaves, H. Handwerker, E.A. Jones, S.E. Libretto, J.C. Szepietowski, and Z. Zylicz. 2003. Itch: scratching more than the surface. *QJM: An International Journal of Medicine* 96, 1 (01 2003), 7–26. arXiv:http://oup.prod.sis.lan/qjmed/article-pdf/96/1/7/4375153/960007.pdf https://doi.org/10.1093/qjmed/hcg002

Appendix B

Scalp scratch detection using a wrist-worn microphone.

Scalp scratch detection using a wrist-worn microphone

Zygimantas Jocys
Sensor Technology Research Centre
University of Sussex
Falmer, UK
z.jocys@sussex.ac.uk

Daniel Roggen
Sensor Technology Research Centre
University of Sussex
Falmer, UK
daniel.roggen@ieee.org

***Abstract*—Itching is a condition caused by different skin diseases and automatically detecting it is useful to assess severity or treatment efficacy. In this work, we evaluate how sound, collected by a wrist-worn microphone, can be used to recognise and distinguish scratch occurrences. We collected a dataset from 4 users, where each user conducted 20 scalp scratches, 10 shoulder scratches and 18 scalp scratches with and without TV noise being played in the background. The classification was done by training a random forest model with engineered features and was tested using leave-one-user-out cross-validation. In a silent environment we achieved sensitivity of 83.75% ± 8.8% and specificity 78.5% ± 4% and in an environment with TV noise present sensitivity decreased to 61%±20.45% and specificity 78%±9.6%. These results indicate that sound may be one modality to include for a future multimodal scratch detection device.**

***Index Terms*—Scratch detection, sound-based, microphone, activity recognition, wearable sensing.**

I. Introduction

In previous work, itch [1] was described as an irritating feeling that causes the desire to scratch. Chronic itch can be a reaction of skin diseases, kidney failure, cancers and neurological disorders. Given that the disease might need urgent medical attention, scratches should be monitored.

Scratches can be observed manually. In one study [2], scratching was monitored on 40 patients in the hospital by the staff. Observing the change of scratch quantity before and after the treatment allows to evaluate how much the treatment reduces the occurrences of scratches. That system is not suitable to be deployed on a large group of people (hundreds). Another approach is to video record nocturnal scratches using an infrared camera. In one experiment, which was done with 7 users [3], the infrared camera was used, so that the user would not be interrupted during the night and the scratches were annotated later. Both of these approaches are invasive and require significant human resources to be scaled. However, wearable sensors allow non-invasive and scalable method for human activity recognition.

Previous work used uncomfortable and hard to install body or bone conduction based microphones (see sec 2). In this work, we explore to which extent a wrist-worn microphone can be used for scratch detection. The novelties are:

- An evaluation of more comfortable wearable wrist-worn microphone's ability to detect scratch.
- An anotated dataset for scratch detection, which incorporates also different levels of ambient noise.
- A frequency analysis of the scratch from the wrist worn microphone data.
- Analysis and visualisation of how the information of the extracted features change when background noise is present and when it is not. We used Mutual Information - based feature selection.
- An evaluation of scratch detection algorithms.

II. Related work

Triaxial accelerometers [4] on the wrist for scratch detection have been successful and achieved high sensitivity (0.96) and specificity (0.92). They recruited 12 peopled and they were instructed when to scratch. The duration of the dataset collection was 140s per user. Each person scratched the back of their head, leg, elbow, walked and rolled. Research was done using the acceleration of wrist-mounted mobile phone [5] to detect scratch, where 90% accuracy was achieved. It was evaluated on three participants with the ground truth captured with a infra-red camera during one night.

A body-conduction based microphone [6] was used to detect nocturnal scratching in acute dermatitis patients. They used a piezoelectric microphone, which was placed on the participants wrist. The sensor needs to be attached tightly to a the hand, to avoid capturing the sound from the movement of the watch. The data was collected from 4 atopic dermatitis patients and 8 healthy patients during the night. They were filmed in order to annotate the scratches. On average, healthy candidates were scratching for 2 minutes per night, while the users with atopic dermatitis scratched for 24 minutes. The system had a correlation coefficient of 0.98 between the automatically annotated software and the ground truth.

Bone-conduction microphone [7] has been used for scratch detection. The sensor measures the vibrations of

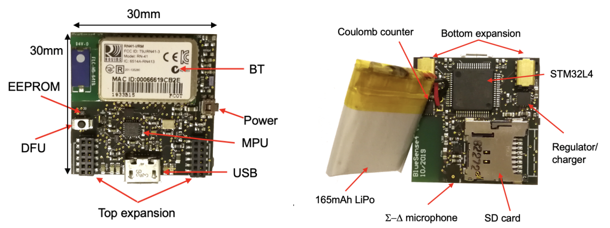
Acitivity ID	Activity	Repetition	Duration	Wait time
1	Low intensity scalp scratch	10	3 s	7 s
2	High intensity scalp scratch	10	3 s	7 s
3	Low intensity scalp scratch while watching TV-series (distance=50 cm, 1m, 2m)	3, 3, 3	3 s	7 s
4	High intensity scalp scratch while watching TV-series (distance=50 cm, 1m, 2m)	3, 3, 3	3 s	7 s
5	Shoulder scratch	10	3 s	7 s

TABLE I: Protocol of data collection used to collect the dataset. Each participant did 30 scalp scratches and 10 decoy scratches while scratching the shoulder rather than the head. The wait time is the between scratches, where a person does nothing in order to visually assess from the signal when the scratch occurs.

the fingernail while scratching. The power spectrum has shown that it can also be used to distinguish the intensity of the scratch. However, for longer data collections this system is not practical because it requires a fixed sensor on the fingernail and wires going from a fingernail to the sensor platform.

In summary, sound can be a powerful predictor for scratch detection. The limitation is that body and bone conduction microphones require a specific set up, which is uncomfortable for long duration data collection (multiple days).

III. Wearable Sensing platform



(a)

(b)

Fig. 1: The data collection platform is 30x30mm in size with connectors for additional expansions (a). The microphone is on the skin-facing side when worn on the body (b).

The platform [8] used for data collection is an in-house wearable sensing research platform. Its primary function is to be an Inertial Measurement Unit (IMU) and a digital microphone, which can be expanded using expansion connectors for sensor research purposes. The device is 30x30mm. It is based on an ARM Cortex M4 processor (STM32L496 from ST), which runs at 20MHz with our default firmware. The platform comprises a 9-axis inertial measurement unit (TDK Invensense ICM-20948), an digital MEMS microphone (ST MP34DT05-A), a micro-SD card, Bluetooth 2 and USB interfaces, a fuel gauge for built-in power measurements (LTC2942), an EEPROM to store configuration (M24128). The processor built-in real-time clock (RTC) is operated from a dedicated 32KHz quartz (10ppm frequency tolerance). The platform operates at 3V from a lithium polymer battery (165mAh) with a LTC3553 voltage regulator. The expansion connectors provide I2C, SPI, UART, ADC inputs and DAC output for expansion boards.

The device firmware has been designed for ease of use. It allows without any programming to acquire the data from the built-in sensors or external ADC inputs through a command line interface. Data can be streamed over Bluetooth or USB, or stored in the SD card. The current firmware allows to acquire in isolation IMU data, sound data, or analog inputs data from the expansion connector. It can also acquire multimodal data: i.e. it can simultaneously acquire IMU, sound and analog inputs in a multiplexed streaming and storage format. This is particularly relevant to acquire data for activity recognition based on a combination of multimodal sensors. The sample rate of IMU, microphone and ADC is fully configurable. The IMU data is also processed by the firmware to obtain the device orientation in quaternions. All data is time-stamped using the internal RTC.

The microphone is omnidirectional. It has a 64dB signal-to-noise ratio and a -26 dbFS sensitivity. It is on the bottom-side of the device (i.e. facing towards the skin if worn on the wrist), but it is not in direct contact with the skin due to the case. The microphone is clocked at 2MHz and provides a 1-bit digital pulse density modulation output, which comes straight from its sigma-delta analog to digital converter. The data is converted to an audio signal using a 3rd order Sinc filter, each with a decimation ratio of 82 which yields a 16-bit 8KHz audio signal with a dynamic range has been experimentally tuned for typical ambient sounds and speech ($\leq dB$).

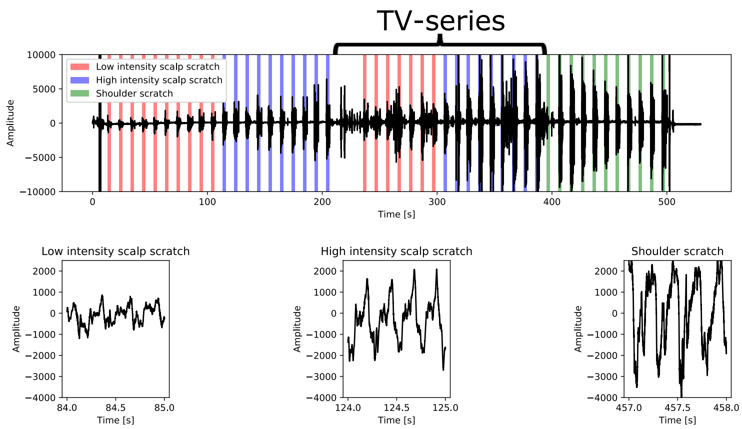
IV. Evaluation of scratch detection

The aim of this work is to evaluate the effectiveness of scratch detection from sound collected by a wrist-worn microphone using three steps: data collection, analysis and classification.

A. Dataset collection

The aim of this protocol (Table I) is to evaluate:

1) if scratching is detectable with the microphone in a silent environment (activities 1, 2 and 5)
2) if scratching is detectable when there is background noise (activities 3,4)
3) if scalp scratching (activities 1, 2) can be distinguished from shoulder scratching (activity 5).



(a) Data collection

(b) The sensor was attached on the wrist using a band.

Fig. 2: The protocol contains 10 low-intensity, 10 high-intensity scalp scratches and 10 shoulder scratches which was recorded without any noise in the background. There were 18 recordings with noise in the background.

The dataset was collected from 4 users, as seen in table II. To have a representative evaluation we collected a dataset from 2 males with different hair lengths, as well as 2 females: one with loose hair and one with a tight pony tail.

User	Gender	Hair type	Hair length
1	Female	Straight (tight pony tail)	40cm
2	Male	Black, straight	10cm
3	Female	Straight loose	50cm
4	Male	Straight loose	20 cm

TABLE II: Characteristics of the participants in the dataset.

For the data collection each user wore a sensor (Figure 1) on their main hand (Figure 2 (b)). As it is common to wear watches and bracelets on the wrist it is also the most adequate place for the device's location. For large scale data collections the device is not more invasive than a smart watch.

The protocol was shown on an Ipad using "Seconds Pro" [9]. This was a constrained recording. It was done to show that scratch can be recognised using wrist-worn microphone and if it can be recognised with ambient noise present. The audio was recorded at 8kHz and was stored to a SD card.

B. Signal Visualisation

The spectogram was created by computing Fast fourier transforms of a sliding window of 1000 samples (0.125 seconds) with an overlap of 50%. The higher size of the window was chosen as the noise originating from body movements is captured by the microphone, as seen in the high intensity scalp scratch in figure 2(a). From the figure 3, it can be seen that there is no observable clusters of dominant frequencies.

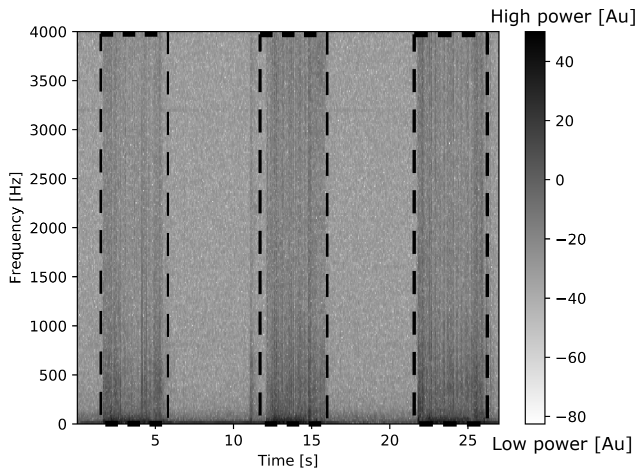


Fig. 3: Spectogram of an excerpt of the recording (user 2). Three scratches (dashed boxes) and the null class are visibile.

C. Feature Analysis

Two of the most common methods used to select the most important features are the filter and wrapper methods. The wrapper method would take an unreasonable amount of time to find the best combination of features from every permutation of the 85 features. Therefore, the filter method that used the mutual information [10] algorithm was used to select the features carrying the highest amount of information.

$$\mathrm{I}(X;Y) = \sum_{y\in\mathcal{Y}} \sum_{x\in\mathcal{X}} p_{(X,Y)}(x,y) \log\left(\frac{p_{(X,Y)}(x,y)}{p_X(x)\, p_Y(y)}\right) \tag{1}$$

where p (X,Y) is the joint probability mass function of X, and $p_{(X)}$ and $p_{(Y)}$ are the marginal probability mass

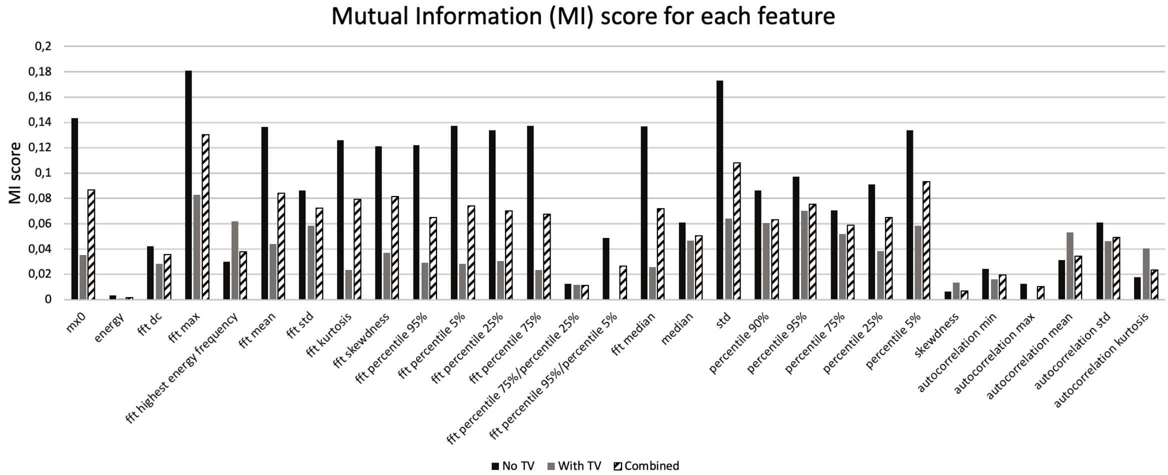


Fig. 4: The Mutual Information score is shown for each feature for data, that was collected under three different conditions: without background noise (no TV), with background noise (with TV) and when these two datasets are combined. It can be seen, that when the TV-series are on, the MI score drops compared to when the scratches occur in a silent environment.

functions of X and Y, respectively, where X is the bin of the feature and Y is the class. For each feature, we created 100 equal-size bins and performed a small correction by adding 10^{-10} to the division so that when there were empty bins, division by zero was not occurring.

We show in figure 4, that the features extracted from the frequency domain (Fourier transform) carries the largest amount of information in the silent background. However, when TV noise are being played in the background (noisy), the amount of information that each feature carries for scratch detection is smaller.

	Silent background		Noisy background	
Order	Feature	MI score	Feature	MI score
1	FFT max	0.181	FFT max	0.083
2	stdD	0.173	75%	0.07
3	mx0	0.143	std	0.064
4	FFT 5%	0.137	fft max f	0.0602
5	FFT 75%	0.137	90 %	0.06

TABLE III: The top features are shown with the associated Mutual Information score. FFT maximum amplitude had the largest MI score in both cases. However, the top features when noise is present have a much smaller MI score compared to the features extracted from the data, which was collected in silence.

In the table III we show, that the top 5 features in the silent and noisy backgrounds. It can be seen, that the MI scores are much lower from the data, which was collected with TV-series in the background, compared to the data collected in a silent background. The top feature FFT maximum amplitude carried the largest amount of information. A key difference in both cases is that when TV-series are not present, top 5 features are from the frequency domain, while when TV-series are not present two (75% and 90% of the window) of the top 5 features are frequency invariant.

D. Classification

Type	Features	Dimension
$\mathcal{F}$	DC component of FFT	1
	Highest FFT value and frequency	2
	Max, Min, Mean, std of the FFT values	2
	Kurtosis, skewness of FFT	2
	Percentiles 5, 25, 50, 75, 90 of FFT	5
$\mathcal{T}$	Mean, std	1
	Percentiles 5, 25, 50, 75, 90	5
	Auto-correlation (min,max,mean,std)	4
	Mean-crossing rate	1
	Kurtosis, Skewness	2
$\mathcal{E}$	Energy of 50 Hz bands	60
	Total number of features	**87**

TABLE IV: We used 85 extracted features from the sliding windows of size 1000 (0.125 s) and 50% overlap. $\mathcal{T}$ are the features extracted from the sliding window, $\mathcal{F}$ - from Fourier transforms, $\mathcal{E}$ - energy of the frequency bands.

We trained a classification task using random forest (100 estimators with 5 max depth) model with Gini impurity and leave-one-user-out cross-validation. To evaluate if a user-independent scratch recognition model can be trained to work on unseen users, we train on User 1, 2, 3 and test on User 4; and then we repeat this with different test users. For the classification task we extracted the features (Table IV) from a sliding window. The sliding window size is 1000 samples (0.125 seconds) with an overlap of 50%. The null class contains everything except scratch. The labels were split into overlapping windows and the dominant label in the window was chosen. We calculated sensitivity or specificity by defining scratch as positives (P) and null class as negatives (N). To evaluate the performance

Test number	Environment	Classes	User 1	User 2	User 3	User 4	Accuracy	$F1_{weighted}$
1	Silent	Accuracy of scratch	72 %	97 %	83 %	83 %	83.75%±8.8%	82%±2.3%
		Accuracy of null	83 %	75 %	74 %	82 %	78.5%±4%	
2	TV-Series	Accuracy of scratch	42 %	87%	40 %	75 %	61% ±20.45%	73.75% ±3.87%
		Accuracy of null	81 %	62%	88%	81%	78%±9.6%	
3	Silent	Accuracy of scalp scratch	51 %	40%	20%	63%	43±15.8%	65.4%±9.23%
		Accuracy of shoulder scratch	30 %	21 %	54%	52%	39.25%±14.13%	
		Accuracy of null	79 %	43 %	76 %	74 %	68% ±14.54%	

TABLE V: The results of the classification show that in a silent environment scratch can be detected with 83.7% accuracy, while the null class with 78%. With TV-series in the background scratch with 61% and null with 78%. Shoulder scratches cannot be distinguished from scalp scratches. The accuracy is rate at which the model predicts the class correctly.

we used sensitivity and specificity (Equation 1), with the $F1_{weighted}$ (Equation 2), where n_i is the class size and N is the total number of samples.

$$sensitivity = \frac{P_{True}}{P},\ specificity = \frac{N_{True}}{N} \tag{2}$$

$$F1_{weighted} = \sum_i 2w_i \frac{precision_i\ recall_i}{precision_i + recall_i},\ w_i = \frac{n_i}{N} \tag{3}$$

For the 1st aim (ref section 4.1) we trained and tested only on scratches (scalp and shoulder scratches (Activities 1, 2, 5) are one class) with no TV-series present. The model achieved sensitivity of 83.75%±8.8% and specificity of 78.5%±4%. $F1_{weighted}$ score was 82%. The worst performance is obtained with User 1, where the sensitivity is 72%.

For the 2nd aim, where we trained on all the data, but tested only on the section where TV-series (Activities 3, 4) are present. This yielded sensitivity of 61%±20.45%, specificity of 78%±9.6% and a $F1_{weighted}$ score of 73.75%.

For the 3rd aim we wanted to evaluate if shoulder scratch can be distinguished from scalp scratches. We performed a classification with 3 classes: scalp scratch, shoulder scratch and null class (Activities 1, 2, 5). The model was trained and tested only on the data where TV-series are not present in the background. We achieved that on average scalp scratch was recognised with an accuracy of 43%±15.8%, shoulder scratch with an accuracy of 39.25%±14.13%, null class 68%±14.54% and the $F1_{weighted}$ score was 65.4%.

V. Discussion

The sensor platform is available to use with multiple sensor modalities. We show that sound is a strong predictor and can be used for sensor fusion. The fusion should improve the overall scratch detection. The dataset was collected from 4 users with different hair styles. We showed that by doing leave-one-user-out cross-validation sensitivity can be 83.75% and specificity 78.5% in a silent environment. Compared to the body-conduction microphone in [6] our performance is lower as our microphone captures the ambient sound whereas the body-conduction microphone is isolated using a gel and needs to be tightly fixed to the hand. Thus, there is a trade-off of comfort and performance.

Moreover, we observed that the MI scores of the features extracted from the data collected with background noise are smaller than from the data collected in silence for scratch detection. It means, if we were to use features selected in the silent environment and were to use them for classification when there is ambient noise, the performance is likely to degrade.

Furthermore, sound made by scratching is dependent on the type of the hair. For example, the lowest performance was achieved on User 1, who had hair tied in a tight pony tail. Another limitation is scratch detection when ambient noise is present. In this case we achieved a sensitivity of 61%±20.45%.

We observed that scratches do not seem to have a particular frequency signature, and thus filtering the ambient sound does not seem practical. When we listened to the recordings, the scratch was audible despite the TV-series playing in the background. It means that the sound of scratching is captured and more sophisticated algorithms, such as Deep Learning could be tried. Microphone beamforming could help to enhance the directional sensitivity of our device.

VI. Conclusion

In this work, we show that sound can be used as a sensor modality to distinguish scratch. In a silent environment the sensitivity and specificity are 83.75% and 78.5%. On the other hand, in a noisy environment the sensitivity and specificity are 59.25% and 77.25%. Moreover, we came to the same conclusion by computing Mutual Information scores to see the importance of the features. The features extracted from the data that was collected with noise in the background had significantly smaller score compared with the data that was collected in a silent environment. With the ambient noise the model with the chosen features is not reliable enough to detect scratches with a high accuracy. It enforces the idea, that additional sensor modalities should be used together to distinguish the scalp scratches, as it will allow to compensate when ambient noise is present.

Ambient noise is a challenge and appropriate techniques should be used to diminish the impact of the noise on the predictions. A microphone would work extremely well for scratches, where there is a limited amount of noise, but for real daily activities it is not accurate enough to be deployed alone. However, it could significantly increase the scratch detection accuracy if it would be fused with other modalities. Our research platform allows us to jointly acquire sound and IMU [8] data and this will be explored in the future.

VII. Acknowledgments

This work has been partly funded by EPSRC grant EP/S513921/1.

References

[1] D. Bautista, S. Wilson, and M. Hoon, "Why we scratch an itch: The molecules, cells and circuits of itch," *Nature neuroscience*, vol. 17, pp. 175–82, 02 2014.

[2] K. Endo, H. Sano, T. Fukuzumi, J. Adachi, and T. Aoki, "Objective scratch monitor evaluation of the effect of an antihistamine on nocturnal scratching in atopic dermatitis," *Journal of Dermatological Science*, vol. 22, no. 1, pp. 54 – 61, 1999.

[3] T. Ebata, H. Aizawa, and R. Kamide, "An infrared video camera system to observe nocturnal scratching in atopic dermatitis patients," *The Journal of Dermatology*, vol. 23, no. 3, pp. 153–155, 1996. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1346-8138.1996.tb03990.x

[4] J. Petersen, D. Austin, R. Sack, and T. L. Hayes, "Actigraphy-based scratch detection using logistic regression," *IEEE Journal of Biomedical and Health Informatics*, vol. 17, no. 2, pp. 277–283, March 2013.

[5] J. Lee, D.-k. Cho, S. Song, S. Kim, E. Im, and J. Kim, "Mobile system design for scratch recognition," in *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, 2015, p. 1567–1572.

[6] Y. Noro, Y. Omoto, K. Umeda, F. Tanaka, Y. Shiratsuka, T. Yamada, K. Isoda, K. Matsubara, K. Yamanaka, E. C. Gabazza, M. Nishikawa, and H. Mizutani, "Novel acoustic evaluation system for scratching behavior in itching dermatitis: Rapid and accurate analysis for nocturnal scratching of atopic dermatitis patients," *The Journal of Dermatology*, vol. 41, no. 3, pp. 233–238, 2014.

[7] T. Okuyama, K. Hatakeyama, and M. Tanaka, "Measurement of human scratch behavior using compact microphone," vol. 45, 01 2014, pp. 731–737.

[8] D. Roggen, "Arm cortex m4-based extensible multimodal wearable platform for sensor research and context sensing from motion sound." *Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers*, vol. 45, pp. 284–289, 2020.

[9] "Seconds interval timer," 2010-2019. [Online]. Available: https://www.runloop.com/seconds-interval-timer-for-iphone

[10] T. M. Cover and J. A. Thomas, *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. USA: Wiley-Interscience, 2006.

Appendix C

N.Coco Stretch: Strain Sensors Based on Natural Coconut Oil and Carbon Black Filled Elastomers

ADVANCED MATERIALS TECHNOLOGIES
www.advmattechnol.de

Coco Stretch: Strain Sensors Based on Natural Coconut Oil and Carbon Black Filled Elastomers

Pasindu Lugoda, Júlio C. Costa, Leonardo A. Garcia-Garcia, Arash Pouryazdan, Zygimantas Jocys, Filippo Spina, Jonathan Salvage, Daniel Roggen, and Niko Münzenrieder*

A biocompatible inexpensive strain sensor constituting of an elastomer filled with natural coconut oil (CNO) and carbon black (CB) is presented here. Strain sensors are widely utilized for applications in human activity recognition, health monitoring, and soft robotics. Given that these sensors are envisioned to be present in a plethora of fields, it is important that they are low cost, reliable, biocompatible, and eco-friendly. This work demonstrates that CNO can be used to create conductive percolation network in elastomers, without the necessity for harmful chemicals or expensive machinery. The sensor has a gauge factor of 0.77 ± 0.01, and the sensing material has a porous morphology filled with an oily suspension formed of CNO and CB. Results indicate that the liquid filled porous structure can improve the reliability of these resistive strain sensors in comparison to sensors fabricated utilizing commonly used non-polar solvents such as heptane. Consequently, the sensor demonstrates a hysteresis of only 2.41% at 200% strain over 250 stretch/release cycles. Finally, to demonstrate the potential of this fabrication technique, a functionalized glove is developed and used to detect wrist motion. These easily manufacturable and cost-effective sensors enable wearable on-skin ergonomic intervention systems with minimal impact on the environment.

Advancements in flexible and stretchable electronics such as deformable transistors,[1–3] circuits,[1,3] sensors,[3,4] and energy harvesters[1,5] are currently paving the way toward innovative wearable systems or soft robots.[6–12] In particular, conformable strain sensors enable a variety of applications such as physical activity monitoring,[13–17] motion capturing,[18–20] telemedicine,[21–23] and artificial skins.[24,25] Moreover, strain sensors have been integrated onto the surface of textiles showcasing their potential to become ubiquitous in smart textile applications.[16,26] This results in the strain sensors coming into direct contact with human skin. Therefore, there has been a significant drive toward the development of biocompatible strain sensors. For instance, researchers have investigated the effects of ionic liquids and hydrogels to fabricate stretchable strain sensors with low hysteresis.[27,28] However, most of these sensors have low gauge factors and high impedances, which in-turn complicates the sensors' readout electronics, making them unsuitable for wearable applications. Although resistive strain sensors with high gauge factors have been fabricated,[29,30] some of these sensors demonstrate a limited strain range of 50%, while others were unable to go back to their original length once stretched. The nature of human motion requires the use of highly stretchable sensors for activity monitoring. This has led to extensive work in developing strain sensing materials and sensors with high sensitivities and wide strain ranges. Most of these sensing materials are fabricated utilizing harmful chemicals,[31–36] expensive equipment,[37,38] sophisticated mechanical designs,[36,39,40] or advanced low dimensional materials.[41–43]

Given the absence of inherently stretchable piezoresistive materials, strain sensors are commonly synthesized by mixing elastomers and conductive materials. In this context Ecoflex is preferred as an elastomer due to its high stretchability, biocompatibility, and easy fabrication. In addition, sensors constructed using Ecoflex showcase a fast response and low hysteresis.[13,20,44] In the case of conductive fillers, a low cost conductive material is highly desirable for a large scale and inexpensive production of strain sensors. In this regard, carbon black (CB) has been preferred[13] in comparison to other fillers such as metal nanowire/nanoparticles,[19,45–47] carbon nanotubes/nanoparticles,[31,48] and liquid metals.[49] However, the poor dispersibility of CB in water complicates the formation of uniform and controllable CB conductive networks without resorting to potentially harmful solvents.[50] In general, non-conductive solvents such as dimethyl dispersion solution,[31]

Dr. P. Lugoda, J. C. Costa, Dr. L. A. Garcia-Garcia, Dr. A. Pouryazdan, Z. Jocys, F. Spina, Prof. D. Roggen, Prof. N. Münzenrieder
Sensor Technology Research Centre
University of Sussex Falmer
Brighton BN1 9QT, UK
E-mail: d.v.lugoda@sussex.ac.uk
Dr. J. Salvage
School of Pharmacy and Biomolecular Sciences
University of Brighton
Lewes Road, Brighton BN2 4GJ, UK
Prof. N. Münzenrieder
Faculty of Science and Technology
Free University of Bozen-Bolzano
Bolzano 39100, Italy
The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/admt.202000780.

DOI: 10.1002/admt.202000780

chloroform,[32] Toluene,[33] Heptane,[34] hexane, and acetone[35] have been utilized during the preparation of CB-based elastomer sensors. These solvents increase the dispersibility of the CB particles in the mixture. Nonetheless, most of these solvents are explosive or toxic and therefore are required to be handled with care when fabricating the sensors. Moreover, after the fabrication, traces of these chemicals will be present in the structure, and this could harm the wearer. Shintake et al.[38] have created strain sensors with a stretchability of 500% and a gauge factor ranging from 1.62 to 3.37 by mixing CB in Ecoflex without the use of these harmful chemical. However, to synthesize these sensors, expensive equipment such as planetary centrifugal mixers was utilized. Non-toxic but synthetic diluting agents such as silicone oil has been utilized in the synthesis of CB silicone strain sensors,[51] where the excess of silicone oil was evaporated after the CB particles were mixed in the silicone rubber. The material exhibited limited stretchability of 155–164% which makes these sensors less preferable for smart textile application where a large working range is required.

Another important factor that affects the performance of elastomeric strain sensors is the constant friction between the solid-state conductors and the elastomeric molecules.[47,52,53] This leads to slippage and detachment among the conductors resulting in low repeatability of the sensor measurements. This can be addressed by utilizing liquid-state conductors, due to the ability of liquids to undergo virtually limitless and instantaneous deformation.[12,52] Therefore, we utilize eco-friendly coconut oil (CNO) instead of volatile solvents to create a partially liquid and conductive percolation network in Ecoflex when mixed with CB. CNO is an inexpensive, non-hazardous, and environmentally friendly natural polymer. It is recognized as a cooking oil fit for human handling and consumption, which is why it is widely used across the food and cosmetic industries.[54] This oil, as with most of the other natural esters, consists of triglycerides.[55] These triglycerides contain different fatty acids that make the CNO comparatively conductive to other oils.[55] Previously, CNO mixed with grape seed oil was used to fabricate a temperature sensor.[56] Large stretchability (≥100%) and great stability are also important factors for sensors that require wide strain detection range and long-term repeated usage.

An easily manufacturable, eco-friendly, and cost-effective fabrication technique for producing strain sensors with a large strain detection range and great stability, is presented in this work. This fabrication technique is well-suited for environments where standard, yet expensive, electronic and chemical fabrication facilities are not available, as is the case of schools and developing countries. These biodegradable devices are easily customizable and can be utilized to develop ergonomic intervention systems that can be used by everyone. Here, CNO was used to disperse CB and create conductive pathways encased in an Ecoflex structure, as shown in **Figure 1**. Furthermore, scanning electron microscopy (SEM) images and energy-dispersive X-ray (EDX) spectral analysis of the fabricated sensors demonstrated that high CNO density areas displace the Ecoflex, resulting in interconnected tunnel-like CNO/CB pathways. These conductive liquid pathways lead to more reliable measurements for different strains, when compared to sensors fabricated using heptane. Moreover, the sensors demonstrated a hysteresis of only 2.41% at 200% strain over 250 stretch/release cycles. The response time and resistance drift with time and temperature of the sensor were also investigated. In addition to the characterization, a coconut strain sensor was attached to a textile glove and utilized to monitor wrist motion.

The composite sensing material created by mixing Ecoflex, CNO and CB has a porous morphology, as shown in Figure 1b. CNO replaces conventional solvent-based methods to create CB conductive pathways within Ecoflex. Unlike solvents, the CNO does not evaporate and remains in the structure when the Ecoflex solidifies. This creates a porous structure with a porosity of 35.3%. The pores were filled with an oily suspension containing CB particles. To further clarify this, the sensing material was pressed to get the oily suspension out. Then an EDX spectral analysis was conducted on the oily suspension to identify its composition. The results given in Figure 1c1 illustrate that the material contained mainly carbon and oxygen. In comparison, the EDX analysis of the frame shown in Figure 1c2 displays high concentrations of silicon. Hence, it can be said that the CB and CNO were left in the pores of the material. Furthermore, it was also vital to identify if the porous structure without the oily suspension contained CB particles. For this reason, the material was pressed and thoroughly cleaned with IPA and water to remove the oily suspension. Afterward the sample was dried, it consisted of only the porous Ecoflex structure without any oil or solvent. This dried sample was dissected, and a high-resolution SEM image of the cross section was obtained which is displayed in Figure 1d. As shown in Figure 1d, the structure contains well dispersed CB particles. This indicates that during the fabrication process some CB particles were also dispersed in liquid Ecoflex and once it cured these CB particles were trapped in the porous Ecoflex structure.

After the development of the individual wet sensors, a functionalized glove was created and utilized for wrist motion detection. To achieve this, a single sensor was attached using Ecoflex to a wrist glove, as shown in Figure 1e. A volunteer was asked to perform four wrist movements, namely, extension, flexion, ulnar deviation, and radial deviation. Each wrist position was held for 30 s and in between each position the hand was returned to its original neutral position and held for 30 s. The data was recorded using a battery powered Bluetooth enabled Bluesense microcontroller, making this system completely standalone.[57] Although these were multidimensional movements the results shown in Figure 1f demonstrate that a single sensor can clearly distinguish extension and flexion from radial and ulnar deviations. Therefore, we proved the feasibility of using the wet sensor for wrist motion detection, which is beneficial for activity recognition and preventing carpal tunnel syndrome.

The simple fabrication technique of the CB/CO sensor is shown in **Figure 2**a, and this process is explained in detail in the Experimental Section. In order to characterize the properties of the composite material, the conductivity of the wet composite material for different concentrations of CB ranging from 0% to 12% was investigated and the I/V graphs are shown in Figure 2b. As observed in the graphs the best conductivity is obtained at 12% CB. At higher concentrations of CB the solution was too viscous to be efficiently mixed using a simple magnetic stirrer. For comparison purposes, 12% CB was mixed with Ecoflex without CNO and its I/V characteristics are displayed in Figure 2c. This material demonstrated a low conductivity

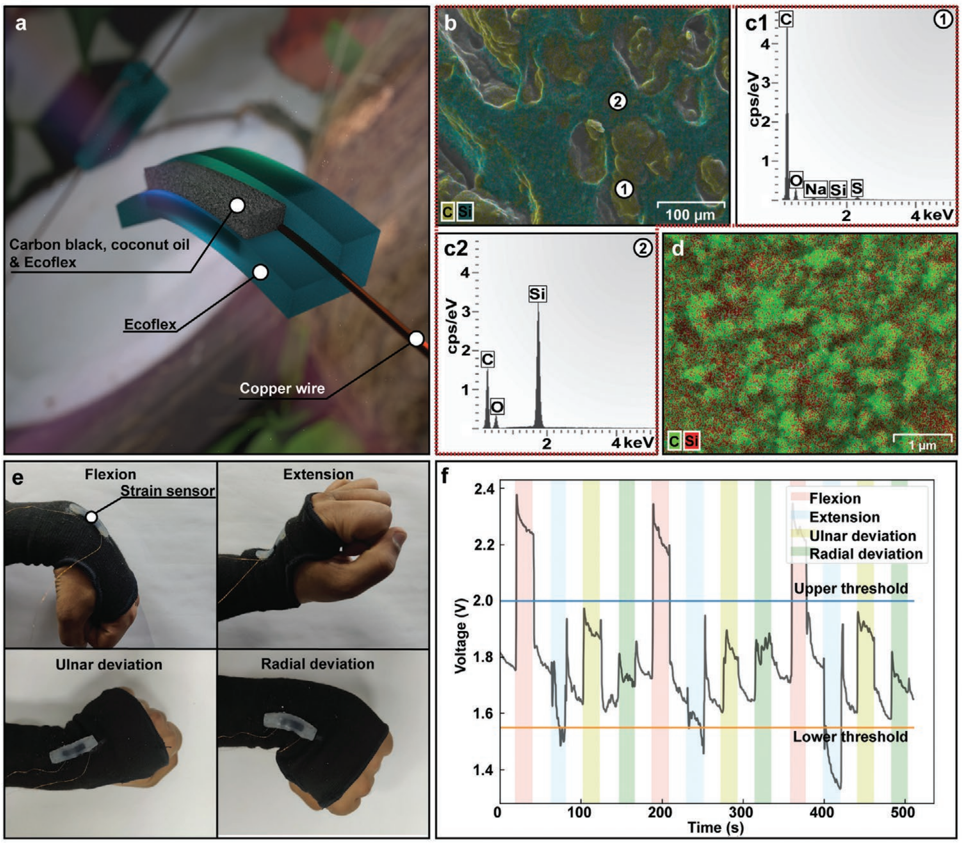


Figure 1. Carbon black dispersed in coconut oil can be used to fabricate sensors for smart textiles. a) Concept figure of a CB/CNO strain sensor encased in an Ecoflex structure. b) SEM image of the cross section of a CB/CNO sensor material. Here the regions with a higher carbon concentration are presented in yellow and cyan shows the regions with a higher silicon concentration. c1) Energy-dispersive X-ray (EDX) spectral analysis done on the oily material present within the pores of the composite material. c2) EDX analysis done on the surface of the sensor. d) Close-up image of the dried material where carbon is shown in green and silicon in red. e) The strain sensor integrated onto a textile glove and utilized for flexion, extension, ulnar deviation, and radial deviation. f) Results from wrist motion experiment where the response due to flexion, extension, ulnar deviation, radial deviation, and neutral position are shown in the red, blue, yellow, green, and white regions of the graph, respectively. The graph was utilized to distinguish between flexion and extension.

and the material did not have uniform electrical properties throughout its structure. Moreover, the electrical characterization of the dried sample is given in Figure 2c. The observed conductivity of the dried sample is due to the CB particles dispersed within the porous structure. As demonstrated in the I/V graphs (Figure 2a,c), the wet sample has a better conductivity in comparison to the dried one. This increase in conductivity is due to the presence of the oily suspension comprising of CB and CNO in the wet sample. The results from the I/V graphs indicate that the 12% CB/CNO wet material has the best conductivity when compared to the rest. Hence, to assess the feasibility of using this CB/CNO wet composite material to measure strain, we fabricated and characterized a CB/CNO resistive strain sensor. The sensors constituted of the CB/CNO material encased and sealed within an Ecoflex structure as illustrated in Figure 1a. The necessity for the Ecoflex encasement was due to the oily exterior of the sensing material, which comprised of an oily suspension in the porous structure. Since the sensors were fabricated using rigorous stirring, it was important to identify if the resistance of the sensors varied significantly in between different devices. Therefore, three different material batches were synthesized, and from each batch three different sensors were fabricated. The sensors had an average resistance of 235 ± 181 kΩ and the resistance of these sensors could be easily customized by changing the CB concentration as shown in the I/V graphs in Figure 2b–d. It was important to identify if stretching the CB/CO material brings about any changes in the surface morphology. Therefore, microscopic images were obtained from the surface of the material when it was relaxed and stretched to 100% strain, and they are displayed in Figure 2e,f. The images show that there are no cracks in the surface due to stretching. This was further clarified by the SEM image of the stretched material displayed in Figure 2g.

Further to the characterization of the composite material, individual CB/CNO strain sensors were fabricated. **Figure 3**a displays a microscopic image of a cross section taken from a CB/CNO strain sensor. From images of several cross sections we extracted a porosity of 35.3% and Figure 3b shows the analysis conducted on Figure 3a, where white areas represent the porous regions. In general, for bulk materials water saturation and water evaporation method are more suitable however, for this material we utilized microscopic images because

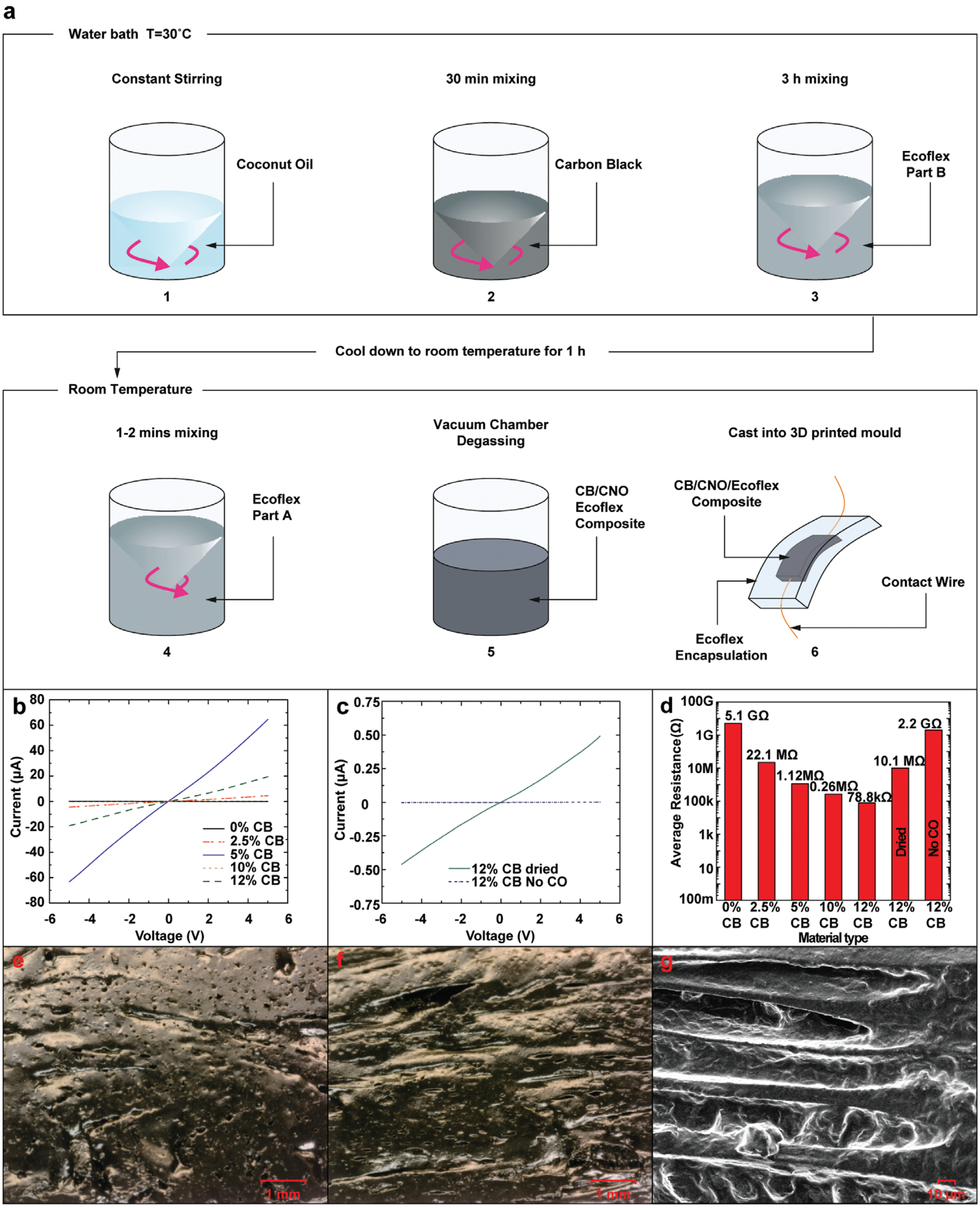
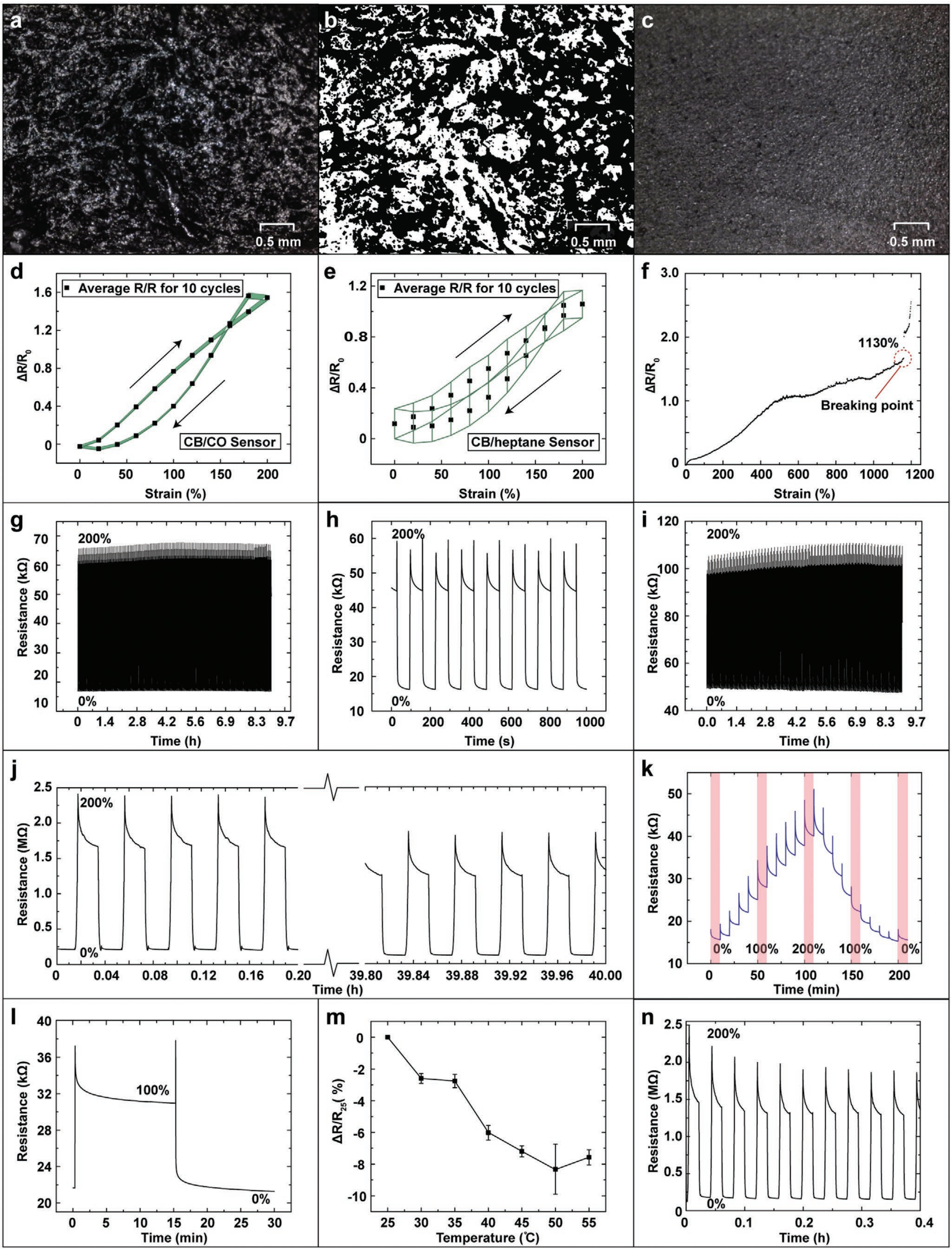


Figure 2. The fabrication technique and *I*/*V* characteristics of the CB/CO composite material along with microscopic images of the material when it is relaxed and stretched. a) Simple fabrication technique of the CB/CO sensors. b) *I*/*V* graphs for different concentrations of CB in CNO and Ecoflex material, from 0–12%. c) The *I*/*V* characteristics of the dried sample that contained 12% CB and the *I*/*V* characteristics of a mixture containing only 12% CB and Ecoflex without CNO. d) The resistance measurements from the different materials compositions. Microscopic images of a surface of the CB/CNO sensor material when e) it is relaxed, f) when it is stretched to 100% strain, captured using a Dino-Lite premier digital microscope (New Taipei City, Taiwan). g) A SEM image of the stretched CB/CNO sensor material at 100% strain.


a
b
c
0.5 mm
d
Average R/R for 10 cycles
CB/CO Sensor
ΔR/R0
Strain (%)
e
Average R/R for 10 cycles
CB/heptane Sensor
f
1130%
Breaking point
g
200%
0%
Resistance (kΩ)
Time (h)
h
Time (s)
i
j
Resistance (MΩ)
k
100%
Time (min)
l
m
ΔR/R25 (%)
Temperature (℃)
n

the material contained sealed pockets of CB/CO within its structure. Moreover, to directly compare these strain sensors with those fabricated using heptane as a solvent, we manufactured CB/heptane sensors with similar dimensions to CB/CNO sensors. The microscopic image of the cross section of a CB/heptane sensor is displayed in Figure 3c and it demonstrates a less porous structure. The performance of strain sensors is commonly characterized by their gauge factor (sensitivity), stretchability, hysteresis, and response time.[36] The response of a resistive strain sensor assuming that the cross section of the electrode layers is uniform (Poisson ratio ≈ 0) can be expressed as follows:[38]

$$R = \rho\left(l_0/w_0h_0\right)\left(\varepsilon_1+1\right)^2 = \left(\frac{\rho}{\rho_0}\right)R_0\left(\varepsilon_1+1\right)^2 \tag{1}$$

where ρ is the electrical resistivity of the conductive elastomer; ρ_0 is the reference resistivity; R_0 is the reference resistance; l_0, w_0, and h_0 are the length, width, and thickness of the electrodes, respectively. When strained, the resistivity of CB/CNO material changes as a result of the breakdown and realignment of CB aggregates which alters the conductive paths within the material.[38,58] This alteration results in a change in resistance and the ratio of this relative change to the mechanical strain is known as the gauge factor, GF_R. The GF_R of a resistive strain sensor is given by the following equation.

$$GF_R = \left(\frac{\Delta R}{R_0\varepsilon_1}\right) = \left(\frac{1}{\varepsilon_1}\right)\left(\left(\frac{\rho}{\rho_0}\right)R_0\left(\varepsilon_1+1\right)^2-1\right) \tag{2}$$

where ΔR is the change in resistance when compared to the unstrained resistance R_0. The strain on the sensor is given by ε_1.

The gauge factor measurements for the presented CB/CNO sensor are shown in Figure 3d. The gauge factor of the sensor ranges from 0.21 ± 0.01 at 20% strain to 0.77 ± 0.01 at 200% strain. The graph also demonstrates the repeatability of $\Delta R/R_0$ for each strain and recovery step over ten cycles. The average standard deviation for each step is only 0.017. The gauge factors of the heptane sensors are 0.77 ± 0.58 at R_{20} and 0.52 ± 0.05 at R_{200}. The measurements over ten cycles for the CB/heptane sensors are shown in Figure 3e. The graph demonstrated a lower repeatability of $\Delta R/R_0$ in comparison to the CB/CNO sensors for strain steps over ten cycles. For the heptane sensor, the average standard deviation of $\Delta R/R_0$ was 0.112 for each step. This is significantly larger than the 0.017 obtained from the CB/CNO sensor. The performance of CB/CNO sensors was superior to the fabricated heptane sensors which used hazardous chemicals. The superior performance of the CB/CNO sensor is due to the oily liquid nature of the material. This reduces the friction in between the CB particles and the elastomer molecules which minimizes the detachment and slippage among these conductive CB particles.[12,47,52]

These sensors when utilized for detecting human motion must be able to withstand large strains. Hence, it is vital to measure the maximum stretchability that the sensor can endeavor. The CB/CNO sensors had a maximum stretchability of 1035 ± 215% which can be regarded more than adequate for detecting human motion. The change in resistance of a CB/CNO sensor when stretched to its rupture strain is displayed in Figure 3f. In real life applications, strain sensors must also maintain their sensing characteristics without fatigue failure and minimal hysteresis. Therefore, the sensors were subjected to 250 stretch/release cycles at 200% strain to examine their long-term stability. The complete experimental results are presented in Figure 3g. A sample of seven cycles of the test is displayed in Figure 3h. The cyclic test verifies that the strain sensor has a good mechanical durability against repeated stretch/release cycles at 200% strain with no signs of fatigue failure. The change in sensor resistance at R_0 after 200 cycles was only −0.28 ± 0.07% and the change in R_{200} was 2.41 ± 0.06%. These percentage changes were similar to the measurements recorded from the CB/heptane sensor, where R_0 and R_{200} changed from −3.52 ± 0.08% and 1.26 ± 0.12%, respectively (shown in Figure 3i). Thereafter, a larger cyclic test of 1000 cycles at 200% strain was conducted on a CB/CO sensor to further understand its durability and stability. The results presented in Figure 3j indicate that these sensors are even functional after 1000 cycles. An overshoot was also observed during the cyclic tests in both the CB/CNO and CB/heptane sensors, as shown in Figure 3g–j. Therefore, an experiment was conducted on CB/CNO sensor to estimate the effects of overshoot on the measurements. As it can be seen in Figure 3k, the sensors' overshoot increases with increasing strain. The average overshoot was calculated as 22.16 ± 4.51%.

For wearable applications, it is also vital to determine the response time of the sensors when subjected to step changes. The response time of the CB/CNO sensors was measured as 1.11 ± 0.21 s at 100% strain. This response time is sufficient for 3D human body shape reconstruction applications,[17,59] ergonomic intervention systems for posture recognition,[16,60,61] and carpal syndrome detection.[62–64] Furthermore, the resistance drift of strain sensors can greatly hinder their usability in wearable applications.[65] Therefore, to observe the drift characteristics, the CB/CNO sensor was stretched to 100% strain and left

Figure 3. Microscopic images of the CB/CNO and CB/heptane sensors along with their performance. a) Close-up of the CB/CNO sensor material using a Dino-Lite premier digital microscope (New Taipei City, Taiwan). b) Image analysis conducted on (a) utilizing Image J, where white areas represent the porous regions. c) Close-up of the CB/heptane sensor material using the digital microscope. d) Change in resistance with response to strain over ten cycles for the sensor fabricated using CNO. e) Change in resistance with response to strain over ten cycles for the sensor fabricated using heptane. f) Resistance change recorded from a representative CB/CNO sensor when stretched up to its rupture strain of 1130%. g) Cyclic tests conducted on the CB/CNO sensor for 250 cycles at 200% strain and h) presents a section of the cyclic test. i) Cyclic tests conducted on the CB/heptane sensor for 250 cycles at 200% strain. j) The first and the last 5 cycles from the stability and durability experiment where a CB/CO sensor was cycled at 200% for 1000 cycles (the complete experimental measurements are given in Figure S2, Supporting Information). k) Increase in strain in steps of 20% up to 200% to identify the overshoot in the sensor. l) Drift measurements from the sensor where the sensor was kept stretched for 15 min. m) Temperature response of the CB/CNO sensor. n) Cyclic measurements at 200% strain from a CB/CO sensor that was cooled to 10 °C.

for 15.0 min. This is shown in Figure 3l. The initial measurements were always obtained 5 min after the step change to ensure that the overshoot was not influencing the measurements. Within a 10 min time window a drift of 2.04 ± 0.06% was recorded from the sensor. In addition, temperature also has an impact on the performance of most resistive strain sensors.[38] Therefore, we investigated the influence of temperature on the unstrained sensor. The maximum total change in $\Delta R/R_{25}$ of the sensor was measured as 8.32 ± 1.58%, when subjected to step changes of temperature from 25 to 55 °C at 5 °C increments. The results are shown in Figure 3m. Moreover, since CNO has a melting point of 24 °C it was important to identify the functionality of the sensors at lower temperatures. Therefore, the same sensor that was utilized for the stability experiments was positioned within a fridge at 10 °C for 1 h. Then the sensor was taken out of the fridge and immediately cyclic tests at 200% strain were performed on the sensor. The results displayed in Figure 3n indicate that cooling the sensor had minimal impact on its performance. The sensor performance is not affected by lower temperatures because CO does not form into a complete solid at 24 °C; it forms into a thick, greasy material. This is due to the triglycerides in CNO consisting of a mixture of different fatty acids, and each of these fatty acids has its own melting point, and at 24 °C only some fatty acids solidify.[66]

This is the first time strain sensors have been successfully fabricated using CNO. The strain sensitive material synthesized had a porous structure filled with an oily suspension containing CB particles. These CB/CNO sensors produced more reliable measurements in comparison to CB/heptane sensors also presented in this work. Heptane is a non-polar solvent that has generally been utilized for mixing CB particles in elastomers.[34] The sensors demonstrated a superior stretchability compared to the once fabricated using silicone oil.[51] Moreover, the gauge factor and temperature response were comparable to the resistive strain sensors manufactured by Shintake et al.[38] using an expensive planetary centrifugal mixer. This paper has demonstrated that CNO can be used to replace harmful chemicals such as heptane and expensive equipment in the creation of strain sensors. These strain sensors can be fabricated using inexpensive and easy-to-get-raw materials with minimal effect on the users' health.

The sensors had a gauge factor of 0.77 ± 0.01, a stretchability of 1035 ± 215%, and a resistance drift of only 0.69% in R_0 and 2.41% in R_{200} after 250 cycles at 200% strain. Ultimately, the strain sensors were integrated onto a glove and successfully utilized for detecting wrist motion. These sensors are a low-cost and easy to fabricate technique that can be utilized in the future to create strain sensors for wearable applications.

Experimental Section

Preparation of the Carbon Black/Coconut Oil Ecoflex Sensor: The sensors fabricated in this paper comprised of CB and CNO, encased in Ecoflex. The sensing material was fabricated by mixing liquid silicone elastomer Ecoflex (00-30 Smooth-on, Pennsylvania, United States), CNO (Pipkin, London, UK), and CB (Vulcan P, Cabot, Boston, Massachusetts, United States) in 10:5:1.2 by weight.

For the preparation of the material initially the CB was added to the CNO and heated over a water bath to 30 °C. It was crucial to maintain this temperature since the melting point of CNO is 24 °C. The solution was stirred for 0.5 h to ensure the release of the agglomerated CB particles. Then part B of Ecoflex was added to the solution and it was stirred rigorously for 3 h. Thereafter, the heating was turned off and the solution was left to cool at room temperature over the course of 1 h, during this time the solution was continuously stirred. This was done to ensure the solution was cooled when part A was added, since higher temperatures caused the Ecoflex to cure faster. Afterward, part A was added and the solution was stirrer for 1–2 min. Then the solution was degassed and poured into an Ecoflex mold for curing. The mold had dimensions of length 20 mm by width 5 mm by height 2 mm. Moreover, copper contact wires were inserted into the structure before it was fully cured. The wires were later used to connect the sensor onto the readout electronics. Thereafter a layer of Ecoflex was added on top to fully encase the sensing material.

Preparation of the Carbon Black/Heptane Ecoflex Sensors: The materials utilized for the CB/heptane sensors were heptane, Ecoflex, and CB (10:10:1.2). These samples were fabricated by stirring CB with heptane for 0.5 h. Afterward, part A of Ecoflex was added and stirred for 3 h. Finally, the part B of Ecoflex was added and stirred for an additional 3 h. As in the previous sample, the structure was degassed and encased in Ecoflex.

Porosity Estimation of the Carbon Black/Coconut Oil Ecoflex Sensor: The porosity of the material was estimated by quantifying the area of the dark spots in the microscopic cross sections of the CB/CNO sensors using the image analysis software tool Image J, as shown in Figure S1, Supporting Information.

Conductivity Experiments on the Composite Materials: The conductivity experiments were conducted using a Keysight B1500A parameter analyzer. The samples tested were positioned on top of two copper tapes positioned 5 mm apart attached onto a glass slide. The copper tapes were utilized as contact pads for the probes instead of the oily material. The thickness and width of the samples were both 5 mm.

Experiments to Measure the Performance of the Strain Sensors: For these experiments, the authors utilized a stretch system comprising of a stepper motor and a MCDC 3006 motion controller (Faulhaber, Schönaich, Germany). Here, one representative sample from each of the two types of sensors (CB/CNO and CB/heptane) was utilized. The measurements from the sensors were recorded using a digital multimeter (34465A, Keysight, Santa Rosa, CA, USA). The sensors were stretched every 10 min by 20% to 200% to determine the gauge factor. Measurements were taken every second. The average resistance measured by the sensor for each step was calculated from the average of the measurements obtained in between 9.5 and 10 min after the previous step. This was done to ensure that a steady state was reached before the measurements were obtained. For the experiment to obtain the maximum stretchability, six devices were stretched until there was complete rupture and infinite resistance between the two ends of the sensor. In the case of the cyclic tests the stretch system was utilized to stretch the sensor to 200% and it was left at that position for 60 s; thereafter it was released to its original length and left for 60 s. The total time for the completion of a single cycle was 131 s. The process was repeated for 250 cycles and the measurements were measured every 0.5 s utilizing the Keysight multimeter. The same procedure was utilized for the durability and stability experiments. In this case a CB/CO sensor was cycled for 1000 cycles at 200% strain. For the response time and drift experiments the stretch system was utilized to stretch the sensors by a step of 100% and relaxed. As in the previous experiment the measurements were measured every 0.5 s using the Keysight multimeter. The response of the sensors to different temperatures was evaluated using an EchoTherm IC50 digital Chilling/Heating Dry Bath (Torrey Pines Scientific Inc., Carlsbad, CA, USA). The temperature of the surface of the dry bath was changed from 25 to 55 °C in steps of 5 °C. In each step the temperature was maintained for 10 min. The average measurement was calculated from measurements obtained in between 9.5 and 10 min after the previous step. This ensured that a steady state had been reached when the measurements were obtained. For the experiment to evaluate the performance of the cooled sensors, the sensor used for the durability and stability experiment was put inside a fridge at 10 °C for 1 h. Thereafter it was taken out and a cyclic test was performed on the sensor at 200% strain. For the wrist motion detection experiment, the sensor was connected to the ADC channel of a Bluesense microcontroller.[57] The change in voltage was measured using an analogue input of the microcontroller and transmitted through Bluetooth onto a computer. The data was processed and analyzed in Python. The experiments involving human subject have been performed with the full, informed consent of the volunteer.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

P.L., N.M., and J.C.C. conceived and designed the experiments; P.L. synthesized the material; P.L. and L.A.G.-G prepared the samples and the sensors; P.L. and J.C.C. performed the experiments; P.L. and J.C.C. analyzed the data; J.S. did the SEM imaging of the materials; Z.S. and P.L. conducted the wrist motion experiment and analyzed the data; N.M. provided expertise in flexible electronics; D.R. provided expertise in data analysis; P.L., J.C.C., L.A.G.-G., and A.P. wrote the paper with input from all authors; D.R. and N.M. supervised the work. The authors will like to thank Dr. Martina Costa Angeli for assisting with the experiments. The authors also thank Kalana Marasinghe for his assistance with the figures. All authors have read and agreed to the published version of the manuscript. This work was partially funded by EPSRC, GCRF, and NIHR, under the contact number: EP/R013837/1 (SmartSensOtics).

Conflict of Interest

The authors declare no conflict of interest.

Keywords

carbon black, cocos nucifera oil, eco-friendly sensors, strain sensors, wearable devices

Received: August 7, 2020
Revised: October 30, 2020
Published online: December 10, 2020

[1] Y. Liu, M. Pharr, G. A. Salvatore, *ACS Nano* **2017**, *11*, 9614.
[2] D.-H. Kim, J. A. Rogers, *Adv. Mater.* **2008**, *20*, 4887.
[3] N. Münzenrieder, G. Cantarella, C. Vogt, L. Petti, L. Büthe, G. A. Salvatore, Y. Fang, R. Andri, Y. Lam, R. Libanori, D. Widner, A. R. Studart, G. Tröster, *Adv. Electron. Mater.* **2015**, *1*, 1400038.
[4] J. Peng, I. Witting, N. Geisendorfer, M. Wang, M. Chang, A. Jakus, C. Kenel, X. Yan, R. Shah, G. J. Snyder, M. Grayson, *Nat. Commun.* **2019**, *10*, 5590.
[5] C. Wang, K. Xia, H. Wang, X. Liang, Z. Yin, Y. Zhang, *Adv. Mater.* **2019**, *31*, 1801072.
[6] M. Baumgartner, F. Hartmann, M. Drack, D. Preninger, D. Wirthl, R. Gerstmayr, L. Lehner, G. Mao, R. Pruckner, S. Demchyshyn, L. Reiter, M. Strobel, T. Stockinger, D. Schiller, S. Kimeswenger, F. Greibich, G. Buchberger, E. Bradt, S. Hild, S. Bauer, M. Kaltenbrunner, *Nat. Mater.* **2020**, *19*, 1102.
[7] H.-L. Park, Y. Lee, N. Kim, D.-G. Seo, G.-T. Go, T.-W. Lee, *Adv. Mater.* **2020**, *32*, 1903558.
[8] L.-Y. Zhou, J.-Z. Fu, Q. Gao, P. Zhao, Y. He, *Adv. Funct. Mater.* **2020**, *30*, 1906683.
[9] A. Nathan, A. Ahnood, M. T. Cole, S. Lee, Y. Suzuki, P. Hiralal, F. Bonaccorso, T. Hasan, L. Garcia-Gancedo, A. Dyadyusha, S. Haque, P. Andrew, S. Hofmann, J. Moultrie, D. Chu, A. J. Flewitt, A. C. Ferrari, M. J. Kelly, J. Robertson, G. A. J. Amaratunga, W. I. Milne, *Proc. IEEE* **2012**, *100*, 1486.
[10] Z. Jiang, M. O. G. Nayeem, K. Fukuda, S. Ding, H. Jin, T. Yokota, D. Inoue, D. Hashizume, T. Someya, *Adv. Mater.* **2019**, *31*, 1903446.
[11] G. Cantarella, C. Vogt, R. Hopf, N. Münzenrieder, P. Andrianakis, L. Petti, A. Daus, S. Knobelspies, L. Büthe, G. Tröster, G. A. Salvatore, *ACS Appl. Mater. Interfaces* **2017**, *9*, 28750.
[12] F. Spina, J. C. Costa, N. Münzenrieder, in *2019 IEEE Int. Conf. Flexible Printable Sensors Systems*, IEEE, Glasgow, UK **2019**, pp. 1–3, https://ieeexplore.ieee.org/xpl/conhome/8784101/proceeding.
[13] J. T. Muth, D. M. Vogt, R. L. Truby, Y. Mengüç, D. B. Kolesky, R. J. Wood, J. A. Lewis, *Adv. Mater.* **2014**, *26*, 6307.
[14] A. Atalay, V. Sanchez, O. Atalay, D. M. Vogt, F. Haufe, R. J. Wood, C. J. Walsh, *Adv. Mater. Technol.* **2017**, *2*, 1700136.
[15] H. Souri, D. Bhattacharyya, *ACS Appl. Mater. Interfaces* **2018**, *10*, 20845.
[16] C. Mattmann, O. Amft, H. Harms, G. Troster, F. Clemens, in *2007–11th Int. Symp. Wearable Computers*, IEEE, Boston, USA **2007**, pp. 29–36, https://ieeexplore.ieee.org/xpl/conhome/4373753/proceeding.
[17] P. Lugoda, L. A. Garcia-Garcia, S. Richoz, N. Munzenrieder, D. Roggen, in *UbiComp/ISWC '19 Adjunct: Adjunct Proceedings of the 2019 ACM Int. Joint Conf. Pervasive Ubiquitous Computing Proceedings 2019 ACM Int. Symp. Wearable Computers*, Association for Computing Machinery, London, UK **2019**, pp. 133–136, https://dl.acm.org/doi/proceedings/10.1145/3341163.
[18] M. Amjadi, Y. J. Yoon, I. Park, *Nanotechnology* **2015**, *26*, 375501.
[19] J. Lee, S. Kim, J. Lee, D. Yang, B. C. Park, S. Ryu, I. Park, *Nanoscale* **2014**, *6*, 11932.
[20] B. Huang, M. Li, T. Mei, D. McCoul, S. Qin, Z. Zhao, J. Zhao, *Sensors* **2017**, *17*, 2708.
[21] M. K. Kim, C. Kantarcigil, B. Kim, R. K. Baruah, S. Maity, Y. Park, K. Kim, S. Lee, J. B. Malandraki, S. Avlani, A. Smith, S. Sen, M. A. Alam, G. Malandraki, C. H. Lee, *Sci. Adv.* **2019**, *5*, eaay3210.
[22] Y. Cai, J. Shen, Z. Dai, X. Zang, Q. Dong, G. Guan, L.-J. Li, W. Huang, X. Dong, *Adv. Mater.* **2017**, *29*, 1606411.
[23] M. Mack, C.-H. Min, in *2019 62nd Int. Midwest Symp. Circuits and Systems*, IEEE, Dallas, USA **2019**, pp. 1195–1198, https://ieeexplore.ieee.org/xpl/conhome/8882388/proceeding.
[24] J. Ge, L. Sun, F.-R. Zhang, Y. Zhang, L.-A. Shi, H.-Y. Zhao, H.-W. Zhu, H.-L. Jiang, S.-H. Yu, *Adv. Mater.* **2015**, *28*, 722.
[25] J. C. Costa, F. Spina, P. Lugoda, L. Garcia-Garcia, D. Roggen, N. Münzenrieder, *Technologies* **2019**, *7*, 35.
[26] K. Cherenack, C. Zysset, T. Kinkeldei, N. Münzenrieder, G. Tröster, *Adv. Mater.* **2010**, *22*, 5178.
[27] J.-Y. Sun, C. Keplinger, G. M. Whitesides, Z. Suo, *Adv. Mater.* **2014**, *26*, 7608.
[28] W. Hou, N. Sheng, X. Zhang, Z. Luan, P. Qi, M. Lin, Y. Tan, Y. Xia, Y. Li, K. Sui, *Carbohydr. Polym.* **2019**, *211*, 322.
[29] J. Zhou, H. Yu, X. Xu, F. Han, G. Lubineau, *ACS Appl. Mater. Interfaces* **2017**, *9*, 4835.
[30] C. S. Boland, U. Khan, G. Ryan, S. Barwich, R. Charifou, A. Harvey, C. Backes, Z. Li, M. S. Ferreira, M. E. Mobius, R. J. Young, J. N. Coleman, *Science* **2016**, *354*, 1257.
[31] X. Guo, Y. Huang, Y. Zhao, L. Mao, L. Gao, W. Pan, Y. Zhang, P. Liu, *Smart Mater. Struct.* **2017**, *26*, 095017.
[32] X. Wang, T. Li, J. Adams, J. Yang, *J. Mater. Chem. A* **2013**, *1*, 3580.
[33] A. S. Kurian, V. B. Mohan, D. Bhattacharyya, *Sens. Actuators A* **2018**, *282*, 206.

[34] J. Ruhhammer, M. Zens, F. Goldschmidtboeing, A. Seifert, P. Woias, *Sci. Technol. Adv. Mater.* **2015**, *16*, 015003.
[35] J. Kim, W. S. Kim, *Sens. Actuators A* **2016**, *238*, 329.
[36] X. Liao, Q. Liao, Z. Zhang, X. Yan, Q. Liang, Q. Wang, M. Li, Y. Zhang, *Adv. Funct. Mater.* **2016**, *26*, 3074.
[37] Y. Liu, H. Fan, K. Li, N. Zhao, S. Chen, Y. Ma, X. Ouyang, X. Wang, *Adv. Mater. Technol.* **2019**, *4*, 1900309.
[38] J. Shintake, E. Piskarev, S. H. Jeong, D. Floreano, *Adv. Mater. Technol.* **2017**, *3*, 1700284.
[39] Y. Jiang, Z. Liu, N. Matsuhisa, D. Qi, W. R. Leow, H. Yang, J. Yu, G. Chen, Y. Liu, C. Wan, Z. Liu, X. Chen, *Adv. Mater.* **2018**, *30*, 1706589.
[40] J. Lee, S. Pyo, D.-S. Kwon, E. Jo, W. Kim, J. Kim, *Small* **2019**, *15*, 1805120.
[41] Y. Yang, L. Shi, Z. Cao, R. Wang, J. Sun, *Adv. Funct. Mater.* **2019**, *29*, 1807882.
[42] Y. Tang, Z. Zhao, H. Hu, Y. Liu, X. Wang, S. Zhou, J. Qiu, *ACS Appl. Mater. Interfaces* **2015**, *7*, 27432.
[43] J. Xu, J. Hu, Q. Li, R. Wang, W. Li, Y. Guo, Y. Zhu, F. Liu, Z. Ullah, G. Dong, Z. Zeng, L. Liu, *Small* **2017**, *13*, 1700651.
[44] S. K. Yildiz, R. Mutlu, G. Alici, *Sens. Actuators A* **2016**, *247*, 514.
[45] M. Park, J. Im, M. Shin, Y. Min, J. Park, H. Cho, S. Park, M.-B. Shim, S. Jeon, D.-Y. Chung, J. Bae, J. Park, U. Jeong, K. Kim, *Nat. Nanotechnol.* **2012**, *7*, 803.
[46] S. Gong, D. T. H. Lai, B. Su, K. J. Si, Z. Ma, L. W. Yap, P. Guo, W. Cheng, *Adv. Electron. Mater.* **2015**, *1*, 1400063.
[47] M. Amjadi, A. Pichitpajongkit, S. Lee, S. Ryu, I. Park, *ACS Nano* **2014**, *8*, 5154.
[48] T. Yamada, Y. Hayamizu, Y. Yamamoto, Y. Yomogida, A. Izadi-Najafabadi, D. N. Futaba, K. Hata, *Nat. Nanotechnol.* **2011**, *6*, 296.
[49] S. Kim, J. Lee, B. Choi, *IEEE Sens. J.* **2015**, *15*, 6077.
[50] X. Wu, Y. Han, X. Zhang, Z. Zhou, C. Lu, *Adv. Funct. Mater.* **2016**, *26*, 6246.
[51] W. Yi, Y. Wang, G. Wang, X. Tao, *Polym. Test.* **2012**, *31*, 677.
[52] D. Y. Choi, M. H. Kim, Y. S. Oh, S.-H. Jung, J. H. Jung, H. J. Sung, H. W. Lee, H. M. Lee, *ACS Appl. Mater. Interfaces* **2017**, *9*, 1770.
[53] J.-H. Kong, N.-S. Jang, S.-H. Kim, J.-M. Kim, *Carbon* **2014**, *77*, 199.
[54] L. Boateng, R. Ansong, W. Owusu, M. Steiner-Asiedu, *Ghana Med. J.* **2016**, *50*, 189.
[55] B. Matharage, M. Fernando, M. Bandara, G. A. Jayantha, C. S. Kalpage, *IEEE Trans. Dielectr. Electr. Insul.* **2013**, *20*, 887.
[56] R. Bhattacharyya, S. Sarma, T. Athauda, N. Karmakar, in *2019 IEEE Int. Conf. RFID*, IEEE, Pisa, Italy **2019**, pp. 1–6, https://ieeexplore.ieee.org/xpl/conhome/8890568/proceeding.
[57] D. Roggen, A. Pouryazdan, M. Ciliberto, in *EWSN'18 Proc. 2018 Int. Conf. Embedded Wireless Systems Networks* (Eds: D. Giustiniano, D. Koutsonikolas, A. Banchs, E. Mingozzi, K. R. Chowdhury), Junction Publishing, Madrid, Spain **2018**, pp. 177–178.
[58] K. Yamaguchi, J. J. C. Busfield, A. G. Thomas, *J. Polym. Sci., Part B: Polym. Phys.* **2003**, *41*, 2079.
[59] O. Glauser, D. Panozzo, O. Hilliges, O. Sorkine-Hornung, *ACM Trans. Graphics* **2019**, *38*, 16.
[60] Q. Lin, S. Peng, Y. Wu, J. Liu, W. Hu, M. Hassan, A. Seneviratne, C. H. Wang, in *2020 19th ACM/IEEE Int. Conf. Information Processing in Sensor Networks*, IEEE, Sydney Australia **2020**, pp. 49–60, https://ieeexplore.ieee.org/xpl/conhome/9108342/proceeding?searchWithin=lin.
[61] J. Ahmad, H. Andersson, J. Sidén, *IEEE Sens. J.* **2019**, *19*, 2055.
[62] L. Padua, D. Coraci, C. Erra, C. Pazzaglia, I. Paolasso, C. Loreti, P. Caliandro, L. D. Hobson-Webb, *Lancet Neurol.* **2016**, *15*, 1273.
[63] T. Kuroiwa, K. Fujita, A. Nimura, T. Miyamoto, T. Sasaki, A. Okawa, *J. Orthop. Surg. Res.* **2018**, *13*, 288.
[64] T. Kuroiwa, A. Nimura, Y. Takahashi, T. Sasaki, T. Koyama, A. Okawa, K. Fujita, *Sensors* **2020**, *20*, 3998.
[65] B. P. L. Lo, H. Ip, G. Yang, *IEEE Pulse* **2016**, *7*, 4.
[66] K. Chaleepa, A. Szepes, J. Ulrich, *Chem. Eng. Res. Des.* **2010**, *88*, 1217.