Emotions and Activity Recognition System Using Wearable Device Sensors

Mikhail Rumiantcev
University of Jyväskylä
Jyväskylä, Finland
mikhail.rumiantcev@jyu.fi

***Abstract*—Nowadays machines have become extremely smart, there are a lot of existing services that seemed to be unexpectable and futuristic decades or even a few years ago. However, artificial intelligence is still far from human intelligence, machines do not have feelings, consciousness, and intuition. How can we help machines to learn about human feelings and understand their needs better? People take their devices wherever they go, what can devices tell us about their owners? Personal preferences and needs are dependent on emotional and situational contexts. Therefore, emotional and activity aware gadgets would be more intuitive and provide more appropriate information to users. Contemporary wearable devices involve wide-ranging sensors. In this paper, I am going to present emotion and activity recognition approaches. The experimental recognition system elaborated during this research, enriched with sensor data collection and machine learning algorithms. It is targeted to guess how users are doing and what they are feeling. Such recognition systems can find applications in different areas such as music recommendations, personal safety or healthcare domains.**

I. Introduction

Nowadays, the world is filled with a big number of devices which provide a lot of various services. The amount of data is growing continuously, machines become more intelligent and can substitute human labor in many different areas. Machines can process large amounts of information and can do it much faster than humans. Even at the end of the last century, a Deep Blue [1] chess machine won the world champion, Garry Kasparov. In 2016 the Alpha Go [2] program won a professional board game Go player. Machine learning algorithms are constantly being improved and machines become more human-like. However, robots are still robots, they are lacking emotions, feelings, and critical thinking. At the same time, some services can process data and retrieve emotional information from various sources, for example, Cloud Natural Language API [3], which retrieves meaningful and structural information from texts, another example is IBM Watson tone analyzer, which can understand the emotions and communication style presented in the text. Along with this, there is a significant development of various sensors in IoT and wearable devices, which can provide more comprehensive information to machines and help to learn more about the environment around them.

Personal preferences and needs are always bound to situational contexts such as activity, weather, time of the year, mood and many others. If we ask someone or ourselves about music preferences, most likely the answer would be similar to this: “Depending on the mood or situation”. Recommendation services and targeted advertisement are present in most contemporary web services. Recent studies showed that a user's personality can provide valuable information and significantly improve the recommendation process. Capturing the behavioral data from social networks in conjunction with collaborating filtering can earn more about personality, emotions, and adjust recommendations to fulfill personal preferences and needs [4]. Almomani et al. [5] in their work elaborate emotional, attentional, and rational models for recommender systems. Tkalvcivc et al. [6] describe the significance of emotions in preferences, recommendations, and consumption chain.

Emotional and activity-based recommendations can have various applications, such as driving safety, music recommendation and health and wellbeing support.

In this work, I present approaches for automated emotions and activity capturing. The motivation is driven by the understanding of the significant importance of emotional and situational contexts in recommendations and the availability of wide-ranging sensors in wearable and IoT devices. General purposes of the emotions and activity capturing system are: achieve maximum personalization in recommendations, monitor emotional transitions while particular activity, adjusting recommendations for particularly emotional and situational contexts.

This research is targeted to describe which data can be gathered from sensors and how to process it and retrieve valuable information to clarify emotional and activity contexts.

The following section shows the research motivation and describes areas where emotion and activity capturing can be applied. The third section is focused on sensors and data processing approaches for activity and emotional recognition and learning personalized contexts. The fourth section illustrates the experimental prototype of the emotion and activity recognition system. The final section of this study shows the results of the research, limitations, possible improvements, and further work.

II. Research applications

Taking into account the importance of the physiological conditions and emotional and activity contexts in everyday life of people and the availability of the wide-ranging sensors in contemporary wearable devices, there is an opportunity and need for algorithms for processing such sensor-based data to assist people and provide better services. In this section I am going to review practical applications of the emotion and activity recognition system, thereby highlighting the research motivation.

1) Driving safety: Traffic accidents happen every day, many people suffer and even die on roads. A lot of accidents happen due to the destruction or tiredness of drivers. The time when cars will be fully automatic and will not require to be managed by humans has not come yet and the topic of the driver's activity monitoring remains highly important and relevant nowadays. Risk situations might be caused when people are distracted on the smartphone or other devices while driving or walking, Zaghetto et al. [7] present machine learning approaches for detection inappropriate smartphone usage in traffic. Sometimes drivers get tired or sleepy, their attention and reaction are decreasing, but they continue driving. Some people forget to use safety belts. All of these cases might lead to dismal consequences. If wearable devices would be able to detect risky situations and warn people in time or in advance, it can help to save health, lives and prevent many accident situations.

2) Music therapy and recommendations: In many cases, our music preferences depend on situational and emotional contexts. Contemporary music platforms offer easy online and offline access to large amounts of music sources. However, it leads to the problem of choice, sometimes it is hard to explore new music and we listen to the same tracks again and again. Large music services such as Spotify [8] offer pretty good music recommendations, however, they are not always adjusted to the current mood or activity context of a particular listener. Emotions and activity-aware recommendation systems would better understand the personal preferences of a user to select appropriate music tracks. Music therapy methods are popular and efficient in diseases and mental health treatments [9]. Automatic emotions recognition in wearable devices in conjunction with music therapy approaches [10], [11] can help to detect and apply preventative intervention in case of mental health problems such as depression.

3) Healthcare and wellbeing support: Innovative and cutting-edge technologies are incorporating and adopting continuously in the healthcare sector and wellbeing support. Diseases related to the cardiovascular, respiratory system, mental health, and vision and hearing problems are widespread in the world. Population aging and raising elderly demographics in many countries affect their socioeconomic in terms of growing wellbeing and healthcare needs and costs [12]. Many diseases can be properly treated or even prevented by continuous monitoring. Contemporary technologies and raised wearable devices and the popularity of smartphones allow performing distant and continuous monitoring of patients. Additionally, the usage of wearable device sensors can help to cover a wider audience and make treatment more efficient.

4) Customer experience improvement: Analyzing of the voice data while customer support calls may significantly improve the support quality. If a support system or specialist is aware of feelings and emotions of a customer at any point of the call, they can adjust recommendations or the dialog flow according and provide better services. Branding and advertisement systems with the real-time emotional and activity user data would be able to adjust their content according to it. From one side it would help to select useful content which is relevant at the moment. From another side these adjustments can help to minimize annoying adds pushing and spamming. Of course, monitoring personal data needs to be transparent for users and with appropriate consents.

III. Data processing

People take gadgets with them everywhere they go and whatever they do (see Fig. 1). Contemporary devices such as smartphones, smartwatches, tablets, and IoT devices have a wide range of sensors that can provide data about motion, rotation, ambient temperature, geolocation, atmosphere pressure, and many others. In this chapter, I want to review and investigate what kind of sensors and data from wearable devices can help to learn more about emotions and activities.

Fig. 1. Wearable devices during activities

The main objective of the emotion recognition system is to gather information from the wearable devices, preprocess it for the classification algorithms, and perform emotional and activity recognition of a user. Fig.2 illustrates the general data processing flow of the emotion and activity recognition flow. A sensor and social-based data are gathered and preprocessed at the first stage. The system should adopt gathered information to a particular context. When the data is preprocessed and machine learning models are trained features can be classified into particular labels. The final result is expected to be represented as activity and emotional labels.

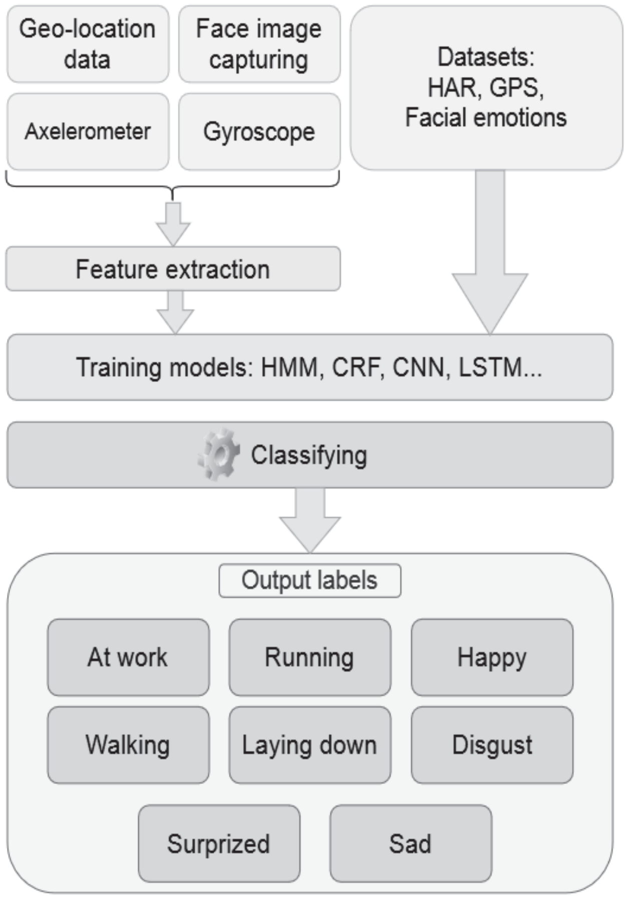

Fig. 2. Activity and emotions classification process

A. Global Positioning System

High-level activity recognition can be enforced with the data gathered from the Global Positioning System (GPS), the majority of contemporary wearable devices have geolocation detection systems. With this system, we can determine the location of the device on a map and define the movement speed. Defining which objects are located near the device we can estimate preliminary activities, for example, some public places such as cafes, cinemas, gyms, museums and many others can narrow the list of potential activities. The location-based on highways and higher movement speed rates can tell that a user is traveling in conjunction with accelerometer and gyroscope data the system can judge if a person is driving or uses public transportation.

Travelling trajectories and modes can be learned from the monitoring and processing the sequential data of the location changes [13].

Driving and traveling activity recognition might be problematic unless GPS coordinates are not captured with higher sampling rates to achieve sufficient prediction accuracy. Luo et al. [14] propose a novel algorithm based on Hidden Markov Model (HMM) which allows performing map and route matching with geolocation data capturing at lower sampling rates. Raymond et al. [15] in their work demonstrated the application of the HMM method to find the sequence of roads which correspond to certain geolocation points. Xie et al. [16] in their study designed an accurate off-line map-matching system for a city network of trajectories based on HMM.

HMM is a statistical model which is represented by sets of hidden states $\{S_1,S_2,S_3,\dots,SN\}$, where N is the number of states, these states randomly generate observations or visible states $\{o1,o2,o3,\dots,oM\}$, where M is the number of observations. Observations represent the output of the model. The process moves from one state to another making a sequence of states (Si1,Si2,…,Sik,…), where Sik is a current state at a time.

$$P(S_{ik} \mid S_{i1},S_{i2},\dots,S_{jk-1}) = P(S_{ik} \mid S_{jk-1}) \quad (1)$$

where P is a probability of each state is dependent on the previous state.

$$A(a_{ij}),\quad a_{ij} = P(S_i \mid S_j),\quad 1 \le i,j \le N \quad (2)$$

where A represents state transition probability when process or system transits from one state to another.

$$B(b_i(v_m)),\quad b_i(v_m) = P(O_m \mid S_j),\quad 1 \le v \le M \quad (3)$$

where B is an observation probability distribution.

$$\pi = (\pi_i),\quad \pi_i = = P(S_j) \quad (4)$$

where π represents the initial state probability vector.

Compact HMM notation can be represented as:

$$M = (A,B,\pi) \quad (5)$$

Therefore, geolocation points are considered as hidden states and the trajectories or road sequences are considered as observations. This method allows establishing matchings between geolocation points of the device and trajectories which can be bound in turn to a particular activity.

A discriminative relational Markov network can be used to determine and label valuable places on a map, Bayesian network can be utilized to detect transportation routings and adjust errors in real-time [17]. Liao et al. [18] in their research defined a framework of the activity recognition with extended Relational Markov Networks.

Conditional Random Fields (CRF) [19] are used to generalize activities according to the duration, time of the day, day of the week, and other attributes (see Fig. 3). It involves location data, time of the day temporal information, spatial information gathered from the geographic databases, and some global constraints representing usual places such as homes, workplaces, etc. Zhao et al. [20] applied probabilistic topic models to the various trip attributes such as destination, time of the day, traveling and destination time, duration of staying, day of the week and others. This approach helps to determine latent activity patterns and traveling purposes.

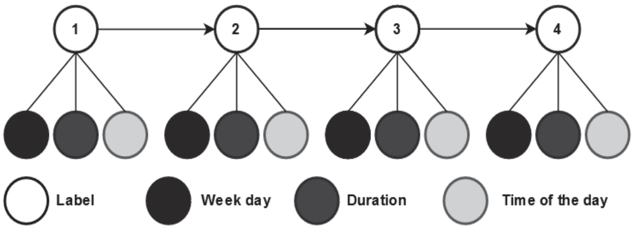

Fig. 3. Conditional Random Fields for activity recognition example

Conditional Random Fields [21] is the framework to create probabilistic models for the labeling the sequential data.

In comparison with HMM which defines joint probability P(O,S), CRF defines conditional probability P(S|O), where S defines states and O observations. HMM has access only to the current time observation while CRF has access to the observation sequence at any given time (see Fig. 4).

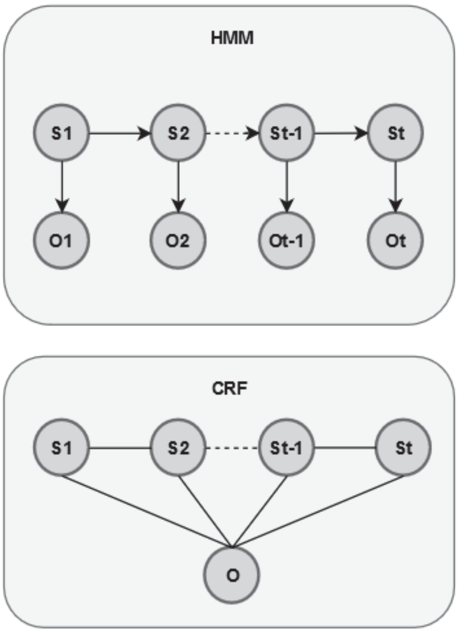

Fig. 4. HMM and CRF comparison for activity recognition example

B. Accelerometer and gyroscope

Gyroscope captures rotations around each axis in radians. The accelerometer measures the acceleration and motion of the device in a three-dimensional coordinate system. It is possible to obtain and visualize such information in graphs (see the example in Fig. 5).

Accelerometers are frequently used to determine physical activity in a health care. A lot of research has been done with respect to supervised and unsupervised machine learning models to process accelerometer data. Montoye et al. [22] compared different supervised learning models with accelerometer data and found that the random forest model had higher accuracy in comparison with others. Zhang et al. [23] in their studies developed algorithms suitable for detection and classifying particular types of activities with wrist-worn accelerometers. Statistical machine learning of sleep and physical activity phenotypes from sensor data research [24] showed that men spend more time in low and high-intensity behaviors and women spend more time in mixed behaviors. The data quality and further prediction accuracy depend much on the data sampling frequency, according to Bonomy [25] optimal data sampling frequency is between 20 and 50 Hz.

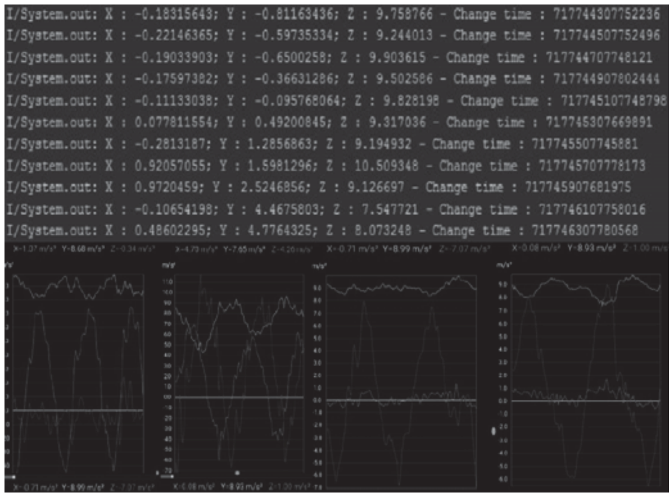

Fig. 5. Accelerometer data visualization

The idea is to classify activities based on the accelerometer data. There are available datasets [26] [27] that can be used for model creation, training, and validation. The dataset [27] is an updated version of the [26] and it has the following features: Triaxial acceleration from the accelerometer and the estimated body acceleration; triaxial angular velocity from the gyroscope; 561-feature vector with time and frequency domain variables; activity label and the subject identifier who acted. The dataset represents the result measurements of the activity observation of the 30 volunteers group from 19 to 48 years. During the experiment each participant wore a waist-attached Galaxy S2 smartphone and performed the following actions: walking, walking downstairs and upstairs, sitting, standing, laying down. Each activity has been performed twice, during first time devices were attached to the left side of the belt, while second-round participants attached the smartphones as they preferred. The available dataset is preprocessed already, preprocessing steps are: applying nose filters, subdividing the data into fixed time windows, and gravitational and body accelerations were separated using Butterworth low-pass filter [28]. For predictive modeling, the dataset can be subdivided into 70% training and 30% validation parts. Multilayer perceptron, Recurrent Neural Network (RNN) with Long Short Term Memory (LSTM), convolutional neural network CNN 2D and their hybrids such as ConvLSTM or CNN-LSTM are suitable for such kind of data prediction. According to Jindong et al. [29], research RNN and LSTM are more suitable for short activities recognition, while CNN better fits long-term repetitive activities. Random Forest Classifier

(RFC) can be employed for the labeling or classifying activities based on triaxial accelerometer data.

Alternative HAR datasets are Activity Prediction and Actitracker which represent measurements collected at laboratory-controlled and real-world conditions respectively [30], [31]. The activity classes of these datasets are: jogging, walking, sitting, walking stairs, standing and lying down.

The task of activity recognition involves a transformation of the sequential raw accelerometer data into feature vectors, which can be subdivided into activity clusters. Structured models such as HMMs and Conditional Random Fields (CRFs) can be employed for such purposes [32], [33]. Unstructured models such as Support Vector Machines (SVMs) can show equivalent performance while processing the sequential data [32].

Normally activities consist of action sequences HMM can be trained for each of the triaxial accelerometer data sequences (x,y,z), then the action probability can be calculated using the weighted sum of these values. Hierarchical HMM [34] is presented in Fig. 6.

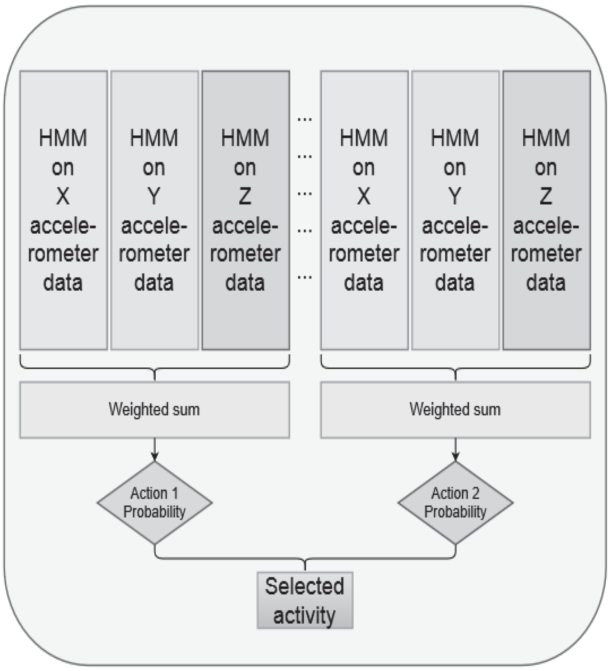

Fig. 6. HMM on accelerometer data for the activity recognition

However, in HMM an observation can depend on a state at a particular time, however in a real-world case an observation can depend on states at different times, for example, different activities might relay on similar motions and physical movements, such as physical work and sport or running in a gym and running to the bus stop. Such dependencies can be modeled with CRFs.

It makes sense to segment the captured raw triaxial accelerometer data and process it in time windows rather than calculate values for each sample. This will reduce much the computation costs for the classification process and decrease the noise uncertainties. Features that can be extracted from accelerometer data are entropy, coherence, energy, standard deviation and correlation in a time window. Then extracted features can be classified and labeled corresponding to an appropriate activity (see Fig. 7).

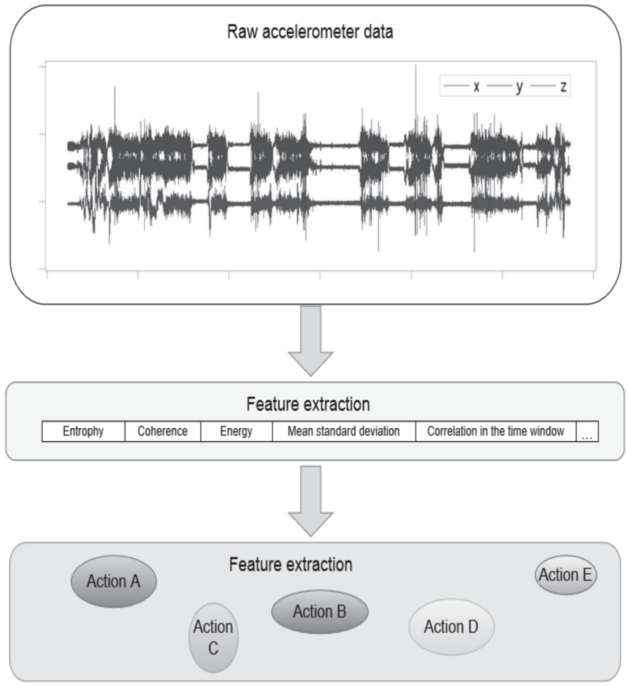

Fig. 7. Feature extraction and activity clustering from the accelerometer data

To continuously extend the knowledge of the model with input data incremental learning approaches can be employed. This will allow to continuously improve recognition performance in real-time and achieve higher personalization by tuning the model according to the particular inbound data. Siirtola and Rönning [35] presented incremental learning-based methods for human activities recognition personalization. Ntalampiras and Roveri [36] in their research proposed an incremental learning mechanism for human activity recognition, using HMMs, based on publicly available datasets of six human activities: walking downstairs, walking upstairs, walking, sitting, standing and laying. Learn++ is an incremental learning algorithm suitable for both supervised and unsupervised learning approaches, such as decision trees NNs or SVMs. Polikar and Upda [37] in their research introduced an incremental learning mechanism based on Learn++ algorithm for training neural networks such as Multilayer Perceptron (MLP), which also allows accommodating new previously unseen classes from the data.

The activity recognition accuracy of neural networks is considerably high when the sensor data is collected among a particular set of users. However, when the audience becomes wider, the accuracy decreases because human movements and motions might be different for the same activities, people have different habits and manners to walk, seat, run and other activities. Ding et al. [38] propose approaches for improvement of deep learning human activity recognition by adopting the process with transfer learning.

C. Camera

Recognition of facial expressions is essential for humans, however, could we pass the same ability to machines? Bagheri et al. [39] in their research described a model for emotion recognition from facial muscle activity. Human emotions which can be determined by facial activity can be classified into various categories, for example, sadness, disgust, neutral, anger, surprise, etc. Changes and transitions in points of facial features can be converted to vector values, Fig. 8 illustrates key points of the smiley and neutral or even sad facial emotions.

Fig. 8. Facial expression key points

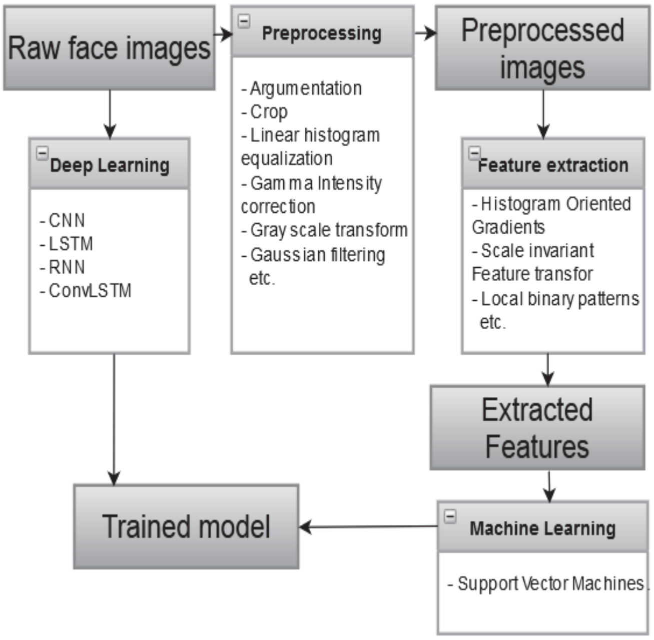

Fig. 9. Facial emotion recognitions

The initial image preprocessing step includes cropping, scaling, applying appropriate filters, and removing the background. After that, more informative parts have to be retrieved by texture, edges and color analyses such as Histogram of oriented gradients (HOG), Scale-invariant Feature Transform (SIFT), Local Binary Patterns (LBP), etc. This stage is also known as facial feature extraction when key points of a face are determined. When facial features are extracted, the model training takes place, let's have a brief look at models that are suitable for this data processing task. Multiclass Support Vector Machines (SVM) is a supervised machine learning algorithm, which performs well in image classification, however, to perform better with facial features, images should be preprocessed more properly, with consistent illumination and appropriate position of the head, which is hard to achieve in a real-world image capturing. CNN performs better in such situations because it applies segmentation of the image and processing it in chunks. Incremental learning such as RNN bases not only on the data extracted from the current image but also on the previous inputs and the data processing experience. Fig. 9 illustrates the whole image processing flow for facial emotion recognition.

IV. Experimental prototype

The working prototype has been elaborated and tailored to this research to validate the approaches described in the previous chapter. The kernel of the system is represented by the web service which feeds the data from the user wearable device application. The mobile application can work in a background mode and send notifications when the emotion or activity can be recognized based on collected data. The idea is to allow normal usage of the smartphone, without annoying surveys and filling forms, in this case, the application looks more natural and does not impose like a research tool.

The activity and emotions recognition system rely on three main data sources: geolocation data, accelerometer and image capturing. Discriminative relational and hidden Makarov models are used to perform map-matching, process latitude and longitude device point transitions and determine valuable labeled points on the map. CRFs are applied to adjust traveling recognitions with respect to the trip duration, time of the day and day of the week. RFC and CNN are selected to process the triaxial accelerometer data for activity recognition. Models are trained with publicly available datasets [26], [27], [30]. Facial emotions recognition relies mainly on image processing as described at Fig.8.

As the application captures sensor-based data from the accelerometer, gyroscope, and global positioning system, appropriate permission, appropriate permissions and consent from the user are required to be configured at application settings. From time to time the application captures images from the frontal camera, when this happens, the user is notified about it and has options to confirm, delete or view the photo which has been captured. Collected data is forwarded to the web service where it is prepared and processed with steps described in the previous section. When the current activity and emotional contexts are predicted, the classified labels are forwarded back to the mobile application, which shows a notification with an appropriate message. At this step the user has options to confirm or reject predicted values, this feedback is sent to the web service and the personalized model is adjusted according to it.

Accelerometer-based activity recognition: after the system was trained on publicly available datasets [26], [27], [30] it was trained and tested using real accelerometer data, by 2 volunteers, while each activity was repeated several times (see Table 1) during 4 days with accelerometers attached to the wrist and on a hip (in a pocket). Two types of devices were used: smartphones and smartwatches. Activities are: walking, standing, lying down, cycling, jumping, sit-ups, running, washing hands, brushing teeth, driving. During experiments, CNN showed an overall recognition accuracy of 83%, while LSTM was a bit lower 71%. CNN performs better for recognition of longer activities with repetitive motions and physical movements, such as driving, cycling, walking, running, while LSTM showed better results on recognition of shorter duration activities.

TABLE I. RESULTS ACTIVITY RECOGNITION, TRAINING MODE, ACCELEROMETER DATA

Activity	Duration per time	Repeating number	CNN recognition accuracy	LSTM recognition accuracy
walking	5 min.	20	93%	88%
standing	2 min.	20	70%	45%
lying down	2-5 min.	20	58%	64%
cycling	~3-5 min.	15	86%	75%
jumping	2 min.	15	91%	85%
sit-ups	2 min	15	90%	77%
running	3 min.	10	91%	79%
washing hands	~1 min.	30	80%	70%
brushing teeth	~1 min	8	79%	73%
driving	~10-15 min.	10	87%	55%

Geolocation based activity recognition: HMM and CRF modelling approaches were used for the rode matching on a map. The process involved the following procedures: considered geolocation transitions at a higher speed (>30km/h), selection of the road candidates located nearby the initial location point of the device (considered as initial state), candidate road selection placed near the set of location points during the trip (state space), geolocation points are considered as observations, calculation of transition and observation probabilities and linking the data to the road sequences with respect to the time of the day, trip duration, intervals between trips and days of the week. Actual trips are: 30 km 2 times per day from Monday to Friday during 3 weeks, 5 trips in a city area around 15 km, 4 trips on a high way, 200 km per each trip. Experiment showed the following outcomes: travelling activity recognition requires GPS data capturing every several minutes, more frequent better accuracy, map matching algorithms have higher performance on a highways during longer trips. The data processing and delays in internet connection might lead to problems in map matching during short distance trips.

TABLE II. RESULTS ACTIVITY RECOGNITION, REAL LIFE MODE

Activity	Emotion	Activity guessed correctly	Emotion guessed correctly	Total notifications
Driving	Neutral	11	7	14
Driving	-	8	-	10
Reading	Disgust	1	2	5
Gym	Happy	6	5	9
Walking	Happy	15	7	15
At work	Neutral	15	14	17
-	Sad	-	5	19
Music listening	Happy	20	17	20

The second stage of experiments was performed by a group of three voluntaries during the period of two weeks. Observation results are presented in the Table 2. During the observation period participants used their smartphones normally. Periodically the system pushed notification with emotional and activity predictions. From the Table 1 we can see the notifications which have been pushed. At the moment overall activity recognition accuracy is 76%, emotions prediction accuracy is 52%. Activity prediction has higher rates, which is caused by the fact that some types of activities are clearer to recognize by sensors. For example, walking or sport activities are based on the data from accelerometer. Work activity and driving is mainly based on the location system, with detecting displacement and coordinates of the work place. Information that user listen to the music is determined by the data fetched from the media players integrations. Emotions predictions are done with processing of the sound data and images captured when user places the phone in front of a face with appropriate consent notifications.

V. Conclusion and future work

Situational and emotional factors play significant roles in our lives, depending on the situation people have particular needs and preferences. Personalization and proper understanding of the user's feelings and surrounding situations allow machines to adapt their performance, select appropriate content, and provide better services in an automated manner. In this paper, I proposed an emotion and activity recognition approaches using the data from wearable device sensors. The solution of this study is targeted to maximize benefits from the gadgets using experience. By processing the data from various sensors such as accelerometer, gyroscope, GPS, and camera, the system attempts to guess the activity and the emotions of the user. At this moment the developed working prototype only notifies the user of what has been predicted. The systematic responsive feedback to the system predictions is provided by the confirmation questions which come together with notifications. Such feedback allows the system to adjust predictions, validate their accuracy, and evaluate the performance of the system. The knowledge base of the recognition system is based on model training on existing datasets. During the usage experience, predictions are tuned according to it, which gives the system a better understanding of each particular user.

The principal data collection and processing approaches for emotion and activity recognition were covered in this paper and the experimental prototype was developed. Results of a basic validation of the recognition system are presented in the previous chapter, however, it is a trial working model, which requires a lot of further elaboration and testing on a wider audience. The accuracy sensor is varying on different device types, the solution needs to be tested on different devices, which will require further adjustments and tunings of the service. Further elaborations for healthcare applications, such as health and wellbeing support and monitoring systems, will require incorporation of the additional data and features to the system such as heart rate and blood pressure sensors. It will also require additional research and collaboration with medical specialists. In further developments the proposed solution can be integrated with recommendation systems to select content suitable for the appropriate situational contexts, it can be used for advertising or music playlist creation systems.

References

[1] Deep Blue - computer developed by IBM, programmed with playing logic of chess game, Web: https://www.ibm.com/ibm/history/ibm100/us/en/icons/deepblue/

[2] Alpha Go playing program for the Chinese traditional game, Web: https://deepmind.com/research/case-studies/alphago-the-story-so-far.

[3] Cloud Natural Language API. Web: https://cloud.google.com/natural-language.

[4] Moscato, Vincenzo & Picariello, Antonio & Sperli, Giancarlo. (2020). An emotional recommender system for music. IEEE Intelligent Systems. pp. 1-1.

[5] Almomani, Ameed & Monreal, Cristina & Sieira, Jorge & Graña, Juan & Sánchez, Eduardo. (2020). Rational, emotional, and attentional models for recommender systems. Expert Systems.

[6] Tkalcivc, Marko & Kosir, Andrej & Tasivc, Jurij & Kunaver, Matevž. (2011). Affective recommender systems: the role of emotions in recommender systems, pp. 9-13.

[7] Zaghetto, Alexandre & da Silva, Mateus Mendelson & Zaghetto, Cauê & Vidal, Flavio. (2019). Detection of Inappropriate Use of Smartphones in Traffic Using Artificial Neural Networks.

[8] Spotify music platform, Web: http://www.spotify.com

[9] Erkkilä, Jaakko & Punkanen, Marko & Fachner, Jörg & Ala-Ruona, Esa & Pöntiö, Inga & Tervaniemi, Mari & Vanhala, Mauno & Gold, Christian. (2011). Individual music therapy for depression – Randomised Controlled Trial. *The British journal of psychiatry: the journal of mental science*

[10] Gold, Christian & Saarikallio, Suvi & Crooke, Alexander & Mcferran, Katrina. (2017). Group Music Therapy as a Preventive Intervention for Young People at Risk: Cluster-Randomized Trial. *Journal of Music Therapy*, pp. 54.

[11] Erkkilä, Jaakko & Brabant, Olivier & Saarikallio, Suvi & Ala-Ruona, Esa & Hartmann, Martin & Letule, Nerdinga & Geretsegger, Monika & Gold, Christian. (2019). Enhancing the efficacy of integrative improvisational music therapy in the treatment of depression: Study protocol for a randomized controlled trial

[12] Al-khafajiy, Mohammed & Baker, Thar & Chalmers, Carl & Asim, Muhammad & Kolivand, Hoshang & Fahim, Muhammad & Waraich, Atif. (2019). Remote health monitoring of elderly through wearable sensors. Multimedia Tools and Applications. 78. pp. 1-26.

[13] Xiao, Guangnian & Cheng, Qin & Zhang, Chunqin. (2019). Detecting travel modes from smartphone-based travel surveys with continuous hidden Markov models. *International Journal of Distributed Sensor Networks*, pp. 15.

[14] Luo, An & Chen, Shenghua & Xv, Bin. (2017). Enhanced Map-Matching Algorithm with a Hidden Markov Model for Mobile Phone Positioning. ISPRS International Journal of Geo-Information.

[15] Raymond, Rudy & Morimura, Tetsuro & Osogami, Takayuki & Hirosue, N. (2012). Map matching with Hidden Markov Model on sampled road network. *Proceedings - International Conference on Pattern Recognition.*

[16] Xie, Yan & Zhou, Kai & Miao, Fang & Zhang, Qian. (2020). High-Accuracy Off-Line Map-Matching of Trajectory Network Division Based on Weight Adaptation HMM. *IEEE Access*. pp. 1-1.

[17] Liao, Lin & Patterson, Donald & Fox, Dieter & Kautz, Henry. (2007). Building Personal Maps from GPS Data. *Annals of the New York Academy of Sciences*. 1093. pp. 249-65.

[18] Liao, Lin & Fox, Dieter & Kautz, Henry. (2005). Location-Based Activity Recognition using Relational Markov Networks. *IJCAI International Joint Conference on Artificial Intelligence*. Pp.773-778.

[19] Liao, Lin & Fox, Dieter & Kautz, Henry. (2007). Extracting Places and Activities from GPS Traces Using Hierarchical Conditional Random Fields. I. J. Robotic Res. 26, pp. 119-134

[20] Zhao, Zhan & Koutsopoulos, Haris & Zhao, Jinhua. (2018). Discovering Latent Activity Patterns from Human Mobility

[21] Lafferty, John & Mccallum, Andrew & Pereira, Fernando. (2001). Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data. *Proceedings of the Eighteenth International Conference on Machine Learning*, pp. 282-289

[22] Montoye, Alexander & Westgate, Bradford & Fonley, Morgan & Pfeiffer, Karin. (2018). Cross-validation and out-of-sample testing of physical activity intensity predictions with a wrist-worn accelerometer. *Journal of Applied Physiology*, 124.

[23] Zhang, Shaoyan & Rowlands, Alex & Murray, Peter & Hurst, Tina. (2011). Physical Activity Classification Using the GENEA Wrist-Worn Accelerometer. *Medicine and science in sports and exercise*.

[24] Willetts, Matthew & Hollowell, Sven & Aslett, Louis & Holmes, Chris & Doherty, Aiden. (2018). Statistical machine learning of sleep and physical activity phenotypes from sensor data in 96,220 UK Biobank participants. *Scientific Reports*. pp. 8.

[25] Alberto G Bonomi. 'Physical activity recognition using a wearable accelerometer'. *Sensing Emotions. Springer*, 2010, pp. 41–51

[26] Human Activity Recognition Using Smartphones Data Set, Web: https://archive.ics.uci.edu/ml/datasets/human+activity+recognition+using+smartphones

[27] Smartphone-Based Recognition of Human Activities and Postural Transitions Data Set. Web: http://archive.ics.uci.edu/ml/datasets/SmartphoneBased+Recognition+of+Human+Activities+and+Postural+Transitions

[28] Butterworth Filter. Web: https://www.electronicshub.org/butterworth-filter/

[29] Wang, Jindong & Chen, Yiqiang & Hao, Shuji & Peng, Xiaohui & Lisha, Hu. (2017). Deep Learning for Sensor-based Activity Recognition: A Survey. Pattern Recognition Letters, pp. 119.

[30] Wisdm activity prediction dataset, web: https://www.cis.fordham.edu/wisdm/dataset.php

[31] Zhang, Mi and A. Sawchuk. "USC-HAD: a daily activity dataset for ubiquitous activity recognition using wearable sensors. (2012) *UbiComp*.

[32] Twomey, Niall & Diethe, Tom & Fafoutis, Xenofon & Elsts, Atis & McConville, Ryan & Flach, Peter & Craddock, I.J. (2018). A Comprehensive Study of Activity Recognition Using Accelerometers.

[33] Garcia Ceja, Enrique & Brena, Ramon & carrasco, carlos & Garrido, Leonardo. (2014). Long-Term Activity Recognition from Wristwatch Accelerometer Data. Sensors. pp. 14.

[34] Lee, Young-Seol & Cho, Sung-Bae. (2011). Activity Recognition Using Hierarchical Hidden Markov Models on a Smartphone with 3D Accelerometer. *Hybrid Artificial Intelligence*.

[35] Siirtola, Pekka & Röning, Juha. (2019). Incremental Learning to Personalize Human Activity Recognition Models: The Importance of Human AI Collaboration. Sensors.

[36] Ntalampiras, Stavros & Roveri, Manuel. (2016). An incremental learning mechanism for human activity recognition, pp. 1-6

[37] Polikar, Robi & Upda, L. & Upda, S.S. & Honavar, Vasant. (2001). Learn++: An incremental learning algorithm for supervised neural networks. Systems, Man, and Cybernetics, Part C: Applications and Reviews, *IEEE Transactions on. 31*, pp. 497 - 508.

[38] Ding, Renjie & Xue, Li & Nie, Lanshun & Li, Jiazhen & Si, Xiandong & Chu, Dianhui & Liu, Guozhong & Zhan, Dechen. (2018). Empirical Study and Improvement on Deep Transfer Learning for Human Activity Recognition. Sensors.

[39] Bagheri, Elahe & Bagheri, Azam & Gomez Esteban, Pablo & Vanderborgth, Bram. (2020). A Novel Model for Emotion Detection from Facial Muscles Activity.