VoiceCare: A Voice-Interactive Cognitive Assistant on a Smartwatch for Monitoring and Assisting Daily Healthcare Activities

Sirat Samyoun*, *Student Member, IEEE*, and John Stankovic, *Fellow, IEEE*

***Abstract*— Following several health activities in daily life (e.g., medication/exercise plans, handwashing, physiological monitoring) properly often requires monitoring and assistance support. Although the emergent smartwatch and wearable technologies have opened great opportunities to monitor these activities in the wild, existing smartwatch-based systems do not interactively guide the user and also lacks comprehensiveness to provide knowledge related to this set of daily activities. To overcome these limitations of the state-of-the-art, we present *VoiceCare*, a wearable cognitive assistant on a smartwatch for daily life healthcare that interactively assists the user and monitors adherence to these activities. We conduct an extensive user study and thorough evaluation on collected data to show that *VoiceCare* effectively brings multiple health sensing solutions into one system while operating with minimal resource usage and helps improve the user conformance to daily health.**

I. Introduction

Healthy individuals as well as patients have to correctly follow several daily health activities (e.g., medications, exercise plans) at home to ensure proper adherence and improved health outcomes. Approximately $100 billion annual healthcare cost is incurred worldwide due to medication non-adherence alone [1]. According to studies [1][2], the factors that lead to daily non-adherence include users' feelings of helplessness, insufficient instructions, forgetfulness, etc. Additionally, since the COVID-19 pandemic, health experts have recommended several daily activities, such as, monitoring physiological vitals, washing hands frequently, and being aware of the daily pandemic situation. However, in many situations, such activities require interactive assistance or background knowledge. For example, people often forget detailed instructions of an exercise, and feel helpless without a physical therapist [2]. Another study shows that people often miss crucial hand areas while washing without knowing it [3]. Therefore, a *cognitive assistance* that can interactively assist the user and provide related knowledge can greatly help to perform required activities correctly.

Recently, smartwatches have emerged as a subtle way to monitor health activities, including COVID-19 symptoms [4]. Such wearables have significant advantages over the state-of-the-art systems that predominately use a video camera or a smartphone [1][5][6][7]. For example, monitoring a pill-intake or a handwashing event using a video camera at home is privacy-invasive and not ubiquitous. Moreover, compared to the smartphones, modern smartwatches offer sophisticated health sensing capabilities (e.g., heart-rate, skin conductance, temperature) and are easier to wear besides being unobtrusive and privacy-preserving.

Despite these significant benefits, the existing smartwatch-based systems [4][8][9][10][11] are not comprehensive enough for providing interactive assistance and related knowledge for a range of health activities. There are several reasons why. *First*, conducting an user interaction is difficult on a smartwatch with any visual input and output due to the small form factor. Moreover, none of these systems have tried to utilize the dialogue exchange capabilities (e.g., speaker, microphone) to facilitate interaction with the user while monitoring the activity. *Second*, although voice assistants (e.g, Siri, Alexa) have dialogue exchange capabilities, such systems do not provide any detailed knowledge or assistance specific to these activities (e.g, which areas of hand weren't cleaned properly or information about a prescribed medication). *Third*, smartwatches have very limited resources (e.g, battery, processing power). So, any system that consumes minimal resources besides supporting this comprehensive set of activities is preferred rather than the existing works that provides separate systems for each activity.

In this paper, we take the first step to overcome these notable limitations. We develop *VoiceCare*, a voice-based cognitive assistant on a smartwatch that integrates an intelligent dialogue exchange system and efficiently utilizes the smartwatch sensors, along with a knowledge-base of health activities and a remote adherence tracker and thereby moves the state-of-the-art toward a more comprehensive healthcare solution. The key contributions of this work are threefold: **First**, we build *VoiceCare*, a cognitive assistant on a smartwatch that demonstrates that we can provide a set of healthcare services (which can easily be extended towards more and more comprehensiveness) for the new normal life by reminding, monitoring adherence, and assisting the user interactively. **Second**, we integrate a knowledge-base related to the health activities and build a dialogue exchange system on the smartwatch for effectively interacting with the user, and also build a cloud based system for remotely monitoring user adherence. **Third**, we conduct a user study and collect data from real-life deployments using *VoiceCare*. Our extensive evaluation and usage analysis shows that *VoiceCare* effectively brings multiple health sensing solutions into one system while guaranteeing minimal resource usage and helps improve the user conformance to daily health.

II. Architecture and Methodology of *VoiceCare*

VoiceCare is a smartwatch-based cognitive assistant that

All the authors are with the Department of Computer Science, University of Virginia, Charlottesville, VA 22903, USA.

*Corresponding author: Sirat Samyoun (`ss8hf@virginia.edu`)

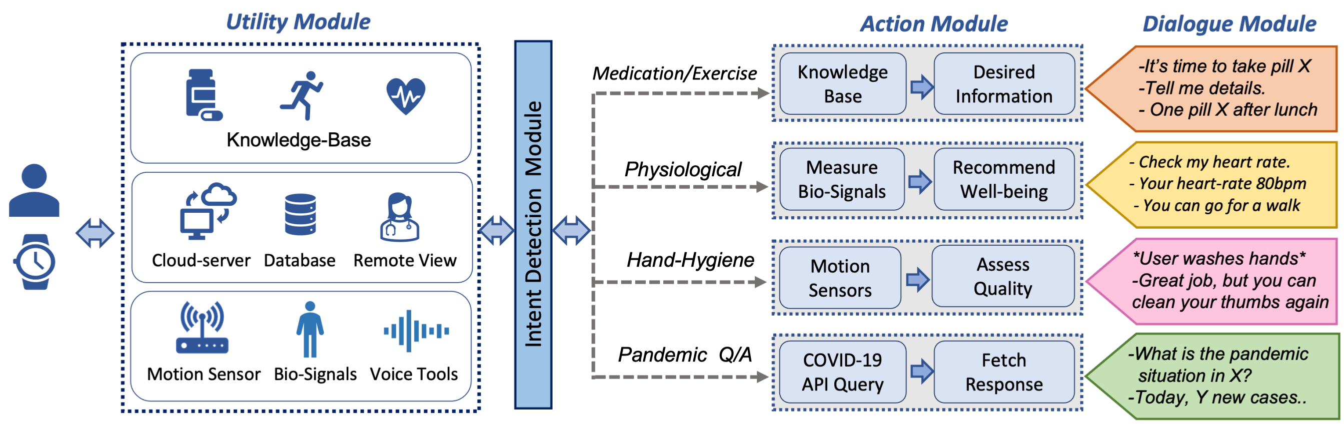

Fig. 1. Architecture of the cognitive assistant *VoiceCare* that runs on the user smartwatch, and the working flow of the four modules.

provides interactive assistance for a set of healthcare activities that are recommended for daily life well-being, such as, prescribed medication and exercise assistance, physiological monitoring, hand hygiene assistance, and pandemic related query answering. The architecture of *VoiceCare* (Fig. 1) consists of 4 functional modules. We describe these modules and the working flow of *VoiceCare* below:

1) Intent Detection Module: This module engages the users in voice interaction while reminding them for an activity. The user can also initiate an interaction from the smartwatch display. The module takes voice input from the user, and converts the speech to text. Next it extracts different entities from the utterance to understand the intent of the user, specifically, the context c, response type r, polarity p and an optional entity t. We provide examples from different interactive responses below:

I have taken it ⇒ *c=pill,r=confirmed, p=positive*
Tell me details ⇒ *c=exercise, r=details, p=positive*
Remind me later ⇒ *c=handwash, r=later, p=positive*
What is the pandemic situation in X today? ⇒ *c=pandemic, r=details, p=positive, t=X*

Next, it maps the extracted entities into a pre-installed knowledge-base, which provides a set of actions for the entities. Specifically, given the tuple $e = < c, r, t >$, it returns the row $KB(e)$ that provides maximum number of entity matching among all entries (E). Moreover, throughout the interaction, if the values {*i=confirmed, p=positive*}, are obtained, it marks the user adherence to the activity as positive. Overall, the possible *actions* are chosen by:

$$actions(c, r, t) = \underset{e \in E}{\text{argmax}}\, KB(e)$$

2) Action Module: This module executes the actions obtained in the earlier step, and produces the results. Overall, there are 4 categories of actions providing the assistance:

- *Medication and exercise assistance*: This module parses different information from a prescription related to a medication or exercise, and stores into the knowledge-base information, such as, *name* (e.g, Apixaban), *prescribed time* (e.g., 3pm), *constraint* (e.g., before breakfast), *quantity* (e.g., repetition count), *purpose* (e.g., diabetics). When the user asks for a specific information, it is provided to the user after searching the knowledge-base. For example:

  "Take *one* pill of *Apixaban after breakfast*"
  "To do *shoulder flexion*, *raise your arm beyond..*"
- *Physiological monitoring*: When the user opts for checking a physiological vital, this module activates the specific bio-signal sensor, takes the readings, and verbally provides the information to the user. Additionally, it passes different recommendations for physical well-being to the user from time-to-time, which are previously stored in the knowledge-base. For example:

  "Your *heart-rate* is *80 bpm*"
  "You can *go out for a walk* or *do some yoga*"
- *Hand-hygiene assistance*: *VoiceCare* periodically reminds the user to wash hands, and provides feedback to the user about the handwashing quality. Specifically, a proper handwash consists of 8 steps from the WHO guidelines performed for at least 20 seconds. It collects data from the motion sensors of the smartwatch, and applies the hand-wash model adopted from [9]. The model is a hybrid CNN-RNN based network that infers the handwash steps present in the data with high accuracy. It also applies a model compression technique to make the model lightweight for running on devices with low resources. Based on the results (e.g., missing steps, proper hand-wash), a verbal feedback message is generated for the user. For example:

  "Congratulations, you washed our hands perfectly"
  "Great job, but you can *clean your thumbs* again"
- *Pandemic related question answering*: When the user intent matches a pandemic related query, this module first identifies the name of the locality (e.g., country, state) from the optional entity t. Next, it makes a HTTP request to a remote web API [12] that provides the COVID-19 updates in a locality, specifically, the *new cases count, total cases count, and death count*. Upon getting the API response, the data is parsed and a message is generated. For example:

  "Today, there are *X* new cases, *Y* total cases,.., in *W*"

3) Dialogue module: This module generates a feedback message for the user based on the results of the *action* module, as per the aforementioned examples for each category.

Fig. 2. *VoiceCare* Apple watch app interfaces during a voice interaction

It then converts the text to speech and passes it to the user.

4) Utility module: This module integrates and serves different utility components and services for the functionalities of the other modules. First, it contains the knowledge-base that stores the associated information of each category. Second, it includes a cloud-based remote monitoring module where it uploads the information from a user interaction, specifically, *time of interaction, assistance type, interaction summary, adherence, and smartwatch battery level*. Third, it utilizes different embedded hardware and software tools as necessary, such as, motion sensors (e.g., accelerometer, gyroscope, magnetometer), bio-signals (e.g., heart-rate, skin conductance, temperature), voice framework tools (e.g., speech to text, text to speech).

III. Evaluation

A. User study and experimental setup

1) Implementation of VoiceCare: We implemented *VoiceCare* using an Apple watch series 4 app in combination with an Amazon S3 cloud, and a MySQL database, and a website for remote monitoring. The *VoiceCare* modules were implemented and installed into the WatchOS app. The app engages the user in voice interaction (Fig. 2) and uploads the interaction summary to the cloud. We implemented a Node.js based server in the cloud. It processes the data and stores it into the database by a unique user identifier without storing any HIPAA identifiers. We also implemented a HTML website which fetches the user data from the server by querying with the identifier. Next, it analyses the data and presents a summary which allows the healthcare provider to remotely monitor the adherence of the user.

2) Dataset collection: We designed and conducted a study protocol in real-life setting using the *VoiceCare* system. The study was approved by the University of Virginia IRB HSR Board (HSR#21012). It included 11 subjects, with 6 female and 5 female persons, all healthy individuals, aged 26-39 years (average 30 years). Each subject used the system for around 2 weeks, and the overall study took around 6 months. For each subject, we generated a daily schedule of health activities consisting of the 4 assistance categories which included 8 daily voice-interactive reminders, 2 for each category. Additionally, 4 daily non-voice reminders were provided, each from one category for evaluation purpose. The subjects having an active medication or exercise plan had the corresponding reminders included along with all necessary background integrated into *VoiceCare*, which were verified by a clinician. In total, 1848 reminders were provided, and responses were collected from the cloud platform.

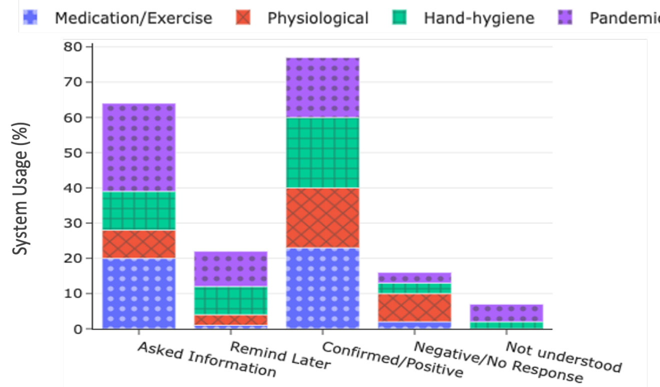

Fig. 3. *VoiceCare* usage analysis from different kinds of user response

B. Deployment Results and Performance Analysis

1) Usage analysis and findings: We analyze the user response during the *VoiceCare* deployments, and present the summary in Fig. 3. The response categories are as follows:

- *Asked information*: User asked for details regarding a medication/exercise, or a physiological vital, or handwashing guidelines, or the pandemic situation.
- *Remind later*: User asked to interact at a later time.
- *Confirmed/positive response*: User confirmed adhering or gave a positive response regarding the activity.
- *Negative/no response*: User did not interact with *VoiceCare* after a reminder, or provided a negative response.
- *Not understood*: User response was not understood.

Our analysis suggest that the users prefer to ask information when the activity steps are difficult to remember or involves knowledge (e.g., medication details or steps of an arm exercise or pandemic update). Moreover, the user preferred to interact *VoiceCare* later rather than instantly at only 5.4% usage time on average. This is mostly for hand washing (10% usage time), as this activity involves some preparation. We also note that the rate of positive response or confirmation of adherence across all activities is high (77% usage time overall), which shows the positive impact of *VoiceCare* among users. On the contrary, the users responded negatively or provided no response at only 4% usage time on average. Among these, users were most reluctant about monitoring the physiological vitals (8% usage time). Moreover, the responses not understood by *VoiceCare*(1.75% usage time only) were mostly for a pandemic update (5% usage time), which is dependent on external APIs, and also due to environmental noise during transcription. The users mostly asked for simple information for medication or exercise (e.g., steps of an exercise, purpose of a pill), which were understood by *VoiceCare*.

2) Evaluation of user conformance: We evaluated the user conformance to the health activities in *VoiceCare* compared to that from the non-voice interactions provided during

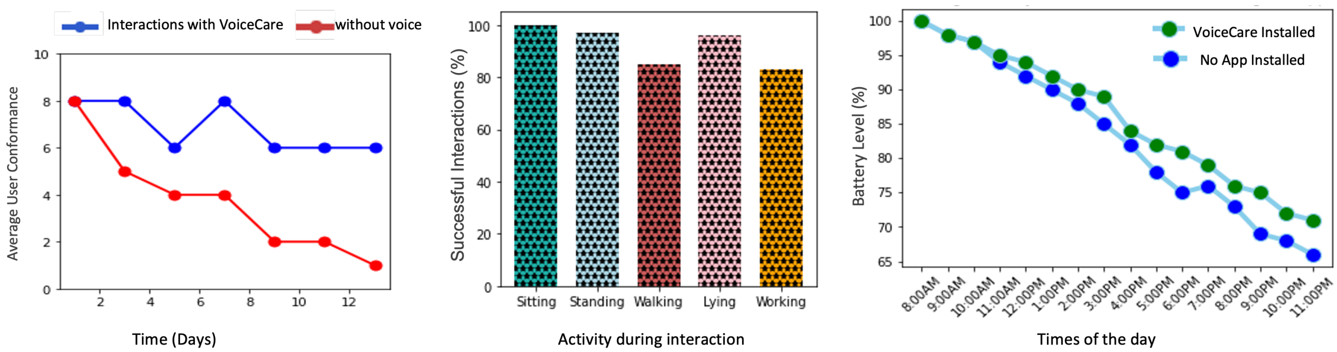

Fig. 4. (a) User conformance evaluation in *VoiceCare*.(b) Performance of interactions during other activities. (c) Battery usage analysis of *VoiceCare*.

the study. User conformance was measured by the user responsiveness to the reminders which resulted in successful user interactions. We present the average results for 2 weeks across all participants in Fig. 4(a). We observe that although user conformance gradually decreases over time, it decreased remarkably without voice interaction for all subjects, as the users had to choose the actions manually on the smartwatch screen. Compared to this, *VoiceCare* causes minimal reduction of user conformance over the time. This also shows great potential of using voice interaction on a smartwatch to improve user conformance over non-voice interactions.

3) Evaluation of interaction during daily activities: An interactive reminder might arrive in a situation where the user is engaged to some other activities. To test the performance of *VoiceCare* in such situations, we provided interactive reminders to a subject while doing different household and outdoor activities, such as sitting, standing, walking, lying, and working with a laptop. Overall, 75 interactions (15 per activity) were performed between the user and *VoiceCare*, and the successful speech-to-text transcriptions were counted from the cloud server. The results, presented in Fig. 4(b) suggests that voice interaction works perfectly while doing steady activities (e.g., sitting, standing), while activities like walking, working slightly hampers the interaction.

4) Evaluation of battery usage: We record the average battery levels of the smartwatch at different times of day for 2 weeks without *VoiceCare* installed on it, and compared to that with *VoiceCare* from the deployment study. The results, presented in Fig. 4(c) shows that with *VoiceCare*, the average battery level drops to 65% at night, where without *VoiceCare*, it drops to around 72%. Our intuition behind this minimal battery usage by *VoiceCare* is that it intelligently uses reminders to interact with users and remains idle otherwise. Furthermore, to preserve battery, it applies a model compression technique on the heavyweight handwashing model, and also communicates with the cloud server periodically instead of doing it regularly.

IV. Conclusion and Future Works

Overall, *VoiceCare* shows that utilizing voice interaction that provides clarifying and assisting information on a smartwatch in combination with the health sensing capabilities can potentially improve daily life healthcare. Additionally, our work provides a scalable and generalized platform which opens the possibility to add further capabilities. Such capabilities can be tested among larger population, specially the elder population who require in-home monitoring support. For example, a possible extension of the exercise assistance part can be detecting the quality of a performed exercise by utilizing the motion sensors and providing guidance to the user. Our results also show that such a cognitive assistant can bring multiple health sensing solutions into one system, can help improve the user conformance to daily health.

References

[1] L. Dayer, S. Heldenbrand, P. Anderson, P. O. Gubbins, and B. C. Martin, Smartphone medication adherence apps: potential benefits to patients and providers, J. of the American Pharmacists Association, vol. 53, no. 2, pp. 172–181, 2013.

[2] F. Sheehan, Patient Self Monitoring to Transfer Physical Therapy Exercise from Clinic to Home - Final Report. (Prepared by the University of Washington under Grant No. R21 HS021733). Rockville, MD: Agency for Healthcare Research and Quality, 2014.

[3] J. S. Wong and J. Lee, The common missed handwashing instances and areas after 15 years of hand-hygiene education, J. of environmental and public health, 2019.

[4] T. Mishra, M. Wang, A. A. Metwally, G. K. Bogu, A. W. Brooks, A. Bahmani, A. Alavi, A. Celli, E. Higgs, O. Dagan-Rosenfeld et al., Pre-symptomatic detection of covid-19 from smartwatch data, Nature biomedical engineering, vol. 4, no. 12, pp. 1208–1220, 2020.

[5] G.-A. Bilodeau and S. Ammouri, Monitoring of medication intake using a camera system, J. of medical systems, vol. 35, no. 3, pp. 377–389, 2011.

[6] J. P. Higgins, Smartphone applications for patients' health and fitness, The American J. of medicine, vol. 129, no. 1, pp. 11–19, 2016.

[7] D. F. Llorca, I. Parra, M. A. Sotelo, and G. Lacey, A vision-based system for automatic hand washing quality assessment, Machine Vision and Applications, vol. 22, no. 2, pp. 219–234, 2011.

[8] A. Henriksen, M. H. Mikalsen, A. Z. Woldaregay, M. Muzny, G. Hartvigsen, L. A. Hopstock, and S. Grimsgaard, Using fitness trackers and smartwatches to measure physical activity in research: analysis of consumer wrist-worn wearables, J. of medical Internet research, vol. 20, no. 3, p. e9157, 2018.

[9] S. Samyoun, S. S. Shubha, M. A. S. Mondol, and J. A. Stankovic, iwash: A smartwatch handwashing quality assessment and reminder system with real-time feedback in the context of infectious disease, Smart Health, vol. 19, p. 100171, 2021.

[10] H. Kalantarian, N. Alshurafa, E. Nemati, T. Le, and M. Sarrafzadeh, A smartwatch-based medication adherence system, in 2015 IEEE 12th Int. Conf. on Wearable and Implantable Body Sensor Networks (BSN). IEEE, 2015, pp. 1–6.

[11] S. Samyoun, M. A. S. Mondol, I. A. Emi, and J. A. Stankovic, iadhere: A voice interactive assistant to improve adherence to medical treatments: demo abstract, in Proc. of the 10th ACM/IEEE Int. Conf. on Cyber-Physical Systems, 2019, pp. 334–335.

[12] The COVID Tracking Project, 2020, https://covidtracking.com/.