Activity Recognition and Explanations for Cancer Health Awareness

Hayley Borck[1,*], Jack Ladwig[1], Joseph B. Mueller[1], Steven Johnston[1], Helen Wauck[1], Ruta Wheelock[1] and Richard G. Freedman[1]

[1]SIFT, 319 1st Ave N, Minneapolis, 55401, USA

Abstract

Cancer patients' activity level and performance is difficult to assess outside the clinical setting. Patients are often not able to communicate subtleties in activity performance that may indicate secondary concerns or true pain levels. There is a need for doctors to monitor patient activity performance at home, especially to identify performance anomalies that may require the doctor's intervention. Many black box activity classification algorithms lack the specificity and succinctness required to alert doctors of degrading health issues, such as suddenly requiring a cane to walk. Additionally, traditional black box classification systems lack the ability to deliver personalized information on activity performance. For instance, in the previous example a doctor does not need to know if a patient who already uses a cane is using a cane during their exercises. By combining deep learning models and symbolic reasoning we have created the Characterizing Human Activities for Cancer Health Awareness (CHA-CHA) system to classify exercises performed at home and alert the doctor of patient specific anomalies in their performance.

Keywords

Case Based Reasoning, Activity Recognition, Health

1. Introduction

All physicians strive to maintain the quality of life (QoL) for their patients. However, cancer patients specifically are often elderly and therefore face additional health risks. A simple way to maintain QoL and reduce other health risks is to remain physically active. A key approach to reduce health risks and preserve/maintain QoL is to remain physically active which is vital to elderly cancer patients in particular as they are at greater risk of being sent to a care facility, drastically reducing their QoL. Unfortunately, a patients activity level and performance is difficult to assess outside the clinical setting. Patients are often not able to communicate subtleties in activity performance that may indicate secondary concerns. This disconnection prevents physicians from providing updated care and monitoring the patients evolving health state. The Characterizing Human Activities for Cancer Health Awareness (CHA-CHA) system recognizes physical activities and activity performance and translates that information into a human-interpretable explanation for physicians to monitor patients while at home and/or between appointments.

ICCBR XCBR WORKSHOP at ICCBR-2022, September, 2022, Nancy, France

*Corresponding author.

hborck@sift.net (H. Borck)

CHA-CHA recognizes activity and performance from video taken on a smart phone. Our system reasons over symbolic information to recognize activities making our system interpretable end-to-end. We enlisted Subject Matter Experts (SMEs) to ensure activities and performance parameters are directly relevant to the cancer health domain and evaluated the explanations provided to the SMEs with a series of usability tests. Preliminary results show the classification accuracy of our system is comparable to state-of-the-art black box systems. CHA-CHA uses the ML-detected objects and skeletal frame to create high-level posture features via Qualitative Spatial Reasoning. A Case-Based Reasoning (CBR) algorithm is then used to classify the activity and activity performance. We currently have a CHA-CHA working prototype, including a physician facing web application.

The rest of the paper is presented as follows: Section 8 discusses previous work in this area. Section 2 provides an overview of the system architecture and how the raw data is processed into the CHA-CHA system. Sections 3 and 4 discuss how we derive the qualitative features. In Section 5 the CBR activity classification and explanation algorithm is presented and Section 6 shows the preliminary results. Section 7 shows the prototype physician interface and discusses SME feedback. Finally we present our conclusions in Section 9.

2. Information Pipeline

The CHA-CHA architecture (Fig. 1) begins after the input video data has been processed. Rather than rely on recognizing activities from the firehose of pixel data that lacks human-interpetable semantics, off-the-shelf software interprets context from the video that becomes the actual input for CHA-CHA. The context that is used for CHA-CHA's activity recognition are patient's posture and the objects in the scene.

Object and pose detection tasks have been well studied in the area of artificial intelligence computer vision, and we took advantage of that research to focus on the qualitative spatial reasoning unique to CHA-CHA. Object detection algorithms record the objects in a picture through a collection of labeled rectangles; the label is the object's name and the rectangle bounds the region of the image where the object is located. We are currently using an Off The Shelf (OTS) version of Darknet YOLOv4 with pre-trained weights (trained on the MS COCO dataset) from [14] for object detection which is both open source and has strong accuracy. Pose detection algorithms provide a stick Fig. representation of people in an image through a collection of labeled points, each representing a joint in the human body (Fig.s 2 and 3). We group these joints into specific pairs that form links in the human body, such as upper/lower appendages, hips, etc. Depending on the choice of off-the-shelf software we use for each of these tasks, the list of recognizable objects and available joints will differ (like people, a machine cannot identify something that it has never seen before). This will impact CHA-CHA's implementation in the future, requiring manual adjustments in the feature detection algorithm. We use the BlazePose [15] for pose detection, which is a Google TFLite model able to give points in 3D and 'light' enough to run on a current generation smart phone. BlazePose accomplishes pose estimation within the two-dimensional image by approximating their relative position in three-dimensional space. In contrast, Darknet YOLOv4 has intensive computational requirements to run on a computer's CPU, therefore, we use Darknet YOLOv4 run on a GPU as a smart phone typically

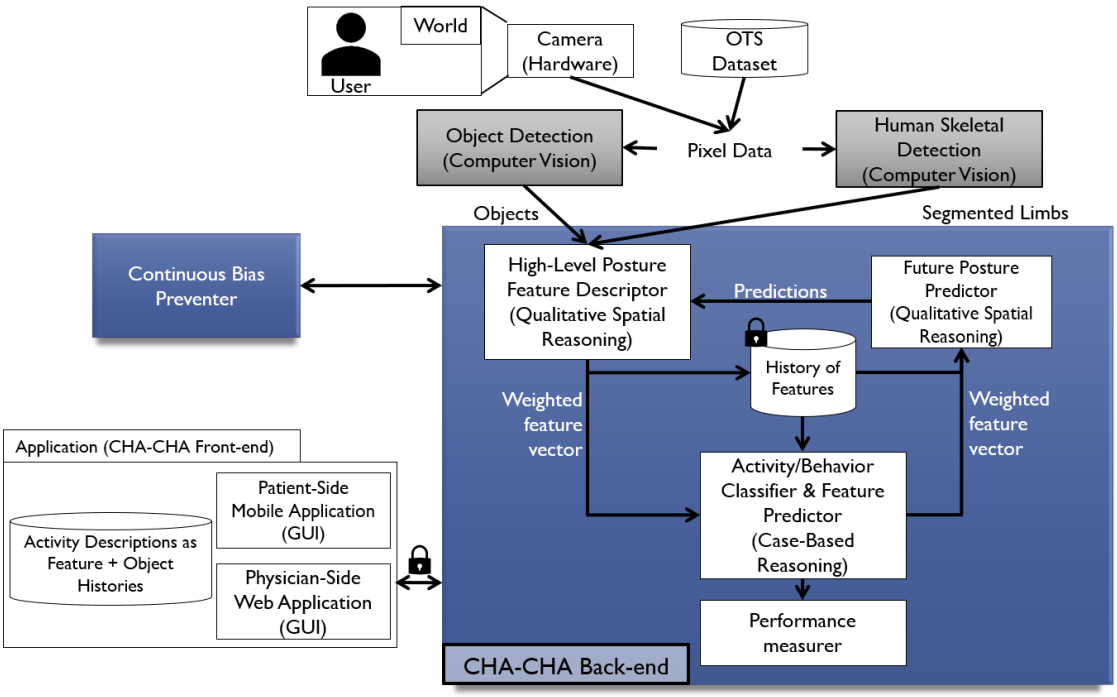


Figure 1: The CHA-CHA system architecture combines Off The Shelf (OTS) object and human skeletal detection with symbolic reasoning to classify and explain activities and performance anomalies to physicians via the physician facing front end.

Figure 2: Pose detection via BlazePose

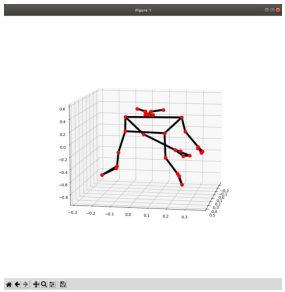


Figure 3: Points in 3-D of pose in Fig 2

has at least one built-in GPU to handle the graphics display.

We implemented an independent front-end component that converts sets of quantitative information into qualitative relations. The back-end implementations generate this quantitative information for a single off-the-shelf program, acting as the connection between the software and the CHA-CHA pipeline. Through this distinct separation in the code, we are able to use any format for detected information as long as we provide a small amount of code that connects it from the back-end implementation to the front-end implementation.

3. Feature Extraction

Once the data from the pose and object has been extracted, we extract the information further into a set of features. The rest of the CHA-CHA system uses these features for reasoning. As an abstraction of the numerical context from various sensor measurements, we developed a set of qualitative features that describe the stick Fig. representation of the human posture. This layer of human knowledge identifies the numerical relations that humans consider when describing activities, while filtering out other numbers and values that could categorize some collection of data by coincidence. Although it is still possible for intelligent systems to find extraneous patterns using the qualitative features, these patterns are in a form that people can interpret and use to assess the system as part of their own decision-making process. For example, if the system displays that a patient is limping because it noticed that their arm is always bent while walking, then a medical practitioner can be aware of the risk that the recognition might be inaccurate –the practitioner would be unable to catch this situation if the system just presents a list of seemingly arbitrary sensor values. More importantly, in the cases that the intelligent system's model is valid, then an explanation like the patient's leg is always bent provides useful context and information that the medical practitioner can apply when deciding how to revise the patient's exercise regimen.

Given the joints available from BlazePose, we decided to use the following qualitative features as the basis of describing human posture. Each feature includes a set of complements specifying its possible states, with transition between states based on how we interpret the numerical information. Additionally, each feature has both an instantaneous (still image) version and a locally temporal (change over the past few frames of motion) version. To maintain uniformity between individuals of various body sizes, the distance features are measured with respect to head lengths, which is relatively consistent in human anatomy if patients have traditional proportions.

- Features which describe the left and right arms and legs:
  - Bent vs. Straight [instantaneous]; Bending vs. Straightening [locally temporal]
  - In Front vs. Behind [instantaneous]; Moving Forward vs. Moving Backward [locally temporal]
  - Raised vs. Lowered [instantaneous]; Raising vs. Lowering [locally temporal]
  - Outward vs. Inward [instantaneous]; Moving Outward vs. Moving Inward [locally temporal]
- Features which describe the head and/or torso:
  - Tilted [instantaneous]; Tilting [locally temporal]
  - Twisted [instantaneous]; Twisting [locally temporal]
- Features which describe pairs of joints: nose, left eye, right eye, left ear, right ear, mouth, left shoulder, left elbow, left wrist, left index finger, right shoulder, right elbow, right wrist, right index finger, left hip, right hip, left knee, right knee, left ankle, right ankle, left heel, right heel, left big toe, and right big toe.
  - Near vs. Far [instantaneous]; Nearing vs. Distancing [locally temporal]

4. Dynamic Action Detection

We first used the features from the previous section for classification, which we found to be too fine grained to give accurate results. We found using more abstracted motions as features, such as 'left knee oscillation', gave much better classification results (the reported results in Sec. 6 only use these abstracted features). The fine grained features from Sec. 3 are used in the explanation, to pin-point activity anomalies more specifically. To extract the courser grained motion features we used a Dynamic Action Detection (DAD) algorithm, designed to detect pre-defined action types from time-series data of skeletal motion. DAD uses sparse identification of non-linear dynamics (SINDy) [17, 18] which has been used effectively to reconstruct diverse types of dynamic models directly from time-series state data.

DAD provides a method of characterization that is complementary to the discrete feature extraction discussed, but still operates with in input/output structure that is compatible with CBR. We focus specifically on the time-varying states of various body parts and attempt to match them to one or more primitive motions. Because DAD provides solutions based on pre-defined motion descriptions, it boasts some versatility. If the types of primitive motions that make up an exercise are known and well-defined, then it may be configured as an expert system without the need for any training data. Alternatively, given a broad enough set of metrics and properly labeled videos from a data set, then it may be used in a machine learning context to learn which metrics are the best descriptors for different exercises. In either case, CHA-CHA is able to use the DAD system to characterize and explain the potential anomalies in the activity; for example, if a user completes a jumping jack with a frozen shoulder. Just like the qualitative features, the primitive motions are human-readable and quickly understandable.

5. Activity Classification and Explanation

CHA-CHA uses case-based reasoning with the extracted qualitative features to classify human activities, find anomalies or deviations within those activities, and report those anomalies to the physician. A set of qualitative features describing previously seen activity instances is the 'problem', and the classification labels (e.g., 'squat', 'jumping jack') are the 'solution'. The case-base keeps a set of cases, which are comprised of previously seen problems and their solutions. The problem representing the currently observed activity q is matched against cases within the case-base (case $c1$, $c2$, ...) to determine which of the previously seen instances of each activity is most similar using a Euclidean distance similarity (Eq. 1). The Euclidean distance similarity measure (Eq. 1) is the combination of the detection of the features $q_f == c_f$ for the m features in the case. The posture features similarity is either equal (1) or not (0). The solution (classification label) from the case that is most similar is used as the solution to the currently observed situation. This approach has been proven to accurately identify actions with partial cases (i.e., predict actions in progress) [19]. Table 1 shows the full set of the features used in the cases where el is the elevation angle, and az is the azimuth angle.

$$sim(q, c) = \frac{\sum_{f=1}^{m} \alpha(q_f == c_f)}{m} \tag{1}$$

The CHA-CHA case base was seeded with cases from the Kinetics-700 human activity dataset

elbows anti sync	elbows in sync	legs oscillation
legs split	hips az anti sync	hips az in sync
hips el anti sync	hips el in sync	hips up anti sync
hips up in sync	left hip az oscillation high	left hip az oscillation low
left hip el oscillation high	left hip el oscillation low	left hip up oscillation high
left hip up oscillation low	right hip az oscillation high	right hip az oscillation low
right hip el oscillation high	right hip el oscillation low	right hip up oscillation high
right hip up oscillation low	right knee and hip el in sync	right shoulder az and hip az in sync
right shoulder az oscillation high	right shoulder az oscillation low	right shoulder el small
shoulders az anti sync	shoulders az in sync	shoulders el anti sync
shoulders el in sync	shoulders up anti sync	shoulders up in sync
left shoulder az and hip az in sync	left shoulder az oscillation high	left shoulder az oscillation low
left shoulder el small	high shoulder oscillation	small knee oscillation
mean knee oscillation small	no knee oscillation	left knee and hip el in sync
left knee oscillation	left knee oscillation high	left knee oscillation low
right knee oscillation	right knee oscillation high	right knee oscillation low
knees anti sync	knees in sync	

Table 1
Features which makeup the 'problem' of a case, captured using the DAD algorithm.

[20]. Cases were created automatically by running the labeled videos through the pose and object detection to get the raw data then the DAD algorithm to determine the primitive motions. There are at least 700 video clips from different YouTube videos for each of the 700 classes. For our prototype we selected a subset of the Kintetics dataset labels representing physical activities defined by our SME's (such as 'jumping jack' and 'squat'). The CHA-CHA CBR similarity function does consider weights as a means of measuring a feature's relevance to the case instance: a weight of 0 simply ignores the respective feature regardless of whether it matches while a weight of 1 emphasizes that the feature must match in order to support the similarity score. Currently all 'on' qualitative features are weighted evenly, however, each activity has certain (and different) body parts which are not involved with the motion of the activity. For example, a person's hands are equally weighted in the squat activity even though it is less indicative of a squat than their leg movement. In future work we will determine the best weighting of the features for continued high accuracy within a larger set of activities. Too many features that were not relevant to the activity being included in the case 'problem' was one of the major contributions to earlier experiments performing poorly with fine grained features. Weighting the features will also aid in the explanation to the CHA-CHA users - as it will prioritize the qualitative features in a way which is more interpretable than a simple list.

Explanations of how and why CHA-CHA classified an activity are generated from features of the most similar case from the CBR algorithm. We have not yet implemented the algorithm to generate explanations of anomalies in performance of an activity, based on previous work by

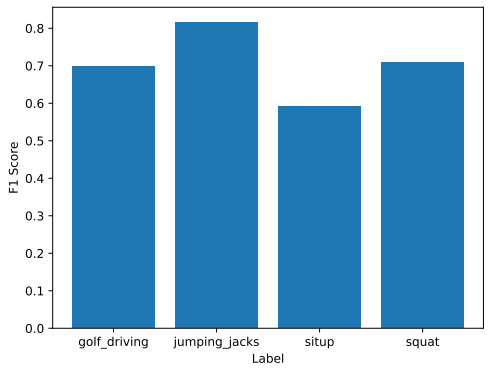
Borck et. al., [19]. When CHA-CHA is fully implemented, the performance will be evaluated by comparing the case representing the patient's activity (the query case q) to the most similar case base case c to determine the features which are indicative of the performance anomaly. Not all features divergences will indicate an anomaly, for example picking up a water bottle and taking a drink during a squat will show as a divergence but is not necessary to alert the physician. Therefore, this search will be guided by SME (oncologist in our domain) feedback, some of which is discussed in Section 7. By adding time-step annotations to the Primitive Motion features and enforcing an order within problem portion of the case we will be able to map the deviation to the fine grained features (from Section 3; ex: '*left arm bent*') for a more description explanation. For example, in the 'jumping jack' activity the motion features are (in order):

1. high shoulder oscillation
2. small knee oscillation
3. no knee oscillation
4. legs oscillation
5. legs split
6. shoulder az anti-sync
7. hips az anti-sync
8. left shoulder az and hip az in sync
9. right shoulder az and hip az in sync
10. left shoulder el small

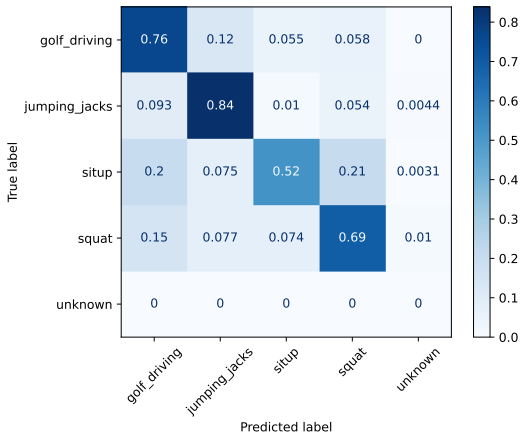
If, however, the patient cannot lift their right arm the '*right shoulder az and hip az in sync*' would be missing from the query feature set. This can then be further mapped to the fine grained features to find that, at that point in time in the activity, the '*right arm raising*' feature was missing. An alert would be sent to the physician in this example (see Sec. 7) which would indicate the activity 'jumping jack' was performed without the right armed being raised.

6. Validation

Preliminary results indicate the activity classification works well and quickly. We validated the activity classification CHA-CHA with a k-fold cross validation experiment (k = 10) on our dataset. Using a k-folds cross validation scheme eliminates selection bias from the experimental results. The algorithm randomly samples data for testing that was not used to train the classification. Cross validation testing tests the classification's ability to make predictions about new data, highlights whether the classification has been overfit, and gives insight into how the classifier will generalize to previously unseen data. Results (averaged over all folds) are shown in Fig. 4: Fig. 4a shows the average F1 score by label (averaged over all folds) and Fig. 4b shows the averaged confusion matrix. Experimentation with our initial prototype shows good classification results over a small subset of domain-relevant labels. Our F1 score for all labels is above 80%. The confusion matrix, using straight accuracy, shows very good results for all labels except situp. We believe this is due to the pose extraction which centers, and locks, the pose points at the hip. This creates an issue with confusing activities in which the person bends at the waist, for example situp and squat which both bend at the waist, start to look very similar. This issue will be alleviated when we add object adjacency features which will allow us to reason on whether the patient is moving their hips (i.e., a squat) or not (i.e., situp).



(a) Average F1 score over all folds by label



(b) Confusion matrix showing accuracy averaged over all folds

Figure 4: Initial experimental results

7. Interface with Physicians

The physicians are the primary users of the explanations generated by CHA-CHA. We designed the physician user interface using an iterative, user-centered approach to ensure the explanations are the right level of detail and succinctness. First, we conducted interviews with medical professional SMEs to understand user needs. Second, we brainstormed and sketched design concepts based on what we learned during the interviews. Then, we conducted four cycles of iterative development and usability testing: two with a low-fidelity prototype and two with a functional, web-based prototype. After each cycle, we refined the prototype based on the results from each usability test.

We conducted interviews with four medical practitioners, to learn more about their interactions with patients and what they would need from the CHA-CHA system and physician-facing web interface to preserve and enhance these interactions. One of the biggest lessons we learned from these discussions was that the exercises can sometimes be daily living tasks outside the home, including walking to the mailbox and pushing a grocery cart. The interviews also revealed that performance assessment is very tailored to the individual and not a standardized practice. Getting patients to try their best and do something rather than nothing is the actual goal, rather than following a strict exercise regimen. Whether amount done is a concern or achievement depends on abilities of the patient. For our domain, geriatric oncology, exercises are often daily living tasks such as sit and stand from a chair, or standing and reaching above their head (i.e., a high cabinet). Oncologist wanted concise reports in medical record and were more interested in patient's last scheduled appointment, their medications, and conditions. They did not want a high level of detail on patients' activities, preferring updates from Physical Therapists (PTs). Whereas PTs wanted to monitor a patient's performance of activity from last appointment.

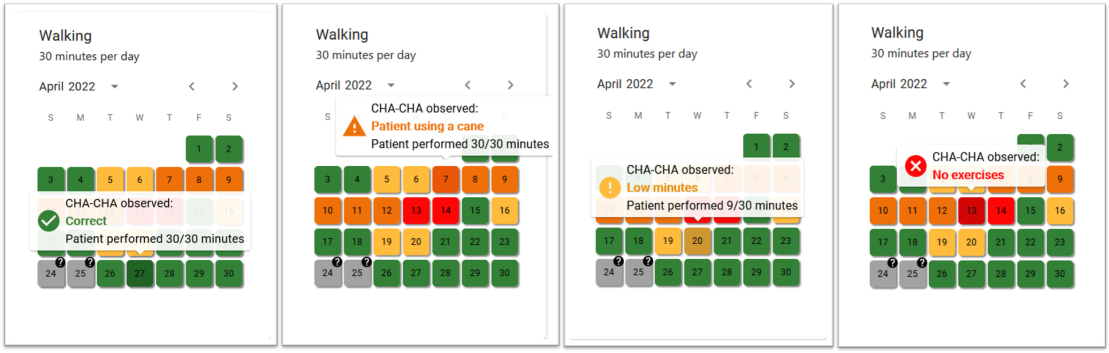


Figure 5: The physician side UI shows each exercise prescribed by the physician in a calendar form indicating whether the patient completed the exercise on each day and any information or alerts. From right to left: the exercise was performed correctly, incorrectly, correctly but not for enough time or reps, and not at all.

8. Related Work

8.1. Feature Extraction

CHA-CHA uses a spatial temporal feature set for Human Activity Recognition (HAR), specifically a body or stick model. The majority of the related work in activity recognition that uses a stick Fig. representation of the human posture still focuses on quantitative aspects when describing posture [1, 2]. Although computational models in artificial intelligence crunch numbers to extrapolate patterns for classification, the lack of context within myriad numbers yields patterns that are accurate with respect to calculations, but uninterpretable to human reasoning [8]. Blindly relying on intelligent systems that do not make sense under-the-hood can be acceptable in low-stakes applications, but recognizing and analyzing human activity for healthcare can have high-stakes consequences such as prescribing the wrong exercises or missing details that could later affect a patient's independent living status. Physicians, therapists, practitioners, caretakers, and the patients should be able to understand an intelligent system's inputs, models, and decision-making rationale.

8.2. Human Activity Recognition

Modalities for activity classification include wearable devices, smart phone sensors, and video. Many state-of-the-art HAR systems use sensor data such as Uddin and Soylu [3] whose system requires on-body sensors attached to patients for longitudinal data recording, deep learning algorithms then process the time series data for classification. A recent survey found "physical contact requires some skills and sophisticated equipment that make them accessible only to experimented users" therefore some researchers are abandoning contact based Human Activity Recognition (HAR) in favor for remote (vision based) HAR [4]. Zin et al. [5] identify regions of interest and UV-disparity from depth data (infra-red camera) which is then fed into a machine learning classifier to classify activities in the elder population. This technique, however, uses a special stereo depth camera. CHA-CHA was designed specifically to be able to be used at home

by patients with a normal smart phone.

Similar to our approach, the SelfBACK system in Wiratunga et. al., [6] uses CBR for personalized HAR for the maintenance of chronic health issues by monitoring patient exercises. However, the SelfBACK system uses accelerometer data whereas CHA-CHA uses solely video data. CBR has been used for monitoring elderly at home on a larger time scale as well (days vs one activity) [7].

8.3. Explainable AI in Health Care

Rudin [8] lays out the case for interpretable models, as opposed to symbolic models which explain black box AI models which may "perpetuate bad practices and can potentially cause catastrophic harm to society" especially in domains such as ours. Our system is designed to be interpretable throughout all of our model. By including both fine and coarse grained features in our reasoning we are able to provide explanations with differing levels of abstraction for different types of users (Oncologists versus Physical Therapists) and throughout many stages of development and research (i.e., from debugging in development to modifications and explanations of the system in user studies). CBR has a long history of uses within health care domains [9, 10] largely because it is so explainable. Lamy et. al., [11] used CBR to provide visual explanations of breast cancer diagnosis. Vásquez-Morales et. al., [12] use CBR as a twin system with Neural Networks (NN) to provide explanations of Chronic Kidney Disease predictions. Keane et. al., also provide theoretical analysis of explainable AI using CBR-NN twin systems [13].

9. Conclusion

We have presented the CHA-CHA system for activity and performance classification within the cancer health domain. Our prototype works on raw video data of activities which are parsed into qualitative features that make up the explanation of the activity classification and any health related alerts the physician may need to be aware of. Qualitative features at various levels of abstraction are fast to compute, accurately describe posture and motion, and are interpretable. We have validated the classification algorithm on a small set of domain relevant activities, taken from the Kinetics-700 dataset, and shown good preliminary results. Our physician facing interface was evaluated by a set of domain expert Oncologists and Physical Therapists who provided insights and feedback on the functionality and modality of the explanations. Overall the interface was met with enthusiasm good feedback. More work is needed to fully integrate the explanation system within the CHA-CHA prototype.

Acknowledgments

This project has been funded in whole or in part with Federal funds from the National Cancer Institute, National Institutes of Health, Department of Health and Human Services, under Contract No. 75N91021C00039

References

[1] R. G. Freedman, H.-T. Jung, S. Zilberstein, Plan and activity recognition from a topic modeling perspective, in: Proceedings of the Twenty-Fourth International Conference on Automated Planning and Scheduling, Portsmouth, New Hampshire, USA, 2014, pp. 360–364.

[2] G. Ercolano, S. Rossi, Combining CNN and LSTM for activity of daily living recognition with a 3D matrix skeleton representation, Intelligent Service Robotics 14 (2021) 175–185. doi:https://doi.org/10.1007/s11370-021-00358-7.

[3] M. Z. Uddin, A. Soylu, Human activity recognition using wearable sensors, discriminant analysis, and long short-term memory-based neural structured learning, Scientific Reports 11 (2021) 1–15.

[4] D. R. Beddiar, B. Nini, M. Sabokrou, A. Hadid, Vision-based human activity recognition: a survey, Multimedia Tools and Applications 79 (2020) 30509–30555.

[5] T. T. Zin, Y. Htet, Y. Akagi, H. Tamura, K. Kondo, S. Araki, E. Chosa, Real-time action recognition system for elderly people using stereo depth camera, Sensors 21 (2021) 5895.

[6] N. Wiratunga, A. Wijekoon, K. Cooper, Learning to compare with few data for personalised human activity recognition, in: International Conference on Case-Based Reasoning, Springer, 2020, pp. 3–14.

[7] E. Lupiani, J. M. Juarez, J. Palma, R. Marin, Monitoring elderly people at home with temporal case-based reasoning, Knowledge-Based Systems 134 (2017) 116–134.

[8] C. Rudin, Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead, Nature Machine Intelligence 1 (2019) 206–215.

[9] A. Holt, I. Bichindaritz, R. Schmidt, P. Perner, Medical applications in case-based reasoning, The Knowledge Engineering Review 20 (2005) 289–292.

[10] N. Choudhury, S. A. Begum, A survey on case-based reasoning in medicine, International Journal of Advanced Computer Science and Applications 7 (2016).

[11] J.-B. Lamy, B. Sekar, G. Guezennec, J. Bouaud, B. Séroussi, Explainable artificial intelligence for breast cancer: A visual case-based reasoning approach, Artificial intelligence in medicine 94 (2019) 42–53.

[12] G. R. Vásquez-Morales, S. M. Martinez-Monterrubio, P. Moreno-Ger, J. A. Recio-Garcia, Explainable prediction of chronic renal disease in the colombian population using neural networks and case-based reasoning, Ieee Access 7 (2019) 152900–152910.

[13] M. T. Keane, E. M. Kenny, How case-based reasoning explains neural networks: A theoretical analysis of xai using post-hoc explanation-by-example from a survey of ann-cbr twin-systems, in: International Conference on Case-Based Reasoning, Springer, 2019, pp. 155–171.

[14] A. Bochkovskiy, C.-Y. Wang, H.-Y. M. Liao, Yolov4: Optimal speed and accuracy of object detection, arXiv preprint arXiv:2004.10934 (2020).

[15] V. Bazarevsky, I. Grishchenko, K. Raveendran, T. L. Zhu, F. Zhang, M. Grundmann, Blazepose: On-device real-time body pose tracking, ArXiv abs/2006.10204 (2020).

[16] B. J. Cohen, S. Chitta, M. Likhachev, Search-based planning for manipulation with motion primitives, in: IEEE International Conference on Robotics and Automation, IEEE, Anchorage, Alaska, USA, 2010, pp. 2902–2908. URL: https://doi.org/10.1109/ROBOT.2010.5509685.

doi:10.1109/ROBOT.2010.5509685.

[17] S. L. Brunton, J. L. Proctor, J. N. Kutz, Discovering governing equations from data by sparse identification of nonlinear dynamical systems, Proceedings of the national academy of sciences 113 (2016) 3932–3937.

[18] U. Fasel, E. Kaiser, J. N. Kutz, B. W. Brunton, S. L. Brunton, Sindy with control: A tutorial, arXiv preprint arXiv:2108.13404 (2021).

[19] H. Borck, S. Johnston, M. Southern, M. Boddy, Exploiting time series data for task prediction and diagnosis in an intelligent guidance system (2016) 132–142.

[20] L. Smaira, J. Carreira, E. Noland, E. Clancy, A. Wu, A. Zisserman, A short note on the kinetics-700-2020 human action dataset, arXiv preprint arXiv:2010.10864 (2020).