
Comparative Techniques Using Hierarchical Modelling and Machine Learning for Procedure Recognition in Smart Hospitals

Shaheena Noor^{1,*}, Muhammad Aamir², Najma Ismat¹
and Muhammad Imran Saleem¹

¹*Department of Computer Engineering, Sir Syed University of Engineering & Technology, Karachi, Pakistan*

²*Department of Telecommunication Engineering, Sir Syed University of Engineering & Technology, Karachi, Pakistan*

E-mail: shanoor@ssuet.edu.pk; maamir@ssuet.edu.pk; nismat@ssuet.edu.pk; isaleem@ssuet.edu.pk

**Corresponding Author*

Received 21 April 2021; Accepted 10 March 2022;
Publication 06 May 2022

Abstract

6G is one of the key cornerstone elements of the futuristic smart system setup – the others being cloud computing, big data, wearable devices and Artificial Intelligence. Also, smart offices and homes have become even more popular than before, because of the advancement in computer vision and Machine Learning (ML) technologies. Recognition of human actions and situations are fundamental components of such systems, especially in complex environments like healthcare, for example at the dentist clinic, where we need cues such as eye movement to distinguish procedures being undertaken. In this work, we compare models based on hierarchical modelling and machine learning to identify the dental procedure. We used the objects seen while

Journal of ICT Standardization, Vol. 10.2, 145–164.

doi: 10.13052/jicts2245-800X.1023

© 2022 River Publishers

following the eye trajectories and focussed on elements including material used for treatment, equipment involved and the teeth conditions i.e. symptoms. Our experiments showed that using Artificial Neural Network (ANN) increased the accuracy of prediction compared to hierarchical modelling. Our experiments show an improvement in accuracy for each of the constituent parameters i.e., symptom (ANN: 95.58% vs. Hierarchical: 45.68%), material (ANN: 86.32% vs. Hierarchical: 45.18%) and equipment (ANN: 92.65% vs. Hierarchical: 59.39%).

Keywords: Procedure recognition, inside-out vision, machine learning, artificial neural network, 6G-enabled applications.

1 Introduction

Wireless technologies have seen a fast rate of growth ever since they first started evolving almost four decades ago. The 6th Generation i.e., 6G of wide area wireless technology is expected to be even more powerful, pervasive and diverse in applications. The advancements in wireless technologies has allowed widespread use of IOT (Internet of Things) devices thus making it possible to have smart buildings like smart home, offices and healthcare including smart hospitals.

A smart city is one in which traditional networks and services are made more efficient via the application of digital and communications technologies, that is beneficial to its residents and companies. Smart cities integrate infrastructure, social capital and digital technology to foster long-term economic growth and provide a welcoming environment for all residents.

By 2035, it is expected 6G to be a fundamental aspect of smart city communications and will rule in different industries like health industry. Many industries, including healthcare, are projected to be transformed by 6G. Healthcare will be totally AI-driven and reliant on 6G connectivity technologies, altering our lifestyle perceptions. Time and space are now the most significant hurdles to health care, but 6G will be able to transcend these obstacles. In addition, 6G will be demonstrated to be a game-changing technology in the field of healthcare. As a result, we foresee a healthcare system for the future of 6G communication technology in this light.

Also, the advances in machine vision research has encouraged us to incorporate these innovations into our daily lives, including households, businesses and healthcare [1, 2]. Especially w.r.t. the first-person view point, machine

vision researchers are also focusing on smart glasses [3, 4] and gaze-directed cameras as part of wearable technology. Since the procedures or scenarios can be broken down into smaller tasks and activities that can be repeated at the level of individual artifacts and simple actions, this principle can be used for the reverse identification of scenarios.

The rise of medical robots has made the healthcare treatments more intelligent [5]. In this paper, we consider the smart dental setup and propose an intelligent solution to procedure recognition using first person vision. Recognizing procedures is important so that the smart system and automated tools (and robots) can help the human medical practitioner in carrying out the tasks efficiently. We present two complimentary approaches – hierarchical modeling and Artificial Neural Networks – to predict the dental procedure in progress, only by looking within the perceptive field of dentist.

The research paper has been organized in following manner: In Section 2, we covered a review of the literature on wireless and first-person vision-based scenario recognition methods and ML solutions. In Section 3, we presented our solution using gaze-based vision and ANN. In Section 4, we give details of the dataset and the experiments using hierarchical model and ANN. In the last section, we present the conclusions.

2 Literature Review

Our eye gaze is directly influenced by our minds and thoughts [6, 7]. Hence, it is increasingly being used in intelligent houses, to identify locations and objects [8], analyse different activities and behaviours [9], research human mind and nervous system [9], and even detect mental disorders. Furnari et al. in [10] predicted user future actions and the objects to be interacted with using first person vision.

First Person Vision is different from the conventional perception of Third Person Vision (TPV) (also refer as outside-in view) as here a wearable camera is used to record the scene. Hence, the actor is not the main figure to be shot but instead the scene is portrayed from his viewpoint of the world around him. Traditionally, in TPV, one or more cameras are mounted to monitor human activity including motion and manipulation. However, a major limitation here is in detecting small movements, which are primarily occluded by actor. Also, mounting cameras at certain locations is not always feasible or possible; thus, it is important to provide a first-person view.

As the wireless and IOT setup becomes more stable and pervasive, a lot of efforts are being spent to provide solutions for activity and

scenario recognition in smart environments. Existing scenario and activity recognition solutions use cameras, PointGrab smartphones, GPS and accelerometers [11] among other technologies. Wang et al. reviewed activity recognition approaches based on radio waves and compared them to approaches based on ZigBee, WiFi and RFID [12]. Lu and Fu [13] used Bayesian network-based fusion engine in an intelligent home setup. They used wireless sensor network to recognize activities using location awareness. Sun et al. [14] created a portable pyroelectric infrared thermal sensing system that helps the visually impaired by providing information of the surrounding environment.

For scenario recognition, some work has been performed using both fixed and wearable cameras, where objects and activities are recorded. In [15], Henderson et al. investigated the importance of eye movement in scenario recognition. In another work, Keith et al. [16] explored the gaze trajectories during scenario execution and their connection with the task being undertaken. Wang et al. [17] introduced behaviour recognition using Channel State Information. In [18], deep learning models are used for object recognition and activity identification. Since eye movements are closely linked to the current environment, job or operation, the objects that are fixated in this trajectory help in recognizing scenarios and tasks at hand.

2.1 Procedure Recognition Using ML Modelling

Human activity recognition is basically a machine learning problem [19] for the purpose of recognizing human activities in a specific setup [20], that consists of extracting features followed by the creation and training of a classification model. A variety of approaches for machine learning have been explored, such as Naive Bayes [21], Support Vector Machines (SVM) [22], Decision Trees [23] and Hidden Markov Model (HMM) [20, 24]. There is also a lot of literature on a comparison of these methods [25, 26]. A number of variants of HMM were used to characterize the human activity. San Segundo et al. [19] trained HMM to identify behaviours and full-body motion using smart phone inertial signals. Such an approach, however, is not suitable in the identification of smaller movement tasks such as dental procedure, where most acts are distinguished by hand movements and do not depend on full body.

With the upsurge of deep learning methods [27], focus has been moved back to using ANN for general classification and in particular, recognition of activities [28].

3 Scenario and Procedure Recognition Using Gaze-based Vision

3.1 Hierarchical Modelling

Information regarding the fixated objects from first person view helps to define the focus of attention. This information is more comprehensive compared to the regular TPV, because it does not have occlusion problems or misidentification issues of small objects. In Figure 1 [29], an external view of a dentist is seen who is performing the treatment on the patient.

As external camera is mounted at a distance therefore the condition of teeth (symptom) cannot be seen easily that creates difficulty to understand scenario. Also, another important aspect i.e., tool or equipment used to examine the patients' teeth cannot be detected from outside-in view, due to similarity in appearance of tools. Consider Figure 2 [30, 31], in which the doctor's first-person view is observed. The equipment used here is clearly visible, along with the teeth conditions a.k.a. symptom. By using wearable, eye tracking cameras, it becomes possible to note the exact objects seen and hence depict the procedure in progress.

Consider the description of an example procedure $P_i = \langle \text{bleaching, scaling, root canal, sealant, crown, filling} \rangle$ from three aspects: "Material", "Symptom" and "Equipment". Each of these aspects is defined in terms of objects seen: O_i , e.g., $O(\text{filling, material})$: $\langle \text{'Resin composites', 'glass}$

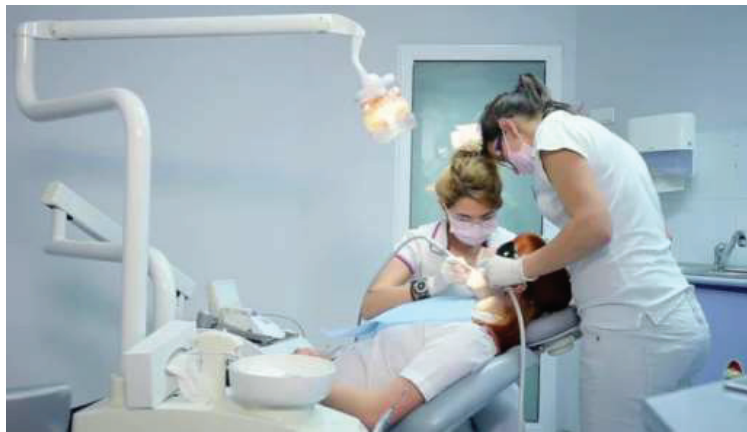


Figure 1 Outside-in view of a Dental surgeon working with a Patient [29]. No information about condition of teeth, equipment or material used can be extracted from the image due to self-occlusion.



Figure 2 Inside-out view of dental surgeon checking and treating the patient [30][31]. Here we can see the equipment used i.e. examination mouth mirror (left) and condition of teeth i.e. finishing and polishing of a restoration (right).

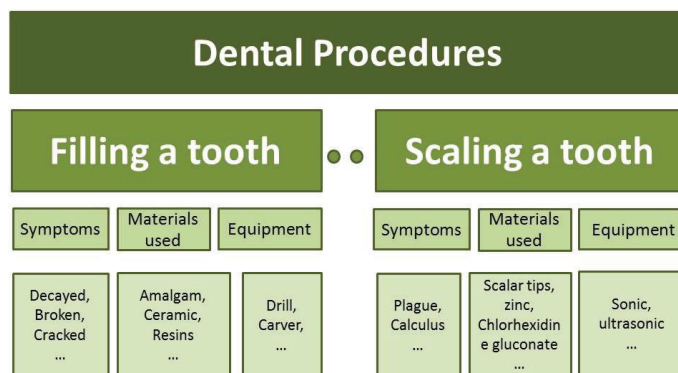


Figure 3 Dental procedures are represented hierarchically based on material, symptom and equipment.

ionomer', ... >, O(filling, symptom): <'decayed', 'cracked', 'broken'> and O(filling, equipment): <'Condenser', 'Carver' ... >. These three parameters are used to depict the procedure being carried out.

Consider Figure 3 that represents the hierarchical form of the dental procedures. Each of the dental procedure is seen as a combination of objects seen while the task is performed.

We may distinguish focused objects through monitoring eye movement and gaze details. Fixations and Saccade are the most common type of eye motion. Fixation refers to the fixed eyes emphasized on a certain point [32]; while on the other hand Saccade refers to the sudden transition between

successive fixation [33]. We regard an object as fixated or seen, in case it appears consecutively in t frames or more – where t is a heuristically- defined threshold. Given $O_E: o_{e1}, o_{e2}, \dots o_{en}$ @ objects focussed by eyes; we use Algorithm 1 to calculate $\{S|P\}$: Possibility score of predicted procedure.

Algorithm 1 Recognition of dental procedure based on objects seen in gaze [34]

Input: $O_G: o_{g1}, o_{g2}, \dots o_{gn} \equiv$ set of objects seen in gaze including parameters: material, symptom and equipment used

Output: $\{S|P\}$: Possible scenario (Procedure) with possibility score

```

1: Initialize Semantic Model individually based on each parameter:  $M_{s_i}^{[O]k_i[P]k_j}$ 
2: Generate parameter-based Inverted object Model:  $\bar{M}_{s_j}^{[\bar{O}]k_i[\bar{P}]k_j}$ 
3: Initialize prediction to Null{ }
4: Initialize probabilities to Null{ }
5: for obj in  $O_G$ 
6:   Compute Number of possibilities = Totalpossibilities (obj,  $M_{obj}$ ,  $P_{obj}$ )
7:   for possible in  $\langle M_{obj}, P_{obj} \rangle$ :
8:     Prob[scenario] = Prob.get(scenario.0) + 1./occurrences
9:   end for
10: end for
11: P:maxx = Get the highest value from the probability hash map
12: S:Prediction = Get items corresponding to maxx above
13: Initialize Semantic Model
14: For each parameter  $\langle Mat, Sym, Equ \rangle$ :
15:  $M_P: \{s_1: \langle o_1, o_2 \dots M_{s_i}^{[O]k_i} \rangle\}$ 
16:  $M_{s_i}^{[O]k_i[P]k_j} \equiv \{s_1: [o_1, o_2 \dots o_{k1}] \dots s_N: [o_1, o_2 \dots o_{kN}]\}$ 
17: Generate Invert Object Model
18:  $\bar{M}_{s_i}^{[\bar{O}]k_i[\bar{P}]k_j} \equiv \{o_1: [s_1, s_2 \dots s_{l1}] \dots o_{K_N}: [s_1, s_2 \dots s_{lN}]\}$ 
19: int TotalPossibilities(object obj, inverted object Model  $M_{obj}$ )
    return len( $M_{obj}$ )

```

3.2 ANN-based Modelling

Based on the literature review in Section 2, we used ANN for machine learning-based procedure recognition. ANNs are typically trained by using a forward cycle where training data is fed into the system. Weights are calculated for the network to make prediction. The difference in network's output vs ground truth is then propagated back to the system that results in updating of weights and training of network. This algorithm is commonly referred to as backpropagation, which was first introduced in 1970s, however gained popularity later after Rumelhart et al. [35] demonstrated its superiority over other learning methods. We will not go into the specifics of general ANN training much because it's out of the scope of our paper. Algorithm 2 provides

an overview of the training for an ANN and the configuration specifics are discussed in detailed in Section 4.1.

Algorithm 2 Training a neural network [36, 37]

```

1: Class Training
2:   Initialize  $W = [W^1, W^2, b^1, b^2]$ 
3:   Repeat for  $i = 1 : m$ 
4:     Perform feedforward pass:
5:     Compute  $\hat{x}^i$ .
6:     Perform backpropagation:
7:     Compute gradients:  $\Delta W J_a(W)$ .
8:     Compute weight change:  $\Delta W$ .
9:     Update weight  $W$ .
10: Class Feature Encoding
11: Compute:  $\hat{x}^i = f(W^i x^i + b^1)$ 

```

4 Experiments & Results

A few dental setups are considered in Figures 4 through 7 interpreting the symptoms, material and equipment in each case.

We start by developing prediction models using three basic parameters – (1) material (2) equipment and (3) symptom. Consider, e.g., the model generated for ‘material’:

- ‘Bleaching’: [carbamide peroxide, hydrogen peroxide],
- ‘Filling’: [Resin composites, glass ionomer, ...],



Figure 4 Treatments (from left): crown [38], filling [39], teeth whitening [40] and sealant [41].



Figure 5 Symptoms (from left): cavities [42], broken [43], plaque [44] and calculus [45].



Figure 6 Materials (from left): Porcelain Fused with Metal [46], stainless steel [47], carbamide peroxide [48], chlorhexidine gluconate [49].



Figure 7 Equipment [50].

- ‘Root Canal’: [guttapercha, Zirconia, ...],
- ‘Crown’: [Stainless steel, Zirconia, ...],
- ‘Scaling’: [chlorhexidine gluconate, scaler tips, ...],
- ‘Sealant’: [acid solution].

We generate an independent model for every parameter. Then, using these models, we construct inverse models based on the items seen in inside-out camera; resulting in the possibility score of the procedure based on objects seen. Consider the following inverse models as example:

- ***symptom:***
 - ‘Weakened’: [Root Canal, Crown],
 - ‘calculus’: [Scaling],
 - ‘broken’: [Filling]
 - ...

– ***material:***

- ‘Ceramics’: [Filling] ...
- ‘Resin composites’: [Crown, Root Canal, Filling],
- ...

– ***equipment:***

- ‘plastic instrument’: [Sealant],
- ‘cement spatula’: [Sealant, Root Canal, Crown],
- ‘profy’: [Bleaching]
- ...

In the final step, we take the fixated objects to measure the probabilities and make prediction.

In Table 1, we present a sample of the findings with a comparative plot in Figure 8. The ground truth is represented in each row while we specify the predicted output in columns. The percentage of correct predictions is mentioned in each cell. We present the results using hierarchical method and the neural network (discussed next) for each case.

To interpret the Table 1, consider a scenario with Ground Truth ‘Crown’ under equipment model, the values for Crown are Crown: 42.31% and Bleaching: 11.54%. This means that 42.31% times the system correctly predicted the Crown procedure. Next, it incorrectly predicted Bleaching as Crown 11.54% times.

4.1 Model for Scenario Recognition configuration

As an alternative approach to hierarchical modelling, we used the EasyNN [51] to develop the ANN – a quick tool for data modelling. Figure 9 illustrate a sample dataset in the tool with the corresponding network in Figure 10.

It took >736725, >2157382 and 115 cycles for the network of material, symptoms, and equipment to finish training. We let the network to continue converging till the average error goes below 0.01. Once the training process is complete, our cumulative error for each of material, symptom and equipment are 0.1571, 0 and 3 respectively; minimum error values are 0 for all three, and average error values are 0.0649, 0.0432 and 0.0096 respectively. We used a learning rate of 0.6 and a momentum of 0.8. The details of the trained network are given in Table 2. The network was regenerated and trained three times

Table 1 Procedure recognition results using equipment, symptom and material

	Hierarchical Method											
	Material					Neural Network						
	Sealant	Crown	Bleaching	Scaling	Filling	Root Canal	Sealant	Crown	Bleaching	Scaling	Filling	Root Canal
Sealant	55.32	0	0	0	0	100	0	0	0	0	0	0
Crown	0	34.62	0	0	11.54	11.54	0	36.36	0	0	27.27	36.36
Bleaching	0	0	40	0	0	0	0	100	0	0	0	0
Scaling	0	0	0	43.75	0	0	33.33	0	0	66.67	0	0
Filling	0	0	0	0	36.73	6.12	7.69	7.69	30.77	0	38.48	15.38
Root Canal	0	0	0	0	2.27	52.27	0	0	0	0	20	80
	Symptom											
	Sealant	Crown	Bleaching	Scaling	Filling	Root Canal	Sealant	Crown	Bleaching	Scaling	Filling	Root Canal
Sealant	40.43	0	0	0	0	100	0	0	0	0	0	0
Crown	0	53.58	0	0	0	0	100	0	0	0	0	0
Bleaching	0	0	60	0	0	0	0	33.33	0	33.33	0	33.33
Scaling	0	0	0	62.5	0	0	0	0	0	100	0	0
Filling	0	0	0	0	30.61	0	0	42.86	0	0	57.14	0
Root Canal	0	0	0	0	0	52.27	0	0	0	0	0	100
	Equipment											
	Sealant	Crown	Bleaching	Scaling	Filling	Root Canal	Sealant	Crown	Bleaching	Scaling	Filling	Root Canal
Sealant	74.47	4.26	0	0	0	4.26	80	0	0	20	0	0
Crown	0	42.31	11.54	0	0	0	0	83.33	0	16.67	0	0
Bleaching	0	0	60	0	0	0	0	0	0	100	0	0
Scaling	0	0	0	62.5	0	0	0	0	0	100	0	0
Filling	0	0	0	0	55.1	0	4.76	0	0	23.81	71.43	0
Root Canal	0	0	0	0	0	56.82	0	0	0	0	0	100

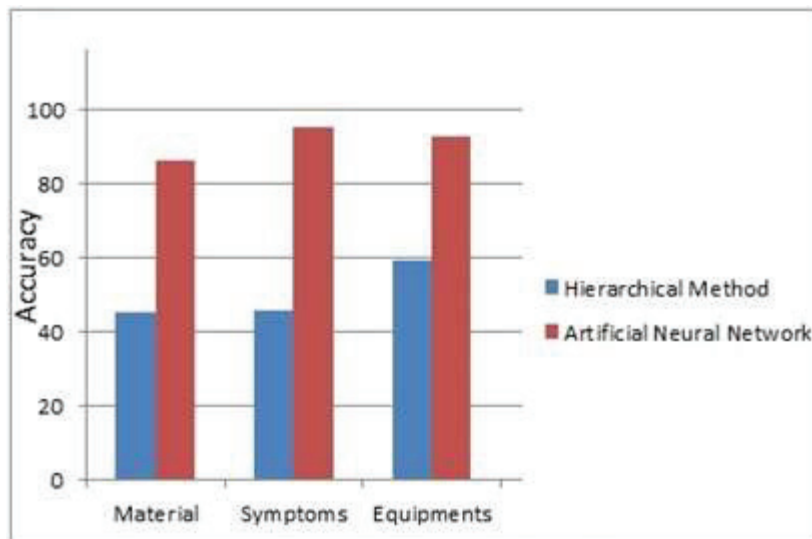


Figure 8 Performance Metrics (Accuracy).

	Selected Input Parameters										Selected Output Parameters						
	burs	handpiece	slow handpiece	profy	ment spat	pression	traction	cc	probe	plastic instrument	uring lig	Filling	Crown	Scaling	Bleaching	Sealant	Root Canal
Subset of Training Data	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
	1	1	0	0	1	1	1	1	1	0	0	0	1	0	0	0	0
	1	1	0	0	1	1	1	1	1	0	0	0	1	0	0	0	0
	1	1	0	0	1	1	1	1	1	0	0	0	1	0	0	0	0
	1	1	0	0	1	1	1	1	1	0	0	0	1	0	0	0	0
	1	1	0	0	1	1	1	1	1	0	0	0	1	0	0	0	0
	1	1	0	0	1	1	1	1	1	0	0	0	1	0	0	0	0
	1	1	0	0	1	1	1	1	1	0	0	0	1	0	0	0	0
	1	1	0	0	1	1	1	1	1	0	0	0	1	0	0	0	0
	1	1	0	0	1	1	1	1	1	0	0	0	1	0	0	0	0
	1	1	0	0	1	1	1	1	1	0	0	0	1	0	0	0	0
Subset of Query Data	0	0	0	0	0	0	0	0	0	0	0	0.9765	0.005	0.0047	0.0047	0.0047	0.0047
	0	0	0	0	0	0	0	0	0	0	0	0.9765	0.005	0.0047	0.0047	0.0047	0.0047
	0	0	0	0	0	0	0	0	0	0	0	0.9765	0.005	0.0047	0.0047	0.0047	0.0047
	0	0	0	0	0	0	0	0	0	0	0	0.9765	0.005	0.0047	0.0047	0.0047	0.0047
	0	0	0	0	0	0	0	0	0	0	0	0.9765	0.005	0.0047	0.0047	0.0047	0.0047
	0	0	0	0	0	0	0	0	0	0	0	0.9765	0.005	0.0047	0.0047	0.0047	0.0047
	0	0	0	0	0	0	0	0	0	0	0	0.9765	0.005	0.0047	0.0047	0.0047	0.0047
	0	0	0	0	0	0	0	0	0	0	0	0.9765	0.005	0.0047	0.0047	0.0047	0.0047
	0	0	0	0	0	0	0	0	0	0	0	0.9765	0.005	0.0047	0.0047	0.0047	0.0047
	0	0	0	0	0	0	0	0	0	0	0	0.9765	0.005	0.0047	0.0047	0.0047	0.0047
	0	0	0	0	0	0	0	0	0	0	0	0.9765	0.005	0.0047	0.0047	0.0047	0.0047
	1	1	0	0	1	0	0	0	0	0	0	0.0024	0.988	0.0024	0.00236	0.00236	0.00236
	0	1	0	0	1	1	0	0	0	0	0	0.0024	0.988	0.0024	0.00242	0.00242	0.00242

Figure 9 Sample parameters & Dataset for ANN Learning.

with inside-out extracted object information – once for each of symptom, material and equipment.

Table 3 shows the result of different models for each configuration of the dataset. It is interesting to note that Machine Learning gives far better result compared to the hierarchical model in all the cases considered.

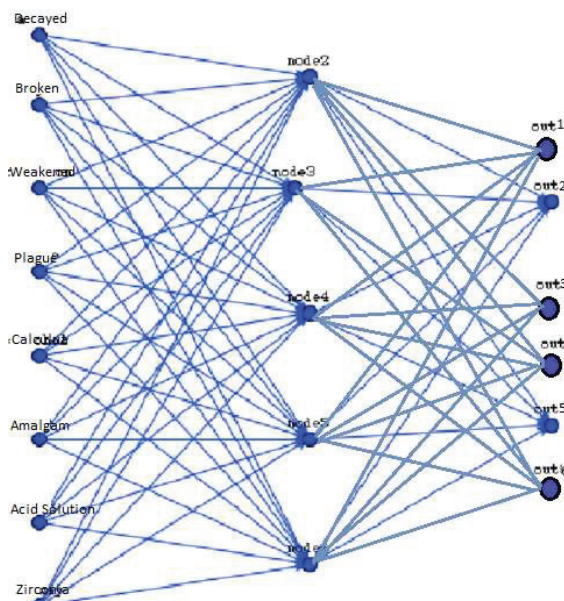


Figure 10 Procedure Recognition using Neural Network from FPV.

Table 2 Details of ANN

	Input Parameter	No. of Hidden Layers (No. of Neurons)	Output Parameters	Learning Cycles
Material	16	2 (H1 = 12, H2 = 9)	6	> 736725
Symptoms	13	1 (8)	6	>2157382
Equipment	19	2 (H1 = 14, H2 = 9)	6	115

Table 3 Accuracy of procedure recognition results

Method	Material (%)	Symptoms (%)	Equipment (%)
Hierarchical Method	45.18	45.68	59.39
Artificial Neural Network	86.32	95.58	92.65

5 Conclusion

We used the dental setup and represented a procedure recognition method using the objects in focus that includes material, symptom and equipment. We have developed probabilistic and machine learning models (Hierarchical and Artificial Neural Network) for understanding dental procedures, which involve several devices/tools, symptoms and materials. Our findings suggest that a given dental procedure can be broken down into sub-processes and

objects in a recursive fashion. Therefore, provided the object information, i.e. symptom of teeth, material or equipment, we can accurately recognize the correct procedure. We found that Machine Learning gives better result compared to the hierarchical model in all the cases considered: symptom (ANN: 95.58% vs. Hierarchical: 45.68%), material (ANN: 86.32% vs. Hierarchical: 45.18%) and equipment (ANN: 92.65% vs. Hierarchical: 59.39%). The findings can be used in smart environments for both industrial and personal robots with a focus in medical setup, which are expected to become even more prevalent as 6G becomes widespread.

References

- [1] J. Gao, Y. Yang, P. Lin, D.S. Park, 'Computer Vision in Healthcare Applications', *J Healthc Eng.* 2018;2018:5157020. 4 March 2018, doi: 10.1155/2018/5157020.
- [2] C. H. Chen. 'Series in Computer Vision', Volume 2, Pages: 412, *Computer Vision in Medical Imaging.* January 2014.
- [3] A. Betancourt, P. Morerio, C.S. Regazzoni, and M. Rauterberg, 'The Evolution of First Person Vision Methods: A Survey', *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 5, pp. 744–760, May 2015.
- [4] M. Dimiccoli. 'Computer Vision for Egocentric (First-Person) Vision', In *Computer Vision for Assistive Healthcare*, Editors: Marco Leo, Giovanni Maria Farinella. Academic Press, 2018, Pages 183–210, ISBN 9780128134450, <https://doi.org/10.1016/B978-0-12-813445-0.00007-1>.
- [5] S. Tian, W. Yang, J. Michael Le Grange, P. Wang, W. Huang, Z. Ye, 'Smart healthcare: making medical care more intelligent', *Global Health Journal*, Volume 3, Issue 3, 2019, Pages 62–65, ISSN 2414-6447, <https://doi.org/10.1016/j.glohj.2019.07.001>.
- [6] G.J. Zelinsky, R.P.N. Rao, M.M.Hayhoe, and D.H.Ballard, 'Eye Movements Reveal the Spatiotemporal Dynamics of Visual Search', *A journal of the association for Psychological Science*, vol. 8, no. 6, pp. 448–453, 1997.
- [7] A.Toet, 'Gaze directed displays as an enabling technology for attention aware systems', *Journal in Computers in Human Behavior*, vol. 22, no. 4, pp. 615–647, July 2006.
- [8] H. Kang, A.E. Alexei, M. Herbert, and T. Kanade, 'Image Matching in Large Scale Indoor Environment', in *IEEE Conference on Computer*

- Vision and Pattern Recognition (CVPR) Workshop on Egocentric Vision, 2009.
- [9] L. Sun, U. Klank, and M. Beetz, 'EYEWATCHME 3D Hand and object tracking for inside out activity analysis', IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2009. CVPR Workshops, pp. 9–16, 20–25 June 2009.
 - [10] A. Furnari, G. M. Farinella, 'What Would You Expect? Anticipating Egocentric Actions with Rolling-Unrolling LSTMs and Modality Attention', International Conference on Computer Vision, 2019.
 - [11] M. A. A. Al-qaness, & F. Li, 'WiGer: WiFi-based gesture recognition system', ISPRS International Journal of Geo-Information, 5(6), 92. 2016.
 - [12] S. Wang, & G. Zhou, 'A review on radio based activity recognition', Digital Communications and Networks, 1(1), 20–29. 2015.
 - [13] C. H. Lu, & L.C. Fu, 'Robust location-aware activity recognition using wireless sensor network in an attentive home', IEEE Transactions on Automation Science and Engineering, 6(4), 598–609. 2009.
 - [14] Q. Sun, J. Shen, H. Qiao, X. Huang, C. Chen, & F. Hu, 'Static human detection and scenario recognition via wearable thermal sensing system', Computers, 6(1), 3, 2017, doi: 10.3390/computers6010003.
 - [15] J. M. Henderson, & A. Hollingworth, 'High-level scene perception', Annual Review of Psychology, 50, 243271, 1999.
 - [16] K. Rayner, T. J. Smith, G. L. Malcolm, & J. M. Henderson, 'Eye movements and visual encoding during scene perception', Psychological Science, 20(1), 6–10, 2009.
 - [17] Y. Wang, X. Jiang, R. Cao, & X. Wang, 'Robust indoor human activity recognition using wireless signals', Sensor, 15(7), 17195–17208, 2015.
 - [18] B. Zhou, A. Lapedriza, J. Xiao, A. Torralba, & A. Oliva, 'Learning deep features for scene recognition using places database', In Z. Ghahramani M. Welling, C. Cortes, N. D. Lawrence, & K. Q. Weinberger (Eds.), Advances in neural information processing systems (Vol. 27, pp. 487–495). Red Hook: Curran Associates Inc. 2014.
 - [19] R.S. Segundo, J.M. Montero, J.M. Pimentel, and J.M. Pardo, 'HMM Adaptation for Improving a Human Activity Recognition System', Algorithms, Volume 9, No. 3, 2016.
 - [20] E. Kim, S. Helal, and D. Cook, 'Human Activity Recognition and Pattern Discovery', IEEE Pervasive Computing, Volume 9, No. 1, pp. 48–53, January, 2010.

- [21] L.C. Jatoba, U. Grossmann, C. Kunze, J. Ottenbacher and W. Stork, 'Context-Aware Mobile Health Monitoring: Evaluation of Different Pattern Recognition Methods for Classification of Physical Activity', 30th IEEE Annual International Conference on Engineering in Medicine and Biology Society, 2008.
- [22] D. Anguita, A. Ghio, L. Oneto, X. Parra and J.L. Reyes-Ortiz, 'Energy Efficient Smartphone-Based Activity Recognition Using Fixed-Point Arithmetic', *Journal of University Computer Science*, 2013.
- [23] U. Maurer, A. Smailagic, D. Siewiorek, and M. Deisher, 'Activity Recognition and Monitoring Using Multiple Sensors on Different Body Positions', *Proceedings of International Workshop on Wearable and Implantable Body Sensor Networks*, 2006.
- [24] K. Shaharyar, J. Ahmad, and D. Kim, 'Depth Images- Based Human Detection, Tracking and Activity Recognition Using Spatiotemporal Features and Modified HMM', *Journal of Electrical Engineering & Technology*, Volume 11, No. 3, pp. 1921–1926, 2016.
- [25] J. Yang, 'Toward Physical Activity Diary: Motion Recognition Using Simple Acceleration Features with Mobile Phones', *Proceedings of 1st ACM International Workshop on Interactive Multimedia for Consumer Electronic*, 2009.
- [26] J.R. Kwapisz, G.M. Weiss, and S.A. Moore, 'Activity Recognition Using Cell Phone Accelerometers', *SIGKDD Explore News Letters*, Volume 12, No. 2, pp. 74–82, March, 2011 (Last Visit: 15 June 2017). [Online]. Available: <https://en.wikipedia.org/wiki/Deeplearning>
- [27] S. Noor, and V. Uddin, 'Using ANN for Multi-View Activity Recognition in Indoor Environment', *International Conference on Frontiers of Information Technology*, pp. 258–263, December, 2016.
- [28] R. Damaevilius, M. Vasiljevas, J. Alkeviilius, and M. Wofniak, 'Human Activity Recognition in AAL Environments Using Random Projections', *Computational and Mathematical Methods in Medicine*, pp. 17, 2016.
- [29] [Last seen on: 13-Apr-2021] <https://motionarray.com/stock-video/dentists-working-on-female-patient-101661>
- [30] [Last seen on: 31-Apr-2021] <https://www.mirror.co.uk/news/uk-news/more-eight-fillings-could-raise-8933885>
- [31] [Last seen on: 31-Mar-2021] <https://www.quora.com/How-do-you-remove-ridges-from-your-teeth>
- [32] [Last seen on: 31-Jan-2020] [https://en.wikipedia.org/wiki/Fixation \(visual\)](https://en.wikipedia.org/wiki/Fixation_(visual))

- [33] [Last seen on: 31-Jan-2020] <https://en.wikipedia.org/wiki/Saccade>.
- [34] S. Noor, H.M. Minhas, M.I. Saleem, V. Uddin and N. Ismat, 'Inside-out Vision for Procedure Recognition in Dental Environment', 2020 Global Conference on Wireless & Optical Technologies (GCWOT), Malaga, Spain 6–8 October 2020, pp. 1–8, doi: 10.1109/GCWOT49901.2020.9391594
- [35] D.E. Rumelhart, G.E. Hinton, R.J. Williams, 'Learning representations by back-propagating errors', *Nat. Int. Wkly. J. Sci.*, 1986, 323, pp. 533–536.
- [36] S. Noor and V. Uddin, 'Using ANN for Multi-view Activity Recognition in Indoor Environment,' 14th International Conference on Frontiers of Information Technology (FIT-2016), 19–21 December 2016.
- [37] S. Noor and V. Uddin, 'Using context from inside-out vision for improved activity recognition', *IET Computer Vision*, vol. 12, no. 3, pp. 276–287, March 2018.
- [38] [Last seen on: 31-Mar-2021] <https://www.newmouth.com/dentistry/restorative/crowns/>
- [39] [Last seen on: 31-Mar-2021] <https://www.preferredentalcaresantarosa.com/cavity-filling-the-procedure-aftercare-and-long-lasting/>
- [40] [Last seen on: 31-Mar-2021] <https://www.knoxvillesmiles.com/teeth-whitening/>
- [41] [Last seen on: 31-Mar-2021] <https://decisionsindentistry.com/article/ovel-technique-placing-sealants/>
- [42] [Last seen on: 31-Mar-2021] <https://www.livescience.com/44223-cavities-tooth-decay.html>
- [43] [Last seen on: 31-Mar-2021] <https://www.southfloridadentalcare.com/category/crackedbroken-teeth/>
- [44] [Last seen on: 31-Mar-2021] <https://wmsmile.com/what-is-tartar/>
- [45] [Last seen on: 31-Mar-2021] <https://www.dentakademi.com.tr/en/what-is-dental-calculus-or-tartar/>
- [46] [Last seen on: 31-Mar-2021] https://www.toothandtips.com/doctor-my-crown-is-chipped-can-we-repairit/?__cf_chl_jschl_tk__=6e6fa31508b62410186d274e4dfc601cd01f997e-1617149475-0-Aekv7u0U6angfNgX8CyPYUF2Y0AacvhyxOrvH5ILRaiZxB23rGokBHWxNuNEe8kJ3rT_PQGVmUI5IxdR_hYw7yptzkeC1SM8G5hrZJqqO4wxa94zuJgIF8oJXbAkV2DuXyR35-UXW9To_o5G04WHjc2A-uhZY7PuSn-50l8HG2EQH7S_LBz2fDZVuw2cWKStbyuBd4iXhEPmVhj0FFGUA7jyrUWtBBnGvZIPXM5DG5wh1PnI8RnjORXxD8CrtIoA5t5rXyF_HqS wz0vyF5EjKMiz2mAHRIFG4HzBIowoBNLyAixOP0HTNYMp3Gz

xngGpKaQc1M8mhGCP8XnaV9cEE2LtoDtHbScOLoO2cKQfFKMJr
m5tpWq-9QVo6Krzwf4m184ECPBQhvC4WtQJ_LGRj5BM3iBxUY
Lrs6TSMbt4zFKXx3vcyOJSqR3mW4vpulxBHakrjvERKaGKIvSt
tMs0jK4RQeS0S1bFqQCrSG4zUIFs0TO-KvUMeSuVR7gg

- [47] [Last seen on: 31-Mar-2021] <http://www.welcomedentistry.net/blog/post/a-stainless-steel-crown-could-extend-the-life-of-a-primary-molar.html>
- [48] [Last seen on: 31-Mar-2021] <https://en.aliradar.com/item/32930788236-original-teeth-whitening-44-carbamide-peroxide-dental-bleaching-system-oral-gel-kit-tooth-whitener-dental-equipment-care-oral-h>
- [49] [Last seen on: 31-Mar-2021] <https://www.myfamilydentistry.com/blog/5-ways-prevent-chlorhexidine-staining/>
- [50] [Last seen on: 20-Feb-2020] <https://www.waterpik.com/oral-health/pro/dental-supplies/dental-instruments/densco-condenser-accessories/>
- [51] [Last Seen on: 27 June 2017] <http://easynn.com/>

Biographies



Shaheena Noor graduated from Hamdard University with a BS degree in Computer Engineering in 2003, the MS degree in computer systems from NED University of Engineering & Technology in 2007, and the PhD degree in Computer Engineering from Hamdard University in 2019, respectively. She is employed as an Assistant Professor at the Department of Computer Engineering, Sir Syed University of Engineering & Technology (SSUET), Karachi, Pakistan. Object recognition, activity recognition and prediction are among her research interest.



Muhammad Aamir was born on 3 July, 1976 in Karachi Pakistan. In 1998, he received BS Electronic Engineering degree and in 2002 his MS degree in Electronic Engineering (with specialization in Telecommunication). He accomplished his PhD in Electronic Engineering from Mehran University of Engineering & Technology. During his PhD studies, he accomplished his research work at the University of Malaga under Erasmus Mundus Scholarship. He has authored and co-authored around 50 research papers and book chapters published in various journals, books and conferences of international repute. For the past 12 Years he is a life member of Pakistan Engineering Council and professional member of IEEE. He was awarded with a grant by the Ministry of Education Spain to teach at the University of Malaga which he successfully availed in May 2012. He is also Member of two separate National Curriculum Revision Committees constituted by Higher Education Commission (HEC) for revision of Electronic Engineering Curriculum and Telecommunication Engineering Curriculum at the National Level. He was served as guest editor for special issue of Springer's Journal with title "Wireless Personal Communication" which had published in November 2016. He is also HEC approved supervisor for Pakistani PhD candidates. He is currently employed as a Professor and Associate Dean in the Faculty of Electrical & Computer Engineering at SSUET. Moreover, he is also Editor-in-Chief of Sir Syed University Research Journal of Engineering & Technology which is HEC-Recognized Research Journal which is published bi-annually.



Najma Ismat is an alumni of SSUET. She has received her post graduate degree in 2018. Dr. Ismat did her undergraduate and graduate degrees in year 1998 and 2002 respectively. Her current research interests are mobility, reliability, connectivity and coverage issues in Underater Sensor Networks, Wireless Sensor Networks, and IoT.



Engr. Muhammad Imran Saleem is currently working as an Assistant Professor in the Department of Computer Engineering at SSUET. He is associated with the university since January 2001. He is doing Ph.D. in Telecommunication Engineering from University of Malaga Spain. He did Masters (M.S) in Computer Engineering with specialization in Computer Network from SSUET. His thesis topic was Differentiated and Integrated services of IP packet. He did Bachelor (B.S) in Electronic Engineering from SSUET.