2020 17th International Multi-Conference on Systems, Signals & Devices (SSD) | 978-1-7281-1080-6/20/$31.00 ©2020 IEEE | DOI: 10.1109/SSD49366.2020.9364244

Hand-gesture-based Touchless Exploration of Medical Images with Leap Motion Controller

Safa AMEUR
Université de Sfax, SETIT- Research Unit, of Sciences of Electronic, Technologies of Information and Telecommunication; 3038, Safx Tunisie;
Université de Sousse, Ecole Nationale d'Ingénieurs de Sousse, 4023, Sousse, Tunisie;
Safa.Ameur@setit.rnu.tn

Anouar BEN KHALIFA*
Université de Sousse, Ecole Nationale d'Ingénieurs de Sousse, LATIS- Laboratory of Advanced Technology and Intelligent Systems, 4023, Sousse, Tunisie;
anouar.benkhalifa@eniso.rnu.tn

Med Salim BOUHLEL
Université de Sfax, Institut supérieure de Biotechnologie de Sfax, SETIT- Research Unit of Sciences of Electronic, Technologies of Information and Telecommunication, 3038, Safx Tunisie;
medsalim.bouhlel@enis.rnu.tn

***Abstract*—Hand gesture recognition has become one of the most interesting means of contactless human-computer interaction. There is significant importance for commanding medical images during surgical procedures by the mean of touchless hand gestures for reducing the time of surgery and the risk of contamination. In this work, we used the Leap Motion Controller as an acquisition device, with different classification methods, to recognize 11 hand gestures dedicated to manipulating medical images through a touchless graphical user interface. This framework was trained and tested on a benchmark dataset called LeapGestureDB. We worked with statistical features calculated from fingers and hand data, then normalized and fed into various classifiers such as the support vector machine, the nearest neighbor, the decision tree, the random forest, the AdaBoost, the linear discriminant analysis and the multi-layer perceptron. The highest accuracy was 91.73% and 89.91% using the cubic SVM and the multi-layer perceptron, respectively. We developed a contactless interface based on the best recognition rate in order to facilitate the way of interaction with medical images in the operating room.**

***Index Terms*—Leap Motion, Human-computer interaction, Touchless interaction, Machine learning, Hand gesture recognition, Surgery.**

I. Introduction and background

Nowadays, the desire for interaction with intelligent systems is growing. In this regard, gesture recognition research has gradually depended on the development of systems capable of recognizing human gestures and interpreting them in order to enhance the user experience [1]–[3], [35]–[37], [39]. Recently with a simple movement of the hand, it would be possible to launch any application or action within an ecosystem of connected objects [33], [38]. In this context, we find the hand gesture device: Leap Motion Controller (LMC). This sensor offers users the opportunity to interact with computers in a contactless way. It is able to recognize and follow hands, fingers, bones and finger-shaped tools. It works in close proximity with great precision higher than 0.01 mm [4] and detects 3D-positions, movements, and some gestures. The LMC consist of three infrared LEDs that mainly illuminates the scene and two infrared cameras, that deliver hand data with a frame rate averaging 100 fps depending on the computer setting [5]. The technology could be interesting for applications that need to interact in a safe way like in an operating room, where surgeons should maintain sterility. Nevertheless, during the surgery, the surgeon may need to consult the scanner or MRI images of the patient's abnormality archived in the computer or make a new image of the affected limb during the operation. In these cases, surgeon needs to manipulate these images on a computer, which is not sterile. Traditionally, if the surgeon wants to browse through these images during surgery, he has two choices: He can either stop the surgery, get to the computer, process the image he needs, decontaminate himself again and return to the patient to continue surgery. Otherwise, he can give very detailed instructions to a medical assistant so that he manipulates the images for him. These ways of working are very laborious and likely to lead to misunderstandings, causing frustration and loss of time [6], [7].

Hence, we opt for a touchless gestural interaction to perform manipulations on the medical images and thus save valuable time. To cope with this problematic, the state of the research systems proposed various modalities such as voice recognition [8], eye movements [9], hand gesture [10], emotion detection [11] and human action interpretation [12] to command computers. The majority of these techniques utilized touchless sensors

Corresponding author at: ENISo, 4023, Sousse, Tunisie. Email address: anouar.benkhalifa@eniso.rnu.tn

like RGB cameras [13], [14], time of flight camera [15], Myo armband [16], accelerometer sensor [17] and depth sensors like Kinect [18]–[20], [34] and Leap Motion [7], [21]–[24], [31].

For the touchless interaction with digital equipment in operating rooms, several studies were presented in the literature: Opromolla et al. [21] studied a hand gesture recognition system based on the LMC to mainly visualize and manipulate medical images throughout the surgical procedures. In addition, Gallo et al. [20] utilized human body parts to explore medical imaging in a contactless way through a Microsoft Kinect. In the same context, Jacob et al. [25] used the whole body skeleton captured by the Microsoft Kinect specifically the torso, the head orientation, and the hands' position.

Regarding the LMC high precision during the tracking process and its reduced dimension, the sensor would be perfectly suitable for an operating room environment. Bizzotto et al. [26] presented their first experience with the LMC to explore a medical image viewer. They worked with a free application named "GameWave App" available in the LMC online store to define their gestures and command the medical images. The same framework was presented by Pauchot et al. [27]. They exploited the LMC and the "Carestream Software" to control images. They propose a detailed configuration of the framework and a description of the test experience. In another way, Ebert et al. [28] put forward a touchless system for manipulating medical images with a two-handed gesture set using the Leap Motion device. Furthermore, Sanguannarm et al. [22] utilized the LMC to classify ten gestures in order to process medical images in surgical procedures. They developed a graphical interface presenting the hand color map, the recognized gesture, and the processed image. Likewise, Cho et al. [23] evaluated a touchless interface using "GestureHook software" to classify five gestures. They recorded good accuracy with Support Vector Machine (SVM) and Naïve Bayes classifiers which make the system adequate for applying in OR. Moreover, Rosa et al. [29] evaluated, during 11 dental surgeries, the effectiveness of a contact-free control of medical image viewer with the Leap Motion sensor . Gestures were about commands to scale, rotate, window and browse medical images.

In this paper, we develop a touchless graphical user interface based on the LMC, which offers a new experience for the surgeon to command medical images named DICOM images (Digital Imaging and COmmunications in Medicine). The framework relies essentially on a strong recognition approach which consists of extracting statistical features like the mean and the standard deviation from the LMC raw data. Then, we train our system on a public dataset composed of 11 gestures dedicated to command DICOM images. After that, we exploit multiple classification methods to evaluate the system performance such as the linear SVM, the cubic SVM, the Gaussian SVM, the Decision Tree (DT), the Random Forest (RF), the k-Nearest Neighbor (KNN), the AdaBoost, the Linear Discriminant Analysis (LDA) and the Multi-Layer Perceptron (MLP). It is noteworthy that all these methods achieve good recognition rates ranges between 74% and 91%. We pick out the best classification model in order to recognize the performed gesture with minimum recognition error.

In a nutshell, our contributions are the following:

- The evaluation and the comparison of a hand gesture recognition approach with the most used classification methods based on the extraction of some statistical features from the LMC raw data.
- Developing a real-time graphical user interface for DICOM images manipulation in surgical procedures according to the best hand gesture recognition model.

The rest of this paper is organized as follows: the system design is described in section II using four subsections. The first one gives an overview of the proposed system. The second subsection describes the dataset utilized. The third and fourth subsections present the feature extraction and the classification methods adopted for the recognition step, respectively. Section III presents the experiment results. We start by the gesture recognition results in the first subsection, then, usability evaluation of the graphical user interface in the second one. Section IV concludes the article.

II. System design

A. System overview

Fig.1 describes the overall structure of the proposed system, starting with the existing problems in operating rooms when the surgeon needs to manipulate DICOM images. Then, it gives a detailed overview of the suggested touchless gesture interface as an alternative to cope with the critical situation. The framework is composed of two main parts based on the Leap Motion as an input device: the software module for hand gesture recognition and the graphical module for data visualization and handling.

B. Study dataset

We use the public LeapGestureDB dataset hosted online in [30] to train and assess our system. It is the unique public dataset gathered with the LMC in the medical field. It is composed of 11 hand gestures that correspond to a set of indispensable commands chosen by the surgical staff from the university hospital Sahloul in Tunisia to manipulate the DICOM images [24].

Gestures description are illustrated in Fig.2. Each of the following 11 gestures was repeated five times by 120

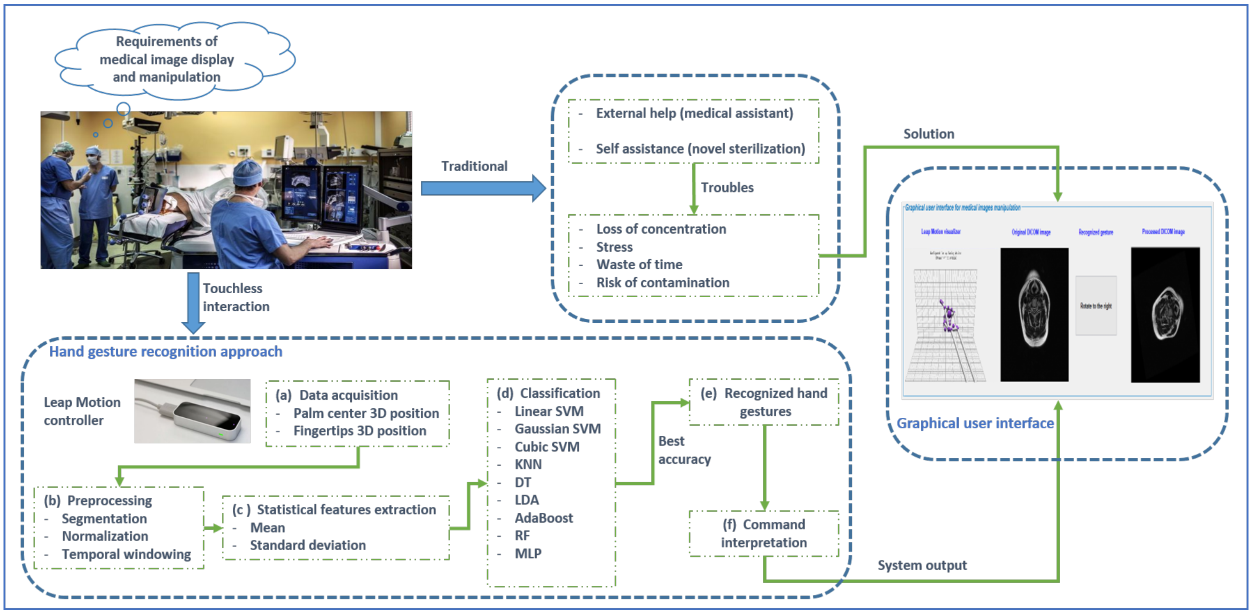

Fig. 1. Flowchart of the proposed framework.

subjects. In total 6600 observations was recorded while collecting the dataset:

- G1: make a thumb up to select something.
- G2 & G3: rotate the image to the left or the right, respectively. It is by drawing a circle in the air in the counter-clockwise direction or in the clockwise direction, respectively.
- G4 & G5: increase or decrease, respectively the image's contrast.
- G6 & G7: zoom in or out, respectively the DICOM image.
- G8 & G9: make a horizontal movement of the image to the left or the right, respectively.
- G10 & G11: display the previous image or the next image, respectively.

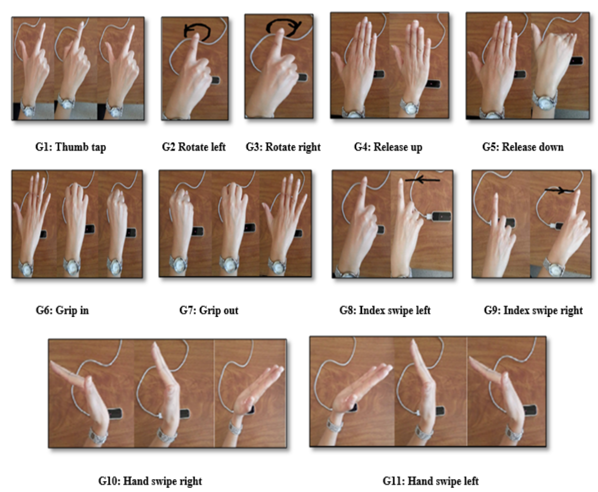

Fig. 2. All gestures in the LeapGestureDB dataset for interacting with DICOM images.

C. Feature extraction

The LMC sensor provides skeleton hand and finger data useful for gesture recognition step during the tracking process. In our previous work [5], [24], we have evaluated various features and even with a fusion of feature sets. We worked with statistical features like the mean, the root mean square, the covariance, and the standard deviation . Furthermore, we have extracted the fast Fourier transform as a frequential feature and the discrete wavelet transform as a spatio-frequential feature. In this work, we chose to work with the six important points on the hand through their coordinates (X, Y, Z), which are the palm center and five fingertips. Those data are then saved in matrices L = [Pa; F1; F2; F3; F4; F5] of six rows and N columns. As a preprocessing step, we make segmentation of 25% from the original acquired signal to eliminate the excess data and ameliorate the accuracy recognition rate. Then, we use the temporal windowing technique 5 window for each repetition). Then, the standard deviation and the mean is calculated in each window. Finally, we got 180 features. This set of features is calculated and normalized in the range [-1, 1], to be fed into multiple classifiers for recognizing the gesture. These features are calculated thanks to the following equations:

$$Mean = \frac{1}{N}\sum_{j=0}^{N} L_j \tag{1}$$

$$Standard\_deviation = \sqrt{\frac{1}{N-1}\sum_{j=1}^{N}(L_j - Mean)^2} \tag{2}$$

D. *Classification*

For the classification step, we conduct several experiments using a variety of classifiers on two bechmark datasets the LeapGestureDB dataset [24] and the RIT dataset [32]. To train and predict the output classes, we compare the performance of most used classifiers in the literature such as the linear SVM, the cubic SVM, the Gaussian SVM, the KNN, the DT, the RF, the AdaBoost, the LDA, and the MLP. The accuracy rates of different methods are summarized in Fig3. It is noteworthy that good recognition results were achieved using all classifiers. It ranges between 74% and 92% for the LeapGestureDB dataset, and between 59% and 87% for the RIT dataset. The cubic SVM classifier outperforms all other classifiers for both datasets.

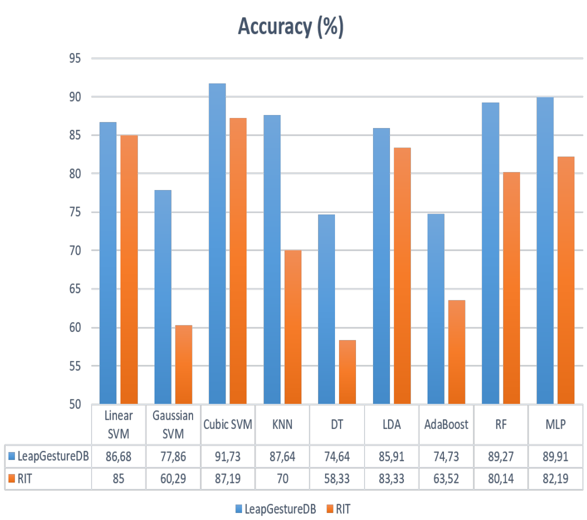

Fig. 3. Performane evaluation with differents classification methods.

III. Experiments

A. *Qualitative hand gesture recognition evaluation*

The best accuracy was reached using the cubic SVM classifier with an accuracy rate of about 91.73% with a processing time equal to 7 ms per gesture. For this reason, we choose to work with the trained cubic SVM model to develop the graphical user interface for the hand gesture recognition step. Fig.4 illustrates the confusion matrix for the prediction of the 11 hand motions with the cubic SVM classifier. We note that the majority of gestures are well recognized with an accuracy above 90%, except for gestures 10 and 11 where the error rate is about 20%. These two gestures present a big confusion between each other and even with other different classes. This is can be explained by the fact that the gestures hand swipe left and hand swipe right involve the use of the whole hand. Similarly, there is still a little confusion between reciprocal gestures especially those performed with the same fingers, like gestures 2, 3, 8 and 9, which essentially use the index finger.

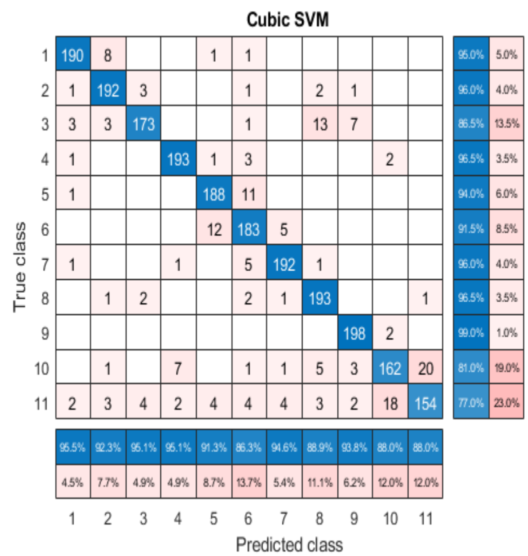

Fig. 4. Confusion matrix for qualitative evaluation for the recognition of 11 hand gestures with Cubic SVM.

B. *Usability evalution of the graphical user inteface*

We develop a simple graphical user interface for manipulation of DICOM images. It is composed essentially of a display zone to visualize the hand skeleton tracked with the LMC sensor, another display zone to show the original image, a learning module based on the cubic SVM as a classifier, for recognizing the adequate gesture, and another display zone depicts the processed DICOM image after the execution of the correspondent command. Fig.5 shows the graphical interface when performing a right rotation of the DICOM image.

In this study, we conduct several experiments to evaluate our interface in all scenarios. The majority of gestures are well recognized and interpreted. Except for browsing images between next and previous ones. An average of 2 out of 10 repetitions failed. This study

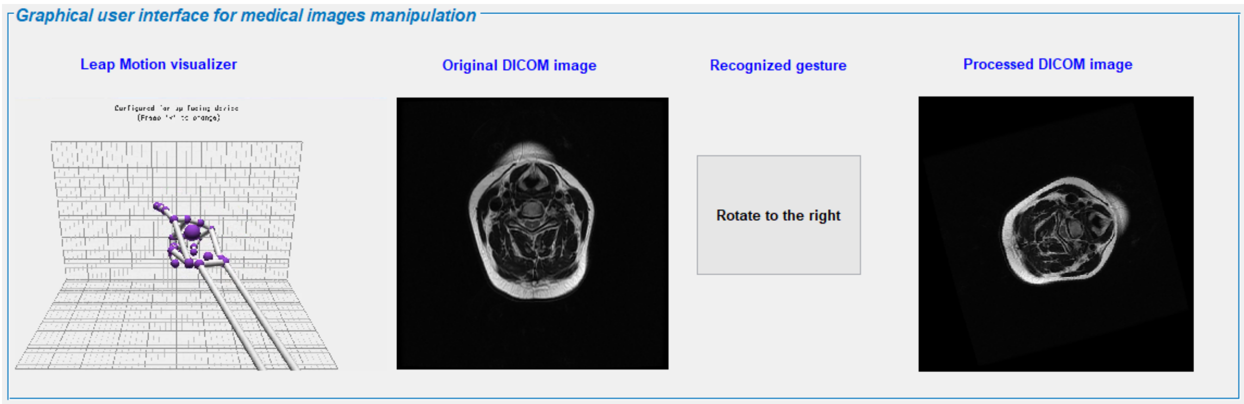

Fig. 5. Graphical user interface for hand gesture recognition and command interpretation for manipulating DICOM images.

presents several limitations such as the extraction of hand skeleton data when occlusion occurs. For example, in the case of a vertical pose of the fingers or when fingers overlap, or the user's hand disappears suddenly. In these scenarios, the extraction of features from the hand skeleton would be difficult. Besides, we note that if the user is not doing any gesture during a period, this pose will be interpreted as a silent moment and no gesture will be recognized.

IV. Conclusion

Touchless interaction through intuitive hand gestures is a great advantage for the surgeon who wants to control medical images during surgical procedures. In this study, we proposed a simple and accurate implementation of the machine learning method for hand gesture recognition tasks. We have evaluated and compared multiple classification methods, to finally choose the best recognition model to develop a touchless real-time graphical user interface for medical image manipulation based on this hand recognition approach. We showed that even our method is simple, it reached high accuracy, which makes it suitable for use in operating rooms. For further study, we propose to implement deep learning techniques for hand gesture recognition with Leap Motion, which may enhance the recognition rates and compensate for the machine learning weakness.

References

[1] I. Jegham, A. Ben Khalifa," Pedestrian Detection in Poor Weather Conditions Using Moving Camera," IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA), pp. 358 - 362, 2017.

[2] W. Lejmi, A. Ben Khalifa, M.A. Mahjoub, "Fusion Strategies for Recognition of Violence Actions," IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA), pp. 178 - 183, 2017.

[3] I. Jegham, A. Ben Khalifa, I. ALOUANI, M.A. MAHJOUB, " Safe Driving : Driver Action Recognition using SURF Keypoints," The 30th International Conference on Microelectronics (ICM2018), pp. 60 - 63, 2018.

[4] F. Weichert, D. Bachmann, B. Rudak, D. Fisseler,"Analysis of the accuracy and robustness of the leap motion controller," Sensors, vol.13, pp. 6380–6393, 2013.

[5] S. Ameur, A. Ben Khalifa, M.S. Bouhlel, "A Comprehensive Leap Motion Database for Hand Gesture Recognition," In 7th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT), pp.514–519. IEEE, Tunisia, 2016

[6] RC. Mackay, K. Matsuno, and J. Mulligan, "Communication problems between doctors and nurses," Qual Assur Health Care, 3(1):119, 1991.

[7] L.D. Tommaso, S. Aubry, J. Godard, H. Katranji and J. Pauchot, "a new human machine interface in neurosurgery: the Leap Motion®. technical note regarding a new touchless interface," 2016.

[8] H.M. Mentis, K. O'Hara, G. Gonzalez, A. Sellen, R. Corish, A. Criminisi, R. Trivedi, P. Theodore," Voice or gesture in the operating room," In: Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems, ACM, pp 773– 780, 2015.

[9] Y. Yanagihara, and H. Hama, "System for selecting and generating images controlled by eye movements applicable to CT image display," Med. Imaging Technol, 18:725-734, 2000.

[10] M.G.Jacob, J.P. Wachs, "Context-based hand gesture recognition for the operating room," Pattern Recognition Letters 36:196–203, 2014.

[11] S. Lalitha, S. Patnaik, T. H. Arvind, V. Madhusudhan and S. Tripathi, "Emotion Recognition through Speech Signal for Human-Computer Interaction," 2014 Fifth International Symposium on Electronic System Design, Surathkal, pp. 217-218, 2014.

[12] S. Jalaliniya, J. Smith, M. Sousa, L. Buthe, T. Pederson," Touch-less interaction with medical images using hand and foot gestures," In: Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication, ACM, pp 1265–1274, 2013.

[13] D.L.M. Achacon, D.M. Carlos, M.K. Puyaoan, C.T. Clarin, Jr. P.C. Naval," Realism: Real-time hand gesture interface for surgeons and medical experts," In: 9th Philippine Computing Science Congress, Citeseer, 2009.

[14] S. Yoshida, M. Tatokoro, M. Yokoyama, J. Ishioka, Y. Matsuoka, N. Numao, K. Saito, Y. Fujii, K. Kihara, "Multitask imaging monitor for surgical navigation: Combination of touchless interface and head-mounted display," Urologia internationalis, 2015.

[15] S. Soutschek,J. Penne, J. Hornegger, J. Kornhuber, " 3-d gesture-based scene navigation in medical imaging applications us-

ing time-of-flight cameras," In: Computer Vision and Pattern Recognition Workshops. CVPRW'08. IEEE Computer Society Conference on, IEEE, pp 1–6, 2008.

[16] J. Hettig, A. Mewes, O. Riabikin, M. Skalej, B. Preim, C . Hansen, " Exploration of 3D Medical Image Data for Interventional Radiology using Myoelectric Gesture Control," In: Proceedings of Eurographics Workshop on Visual Computing for Biology and Medicine, The Eurographics Association, pp 177–185, 2015.

[17] A. Mimouna, A. Ben Khalifa, N. Essoukri Ben Amara, "Human action recognition using triaxial accelerometer data : selective approach," 15th International Multi-Conference on Systems, Signals & Devices (SSD'18), pp. 467-472, 2018.

[18] R. Wipfli,V. Dubois-Ferriere, S.Budry, P. Hoffmeye, C. Lovis, "Gesture-controlled image management for operating room: A randomized crossover study to compare interaction using gestures, mouse, and third person relaying," PloS one 11(4):e0153,596, 2016.

[19] E. Nasr-Esfahani, N. Karimi, S. M. Soroushmehr, M. H. Jafari, M. A.Khorsandi, S. Samavi, and K. Najarian," Hand Gesture Recognition for Contactless Device Control in Operating Rooms," arXiv preprint arXiv:1611.04138, 2016.

[20] L. Gallo, A. P. Placitelli and M. Ciampi, "Controller-free exploration of medical image data: Experiencing the Kinect," 2011 24th International Symposium on Computer-Based Medical Systems (CBMS), Bristol, pp. 1-6, 2011.

[21] A. Opromolla, V.Volpi, A. Ingrosso, S. Fabri, C. Rapuano, D. Passalacqua, and C.M. Medaglia, " A Usability Study of a Gesture Recognition System Applied During the Surgical Procedures," HCI, 2015.

[22] P. Sa-nguannarm, T. Charoenpong, C. Chianrabutra and K. Kiatsoontorn, "A Method of 3D Hand Movement Recognition by a Leap Motion Sensor for Controlling Medical Image in an Operating Room," 2019 First International Symposium on Instrumentation, Control, Artificial Intelligence, and Robotics (ICA-SYMP), Bangkok, Thailand, pp. 17-20, 2019.

[23] C. Yongwon, L. Areum, P. Jongha, K. Bemseok, K. Namkug, "Enhancement of gesture recognition for contactless interface using a personalized classifier in the operating room," Computer Methods and Programs in Biomedicine,Volume 161, pp.39-44, 2018.

[24] S. Ameur, A. Ben Khalifa, M.S. Bouhlel, " LeapGestureDB: A Public Leap Motion Database Applied for Dynamic Hand Gesture Recognition in Surgical Procedures," 8th International Workshop on Soft Computing Applications (SOFA'18), Arad-Romania, pp. 1-14, 2018.

[25] M. Jacob, C. Cange, R. Packer, J.P. Wachs," Intention, context and gesture recognition for sterile MRI navigation in the operating room," In: Alvarez, L., Mejail, M., Gomez, L., Jacobo, J. (eds.) CIARP 2012. LNCS, vol. 7441, pp. 220–227. Springer, Heidelberg, 2012.

[26] N. Bizzotto, A. Costanzo, L. Bizzotto, D. Regis, A. Sandri, B. Magnan, " Leap Motion Gesture Control With OsiriX in the Operating Room to Control Imaging," Surgical Innovation, 21(6), pp.655–656, 2014.

[27] J. Pauchot, L. Di Tommaso, A. Lounis, M. Benassarou, P. Mathieu, D. Bernot, S. Aubry," Leap Motion Gesture Control With Carestream Software in the Operating Room to Control Imaging," Surgical Innovation, 22(6), pp.615–620, 2015.

[28] L. Ebert, P. Flach, M. Thali, S. Ross," Out of touch–a plugin for controlling osirix with gestures using the leap controller," Journal of Forensic Radiology and Imaging, 2(3):126–128, 2014.

[29] G.M. Rosa, M.L. Elizondo," Use of a gesture user interface as a touchless image navigation system in dental surgery: Case series report," Imaging science in dentistry 44(2):155–160, 2014.

[30] LeapGestureDB dataset Main page. https://sites.google.com/view/leapmotiondatabase/accueil

[31] M. Mohandes, S. Aliyu and M. Deriche, "Prototype Arabic Sign language recognition using multi-sensor data fusion of two leap motion controllers," 2015 IEEE 12th International Multi-Conference on Systems, Signals & Devices (SSD15), Mahdia, pp. 1-6, 2015.

[32] McCartney, R., Yuan, J., Bischof, H.P, " Gesture Recognition with the Leap Motion Controller," In the International Conference on Image Processing, Computer Vision, and Pattern Recognition, 2015.

[33] Imen Jegham, Anouar Ben Khalifa, Ihsen Alouani, Mohamed Ali Mahjoub, "Vision-based human action recognition: An overview and real world challenges", Forensic Science International: Digital Investigation, Volume 32, pp. 200901, 2020.

[34] Ben Haj Khaled A., Khalfallah A., Bouhlel M.S.,"Fall Prevention Exergame Using Occupational Therapy Based on Kinect," In: Proceedings of the 8th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT'18), Vol.1. SETIT 2018. Smart Innovation, Systems and Technologies, vol 146. Springer, Cham, 2020.

[35] N. Jaouedi, N. Boujnah, O. Htiwich and M. S. Bouhlel, "Human action recognition to human behavior analysis," 7th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT), Hammamet, pp. 263-266, 2016.

[36] N. Jaouedi, N. Boujnah,and M. S. Bouhlel, "Deep Learning Approach for Human Action Recognition Using Gated Recurrent Unit Neural Networks and Motion Analysis," Journal of Computer Science, Vol.15 (7), pp. 1040-1049, 2019.

[37] N. Jaouedi, N. Boujnah,and M. S. Bouhlel, " A new hybrid deep learning model for human action recognition," Journal of King Saud University - Computer and Information Sciences, 2019.

[38] S. Ammar, N. Zaghden, M. Neji, " An Effective Approach Based on a Subset of Skeleton Joints for Two-Person Interaction Recognition," In: Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications. CIARP 2018. Lecture Notes in Computer Science, vol 11401. Springer, Cham, 2019.

[39] M. Farhat, A. Khalfallah, and M. Bouhlel, "A Real-Time Court Detection and Tracking System to Tennis Videos," Journal of Testing and Evaluation 47, no. 4, pp.3103-3122, 2019.