**A Gesture-based Tool for Sterile Browsing of Radiology Images**

**Literature Survey**

1. **A Non-Contact Mouse for Surgeon-Computer Interaction.**

Grätzel, C & Fong, T & Grange, Sebastien & Baur, Charles. (2004). A Non-Contact Mouse for Surgeon-Computer Interaction. Technology and health care : official journal of the European Society for Engineering and Medicine. 12. 245-57. 10.3233/THC-2004-12304.

The use of doctor-computer interaction devices in the operation room (OR) requires new modalities that support medical imaging manipulation while allowing doctors' hands to remain sterile, supporting their focus of attention, and providing fast response times. This paper presents "Gestix," a vision-based hand gesture capture and recognition system that interprets in real-time the user's gestures for navigation and manipulation of images in an electronic medical record (EMR) database. Navigation and other gestures are translated to commands based on their temporal trajectories, through video capture. "Gestix" was tested during a brain biopsy procedure. In the in vivo experiment, this interface prevented the surgeon's focus shift and change of location while achieving a rapid intuitive reaction and easy interaction. Data from two usability tests provide insights and implications regarding human-computer interaction based on nonverbal conversational modalities.

1. **Dynamic Sign Language Recognition Based on Video Sequence With BLSTM-3D Residual Networks**

Y. Liao, P. Xiong, W. Min, W. Min and J. Lu, "Dynamic Sign Language Recognition Based on Video Sequence With BLSTM-3D Residual Networks," in IEEE, vol. 7, pp. 38044-38054, 2019, doi: 10.1109/ACCESS.2019.2904749.

Y. Liao, et al. presents a multimodal dynamic sign language recognition technique based on 3-dimensional Residual ConvNets and bi-directional LSTM. First, the position of the hands captured from the video by utilizing Fast R-CNN, and the captured video frames were sent to the video sequence feature extraction module for spatiotemporal feature extraction. The feature vectors were trained using the B3D ResNet model. These feature vectors were then provided to the dynamic sign language recognition module in order to predict the dynamic sign language. As a result, it was able to distinguish complicated hand gestures from longer video sequences with greater accuracy, with an accuracy of 89.8% on the DEVISIGN D dataset and 86.9% on the SLR Dataset. Future studies could go towards signing language detection for real-world video in complex environments.

1. **A CNN based human computer interface for American Sign Language recognition for hearing-impaired individuals.**

Ahmed KASAPBAS, Ahmed Eltayeb AHMED ELBUSHRA, Omar AL-HARDANEE, Arif YILMAZ, "DeepASLR: A CNN based human computer interface for American Sign Language recognition for hearing-impaired individuals", Computer Methods and Programs in Biomedicine Update, Volume 2, 2020.

Ahmed KASAPBAS, et al. developed a dataset and a Convolutional Neural Network-based sign language interface system for interpreting sign language gestures. ASLA was represented in the dataset by images and corresponding letters. The datasets were split into 70% for training sets, 15% for validation sets and 15% for testing sets. The OpenCV library was used for video capture due to its performance and robustness. The image capture window, which had dimensions of 196 x 196 pixels Signs was to be performed within this predefined region. Images are captured with a resolution of 64 \* 64, which was necessary for the CNN algorithm. The threshold image was displayed in the preview window after an image is captured from the predefined region. The threshold display capture images were saved and fed to the CNN model for classification. Finally, the predicted character appeared in the image capture window's preset region. The outcomes of the experiments revealed that datasets have a greater accuracy. This work could be enhanced by include additional images in the dataset for more letters and words.

1. **A Modified LSTM Model for Continuous Sign Language Recognition Using Leap Motion.**

Mittal, P. Kumar, P. P. Roy, R. Balasubramanian and B. B. Chaudhuri, "A Modified LSTM Model for Continuous Sign Language Recognition Using Leap Motion," in IEEE Sensors Journal, vol. 19, no. 16, pp. 7056-7063, 15 Aug.15, 2019, doi: 10.1109/JSEN.2019.2909837.

Mittal, et al. [2], the purpose of this paper was to propose an LSTM-based neural network architecture for recognizing sign language using a leap motion sensor. This system used the sensor's API to extract the 3D positions of the fingertips. A modified LSTM model mapped the input sentence to the output sentence. The dataset consisted of 35 isolated signed words and 942 signed sentences. The proposed model was found to have an accuracy of about 72.3% for signed sentences and 89.5% for isolated signed words. In the future, this model can be improved by training more data.