

1. Wachs, Juan P., et al. "A gesture-based tool for sterile browsing of radiology images." *Journal of the American Medical Informatics Association* 15.3 (2008): 321-323.

- Existing problem is the use of keyboards and mice by doctors and nurses in ICUs leading to spreading of infections. No touchless or build up free surfaces available for image manipulation in ORs.
- This paper proposes a system called "Gestix" which does hand gesture capturing and recognition and interprets the gestures into an action to be performed on MRI images (Magnetic Resonance Imaging). This paper was the first instance of hand gesture recognition system was successfully implemented in a neurosurgical biopsy.
- The architecture involved a Camera with IR remote to control room, a monitor with camera attached to it and video capturing equipment. The 2 layers involved were for gesture tracking and image recognition with "Gibson".
- The hand motions are first segmented following a color model building.
- The "Gibson" image browser is a 3D visualization medical tool that can be used for all medical images.
- Intuitive left/right/up/down/sleep/wake/zoom in/out gestures were chosen
- A questionnaire and interview was used to evaluate satisfaction metrics.

2. Ameer, Safa, Anouar Ben Khalifa, and Med Salim Bouhlef. "Hand-gesture-based touchless exploration of medical images with leap motion controller." *2020 17th International multi-conference on systems, signals & devices (SSD)*. IEEE, 2020.

- The Benchmark dataset (LeapGestureDB where 6600 observations were recorded) was used along with Leap Motion Controller as an acquisition on different classification methods, to recognize 11 hand gestures.
- ML algorithms utilized were 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 multilayer perceptron.
- Leap Motion Controller is used to recognize and follow hands, fingers, bones and finger-shaped tools with IR LEDs and cameras to illuminate the place and capture the images.
- The system is comprised of 2 modules – one for gesture recognition and one for image manipulation.

- Six features are first extracted namely the coordinates (X, Y, Z) of the hand's palm center and 5 fingertips. Segmentation is followed by calculating mean and Standard deviation of each 5-window gap.
- 180 features extracted are then fed into ML models mentioned above.

3. Lin, Hsien-I., Ming-Hsiang Hsu, and Wei-Kai Chen. "Human hand gesture recognition using a convolution neural network." *2014 IEEE International Conference on Automation Science and Engineering (CASE)*. IEEE, 2014.

- This paper proposes a Convolutional Neural Network to recognize the hand gestures from images. A Gaussian Mixture model is used to filter out non skin image tones since lighting affects the performance of the system. GMM is shown to perform better than binary segmentation based on thresholds for segmentation task.
- 7 test subjects were evaluated with seven hand gestures with average recognition accuracies around 95.96%
- The principal axis of the hand and the center calibrates the images and the images are rotated into a neutral position which in turn is fed into the CNN for recognition.
- The images were taken by a XBOX Kinect camera and were 200*200. They get rescaled to 28*28 in the preprocessing stage.
- For each gesture type of the 7, 600 and 200 images from a subject were taken to train and test.

4. Amit Gupta, Vijay Kumar and Sehrawat Mamta Khosla. "FPGA Based Real Time Human Hand Gesture Recognition System". *2nd International Conference on Communication, Computing & Security , ICCCS,2012*.

- The paper proposes a real -time hand recognition system for human computer interaction, using FPGA.
- FPGAs are Integrated Circuits that can be customized to the needs of the application being used in. This paper utilizes FPGA to store the image obtained from the user, re-size it to a different resolution to adjust it rightly at the region of interest. It also claims that FPGAs attribute to fast processing of the image.
- It can classify 10 different hand gestures based in shape based features, at a faster rate with a reasonable accuracy.
- The process is broken into 3 phases, image acquisition, image pre processing and gesture recognition.

- In the phase of image acquisition, DDR RAM mounted on the FGPA is used. It is then re-sized .
- In the phase of pre-processing, illumination compensation, to recognize the gesture in altering background lightings and skin color segmentation, to minimize the chances of false recognition are done.
- Finally, in the phase of gesture recognition, features orienting to the shape of the hand, such as perimeter, area, presence of thumb, etc are notified.
- The classification algorithm used is not explicitly mentioned in the paper.
- The model has shown an accuracy of 94% when tested with a real time database, comprising of 25 images of hands of different people.
- The limitation with this paper includes the need of a hardware component FGPA. While there are several other methods to implement real-time systems with faster processing rate without the usage of an explicit IC.

5. Abhishek B1 , Kanya Krishi2 , Meghana M3 , Mohammed Daaniyaal4 , Anupama H. “Hand gesture recognition using machine learning algorithms”. Vol. 1, No. 3, November 2020, pp. 116~120 ISSN: 2722-3221, DOI: 10.11591/csit.v1i3.p116-120.

- This paper proposes a gesture based recognition system for human computer interaction using 3-D CNN.
- The image is pre-processed, identified of the gesture, and a corresponding action - zooming/scrolling/switching of pages is made.
- The model is divided into 3 sub parts:
 - o **Hand and motion detection:** The user video of hand movement is sent to tensorflow and openCV object detector as an input. Edge detection and skin detection are done to recognise the hand and is fed into a 3-D CNN.
 - o **Dataset :** Dataset is used for training the 3D CNN. Two types of datasets are being used – one for the hand detection and the other for the motion or gesture detection. Hand detection uses EGO dataset, Motion or Gesture Recognition uses Jester dataset.
 - o **3-D CNN :** CNN are convolutional neural networks that assist in the classification of images. It consists of several layers – input layer, hidden layers and output layer. It performs back propagation for better accuracy and efficiency.
- The paper progresses in multiple ways. It uses temporal modelling or motion based gestures instead of spatial modelling or single hand gesture, This makes the interaction human- friendly.

- It is also a real time recognition system, unlike other systems which use a pre-captured image as an input.

6. Md. Zahirul Islam, Mohammad Shahadat Hossain, Raihan UI Islam. “Static Hand Gesture Recognition using Convolutional Neural Network with Data Augmentation”.

- This paper proposes a model to recognize hand gestures based on CNN.
- The input image taken from the user is initially pre-processed. This paper implements minimal pre processing, in order to keep the computation fast, without compromising the output accuracy. The background is removed and the image is converted to grayscale. Grayscale images have only 1 channel and is thus easier for CNN to learn. Median filter is applied for noise removal.
- Next, the CNN is configured. It consists of 2 convolutional layers, 2 max-pooling layers, 2 fully connected layers and 1 output layer. ReLU activation function which promotes non linearity is used, since it has better performance rate than sigmoid or tanh functions.
- The output layer has 10 nodes, each corresponding to 1 gesture.
- The paper utilizes two different datasets - base and augmented.
- Base dataset refers to the set of hand gestures taken as pictures. Data augmentation is a technique to increase the number of data by applying zoom, shear, rotation, flip etc. This process not only increases the data but also brings variation in dataset which is essential for CNN to learn sophisticated differences of images. The augmented dataset consists of all images to which augmentation technique is done.
- While testing, it was observed that the model trained with base dataset produces an accuracy of 92.87% and with augmentation, 97.12%.
- Hence, clearly, augmentation of data made a huge impact in learning.
