IDEATION PHASE

Literature survey

Date	19 september 2022
Team ID	PNT2022TMID00221
Project name	A GESTURE-BASED TOOL FOR STERILE BROWSING OF RADIOLOGY IMAGES.

A gesture-based tool for sterile browsing of radiology images

OBJECTIVE :

Humans are adapt at reading body language and signs. This is conceivable because of how vision and synaptic contacts developed throughout the course of brain development. There are a few issues that need to be resolved in order to reproduce this ability in computers, including how to distinguish things of interest in photographs and which image capture method and classification techniques are more suitable.

In this gesture-based desktop automation project, the model is first pre-trained using photos of various hand motions, such as showing the numbers 1 through 4 with the fingers. The video frame is captured by this model using the built-in webcam. The image of the gesture captured in the video frame is compared with the Pre-trained model and the gesture is identified. If the gesture prediction is 1 then images is blurred;2, image is resized; 3, image is rotated etc…

INTRODUCTION:

When operating presentation displays, such as x-rays and transmitted images in the medical industry, interactive presentation systems (IPS) employ modern Human Computer Interaction (HCI) techniques to offer a more practical and user-friendly interface. These methods greatly enhance the new experience compared to the old mouse and the keyboard interface for human-machine control. A hand gesture can be used in many different contexts. In this work, we use it to build a user-friendly interaction interface for an interactive presentation system.

PURPOSE OF THE SYSTEM:

To maintain sterility while browsing radiology images using hand gestures.

EXISTING SYSTEM:

In the twenty-first century, there has been a close interaction between social life and information technology. Future consumer electronics items will feature increasingly complicated interfaces with a wide range of functionalities. How to create a practical human machine. The user interface of every consumer electronics product is now a crucial concern. The most popular form of interaction is still using a mouse, keyboard, or joystick, which are all classic examples of electronic input devices. It does not imply, however, that these gadgets are the most practical and natural input methods for the majority of users. Gestures have been an important form of human contact and communication since ancient times. Before the development of language, people could simply convey their ideas through gestures. Many individuals still use gestures in everyday life, and deaf people in particular find that gestures are the most natural and important form of communication. The gesture control method has recently emerged as a new trend in the development of many human-based electronics products. Development of a real-time hand gesture identification system based on an adaptive colour HSV model and motion history image is the goal of this research (MHI). The influences of lighting, environment, and camera may be significantly minimised by using an adaptive skin colour model, and the robustness of hand motion detection might be significantly increased.

LITERATURE SURVEY:

First, there are a number of approaches utilized in the vision-based approach for hand detection, training gestures, background subtraction, and finger tip detection, which are reviewed as follows: Scale invariant feature transform [1][3] based hand detection as well as the feature-based hand detection methods utilized by Viola and Jones detector have been implemented. These algorithms deliver very accurate results but are more prone to background noise. The second method is picture segmentation, which instead of using the RGB color model to determine the color of human skin, employs the HSV color space model. This approach provides better results for background separation and region boundaries but is unable to identify objects with similar-colored backgrounds that have different skin tones [1]. The third method incorporates data from the same category of items using learning-based gesture recognition in the Adaptive Boosting algorithm. By integrating all weak classifiers into one strong classifier, it trains the network. The best weak classifier is chosen from a set of both positive and negative image examples by the AdaBoost learning algorithm [2]. This technique delivers results more quickly and accurately, although occasionally the network training process takes longer. Another method is to look for convex hulls. For palm detection, there are so many algorithms available. Some of the existing algorithms that are used in our suggested technique will be detailed in this section.

The Divide and Conquer algorithm, the Quick Hull algorithm, the Jarvis March or Gift-wrapping algorithm, and the Graham's Scan algorithm. Any given set of points' convex hull can be calculated using Graham Scan. There is no requirement for the user to touch or carry a peripheral device in order to deploy the system for real-time hand tracking and basic gesture detection. By comparing the results, we can draw the conclusion that using just one detection strategy is insufficient because several techniques can be used to address various detection and recognition issues. AdaBoost, support vector machine technique, hidden Markov model, and principal component analysis are some of the machine learning techniques that are now accessible for training classifiers. The boundary of the hand region may also have a different convex hull and contour detection. Using the adaptive boosting algorithm for hand detection and the Haar classifier algorithm to train the classifier, we will develop the system based on all of these techniques. Convex hull technique is used to form a contour around the palm and finger tip detection in this case, along with the HSV colour model for background reduction and noise removal.

While there are several image processing software programme available, OpenCV (Open Source for Computer Vision) is particularly well-liked for real-time image processing tasks like gesture and object recognition. The ability to simply combine the code with hardware is the main benefit. We put the suggested approach into practise using the Linux-based OpenCV library. Six 2D convolution layers make up the network, and a max-pooling operator comes after each layer. Fig displays the volumes at each layer, the convolution kernel sizes, and the pooling operators. A fully linked network with nine layers receives the output of the sixth convolution layer as input.

Except for the last output layer, which includes nine neurons—one for each of the nine hand gestures—each layer comprises 512 hidden neurons. In the output layer, a sigmoid activation function is applied. The remaining eight levels employ the Tanh activation function. The acquisition of a big dataset for each subject in the context of this article would be time-consuming and impractical when taking into account real-life applications, as a user would frequently not put up with hours of data recording for each training.

The initial stage for any system is to gather the information required to carry out a certain operation. Input data for hand posture and gesture recognition systems is collected using a variety of technologies. This project can meet the requirements of the user's need by tracking the motion or movement of the hand. At the Washington Hospital Center in Washington, DC, the gesture interface's functionality was evaluated. A hand gesture system's applicability was learned from watching two operations in the hospital's neurosurgery department. This is the first instance that we are aware of where a hand gesture recognition system has been effectively used in a "in vivo" neurosurgical biopsy. Batch Normalization [4] is used to solve this overfitting problem in more depth.

[5] A sterile human—machine interface is of supreme importance because it is the means by which the surgeon controls medical information avoiding contamination of the patient, the OR and the surgeon.

[6]Hand gestures are a fundamental type of nonverbal communication. Psychological research has shown that before they can speak, young children communicate through gestures. When people talk to one another about an object, manipulation is a common form of gesticulation. There are several compelling reasons to switch from the current interface technology (such as the keyboard, mouse, and joystick) to more natural interfaces, including naturalness of expression, unrestricted interaction, intuitiveness, and high sterility. At the Washington Hospital Center in Washington, DC, the gesture interface's functionality was evaluated. A hand gesture system's applicability was learned from watching two operations in the hospital's neurosurgery department.This is the first instance that we are aware of where a hand gesture recognition system has been effectively used in a "in vivo" neurosurgical biopsy. A sterile human-machine interface is crucial because it allows the surgeon to control medical information without contaminating the patient, the operating room, or themselves.

[7]More people are getting cancer, and patients are older. Often, invasive surgery, or the removal of the lesion, cannot cure patients, especially those who are elderly and have several diseases.

The radiologist percutaneously inserts energy applicators with high precision into the tumour location with the aid of imaging tools like CT scanners. However, because the majority of CT systems were created as diagnostic tools, they lack the support and interaction paradigms necessary for the interventional workflow. The radiologist must place the X-Ray tube and detector at the specified angle for radiography when using flat panel CTs to capture images that show the positions of the instrument and risk structures. Due to sterility difficulties caused by the usage of touchscreens or haptic buttons wrapped in a sterile drape, observations showed that the current workflow is not appropriate.

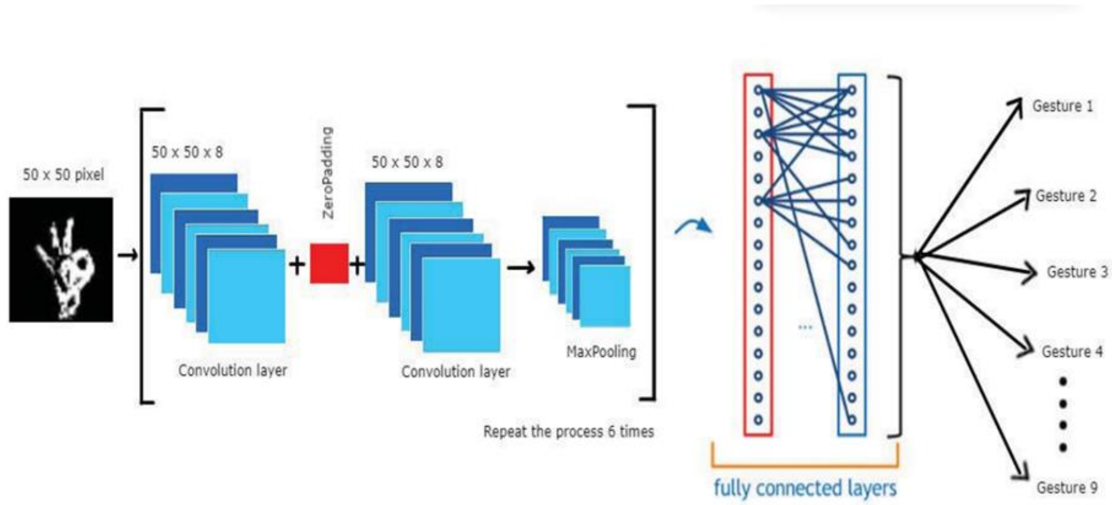


Fig :The network consists of 6 convolutional + max pooling layers, output of the 6th layer is given as input to a fully connected neural network with 9 hidden layers. Each hidden layer has 512 neurons, except the output layer which has 9 neurons, one each for each hand gesture.

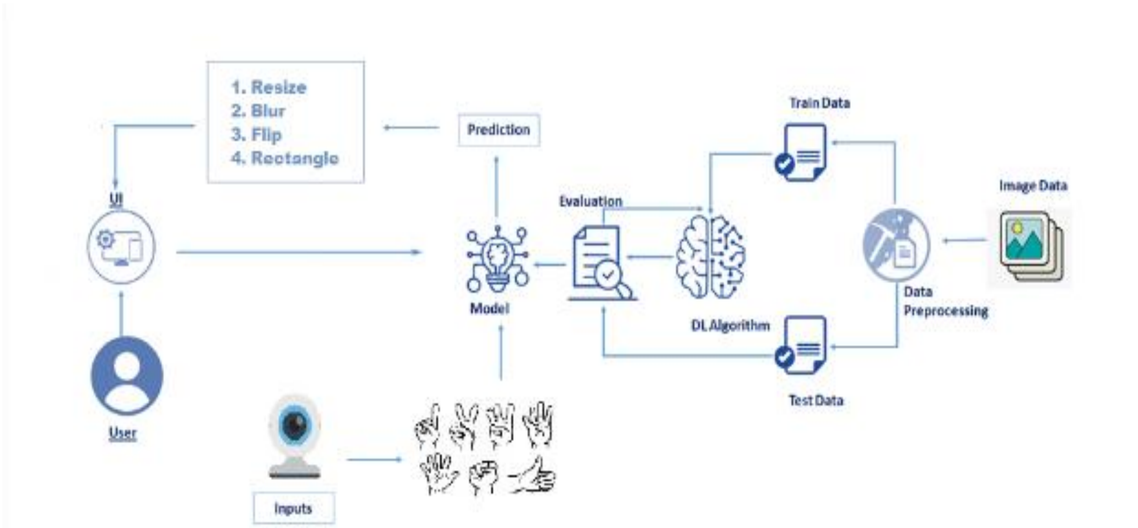
PROPOSED SYSTEM:

Physicians face a difficult problem when interacting with interventional imaging devices in a sterile setting. Because of sterility requirements and workspace limitations, direct physician-machine interaction during an intervention is very restricted. During computed tomography (CT)-based interventions, we provide a way of gesture-controlled projection display that enables a direct and natural physician-machine interface. Therefore, a radiation barrier in front of the doctor is projected with a graphical user interface. Using a jump motion controller, hand motions made in front of this display are recorded and categorized.

We provide a gesture set to operate intervention software's fundamental features, such as motions for selecting and manipulating 3D objects and 2D objects. In a user research with 12 participants that was clinically focused, our approaches were assessed. The outcomes of the user study show that clinical users accept the display and the underlying interaction paradigm. Although there is room for development, the gesture recognition is robust. Less than 10 minutes are spent on gesture training, however there are significant differences across research participants. The designed gestures are simple to use and rationally coupled to the intervention programme.

Contrary to conventional wisdom, the suggested gesture-controlled projection display allows the radiologist total control over the intervention programme. It offers fresh opportunities for direct physician-machine communication during CT-based interventions, particularly those performed during surgeries performed in an intensive care unit, and is ideally suited to become a crucial component of next interventional suites.

TECHNICAL ARCHITECTURE



REFERENCE

[1] Qing Chen Nicolas, D. Georganas, and Emil M. Petriu "Hand Gesture Recognition Using Haar-Like Features And A Stochastic Context-Free Grammar" IEEE ,Vol. 57, No. 8, August 2008

[2] Anupam Agrawal, Rohit Raj and Shubha Porwal "Vision-based Multimodal Human-Computer Interaction using Hand and Head Gestures" IEEE Conference on Information and Communication Technologies ICT 2013

[3] Kenji Oka and Yoichi Sato "Real-Time Fingertip Tracking and Gesture Recognition" IEEE proceeding on Computer Graphics and Applications Nov/Dec 2002

[4] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in International Conference on Machine Learning, 2015, pp. 448–456.

[5]Juan Wachs, Helman Stern, Yael Edan, Michael Gillam, Jon Handler, Craig Feied, Mark Smith

[6]Professor. Juan P. Wachs,

[7]Professor. Benjamin Fritsch,