**Literature Review**

In our modern society, people have become accustomed to simple human-computer interaction (HCI), primarily involving input devices such as joysticks, mice and keyboards. These are commonly used in robotics and automation. However, as technology evolves, these user interfaces (UI) need to advance in parallel. As a result, human-robot interactions (HRI), using hands as input devices. This offers more intuitive control, improves safety, accuracy and responsiveness in complex tasks. This review will explore papers containing research and applications on how the implementation of gesture-based systems can be applied across various industries.

Body

**How can the implementation of this system improve certain industries?**

The use of gesture control can benefit industries where safety and precision are vital instead of relying on simple controls. An example of these industries is construction, specifically regarding the decommissioning of a nuclear facility [1]. Human operators work remotely using a Head Mounted Display (HMD) to provide a video-see-through ability. The robots implemented with these systems have multiple cameras installed, allowing users to have a full 360 view of the environment. This setup provided workers with better visibility and enhanced both safety and control. Such systems can also be applicable in situations, such as search and rescue missions, inspections in hazardous environments, minimally invasive surgery, and teleoperations, where minimising human exposure to danger is essential.

Through these examples, the capability to assist multiple industries is demonstrated, and discussions will proceed towards the technical aspects of HRI development and what programs are used.

**What was used in the creation of vision-based systems?**

Several Vision-based gesture systems utilise RGB images and have been implemented in the creation of vision-based gesture systems, such as Hand gesture recognition (HGR) [3]. This method works by exploiting directly RGB images or by extracting skeletal information from RGB sequences and then applying it to a skeletal-based algorithm. The information is then used in a hand pose estimation method deployed using tools such as Mediapipe. The detectors and classifiers will then be built using a TD-Net architecture (Triple Feature Double Motion), Human-robot interaction (HRI), which uses Opencv library and MATLAB to access the programming for hand recognition [4], with human hands used as the input device. The shape and orientation of a person's hand movement can be recognised in the system, which can be translated into commands for the robotic hand to perform. Researchers were able to develop a video-based sign [4], another group was able to develop a real-time recognition system for 14 motions based on vision and another system was created to identify hand motions by identifying bending of the hands five individual fingers, utilising real-time object recognition software YOLOv4 to assist in the detection and estimation of objects from RGB images and use Particle filter algorithm. As for the object's pose and estimation model, DOPE (Deep Object Pose Estimation) was used; the control system was then implemented using ROS1 noetic on a computer installed with Ubuntu 20.04. [2]. This is a similar method used by another paper that uses RGB images. The only difference is that in that project a robotic hand was controlled [3], and for this project an entire robotic arm was controlled [2], employing Virtual Reality (VR) Head Sets and Stereo lab Zed stereo cameras, which used Unity 3d, which was using Windows 10. The values from the computer were then sent to the joints through a TCP/IP socket server to an Ubuntu computer to control the robot arm [1]. Two separate joint controls were used, specifically position and velocity control were implemented utilising C++ and the libfranka API.

The use of these different methods of detection and tracking is important, but requires advanced hardware to be used to its fullest.

**What motion capture technologies were used?**

As important as the software is in the programming and identification of a human hand/arm, the hardware used is as important, so that when the data is being transferred to the arm, the motions are as smooth as the human hand controlling it. As stated before, one of the projects uses VR to give the operator a see-through environment to allow for better visibility, but in conjunction with this, a control system was developed for a 3DoF arm to assist in minimally invasive surgery, which is reported to have a 2-second latency which can prove critical in such situations [1]. In one of the projects, a basic robotic hand was used, made with a combination of an Arduino UNO and five servo motors, which controlled each finger. To control it, a live camera was used to detect up to five different possible hand gestures, which are translated as commands to serve as a control unit for the robotic hand [4]. While this project focused on controlling a hand, another focused on controlling an entire robotic arm, namely a 7-Dof robotic arm with three coupled tendon-driven joint modules that make up the whole arm. The shoulder joint has a 3-DOF, a 2-DOF joint module for the elbow and another 2-DOF acting as a wrist joint. This arm allows for the same degree of movement as a human arm and is approximately the same shape as a human hand, enabling an automatic control system to be used at the shoulder as a prosthetic [2]. Of course, controlling robotic arms isn’t the only thing that this type of technology can control; the use of UAVS has become more widespread, being used for multiple purposes ranging from a simple hobby, to racing drones, security and military for both reconnaissance and assault. With the ability to express their desire with hand gestures, commands, even complex ones, can be given with these signals. In terms of the human operators, several instruction sets are not necessary and by removing them, operators can command UAVS in more natural ways, which can be of use during emergencies [5].

With the combination of both hardware and software related to the tracking and control of hand gestures, there are still some sections that are limited by the world's current technological development.

**Challenges and Limitations**

Despite the potential of these systems face several technical hurdles before they can properly replace traditional input devices. An example is the problems with reach-to-grasp movement for robotic arms. For the arm to accurately grab the object with a reasonable amount of strength, the technologies used to estimate the target object's structure, calculate the point for contact and plan the robotic arm's trajectory to reach it are a necessity. During tests involving pick-and-place movements where the robot had to pick up a box and place it somewhere else, repeating the process to find any possible errors. One of these errors occurred when the robotic arm had collided with the box, shifting its position, which caused it to fail. The main cause was identified to be a wire in the tendon-driven which appeared to be elongated and thus caused the error because of this, another series of tests were conducted at a position where the arm was not able to reach the object and it had difficulty finding an accurate reachable path with a searchable algorithm that is random [2]. Another problem stems from isolating the hand gesture recognition, as it cannot recognise rapid hand gestures. This also aligns with another problem, that continuous gesture recognition is gesture-spotting, meaning that the program must understand when the gesture starts and when it ends to receive another gesture. This was named temporal segmentation [3]. Of course, other elements such as lighting play a part in detection. The hand gestures were captured through a live camera. Through this operation, a problem occurred when the program ran under different light conditions same problem occurs when it comes to hand gestures with different skin colours. [4]

Once these restrictions can be solved. This technology can be utilised with no complications and will continue to improve as humanity develops.

**Conclusion and possible future opportunities**

Gesture-based systems show promise but require further development before it can replace traditional controls in safety-critical roles. In the future, the use of Artificial Intelligence (AI) can be used to assist in the identification and tracking of a person's arm. Sites such as roboFlow allow users to annotate certain parts of an image and name it, for identification, such as a hand. This process requires the processing of thousands of images to train the dataset to identify what a hand is and what it is doing. The dataset can then be given to YOLO to run the program and verify its success. At the same time, the field of robotics will also have to advance to the point where its movement can copy ours accurately to perform said tasks as well as a human could without any of the risk. These robots will have to be built to do these specific jobs, Surgery robots will need to be small, nimble and accurate; construction robots will need strong servos, be resistant to the elements and be durable, etc. They can also be modular, allowing for adaptation to specific situations.