

MACHINE LEARNING ASSISTED AUTONOMOUS VEHICLE IN AN IoT ENVIRONMENT

Suma Sira Jacob^{1✉}, Lijo Jacob Varghese², Jaisiva Selvaraj²,
Sathish Kumar Shanmugam³

Received on November 2, 2023

Presented by Ch. Roumenin, Member of BAS, on January 30, 2024

Abstract

The article presents the design of a controlled autonomous vehicle intended to perform specific job functions for impaired people by integrating Machine learning technique in an IoT environment along with hand-controlled gestures. In this paper, a novel multi-control autonomous vehicle is specially designed to cater the needs of disabled patients and senior citizens. The operation of this vehicle is fully motorized and gesture-controlled which involves physical implementation of hardware device with software for integrating, coding, interfacing and testing. This Machine learning based autonomous vehicle is completely designed with sensors and interactive cameras controlled by a smart processor unit that can control the various tasks which are designed to perform for the physically impaired people.

Key words: gesture control, machine learning, autonomous vehicle

Introduction. Studies show that 2 out of 10 adults globally have a disability, and about 30.6 million people have difficulty walking. Physical disabilities are a result of temporary or permanent injury, leaving the people either handicapped or partially paralyzed. About 8.5 percent of people worldwide (617 million) are aged 65, and about 35% face difficulty in walking. In both cases, people are bound to depend on a wheelchair for restricted movement, which is impossible to operate independently [1]. Using an autonomous vehicle to be operated for their needs makes them less dependent on another person for their movement.

A Human-Autonomous vehicle Interaction Model is essential in designing an autonomous vehicle. Utilizing a reinforcement learning technique [2] ensures efficient interaction between the user and autonomous vehicle, where sensor feedback under various environments is used to control signals for the wheel motors in the navigation of the mobile robotic base.

Overcoming these difficulties of operating power wheelchairs led researchers to utilize technologies for innovative development of mobile robots in creating an autonomous vehicle [3,4]. An Autonomous vehicle is made up of an essential power wheelchair base and a computing control with a combination of the seat mounted on a mobile robot [5]. Since some people with disabilities cannot use a traditional joystick to navigate their power wheelchairs, they use alternative control systems like head joysticks [6], where the movement of the head is sensed as different navigation controls. The next on the list is the chin joysticks [7], in which the steering control of the wheelchair is based on the movement of the chin. Tongue control [8] is based on an intra-oral device consisting of several inductive sensory coils connected externally to a radio chip controlled by a microcontroller and powered by a battery. Sip and puff [9] is a type of navigation control that utilizes the inhalation and exhalation of the user on a tube or a wand to send motion control commands. Researchers also have evolved with a Face/gaze control [10]. Direction control also utilizes electrical impulses of the brain, which are interpreted through the electrodes attached to the scalp, termed Electroencephalography or thought control [11]. Voice-assisted navigation control [12] using voice commands as inputs could be employed as a control methodology for Autonomous vehicles. This is based on recognition algorithms that already train voice commands and are stored in the control database, which gets executed by the appropriate commands. Human Machine Interface can also be done by a motion control system based on hand gestures [13]. This system employs an accelerometer that could effectively translate hand gestures into motion commands controlling the navigation of the Autonomous vehicle.

In an improvised version of the latter, an adaptive shared control system [14] proposes a system where the user's involvement determines the control assistance, which analyses the users' capability to continue with the appropriate action in line with its safety measures [15].

Proposed methodology. The concept of independent operation utilizing an autonomous vehicle could be realized by enabling the commands of an autonomous vehicle using hand gestures. On the contrary, in machine learning, any number of custom gestures could be taken, and by training these gestures, the system could recognize any of them faster with reasonable error tolerance and a swifter response. Apart from the gesture controls, the autonomous vehicle in this system can be controlled using a smartphone for navigation control commands that include forward, backward, stop, turn left, and turn right utilizing an Android Application. A provision is also included for the person to interact with

a keyboard interface for taking out these navigational commands. These commands are responded to accordingly by the robotic autonomous vehicle when the commands are issued.

T a b l e 1
Gesture and command interpretation

Gesture	Position of the palm	Command
Upward	Palm creates ($\cong$) $+20^\circ$ to $+90^\circ$ with Y axis	Chair moves backward
Downward	Palm creates ($\cong$) -20° to -90° with Y axis	Chair moves forward
Horizontal	Palm is parallel with horizontal	Chair stops
Left	Palm creates ($\cong$) -20° to -90° with X axis	Chair moves left
Right	Palm creates ($\cong$) $+20^\circ$ to $+90^\circ$ with X axis	Chair moves right

Gesture control by machine learning. An autonomous vehicle cannot be made intelligent by assembling hardware but by integrating exclusive computer algorithms that provide the artificial intelligence required for split-second decision-making. The gesture controls are assisted by a machine learning paradigm for taking up gesture inputs as commands for motion control of the autonomous vehicle [16]. This is a backup feature for the accelerometer-based gesture recognition, which has accuracy disadvantages compared to the AI-based gesture recognition technique. The number of custom gestures that could be processed in an accelerometer-based gesture control is smaller than its counterpart, which can take up any number of custom gestures. The processing time taken for identification of the gesture patterns is more in accelerometer-based control, while the response time is very much swifter in AI-based control. Also, the complexity of control increases in an accelerometer control when the number of custom gestures increases, but it is less complex in an AI-based control.

Using the mode selection module, the type of control could be selected for giving the input commands. A second camera that is used for this purpose is a PiCam [17]. Implementing the machine learning aspects of a neural network is best suited for detecting custom gestures. As the training of these gesture patterns is done with neural networks, it is required to have a module with reasonable computing power, like a Graphics Processing Unit (GPU) or Central Processing Unit (CPU), for its hardware implementation [3, 5].

Convolutional neural networks. Convolutional Neural Networks are built with a certain number of neurons where each one has learnable weights and biases. Every neuron acquires images as inputs. They dot product these images that sometimes result in non-linearity while a differentiable score function is given as output where pixels of the raw image is converted to class scores. Convolution is generally used in Digital Signal Processing, which combines two-time functions

$f(t)$ that can be expressed mathematically in a discrete domain of one variable as

$$(1) \quad y(n) = x(n) * h(n) = \sum_k^1 x(k) * h(nk),$$

where $k \in [-\infty, +\infty]$.

For a discrete domain of two variables:

$$(2) \quad (f * g)(x, y) = \sum_{u=-\infty}^{\infty} \sum_{v=-\infty}^{\infty} f(u, v) \cdot g(x - u, y - v).$$

Hence, convolution is the sum of point-wise products of function values, subject to traversal. Although the operation carried out on image inputs with CNNs is called convolution, it is slightly different because one of the inputs is time-reversed and is called cross-correlation.

$$(3) \quad y(n) = x(n) * h(n) = \sum_k^1 x(k) * h(n + k),$$

where $k \in [-\infty, +\infty]$.

The accurate method for convolving a Kernel (16×16) with input volume (RGB image of size $256 \times 256 \times 3$) requires taking into account the patches from the input image of size equivalent to the kernel and computing the dot product between the values found in the patch and the values in the kernel matrix. The convolved value obtained by summing the resultant terms from the dot product forms a single entry in the activation matrix. The convolutional layer is a 3-dimensional arrangement consisting of 2-dimensional layers of neurons that are stacked together, one for each channel depth. It plays a crucial role in training and subsequently firing the neurons present in the network by performing the convolution operation over the input image [18]. Each neuron is connected to the receptive field. For an input image of dimensions $30 \times 30 \times 3$ and a receptive field of 4×4 , each neuron in the Convolutional layer is connected to a region of $4 \times 4 \times 3$ in the input volume. Hence, each neuron will have 48 weighted inputs. For a specific R -value, we have a cross-section of neurons called the ‘depth column’ extending the entire depth of the Convolutional layer committed to taking inputs from this region. Shared Weights are used to optimize the Convolutional layer by reducing the number of unique weights to train and perform matrix calculations.

$$(4) \quad f(x) = 1 + e^x.$$

The Pooling layer lowers the spatial dimensions (width $\times$ height) of the Input image for the next Convolutional Layer and does not affect the depth dimension of the volume. The derivative of the Softplus function is a sigmoid function.

Lowering of size results in information loss, and hence, the operation is called down sampling. It reduces the size and lessens computational complexity. The Fully Connected layer is completely connected to the output of the previous layer.

$$(5) \qquad f'(x) = \frac{1}{1 + e^{-x}}.$$

Results and discussion. The objective of the work was to design and implement an intelligent wheelchair operated on gestures. Various results in simulating the hardware design, the developed hardware design, and the test results

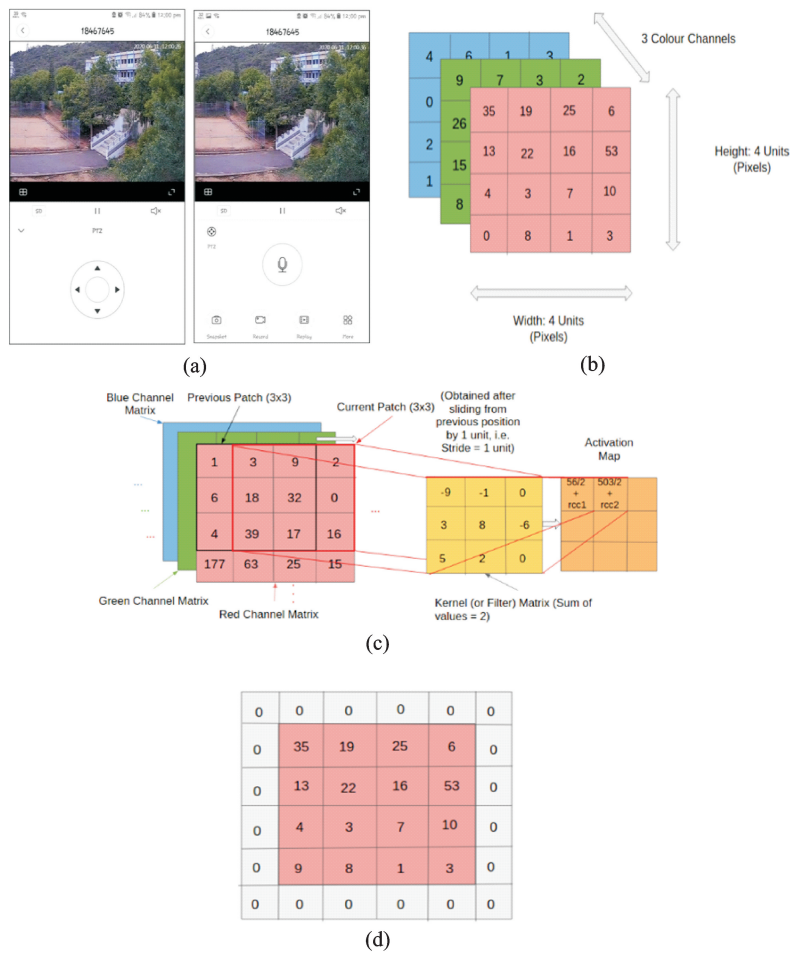


Fig. 1. (a) Camera module accessibility using the mobile App; (b) The cross-section of an input volume (RGB image) of size comprised of 3 colour channel matrices of the input image; (c) The convolution value calculation by the dot product of the values in the Kernel and their channel matrices; (d) A zero-padded 4 × 4 matrix becoming a 6 × 6 matrix

after working with several samples are discussed here. Figure 1 shows the clear vision of the proposed topology, which shows a better operation of the hardware model.

Similar testing for response and accuracy was conducted for the CNN-based Machine Learning Control, and the number of input gestures that are recognized correctly for the different gestures are presented in Table 2. Fifty samples were tested for each gesture for its correctness, and it was found that the gestures backward and stop had relatively lesser response. Figure 2 shows the architecture model of the proposed system.

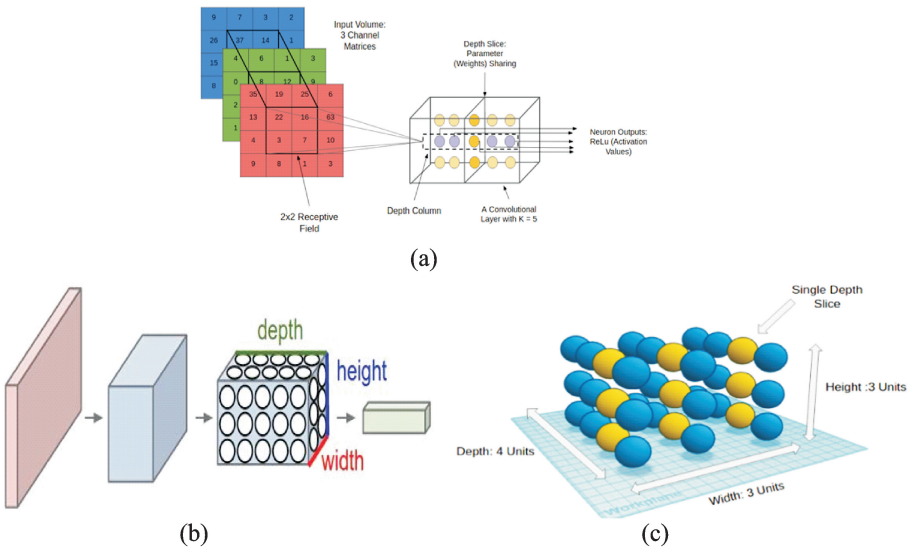


Fig. 2. (a) The concept of receptive field; (b) A ConvNet architecture with its essential building layers; (c) A 3D representation of the Convolutional layer with size $3 \times 3 \times 4 = 36$ neurons

A comparison of accuracy is made concerning recognition capability for various gestures by Accelerometer Control and Machine Learning Control and presented in Table 2. Table 2 clearly enumerates the overall accuracy in gesture recognition by both Accelerometer Control and Machine Learning Control for an overall testing sample of 250 each for individual controls inclusive for all gestures.

Though the Accelerometer Control is reasonably accurate in hand gesture recognition, the accuracy in pattern recognition of the CNN-based Machine Learning Control yielded far better results, as observed from Table 2.

Conclusion. This article focuses on the design of an autonomous vehicle control based on hand gesture inputs of the user results in the motion control command for the autonomous vehicle for navigation in the desired directions.

T a b l e 2
Gesture and command interpretation

Recognition of gestures for accelerometer control						
Recognized gestures	Experimental results	Input gestures				
		Backward	Forward	Stop	Left	Right
	Backward	43	3	2	2	3
	Forward	4	45	4	1	1
	Stop	3	2	44	2	2
	Left	0	0	0	45	0
	Right	0	0	0	0	44
Recognition of gestures for machine learning control						
Recognized gestures	Experimental results	Input gestures				
		Backward	Forward	Stop	Left	Right
	Backward	47	0	1	0	0
	Forward	1	48	2	1	1
	Stop	2	2	47	1	1
	Left	0	0	0	48	0
	Right	0	0	0	0	48
Comparison of accuracy for various gestures						
Methodology		Backward	Forward	Stop	Left	Right
Accelerometer control		86.0	90.0	88.0	90.0	88.0
Machine learning control		94.0	96.0	94.0	96.0	96.0
Comparison of overall accuracy in gesture recognition						
Methodology		Average accuracy (%)		No. of testing samples		
Accelerometer control		88.4		250		
Machine learning control		95.2		250		

Both the controls were implemented and developed into a hardware prototype. The testing of the hardware concerning its accuracy in recognition of the hand gesture inputs was made, and it was found that the machine learning control had reasonably better accuracy in pattern identification. Thus, the implementation of the autonomous vehicle takes various aspects of people with disabilities and the elderly into consideration with regard to the wellbeing of the user in terms of providing various choices of control, an interactive camera for navigation and two-way communication, access to change angle of camera rotation, LCD indicating the mode and the gesture. The implementation and testing proved to be satisfactory in terms of results, and this system must benefit society, where 30.6 million people globally have difficulty walking.

REFERENCES

- [1] SIMPSON R. (2005) Smart wheelchairs: A literature review, *J. Rehabil. Res. Dev.*, **42**(4), 423–436.
- [2] REIS L. P., B. M. FARIA, S. VASCONCELOS, N. LAU (2015) Multimodal interface for an intelligent wheelchair, *Informatics in Control, Automation and Robotics*, **325**, 1–34.
- [3] SINYUKOV D., R. DESMOND, M. DICKERMAN et al. (2014) Multi-modal control framework for a semi-autonomous wheelchair using modular sensor designs, *Intelligent Service Robotics*, **7**(3), 145–155.
- [4] RATHORE D. K., P. SRIVASTAVA, S. PANDEY et al. (2014) A novel multipurpose autonomous vehicle, *Proceedings of IEEE Students' Conference on Electrical, Electronics and Computer Science*, 1–4.
- [5] SIMPSON R., E. LOPRESTI, S. HAYASHI et al. (2004) The Smart Wheelchair Component System, *J. Rehabil. Res. Dev.*, **41**(3B), 429–442.
- [6] MAZUMDER O., A. S. KUNDU, R. CHATTARAJ et al. (2014) Holonomic autonomous vehicle control using EMG signal and joystick interface, *Proceedings of Recent Advances in Engineering and Computational Sciences (RAECS)*, Chandigarh, India, 1–6.
- [7] GUO S., R. A. COOPER, G. G. GRINDLE (2004) Development of Head-Operated, Isometric Controls for Powered Mobility, *Proceedings of RESNA 27th International Annual Conference*, Orlando, Florida, 2004.
- [8] LUND M. E. (2010) Inductive tongue control of powered autonomous vehicles, *Proceedings of Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Buenos Aires, Argentina, 3361–3364.
- [9] MOUGHARBEL I., R. EL-HAJJ, H. GHAMLOUCH et al. (2013) Comparative study on different adaptation approaches concerning a sip and puff controller for a powered wheelchair, *Proceedings of IEEE Science and Information Conference (SAI)*, London, 597–603.
- [10] MATSUMOTO Y., T. INO, T. OGSAWARA (2001) Development of intelligent autonomous vehicle system with face and gaze based interface, *Proceedings of 10th IEEE International Workshop on Robot and Human Interactive Communication*, Bordeaux, Paris, France, 262–267.
- [11] MILLÁN J. DEL R. (2014) BMI: Lessons from tests with impaired users, *International Winter Workshop on Brain-Computer Interface (BCI)*, Jeongsun-kun, 1–1.
- [12] SURYAWANSHI S. D., J. S. CHITODE, S. S. PETHAKAR (2013) Voice Operated Intelligent Autonomous vehicle, *International Journal of Advanced Research in Computer Science and Software Engineering (IJARCSSE)*, **3**(5), 487–490.
- [13] PAULOSE S., M. P. F. ANOODA, G. MOHAN et al. (2014) Automatic Autonomous vehicle using Gesture Recognition along with Room Automation, *Transactions on Engineering and Sciences*, **2**(5), 40–43.
- [14] PHILIPS J. (2007) Adaptive Shared Control of a Brain-Actuated Simulated Autonomous vehicle, *Proceedings of IEEE 10th International Conference on Rehabilitation Robotics, ICORR*, Noordwijk, Netherlands, 408–414.
- [15] MAZO M. (1995) Autonomous vehicle for physically disabled people with voice, ultrasonic and infrared sensor control, *Autonomous Robots*, **2**(3), 203–224.

- [16] XU R., S. ZHOU, W. J. LI (2012) MEMS Accelerometer Based Nonspecific-User Hand Gesture Recognition, *IEEE Sensors Journal*, **12**(5), 1166–1173.
- [17] ALAM M., M. A. YOUSUF (2019) Designing and Implementation of a Wireless Gesture Controlled Robot for Disabled and Elderly People, 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), Cox'sBazar, Bangladesh, 1–6, <https://doi.org/10.1109/ECACE.2019.8679290>.
- [18] ROFER T., A. LANKENAU (2000) Architecture and applications of the Bremen autonomous wheelchair, *Information Sciences*, **126**(1–4), 1–20.

¹*Department of Artificial Intelligence and Data Science,
Sri Krishna College of Technology,
Coimbatore-641 042, Tamilnadu, India
e-mail: sumasarajacob@gmail.com*

²*Department of Electrical and Electronics Engineering,
Sri Krishna College of Technology,
Coimbatore-641 042, Tamilnadu, India
e-mails: lijojacobvarghese@gmail.com, jaisiva1990@gmail.com*

³*Department of Electrical and Electronics Engineering,
M. Kumarasamy College of Engineering,
Karur-639 113, Tamilnadu, India
e-mail: sathishphd2k17@gmail.com*