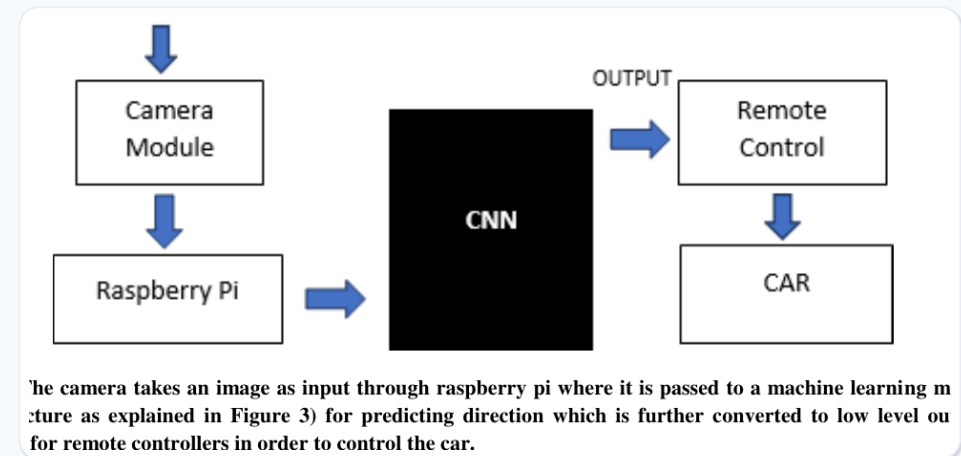


# Smart Autonomous Vehicle Using End-to-End Learning

A low-cost RC-car platform that maps camera pixels directly to driving commands through a CNN.

Pranav Gupta • Vaibhav Singh • Anubha Parashar

Journal of Innovation in Computer Science and Engineering, Vol. 9(2), Jan–June 2020



The camera takes an image as input through raspberry pi where it is passed to a machine learning model as explained in Figure 3) for predicting direction which is further converted to low level output for remote controllers in order to control the car.

# Why end-to-end learning?

Traditional self-driving systems rely on several separate modules and many sensors. The paper proposes a simpler path: use a single front-facing camera and learn the steering decision directly.

1

front-facing camera

3

driving commands

98%

reported accuracy

## Core motivation

- Reduce sensor complexity and cost
- Avoid hand-engineered perception and lane modules
- Enable a reproducible autonomous RC-car setup

### Smart Autonomous Vehicle Using End to End Learning

Pranav Gupta<sup>1</sup>, Vaibhav Singh<sup>2</sup>, Anubha Parashar<sup>3</sup><sup>1</sup>Scholar, Manipal University, Department of Computer Science and Engineering, Jaipur, Rajasthan, India<sup>2</sup>Scholar, Manipal University, Department of Information Technology, Jaipur, Rajasthan, India<sup>3</sup>Scholar, Manipal University, Department of Computer Science and Engineering, Jaipur, Rajasthan, IndiaEmail : [pranavg387@gmail.com](mailto:pranavg387@gmail.com), [vaibhavsingher007@gmail.com](mailto:vaibhavsingher007@gmail.com), [anubhaparashar1025@gmail.com](mailto:anubhaparashar1025@gmail.com)

**Abstract** - Autonomous Driving has passed the point of being called the next big step, as the smart car revolution is already taking shape around the world. Self-driving cars are relevant if not prevalent and the main obstacles to reach mass adoption are customer acceptance, cost, infrastructure and the reliance on several onerous algorithms that include perception, lane detection, path planning and variation in pathways. The objective of this research paper is to tackle the mentioned problems with a straightforward, reproducible and cost-effective solution, using end to end learning and replacing the numerous sensors with a camera. These were optimized directly by the proposed system with limited background processing. In this research paper authors achieved this by mapping pixels from only a single front-facing camera to direct driving instructions. The results obtained were better than state of the art and achieved the aim of the study proficiently.

**Keywords** - Autonomous Driving, Convolution Neural Network, Transfer learning.

#### 1. INTRODUCTION

A smart car is generally made of 2 controlling units that dictate its actions, High-level controller, and Low-level controller. The high-level controller takes input from its components i.e. the driver (in this case the camera), its surroundings such as the traffic and the sensors in place, it then after deducing the correct actions to be taken sends signals to the low-level controller that controls the brakes, steering, engine and throttle [8][9][12]. To do this successfully it needs to understand driver psychology, how and when a specific maneuver is necessary, which changes with the terrain and the area the car drives in since driver temperament and driving style cannot be universally generalized[6][7][10]. To understand and correctly predict such outcomes is precarious. Studies show that if the driver is given even a half a second extra before a collision, 60% of accidents can be avoided and this percentage increases to 90% if one second of warning time can be provided [4]. Such results stress the importance of timely and correct decisions that face problems with the conventional architecture of a smart car system. It is also

difficult for cars relying on so many sensors to be able to adapt to new surroundings and to reconfigure the system to achieve a different goal based on learning that occurs on the move, especially when noisy sensor data is received [5]. The foundation of the end to end approach was laid almost 14 years ago in an effort by Defense Advanced Research Projects Agency (DARPA) known as DARPA Autonomous Vehicle (DAVE) [10] where a radio control (RC) car drove through a junk- filled alley [9]. It demonstrated that this approach was viable and adequate for the functioning of a smart car. The end to end approach thus provided alternatives which were uncomplicated and also easier to test [8]. With the advancement in technology, the study achieved similar functionality more skillfully and with limited computational power. At the centre of this end to end learning is the Convolutional Neural Network which is trained based on the actions of the driver while driving. The data collection was automated which made it possible to do rigorous experimentation as the system itself could be reproduced efficiently at a low cost. Also, it can be shared among multiple systems in order to make the learning faster and for better adaptation in different terrains.

#### 2. LITERATURE REVIEW

This journey to make a Smart Autonomous vehicle began with the invention of modern cruise control in 1948. Many dedicated steps have been taken in this direction to make self-driven cars a reality. ALVINN (An Autonomous Land Vehicle in a Neural Network) diversified the field by combining End to End learning (Shibata et al. 1997) with a Neural Network to resolve the problem (Dean A. Pomerleau 1989). DARPA (Defense Advanced Research Projects Agency) then gave perspective to what can be achieved with the technical know-how of the time (Net-Scale Technologies Inc. 2004). In the recent past when a convolutional neural network was trained to map unaltered pixels from a camera to be interpreted directly into steering commands (Mariusz Bojarskiet al. 2016), it demonstrated the reliability of this approach in contrast to the more modular alternatives[1].

# System idea: pixels to driving commands

The full pipeline converts camera frames into low-level commands for the RC vehicle.



**Input**

Image frames from Pi camera mounted on top of the RC car.

**Decision**

VGG16-based CNN predicts left, right, or forward.

**Output**

Raspberry Pi triggers GPIO-based remote commands.



# Data collection and training set

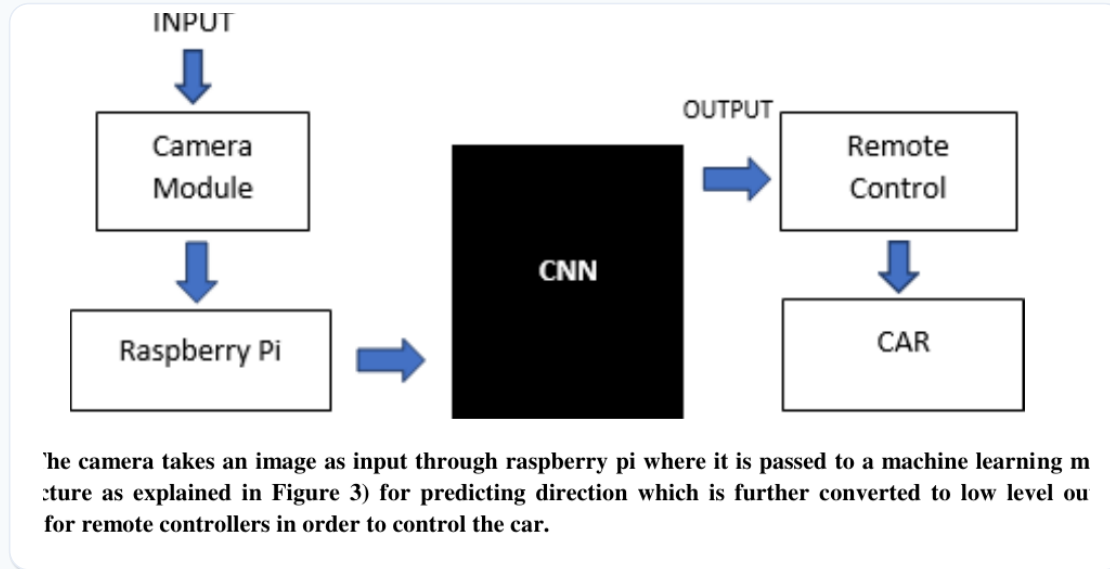


Figure 1: capture → inference → movement

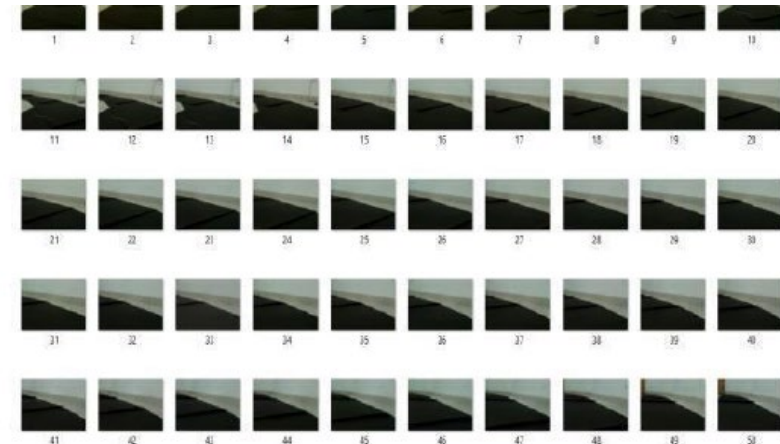


Fig. 2. Sample Section of the Collected Dataset

Processing unit (Raspberry Pi)

Figure 2: sample collected frames

- Udacity simulator validated the plan before hardware testing.
- Custom dataset was collected from self-laid tracks using the Pi camera.
- Frames were time-synchronized with the control command: left, right, or forward.

# Pre-processing made the model robust

The paper emphasizes lightweight feature preparation before training the neural network.

## Crop + resize

Remove unnecessary image area and reduce memory.

## YUV color space

Efficient representation for vision tasks.

## Flip turns

Augment left/right turning data.

## Brightness jitter

Handle varied lighting conditions.

## Gaussian blur

Smooth noise in camera frames.

## Normalize classes

Reduce straight-driving bias.

*Goal: convert small camera frames into reliable driving decisions with limited computation.*

# CNN architecture: transfer learning with VGG16

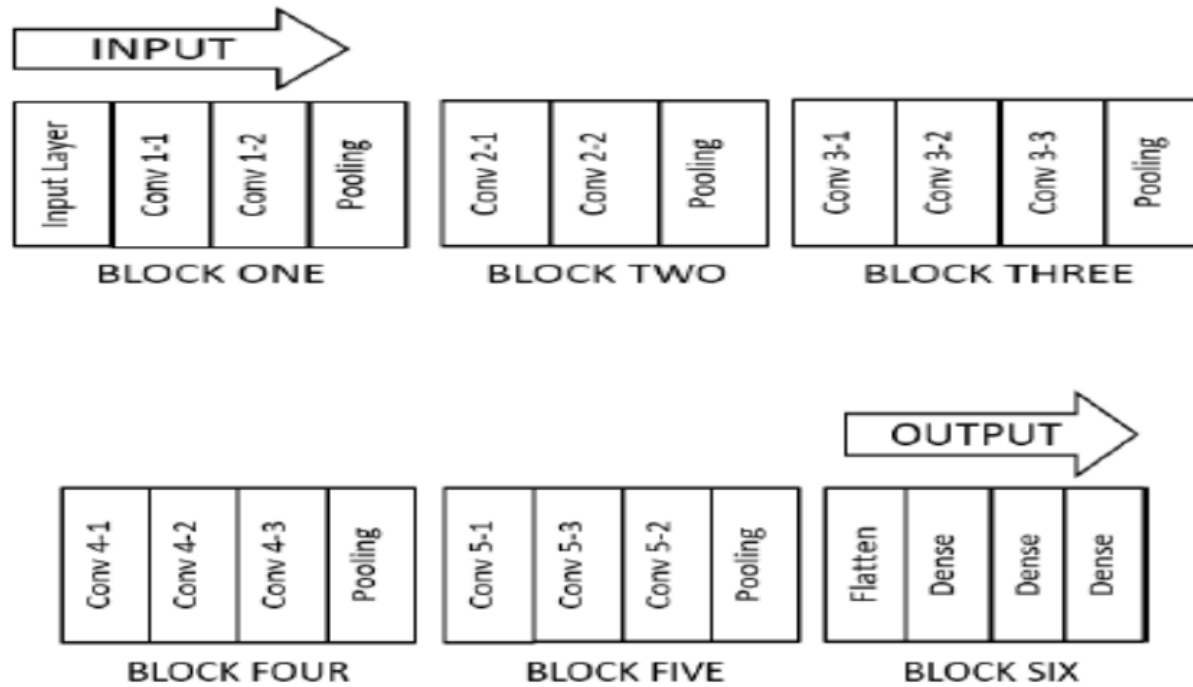


Fig . 3 The architecture of the neural network implemented for the project

## Architecture summary

- Input:  $3 \times 224 \times 224$  image
- Pre-trained VGG16 used for feature extraction
- Top fully-connected layers replaced
- Two dense layers of 128 units each
- Output: 3 classes — left, right, forward

## Training setup

Adam optimizer, learning rate 0.01, categorical cross-entropy loss.

# Hardware deployment on an RC car

The prototype connects perception, processing, and actuation through Raspberry Pi and GPIO-driven remote control.



## What the Pi does

Receives frames, runs neural-network inference, and sends the predicted command.

## How actuation works

The Pi simulates remote button presses by writing HIGH/LOW signals to GPIO pins.

## Why it matters

The setup is low-cost, reproducible, and suitable for controlled-environment prototyping.



# Result: strong reported accuracy

The proposed model reported the highest accuracy in the comparison table.



98% proposed system accuracy

Loss trends down across epochs and the comparison table places the proposed system above prior end-to-end driving studies cited in the paper.

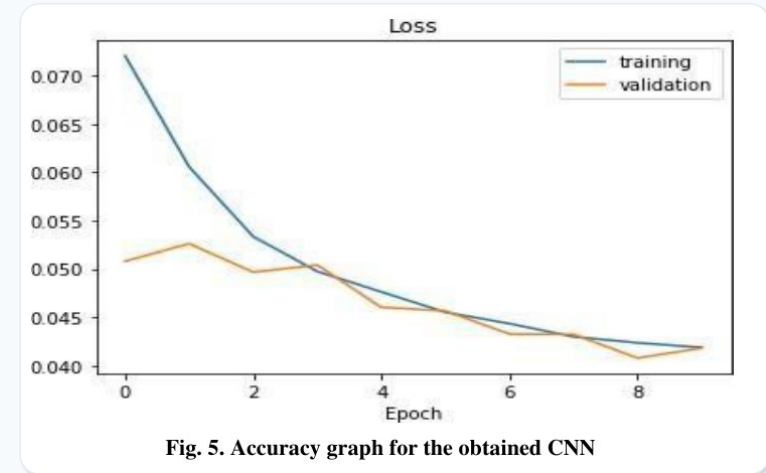


Fig. 5. Accuracy graph for the obtained CNN

Muller, U (2006)	Off-Road Obstacle Avoidance through End-to-End Learning	75 %
Mori, K (2019)	Visual Explanation by Attention Branch Network for End-to-end Learning-based Self-driving	79 %
Xu, H (2017)	End-to-end Learning of Driving Models from Large- scale Video Datasets	84.6 %
Bojarski, M (2016)	End to End Learning for Self-Driving Cars	88 %
Proposed System	Smart Autonomous Vehicle using End to End Learning	98 %

# Conference recognition

---

The work was presented at ICICSE-2020, hosted virtually by Guru Nanak Institutions Technical Campus and Guru Nanak Institute of Technology.

## Paper title

Smart Autonomous Vehicle Using End To End Learning

## Conference

8th International Conference on Innovations in Computer Science & Engineering

## Dates

August 28–29, 2020 • Virtual conference

# Conclusion & future scope

---

## Main takeaway

End-to-end learning is a viable alternative to hand-engineered autonomous-vehicle pipelines in controlled environments.

## Practical value

The setup is low-cost, camera-based, reproducible, and deployable on an RC-car prototype.

## Future work

Improve hardware actuation with precise angle control and test on more complex road conditions.

Thank you

Questions?