ISBN : 978-93-94056-72-5

VOLUME 2 – ISSUE 1 - 20 March 2024

Computer Science Research Reports

A YEARLY PUBLICATION OF
U.G. PROJECT REPORTS

Chief Editor
Muneer V.K.

Editors
Dr. Shameem Kappan
Rizwana Kallooravi Thandil

PG & RESEARCH DEPARTMENT OF
COMPUTER SCIENCE
Sullamussalam Science College, Areekode

VOLUME 2 – ISSUE 1
20 March 2024

Computer Science
Research Reports

A YEARLY PUBLICATION OF U.G. PROJECT REPORTS

Chief Editor:
Muneer V.K

Editors:
Dr. Shameem Kappan
Rizwana Kallooravi Thandil

ISBN : 978-93-94056-72-5

PG & RESEARCH DEPARTMENT OF COMPUTER SCIENCE
Sullamussalam Science College, Areekode
Malappuram Dt, Kerala – 673639
www.sscollege.ac.in

Computer Science Research Reports
(2024)

Editors

Muneer V.K : Assistant Professor
PG & ResearchDepartment of Computer Science
Sullamussalam Science College, Areekode
Malappuram, Kerala, India 673639

Dr. Shameem Kappan : Associate Professor
Head, PG & Research Department of Computer Science
Sullamussalam Science College, Areekode
Malappuram, Kerala, India 673639

Rizwana Kallooravi Thandil : Assistant Professor
PG & Research Department of Computer Science
Sullamussalam Science College, Areekode
Malappuram, Kerala, India 673639

Edition 1
Cover Design: Ashfar Sadath
Typesetting & Layout: Muneer V.K

ISBN : 978-93-94056-72-5

Computer Science Research Reports
(2024)

Preface

Welcome to the inaugural edition of "Computer Science Research Reports"! This journal marks a significant milestone in the academic journey of final year graduate students of Computer Science at Sullamussalam Science College, an esteemed institution affiliated with the University of Calicut, Kerala.

Within the pages of this journal lie the result of months of dedicated research, exploration, and innovation undertaken by our talented students. In this edition, we are happy to present nine insightful articles that explore various facets of contemporary computer science, reflecting the diverse interests and expertise of our budding scholars.

The field of computer science is ever-evolving, and this collection of research papers underscores the latest trends and advancements in areas such as machine learning, object detection, natural language processing, AI chatbots etc. Each article represents a unique contribution to the ongoing dialogue within the discipline, offering fresh perspectives, novel approaches, and valuable insights that push the boundaries of knowledge and technology.

As educators, mentors, and facilitators, we take immense pride in the accomplishments of our students. Their dedication, passion, and intellectual curiosity shine through in the rigor and depth of their research endeavors. Moreover, this journal serves as a testament to the collaborative spirit and vibrant academic community fostered at Sullamussalam Science College.

We extend our heartfelt appreciation to all the contributors, reviewers, research project supervisors and editorial team members who have played pivotal roles in bringing this journal to fruition. Their commitment to excellence and scholarly rigor has been instrumental in ensuring the quality and integrity of the published work.

It is our sincere hope that "Computer Science Research Reports" will serve as a valuable resource for researchers, practitioners, and enthusiasts alike, inspiring further exploration and innovation in the dynamic field of computer science. We invite you to delve into these pages, engage with the cutting-edge research presented herein, and join us in celebrating the intellectual achievements of our students.

Sincerely,

Editor-in-Chief

Table of Contents

Computer Science Research Reports (2024), 2-1, 1-13
ISBN : 978-93-94056-72-5

RESEARCH ARTICLE

PROJECT-U: ENHANCING PERSONAL SAFETY WITH PORTABLE EMBEDDED TECHNOLOGY FOR STRAY DOG DETERRENCE

UA.Samshar[1*], Ramzy Hisham[2], Mishab K[3], Muneer VK[4]

[1,2,3,4] *PG & Research Department of Computer Science, Sullamussalam Science College, Areekode.*

*Samsharsachu1@gmail.com

Abstract: *Project-U stands as an innovative embedded system crafted to enhance personal safety, especially in areas prone to encounters with stray dogs. By integrating cutting-edge technologies, including the MPU-6050 gyroscope for motion sensing and the HCSR04 ultrasonic sensor for transmitting 40kHz ultrasonic waves, Protect-U takes a proactive approach to potential threats. The inclusion of the KY-038 sound sensor empowers the device to intelligently detect and distinguish dog barking, amplifying its ability to recognize and respond to varying threat levels. This compact and user-friendly device prioritizes the safety of individuals, encompassing diverse demographics such as children, the elderly, women, and those with disabilities. Employing a sophisticated interplay of motion sensing, sound detection, and ultrasonic technology, Protect-U represents a significant stride toward creating a secure environment in areas marked by the challenges presented by stray dog encounters.*

Keywords: Stray Dog Deterrence, MPU-6050, KY-038, HCRSO4, Embedded Technology.

1. Introduction

The fear of encountering aggressive dogs is a significant public health concern, deterring many individuals, particularly those from vulnerable populations, from enjoying the physical and mental benefits of outdoor walks. This anxiety can transform a potentially relaxing activity into a stressful experience, ultimately contributing to a more sedentary lifestyle. Current safety devices often lack sophistication, relying on basic motion detection or loud alarms. These approaches can be disruptive to both the user and the surrounding environment, while also creating a false sense of security as they may not effectively distinguish between genuine threats and harmless situations. Project U emerges as a novel solution, offering a more nuanced and user-centric approach to personal safety during walks. This paper introduces Project-U, a wearable device that utilizes a combination of advanced sensors and intelligent algorithms to create a comprehensive understanding of the user's environment and their response to potential threats. Unlike bulky and intrusive safety devices, Project-U is designed to be compact, lightweight, and discreet, integrating seamlessly into the user's walking routine without compromising comfort or style. This unobtrusive design is crucial for promoting user adoption and ensuring long-term adherence to safety precautions.

This paper will delve into the core functionalities of Project-U, highlighting its key differentiators that contribute to a more effective and user-friendly safety experience. Firstly, we will explore the device's multi-sensor approach, which goes beyond basic motion detection. Project-U employs a sophisticated combination of sensors, including, A sound sensor to detect the presence and intensity of dog barks, allowing for differentiation between playful barks and aggressive ones. A

gyroscope sensor to identify sudden changes in user movement, such as a jump or stumble, which might indicate an unexpected encounter with a dog. An ultrasonic sensor acts as a safe and humane deterrent when both the sound and movement sensors suggest a potential threat. This approach minimizes unnecessary disruptions for both the user and the surrounding environment.

Secondly, we will examine the intelligent algorithms that analyze the collected sensor data. This processing power ensures that Project-U reacts only to genuine threats, minimizing false alarms and unnecessary disruptions. The paper will explore the specific criteria used by the algorithms to differentiate between situations that warrant a deterrent and those that do not. Finally, we will discuss the device's commitment to humane deterrence methods, prioritizing the safety of both humans and animals. The subsequent sections will provide a deeper technical analysis of Project-U's components and their functionalities. We will explore the specific capabilities of each sensor and how they work together to create a comprehensive picture of the user's surroundings. Furthermore, we will examine the algorithms employed by Project-U, illustrating their role in differentiating between genuine threats and non-threatening situations. Finally, the paper will explore the potential for future advancements in Project-U technology and its potential applications beyond personal safety during walks, considering its potential integration with other wearables or smart home systems for broader safety applications.

2. Literature review

The recent increase in attacks of animals like dogs against humans has led to a necessity for a solution. Studies conducted have proved that these animals can be repelled by a specific frequency of sounds that is sounds above 20KHz that only the animals can hear. These high-frequency sounds are irritating for them. Many devices have been made into the market that produce these high-frequency sounds, but these devices are not portable sizes. These devices are operated manually. It makes these devices difficult to use when an animal attacks. The device we intend to produce is an automatic, small portable device that can be carried in a handbag or pockets. It is a device that can be used by just switching it on and it automatically senses the presence of an animal switch on the ultrasonic sound sensor and produce an ultrasonic sound.

In our device, we are using an ultrasonic sensor (HC-SR-04), gyroscope sensor (MPU-6050), and sound sensor (Ky-038). The ultrasonic sensor automatically emits the ultrasonic sounds by getting triggered by the other two sensors, a sound sensor that senses the dog’s sound, and a gyroscope sensor that senses the movement in the human body. Studies have proven that humans can hear between 4 Hz 17,600 kilohertz and up to 20 kilohertz. These frequencies are called audio. Whereas, the frequencies that exceed 20 kilohertz, which are called ultrasound frequencies cause inconvenience in dogs, as it was known in the past, as special whistles were used to expel dogs without the human being hearing them [1].

Tests were conducted on the concern that the numerous ultrasonic sensor devices that are designed to repel animals and pests from gardens will influence the behavior of guide dogs resulting in safety implications for their owners. These devices were operating at a range between 67 Hz and 44 kHz. However, it is suggested that guide dogs are not affected by these sounds [2]. A solar-powered multi-tone ultrasonic rodent repeller was designed and constructed to complement the use of conventional methods of trap and chemical (poison). Four frequencies are generated (35kHz, 38kHz, 40kHz, and 50kHz) using a multivibrator mode. This solar-powered multi-tone ultrasonic rodent Repeller is repelled. It was an effective and best method for eliminating or eradicating rodents in a particular environment and hence the multiple tones do not create room for rodents to be accustomed to a particular sound [3]. A prototype of the ultrasonic sound generator was made that could be used to warn the stray animals entering the road. Hundreds of people are injured when their vehicles come in contact with stray animals, and it is very difficult to file a claim against these animals. The ultrasonic sound generator is energized with solar power [4]. The use of ultrasound frequencies is safe in terms of environment and health, as these frequencies have been used in the treatment of soft tissues, as well as

superficial and deep wounds, by using high frequencies of (1-3 megahertz). These frequencies reduce and treat swelling, infections, and pain relief in addition to speeding up wound healing [5, 6].

Studies have found that various animals use ultrasonic sounds to communicate with each other of their kind. These studies were conducted on several rats. Rats emit and respond to sounds in the ultrasonic (US) frequency range (30-100 kHz). This study showed that humans and animals have specific methods of hearing. Sound waves travel through the ear canal and vibrate a thin membrane like a drum. These vibrations are then passed on to tiny bones that act like a lever, making the vibrations stronger. Finally, the vibrations reach a fluid-filled chamber that sends signals to the brain through a nerve. The brain interprets these signals as the sounds we hear [7].

3. Methodology

Project U requires multiple stages to accurately function. It needs to detect the sound of a dog while it attacks. We calculate the threshold of a sound using a sound sensor and calculate inner body movement using a gyroscope sensor. Then, we need to activate the ultrasonic sensor to emit ultrasonic waves to deter the dogs. The operation of this stray dog deterrent machine is as follows: From Figure 1, you can understand the workings of Project-U.

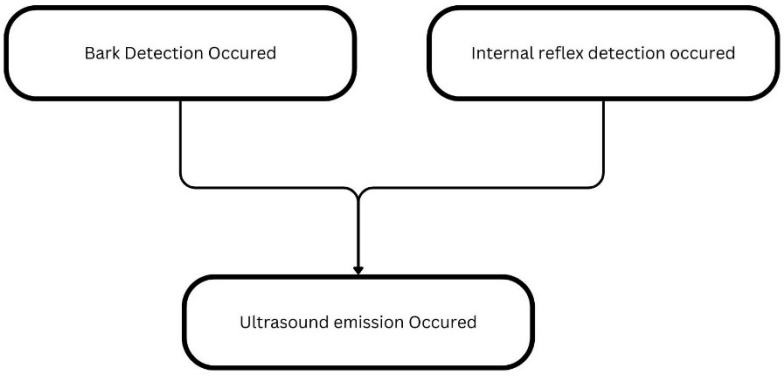


Fig 1. Block diagram of Project-U working

Firstly, there are already many devices that work on the principles of Project-U. A device that repels dogs and cats using ultrasonic sound already exists. In Figure 2, you can see the device that works like that, but this device is mainly used in gardens, grounds, or farms to avoid their nuisance. It is not portable; we need to place it aside on the ground or farm. Figure 3 represents the functional circuit of the device mentioned above.

Fig 2. Animal & insect deterrent device

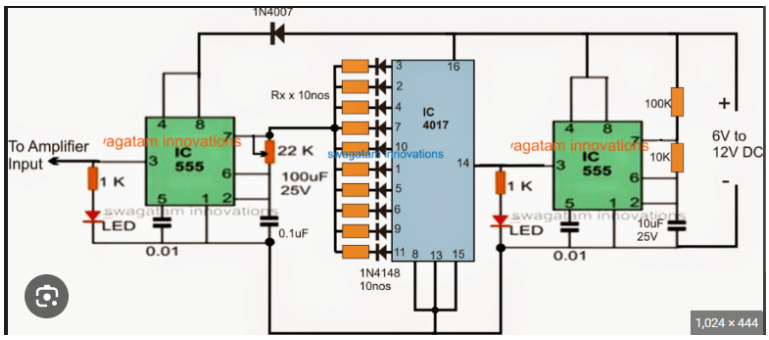


Fig 3. Circuit diagram of device 1

The Yard Sentinel utilizes **ultrasonic sensors** to deter unwanted animals from entering your property. These sensors emit sound waves at a frequency of **20 kHz,** which is outside the human hearing range but **perceptible to most animals** like dogs, cats, and rodents. The Yard Sentinel operates in a **"switched mode."** This means the device **periodically activates** and emits ultrasonic sounds for short bursts throughout the day. In Figure 4, you can see the modified version of it, which is used in cities and neighborhoods to avoid dog attacks.

Fig 4. Modified version of device 1

In the case of both devices, they are not portable, and they are switched on all the time. So, the ultrasonic sounds will always emit from them. It affects non-harmful dogs and makes them more dangerous. Here, our Project-U differs. Our device is portable and will only switch on when a dog attack is detected using our sensors and algorithm.

3.1 Sounds and its features:

Sound can be divided into three main types: human audible sound, infrasound, and ultrasonic sound. Let's break them down. Human audible sound is just what it sounds like - the kind of sound that our ears can pick up. It ranges from 20Hz to 20kHz in frequency. Picture it like this: 20Hz is the lowest rumble you can hear, like distant thunder, while 20kHz is at the upper end, like the high-pitched whine of a mosquito. Anything beyond 20kHz might not bother us, but it could cause discomfort.

Now, let's talk about infrasound. These are the sneaky sounds that slip below our hearing range, typically between 0 and 20Hz. Humans can't detect them, but animals like elephants and moles are tuned in to these low-frequency vibes. Lastly, there's ultrasonic sound, which is like the secret language of animals. It's beyond our hearing range, starting at 20kHz and going up. Dogs, cats, bats, and dolphins are among the creatures that can tune in to these high-pitched sounds, even though we humans are clueless about them. Figure 5 will explain a clearer picture of how these sound classifications work.

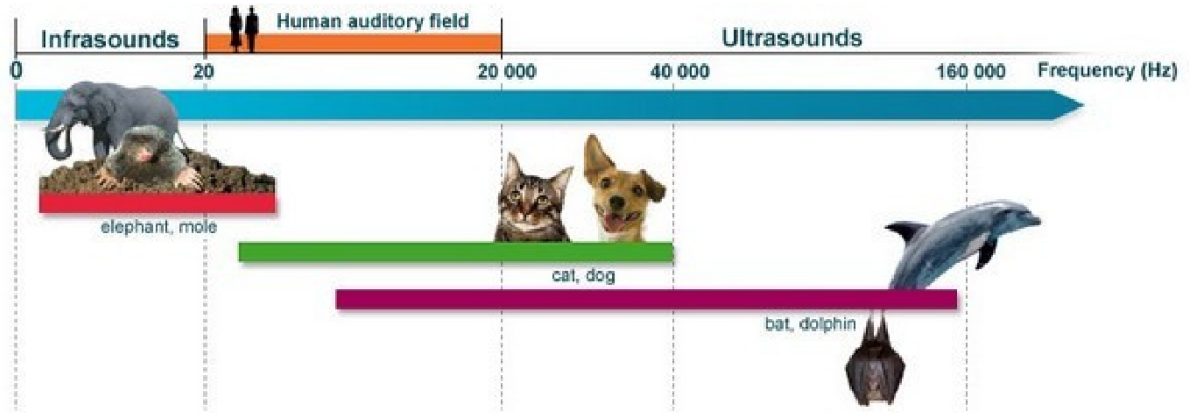


Fig 5. Sounds classification

3.1.1 Ultrasonic sound

In our quest to build an effective device, we delve into the fascinating world of sound, specifically ultrasonic sound. Why ultrasonic, you might ask? Well, studies have shown that ultrasonic sound is a dog's kryptonite. You see, while we humans are blissfully unaware of these high-pitched frequencies, dogs and other animals perk right up at the sound of them. Now, you might think, "Wouldn't high-frequency sounds bother dogs like they bother us?" Interestingly, that's not the case. Dogs have a remarkable hearing range, from a low rumble of 40Hz to a high-pitched squeal of 40kHz. While these ultrasonic frequencies might not tickle our ears, they catch the attention of our furry friends.

Imagine it like this: if we humans had to endure a constant high-pitched whine, we'd be out of there in a heartbeat. Similarly, dogs find ultrasonic sounds in the range of 30kHz to 40kHz seriously bothersome. It's like an invisible force field they just can't stand being around. So, armed with this knowledge, we've crafted our device to emit precisely these ultrasonic frequencies. When a stray dog decides to make a move, our device kicks into action, emitting those high-pitched sounds that send them packing. It's like having a superhero power against dog attacks.

3.2 Detection of dog attack

In the world of gadgets designed to keep dogs at bay, we've seen quite a few options. But let's be real here – most of them aren't exactly user-friendly. They're big, bulky, and always buzzing with ultrasonic waves, which isn't exactly inviting for our furry friends just minding their own business. It's like having a noisy neighbor who never turns down the volume. Enter Project-U, with a heart for both people and pups. Our device isn't just another clunky machine. Nope, it's got a brain of its own. It only jumps into action when there's trouble brewing – like when a stray dog decides to stir up some mischief. How do we know? Well, we've got some pretty smart sensors onboard, like the KY-038 sound sensor, MPU-6050 gyroscope sensor, and HC-SR04 ultrasonic sensor. These little heroes team up with our clever programming to spot a doggy dilemma before it turns into a full-blown commotion. But here's the best part – we've made things easy as pie. We're using an Arduino Uno board Which is given as in Figure 6.

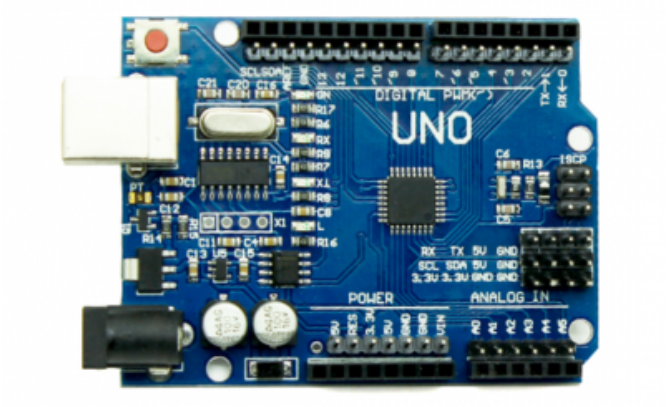


Fig 6. Arduino Uno

3.2.1 KY-038 Sound sensor

Now let’s discuss how we detect animal attacks, there comes the duty of our sensors, just imagine the situation, we just walk through a road peacefully, and suddenly a stray dog comes across us and tries to attack us, that time the KY-038 sound sensor in our device will detect the barking of the sound using the intensity of the sound.ky-038 (figure 7) is a normal sound sensor that has a tiny but crucial component of Project-U and acts like your watchful ear on walks. This sensor can detect sound waves, including the distinctive bark of a dog. Think of it as having a friend who can hear everything around you. But it's not just any ear; the KY-038 is adjustable, allowing you to fine-tune its sensitivity to focus on specific sounds you care about. When it detects a sound, it sends a signal indicating the sound's intensity. Project-U's system then analyses this signal to determine if it's a potential threat, ultimately working to keep you safe. While it doesn't measure exact decibel levels and

requires additional components to function within Project-U, the KY-038 plays a vital role in the initial detection of potential danger, offering you peace of mind and a safer walking experience.

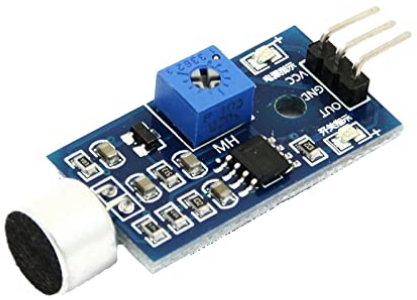


Fig 7. Ky-038 sound sensor

Also, remember that the ky-038 didn't accurately and specifically detect the dog's bark, it detected a pitch sound and assumed that it was a dog's barking using our programming algorithm

3.2.2 mph-6050 Gyroscope sensor

Now imagine that when the barking and the approach of dogs make you fearful, you just shake a little, we all know that that is normal to a human, a reflex will happen when a sudden and fearful situation occurs, there comes our MPU-6050 gyroscope sensor in Figure 8. This tiny sensor is more than just a motion detector. It acts like a friend who understands your body language. It doesn't just detect movement; it can sense the subtle changes in your body's acceleration and velocity, like the slight trembling that often accompanies fear. Here's how it works: Project-U sets a threshold for these internal movements. If the sensor detects changes that exceed this threshold, it interprets it as a potential threat situation, considering the sudden bark from the sound sensor as well. This could be a dog approaching too quickly, causing you to instinctively take a small step back or jump to avoid it. Remember: Project-U doesn't jump to conclusions. The combined information from the sound sensor and the gyroscope is crucial for taking informed action. Once Project-U identifies a potential dog attack, it activates the ultrasonic sensor, the HC-SR04, which emits a high-frequency sound that deters dogs without harming them. This sound is inaudible to most humans, so it won't disrupt your walk or the surrounding environment. Project U is designed to be intelligent and responsible. It only triggers the deterrent when necessary, ensuring your safety and minimizing any unnecessary discomfort to both humans and non-threatening animals. It's like having a friend who understands both your words and your body language, always ready to step in and keep you safe, even from the unexpected.

Fig 8. MPU-6050 Sensor

3.2.3 HCSRO4 Ultrasonic sensor

Now the attack of the stray dog was detected using a ky-038 sensor and MPU6050 gyroscope sensor, we need to emit the ultrasonic sound and create a sound barrier around the human, there comes our Sensor HCSRO4 ultrasonic sensor. The HC-SR04 (Figure) works by utilizing ultrasound. These are sound waves at frequencies beyond the human hearing range, but they're perfectly audible to dogs. This makes them a safe and targeted way to deter unwanted attention. Project U doesn't rely solely on the sound sensor to activate the HC-SR04. Remember, the MPU-6050 gyroscope also plays a crucial role. This sensor detects sudden movements in your body, which could indicate a situation where you might feel startled or fearful, like encountering a dog unexpectedly. Here's where the magic happens: When both the sound sensor detects a bark and the gyroscope detects a sudden movement, Project-U interprets this as a potential threat. Only then, and only under these combined circumstances, does it activate the HC-SR04. This ensures that the ultrasonic sound isn't wasted on everyday noises or triggered accidentally, minimizing unnecessary disruption.

Once activated, the HC-SR04 emits a burst of ultrasonic waves. These waves create a discomforting sensation for dogs, similar to how loud noises might affect us. This discomfort effectively deters them from approaching, keeping you safe without causing them any physical harm. It's important to remember that the HC-SR04 acts as part of a larger system within Project-U. It doesn't directly detect dogs but rather relies on the combined information from other sensors and Project-U's intelligent algorithms to make informed decisions. Additionally, the ultrasonic sound waves have a limited range, making them most effective for close encounters. Overall, the HC-SR04 plays a vital role in Project-U's mission to provide peace of mind and safety during walks. It's a testament to the power of intelligent design, using safe and targeted technology to ensure a harmonious coexistence between humans and animals.

Fig 9. HCSRO-4 Sensor

3.2.4 Battery life

When we go outside with our project u, the system must have battery life so we use a Polymer lithium-ion battery (Figure 10) for this purpose. It's lightweight, so you won't feel weighed down during your walks, and its flexible design seamlessly integrates into the device without adding bulkiness. Safety is a top priority – you can trust it to deliver reliable power without the risk of overheating or explosions. Plus, with its long lifespan and ample charging cycles, you can count on Project-U to keep you safe on countless walks for a long time to come. Charging time is 1 Hour. A

typical charge or use cycle for a lithium-ion battery is 8 hours of use, 1 hour to charge, and another 8 hours of use.

Fig 10. Polymer lithium-Ion battery

3.2.5 Automatic and Manual working

Project-U operates with user-friendly manual controls, ensuring ease of use and flexibility in various situations. Equipped with intuitive switches, users can activate specific functionalities as needed. For instance, the sound detection switch enables the device to monitor ambient noise levels, while the motion detection switch triggers the gyroscope sensor to detect sudden movements. Additionally, users can manually activate the ultrasonic deterrent through a dedicated switch, emitting waves to deter approaching dogs. This manual control system offers users direct command over Project-U's protective features. Here is the fully functioning circuit diagram of the product (Figure 11).

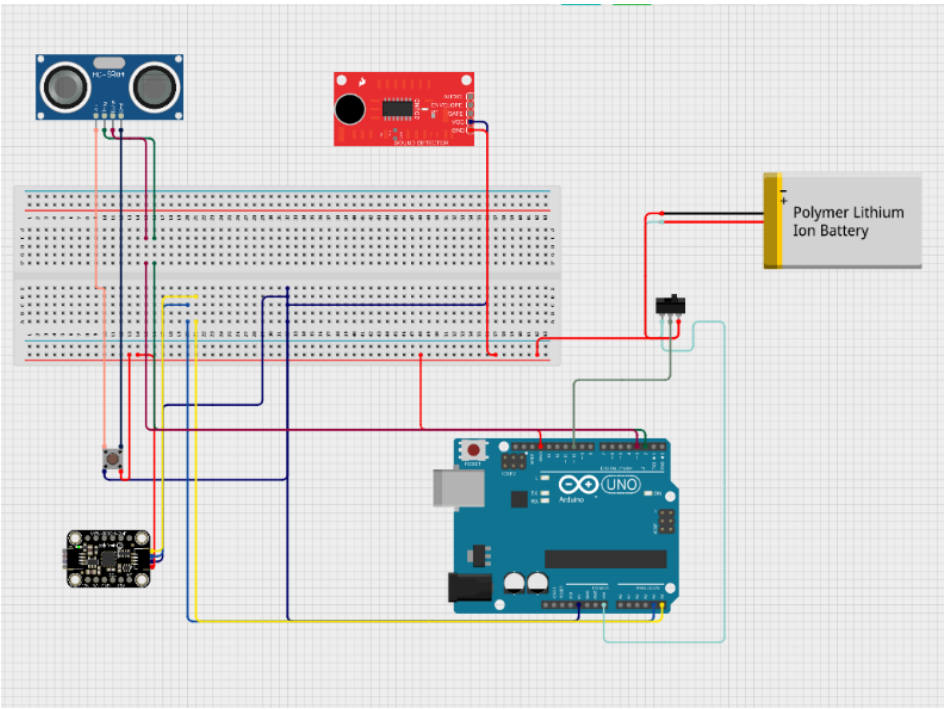


Fig 11. Complete circuit diagram

4. Experimental result

The experiment revealed a successful collaboration between the KY-038 sound sensor, the MPU-6050 gyroscope, and the ultrasonic sensor. The KY-038 accurately measured sound intensities,

allowing us to distinguish crucial sounds from background noise by adjusting the threshold value. The MPU-6050, acting like a human's vestibular system, detected sudden movements, helping differentiate intentional actions from potential startles. Finally, the ultrasonic sensor served as a silent communicator, emitting deterrent sounds triggered by the combined data from the other sensors. Notably, the synergy between the sensors reduced false positives by requiring both a loud sound and a sudden movement to activate the deterrent, ensuring it only responded in necessary situations. Overall, the experiment showcases the effectiveness of each sensor and their combined potential for future safety application advancements.

4.1 Bark sound detection and accuracy

The KY-038 Sound Sensor detects the barking sound by capturing the sound waves in the environment and calculating its intensity. It compares this intensity to a pre-defined threshold value. If the sound intensity exceeds the threshold, the device interprets it as a potential dog bark. During unit testing within the Arduino IDE's serial monitor, you can observe the sound intensity values. If the displayed value surpasses the pre-defined threshold, a message will be displayed stating "Bark detected.". From Figures 10 and 11, we can observe the output on the Arduino IDE serial monitor. In the program we defined the rh threshold value as 560, if the intensity read by the sound sensor is below 560, the serial monitor did not display anything in Figure 12. If the intensity of the sound exceeds the threshold value, the serial monitor will display the message “Sound detected”, and as the intensity decreases the message will disappear Figure 12. When the high threshold sound is detected by the sound sensor, the sensor will assume that it is a dog’s bark and it passes to the program.

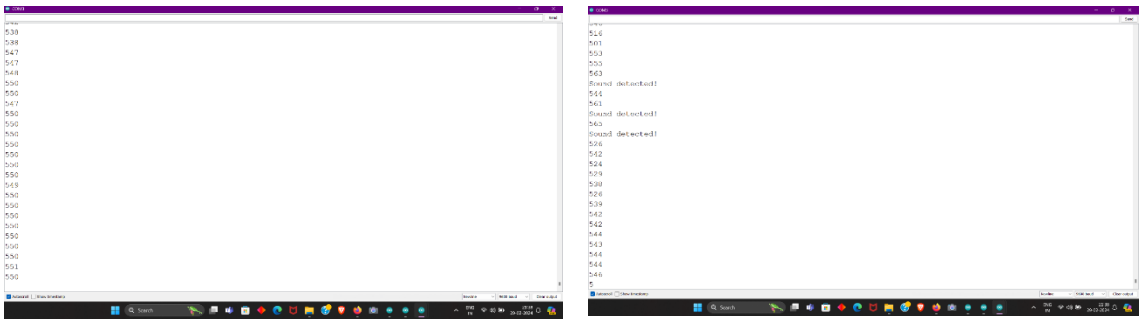


Fig 12. OUTPUT of Serial Monitor for sound sensor_1

4.2 Movement detection MPU-6050 sensor

The MPU-6050 gyroscope sensor will detect the reflex on the body by calculating the internal movement of the device using the acceleration and angular velocity itself. We will define a threshold value of 1000 and if the movement calculated by the gyroscope is greater than it, the system will assume that it is a sudden reflex due to fear And pass the information into the program. From Figures 13 and 14 we can observe the output obtained from the gyroscope sensor. In the unit test of the MPU-6050 gyroscope sensor, If the total movement is less than that threshold value, it doesn't display anything, but if the total movement intensity exceeds the threshold value, the Serial monitor will display a message "Device is shaking".

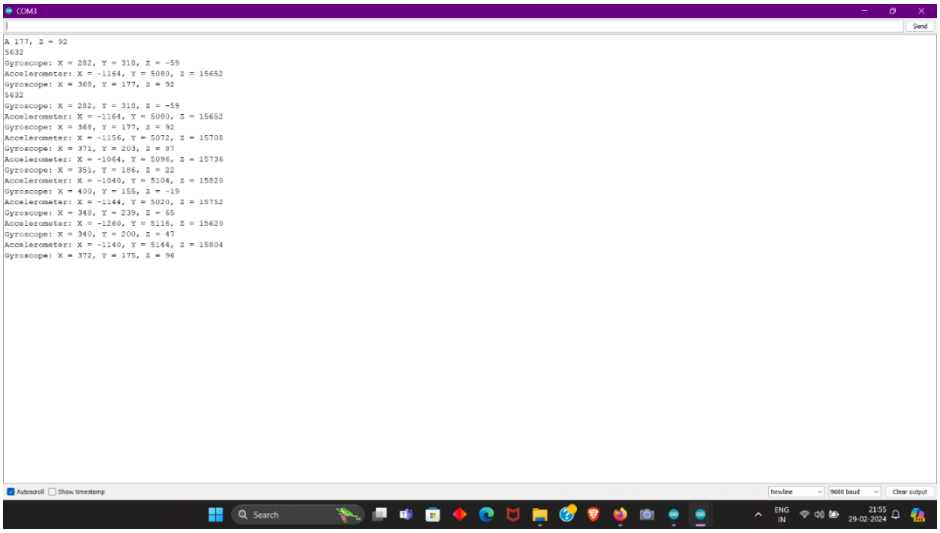


Fig 13. OUTPUT on Serial Monitor for gyroscope sensor_1

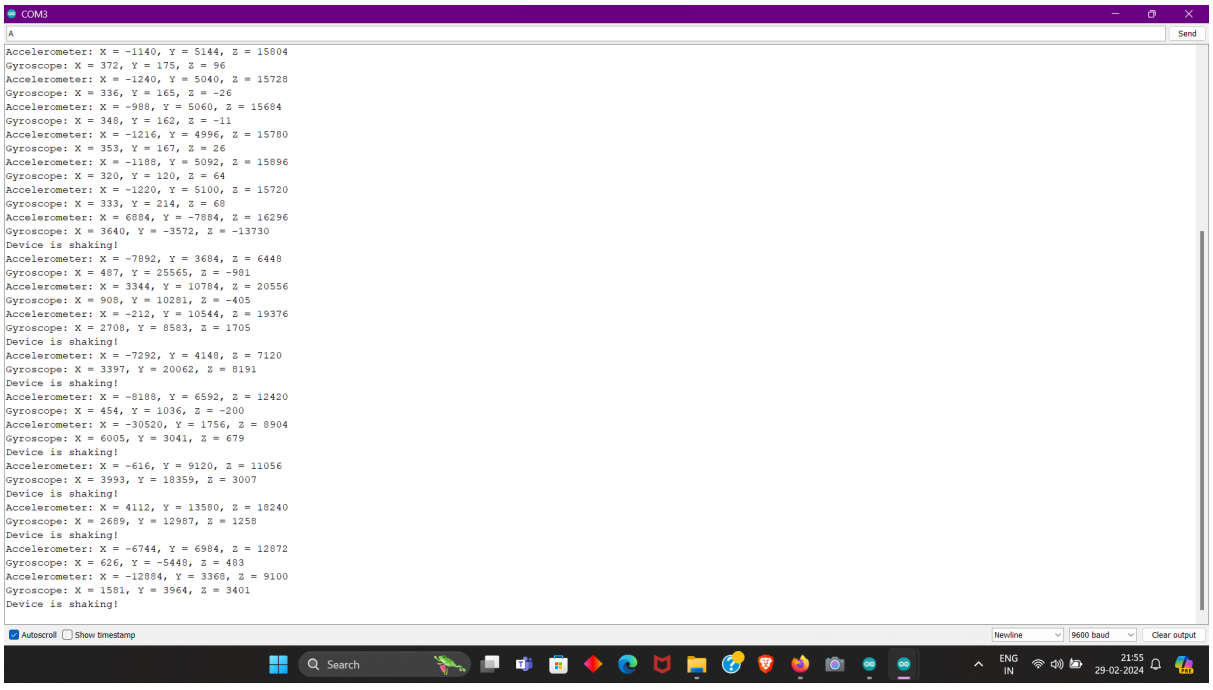


Fig 14. OUTPUT on Serial Monitor for gyroscope sensor_2

4.3 Ultrasonic sound emission using HCSRO4

When the ky-038 sensor detects the high-intensity sound and the MPU-6050 sensor detects the highly intensive movement simultaneously, the device will assume it has a stray dog attacking situation, and this information will pass to the HCSRO4 sensor, then the sensor will activate the trig pin of the sensor and an Ultrasonic sound of frequency 40khz will emit through it. The ultrasonic sounds will act as a barrier around the human and the stray dog cannot enter into that barrier due to the property of ultrasonic sounds. From the graph (Figure 15) below we can observe the change in the frequency while the HCSRO4 sensor triggers the ultrasonic waves

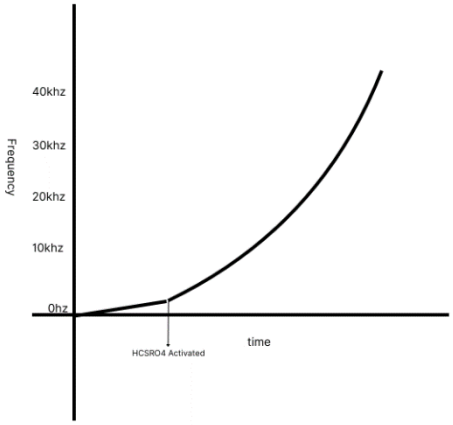


Fig 15. Graphical diagram of frequency change

By Analysing this graph, we can say that when the HCSRO4 sensor, the frequency range in the atmosphere changes gradually and it gradually changes from 147hz to 40khz.From that, we can conclude that the frequency acquired from the ultrasonic sensor will disturb the dogs and repel them dogs from attacks. From Figure 16 we can observe the image of a fully functional prototype of Project.

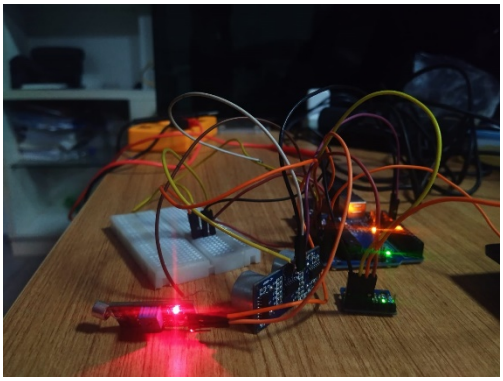


Fig 16. Prototype image

5 Discussion

Project U presents a promising approach to deterring dog attacks while promoting safety for both humans and animals. However, it's crucial to consider the following points for discussion and future development. While Project-U utilizes multiple sensors for increased accuracy, it's important to acknowledge that the KY-038 sensor cannot definitively identify the source of the sound. This means there's a possibility of false positives triggered by loud noises other than dog barks. Future iterations could explore advanced sound recognition algorithms or incorporate additional sensors like directional microphones to enhance accuracy. The HC-SR04 ultrasonic sensor has a limited effective range. This might not be suitable for all situations, especially in open spaces. Future development could explore the use of alternative ultrasonic emitters with a wider range or investigate the integration of additional deterrent measures like visual deterrents activated at a larger distance.

5.2 Future Integration:

Project-U's core functionalities could potentially be miniaturized and integrated into existing wearable devices like smartwatches or smartphones. This would enhance portability and make it even more accessible to users. Continued research and development of the algorithms used in Project U could further refine its ability to differentiate between dog barks and other sounds, leading to fewer false positives.

6. Conclusion

Project-U isn't just about circuits and sensors; it's about giving you, the person enjoying your walk, the peace of mind you deserve. It's a silent guardian walking by your side, ready to step in when you need it most. Sure, there's room for improvement. The technology, as with anything new, has limitations. But the idea behind Project-U - promoting safety for both humans and animals - is one we can all get behind. Imagine a future where Project-U or its even cooler, miniaturized versions, becomes a familiar companion on your walks, integrated seamlessly into your smartwatch or smartphone. Imagine a future where smart homes can deter unwanted animal encounters, keeping your gardens and yards safe. This future is within reach, but it requires continued collaboration and responsible development. We need to keep listening to users, refining the technology, and ensuring animal welfare remains a top priority. Project U is more than just a conclusion to this paper; it's a starting point for a conversation about how technology can create a safer and more harmonious world, where humans and animals can coexist peacefully.

References

[1] Maysoon Mahdi Atheeb et al., Repellent of the stray animals. IOP Conf. Series: Earth and Environmental Science 735 (2012): p (3).

[2] Rachel Moxon et al., Effect of Ultrasonic Devices on the Distraction Behaviour of Guide Dogs. International Journal of Orientation & Mobility. Volume 3, Number 1, 2010.

[3] Nweke F. U., Design Construction and Characterization of a Solar-Powered Multi-tone Ultrasonic Rodent Repeller. International Journal of Science and Research (IJSR) Volume 4 Issue 2, February 2015.

[4] Vijayaraghavan Sundararaman et al., Ultrasonic sensor animal safety system. International Conference in Recent Advances and Innovations in Engineering (ICRAIE-2014), 1-2, 2014.

[5] Olyaie, M., et al., High-frequency and noncontact low-frequency ultrasound therapy for venous leg ulcer treatment: a randomized, controlled study. Ostomy Wound Manage, 2013. 59(8): p. 14-20.

[6] Taradaj, J., et al., The use of therapeutic ultrasound in venous leg ulcers: a randomized, controlled clinical trial. Phlebology, 2008. 23(4): p. 178-83.

[7] Warfield, D., The study of hearing in animals, in Methods of Animal Experimentation. IV, G. W, Editor. 1973: Academic Press; London. p. 43–143

Computer Science Research Reports (2024), 2-1, 14-25
ISBN : 978-93-94056-72-5

RESEARCH ARTICLE

AN INTELLIGENT CATALOGUE USING OBJECT DETECTION FROM CAMERA FOOTAGE

Muhammed Shahil A K[1*], Jalwa V P[2], Afnan M K[2], Rababa K[4], Muneer V.K[5]

[1,2,3,4,5]*Department of Computer Science, Sullamussalam Science College, Areekode.*

[*]*mdshahilak@gmail.com*

Abstract: *Security and wildlife monitoring are dynamic fields, and the addition of CCTV surveillance cameras has certainly improved our capacity to protect a variety of habitats. However, the ongoing difficulty of effectively reviewing large amounts of video in theft or animal contact instances demands creative solutions. This research presents a novel "Automated Catalogue System using Object Detection," utilizing Modern technology such as OpenCV, MobileNetV3, and Single Shot Multibox Detector (SSD). The system uses a Camera to operate in real time and catalogues things that are recognized. It records important information including the category name, date, time, and confidence score of the detected object. In addition to enhancing security, the project tackles wildlife-related issues in a novel way, offering a remedy for situations in which untamed animals intrude into human communities. The project takes pictures and records them using an automatic categorization method, providing an accessible record through a user-friendly web application. This complete approach gives communities a dependable tool for wildlife monitoring and awareness while also eliminating confusion about animal interactions. This project represents a significant contribution to technological advancements, as it integrates real-time object detection, image capture, and detailed cataloguing. It is an adaptable system that can be applied to both security surveillance and wildlife management.*

Keywords: CCTV, MobileNetV3, Single Shot Multibox Detector, OpenCV, Catalogues.

1. Introduction

In modern security and wildlife monitoring scenarios, the widespread use of CCTV surveillance cameras has undoubtedly enhanced the ability to observe and safeguard various environments. However, a significant challenge arises when incidents of theft or other unusual occurrences demand scrutiny of extensive footage, leading to time wastage and inefficiency. To solve this issue, our project aims to revolutionize surveillance methodologies by implementing an automated cataloguing system using real-time object detection through a webcam. The system employs the Single Shot Detection (SSD) algorithm, specifically utilizing the MobileNetV3. The goals of the initiative are Firstly, it seeks to accurately recognize and categorize in real-time each thing that appears in front of the webcam, including people, animals, and vehicles. Second, the system records and saves real-time photos in a database if an object is a person or an animal is identified with a confidence level higher than 70%. Details like the type of object discovered, the degree of confidence, the date, and the time of detection are also included in these photographs. This data makes it possible to retrieve specific events quickly and easily, reducing the need to watch the entire film. The system makes use of Firebase as the database, ReactJS for the web application interface, OpenCV as the framework, and Python for program development.

Our project not only improves security monitoring but also expands its use to address wildlife-related issues, especially in locations close to forests. When wild animals come into villages, it usually makes people afraid and confused. The program can identify and take pictures of these species using our automated cataloguing system, and it will make the records available via the online application. In addition to removing all doubts, this gives communities a solid tool for wildlife awareness and monitoring. Our project is positioned as a valuable solution with applications in security surveillance and wildlife management due to its combination of real-time item detection, image capture, and comprehensive cataloguing.

2. Literature Review

Jisoo Jeong, Hyojin Park, and Nojun Kwak present a novel approach to enhance the accuracy of the Single Shot Multibox Detector (SSD) in object detection. The proposed network exhibits superior results on the Pascal VOC 2007 test set, achieving 78.5% mean average precision at 35.0 FPS with a 300 x 300 input size. This outperforms conventional SSD, YOLO, Faster-RCNN, and RFCN, establishing a state-of-the-art mAP while maintaining faster speeds than Faster-RCNN and RFCN, showcasing its effectiveness in object detection [1]. Reagan R. Galvez, Argel A. Bandala, Elmer P. Dadios, Ryan Rhay P. Vicerra, and Jose Martin Z. Maningo explore the pivotal role of vision systems in mobile robotics for tasks like navigation, surveillance, and explosive ordnance disposal. A comparative analysis between the Single Shot Multi-Box Detector (SSD) with MobileNetV1 and Faster Region-based Convolutional Neural Network (Faster-RCNN) with InceptionV2 reveals a trade-off: one model excels in real-time applications due to speed, while the other provides more accurate object detection, offering valuable insights for diverse mobile robot applications [2].

Ayesha Younis, Li Shixin, Shelembi Jn, and Zhang Hai present a compelling paper on object identification, leveraging MobileNet within the Single Shot Multi-Box Detector (SSD) framework. The proposed algorithm demonstrates real-time detection, specifically applied to webcam feeds for video stream object identification. By integrating MobileNet and SSD, the research emphasizes a swift and efficient deep learning-based approach to object detection. This research enhances accuracy for real-time detection, addressing the requirements of daily monitoring in indoor and outdoor environments [3]. Sheping Zhai, Dingrong Shang, Shuhuan Wang, and Susu Dong propose a novel improvement to the Single Shot MultiBox Detector (SSD) for object detection, introducing DF-SSD based on Dense Convolutional Network (DenseNet) and feature fusion. DF-SSD achieves notable results, outperforming SSD on PASCAL VOC 2007 by 3.1% mAP with reduced parameter requirements, showcasing its advanced detection capabilities for small objects and those with specific relationships [4].

Sheping Zhai, Dingrong Shang, Shuhuan Wang, and Susu Dong present SSD-MSN, an advanced multi-scale object detection network derived from the Single Shot Multibox Detector (SSD). Introducing a valid dividing image strategy further enhances performance through data augmentation, generating 3x3 clipped areas. Experimental results on PASCAL VOC and COCO datasets showcase SSD-MSN's state-of-the-art detection capabilities, particularly in improving multi-scale object detection, demonstrating its effectiveness in complex scenarios [5]. In their recent work, Abhinav Juneja, Sapna Juneja, Aparna Soneja, and Sourav Jain contribute to the evolving field of object detection, focusing on real-time applications. Utilizing the ssd_v2_inception_coco model, their proposed Single Shot Detection (SSD) approach exhibits substantial improvements in accuracy. This model not only detects numerous objects simultaneously but also demonstrates significant accuracy enhancements over established methodologies, making it a valuable asset for applications such as parking lots, human identification, and inventory management [6].

Md Alamin Feroz, Marjia Sultana, Md Rakib Hasan, Aditi Sarker, Partha Chakraborty, and Tanupriya Choudhury address the challenging task of real-time object detection and recognition in uncontrolled environments. Their research introduces an improved technique using Single Shot Detector (SSD) and You Only Look Once (YOLO) models, exhibiting promising results. Utilizing Convolutional Neural Networks (CNN), the proposed technique achieves real-time performance with satisfactory detection and classification results, offering enhanced accuracy. The investigated model attains an impressive accuracy range of 63–90% in object detection and classification, showcasing its effectiveness in challenging environments [7]. Qianjun Shuai and Xingwen Wu introduce enhancements to the Single Shot Multi-Box Detector (SSD), a prominent deep learning-based object detection algorithm. Incorporating Batch Normalization (Batch Norm) operations into the network aims to enhance generalization and accelerate training. The paper extends SSD's capabilities to include object counting in image recognition. Users can interact with the system on the front-end page, selecting data for real-time detection, and viewing results, including object types and counts, contributing to improved accuracy and efficiency in object detection [8].

Xin Lu, Xin Kang, Shun Nishide, and Fuji Ren address challenges in object detection and recognition, especially for small and diverse datasets of dangerous goods. They propose an innovative approach by integrating the Single Shot Multi-Box Detector (SSD) with a ResNet101 network, aiming to enhance learning efficiency and accuracy. The proposed model surpasses other neural networks, showcasing improved performance in learning efficiency and accuracy, particularly beneficial for applications involving small and diverse datasets of hazardous

materials [9]. Kanishk Wadhwa and Jay Kumar Behera delve into the realm of object detection, focusing on Real-Time Object Detection using the Single Shot Detector (SSD) algorithm. The paper aims to analyze and implement SSD for real-time object detection, showcasing its speed and accuracy, ultimately contributing to the evolving knowledge in this domain [10].

G Chandan, Ayush Jain, and Harsh Jain underscore the transformative impact of deep learning in the realm of Artificial Intelligence paper advocates the synergistic fusion of SSD and Mobile Nets, ensuring efficient object detection and tracking without compromising performance. This amalgamation contributes to the ongoing advancements in the field of deep learning-based object detection [11]. Qunjie Yin, Wenzhu Yang, Mengying Ran, and Sile Wang present an innovative solution to the challenge of detecting small objects in computer vision with their proposed method, Feature Fusion and Dilated Convolution-based Single Shot multi-box Detector (FD-SSD). Utilizing VGG-16 as the backbone network, the model incorporates a multi-layer feature fusion module and a multi-branch residual dilated convolution module. Experimental results on PASCAL VOC2007 and MS COCO datasets showcase FD-SSD's effectiveness, achieving significant improvements in small object detection [12].

Dr. S. V. Viraktamath, Ambika Neelopant, and Pratiksha Navalgi delve into the realm of object recognition within computer vision, a pivotal advancement facilitating the identification and localization of entities in digital images and videos. The paper comprehensively surveys various algorithms, quality metrics, speed-size trade-offs, and training approaches. Focusing on the You Only Look Once (YOLO) and Single Shot Multi-Box Detector (SSD) for single-step detection, as well as the Faster R-CNN for two-step detection, the study provides a holistic exploration of object recognition and its applications [13]. Hyeong-Ju Kang introduces a Convolutional Neural Network (CNN) accelerator designed for real-time object detection, addressing the challenge of high computational requirements. Employing a single-shot multi-box detector (SSD) with VGG16, the CNN accelerator implements a pruning scheme to reduce weight storage. The proposed design achieves a remarkable 42 frames per second (FPS) on the XC7VX690T FPGA, showcasing a VOC07 test mean Average Precision (mAP) of 78.13% [14]

.Mohit Phadtare, Varad Choudhari, Rushal Pedram, and Sohan Vartak explore the challenges of object recognition in aerial surveillance using camera-equipped drones. The paper delves into a comparative analysis of systematic approaches for detecting vehicles, people, and animals in images and real-time monitoring. Emphasizing the significance of precision, the research contributes valuable insights into object detection methodologies, particularly in the context of aerial surveillance, where robust performance is essential for reliable results [15]. Shuren Zhou and Jia Qiu introduce an improved Single Shot MultiBox Detector (SSD) method, termed MA-SSD, addressing limitations in quickly recognizing significant object areas in images. MA-SSD incorporates an attention mechanism to generate multi-scale attention features, enhancing the feature representation capability and boosting detection accuracy. Demonstrating competitive performance on the PascalVOC dataset, MA-SSD offers an innovative approach to refining object detection through interactive multi-scale attention features within the SSD framework [16].

Tang Cong, Ling Yongshun, Zheng Kedong, Yang Xing, Zheng Chao, Yang Hua, and Jin Wei present an innovative approach to address the limitations of classical Single Shot Multibox Detector (SSD) in small object detection. Their proposed method, Multi-View SSD, utilizes insights from convolutional receptive fields and default box scale mapping. Through extensive testing on a dataset, the study demonstrates that Multi-View SSD outperforms classical SSD in small object retrieval ability and detection precision, achieving notable improvements in Average F-measure (AF) and mean Average Precision (mAP) [17]. In the study by Tiagrajah V. Janahiraman and Mohamed Shahrul Mohamed Subuhan, traditional machine learning methods for traffic light detection are replaced by advanced deep learning techniques. This research demonstrates the effectiveness of deep learning, particularly Faster-RCNN, in robustly detecting traffic lights and highlights the potential for real-time applications in traffic management [18].

The article by Zhe Huang, Zhenyu Yin, Yue Ma, Chao Fan, and Anying Chai addresses the challenge of low accuracy in detecting small objects in mobile phone component images. They introduce an improved SSD algorithm, CSP-SSD (Cross Stage Partial SSD), which enhances the network structure using gradient flow information, deconvolution, and multi-scale transformation. The proposed CSP-SSD algorithm demonstrates high accuracy, particularly for small objects, addressing robustness concerns in mobile phone component image detection [19]. The paper by Khusni, U., Arymurthy, A.M., Susanto, H. focuses on the critical challenges in traditional object detection methods, emphasizing the limitations in accurately detecting small objects with varied backgrounds. This modification aims to address the accuracy issues associated with small object detection.

Experimental results demonstrate improved accuracy, with SSD (ResNeXt101) achieving 67.17%, surpassing the previous SSD framework with ResNet101, which had an accuracy of 66.09%. The study contributes valuable insights for enhancing object detection in real-world scenarios [20].

3. Methodology

The proposed methodology for this real-time object detection project is diving into the capabilities of the OpenCV framework alongside the Single Shot Multibox Detector (SSD) model. The project is initiated by configuring the webcam through the CV2 library to capture live video frames. Then, the selected SSD model is loaded into the system and configured with essential parameters such as input size, scale, mean, and colour swapping. To facilitate object identification, class names are loaded from the coco. names file. The system processes each video frame utilizing the SSD model, implementing a confidence threshold to refine the object detections. Timestamps are generated with the DateTime library, and pertinent data, including the class name, date, time, confidence level, and image path, is stored in a CSV file for subsequent analysis. Importantly, our methodology throws in an image capture twist using the os library, selectively storing images of detected animals or people. The resulting system is set to make a significant contribution to security applications, monitoring, and aiding in later clarifications, aligning with the project's overall objectives. The architecture of the proposed syatem is given in Figure 1.

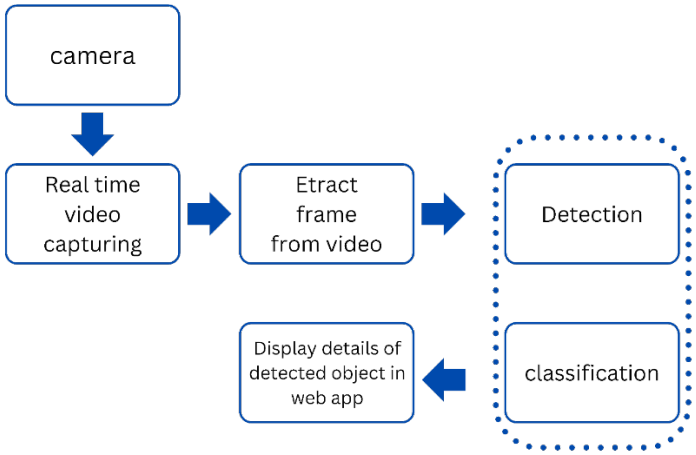


Fig. 1. Proposed System Architecture Diagram

3.1 Single Shot Multibox Detectors(SSDs)

A Single Shot Detector (SSD) is an innovative object detection algorithm in computer vision. It stands out for its ability to swiftly and accurately detect and locate objects within images or video frames. What sets SSD apart is its capacity to accomplish this in a single pass of a deep neural network, making it exceptionally efficient and ideal for real-time applications. SSD achieves this by employing anchor boxes of various aspect ratios at multiple locations in feature maps. These anchor boxes enable it to handle objects of different sizes and shapes effectively. Moreover, SSD uses multi-scale feature maps to detect objects at various scales, ensuring that both small and large objects in the image are accurately identified. With its proficiency in detecting multiple object classes simultaneously, SSD is a valuable tool for tasks that involve numerous object categories in a single image. The architecture of SSDs is given below in Figure 2.

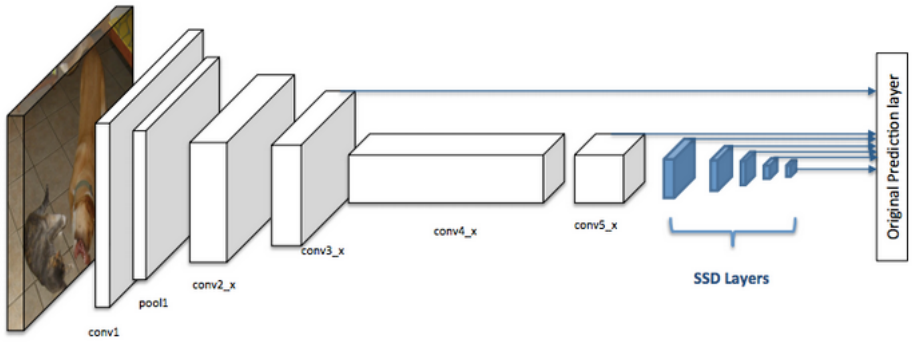


Fig 2. SSDs Architecture.

3.2 MobileNet

MobileNet is a CNN architecture model for Image Classification and Mobile Vision. There are other models as well but what makes MobileNet special is that it has very little computation power to run or apply transfer learning. This makes it a perfect fit for Mobile devices, embedded systems, and computers without GPU or low computational efficiency while compromising significantly the accuracy of the results

3.2.1 MobileNetV3

MobileNetV3 is a smarter and faster way of designing neural networks for tasks like image recognition. MobileNetV3 uses AutoML, a type of automated learning, to find the best network design without human intervention. It combines two techniques, MnasNet and NetAdapt, to first figure out a rough design and then fine-tune it for better performance. One feature of MobileNetV3 is the addition of "squeeze-and-excitation" blocks in its core structure. These blocks help improve the quality of the information the network learns by focusing on the most important parts and ignoring less useful details. This makes the network smarter in understanding and recognizing objects. The block of MobileNetV3 is given below in Figure 3.

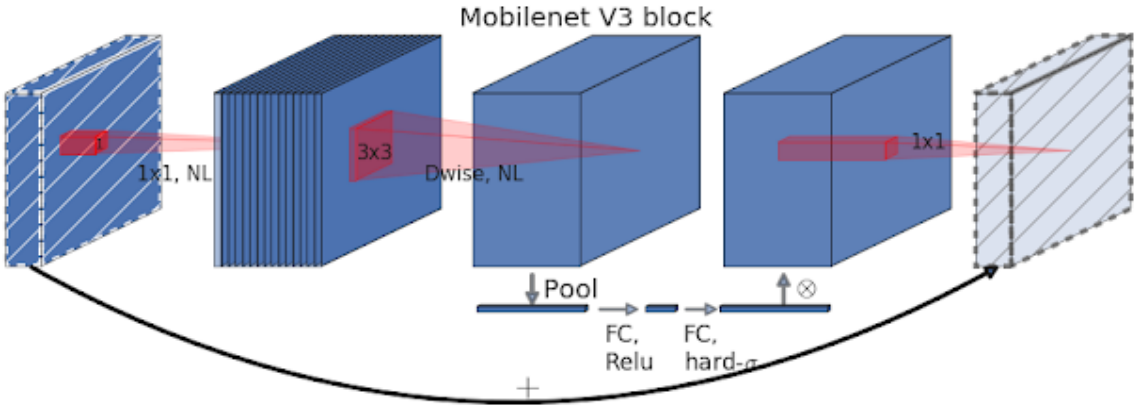


Fig 3. Block of MobileNetV3

Another clever thing about MobileNetV3 is how it optimizes some of the complex parts of its architecture. By doing this, it removes three complicated layers without losing accuracy, making the network faster. In tests, MobileNetV3 has shown a 25% decrease in the time it takes to recognize objects compared to its older versions while keeping the same level of accuracy. MobileNetV3 is a more efficient and intelligent way of building neural networks for tasks like recognizing objects in images. It uses automated learning, clever design choices, and optimizations to make the network faster and smarter.

3.3 Firebase

Google Firebase acts as a one-stop solution for developers, providing a unified platform to handle critical aspects of app development, from user authentication to real-time data management and performance monitoring. With Firebase, developers can enhance user experiences, expedite workflows, and ensure the reliability of their mobile applications. The view of data in the database is given below in Figure 4 and Figure 5.

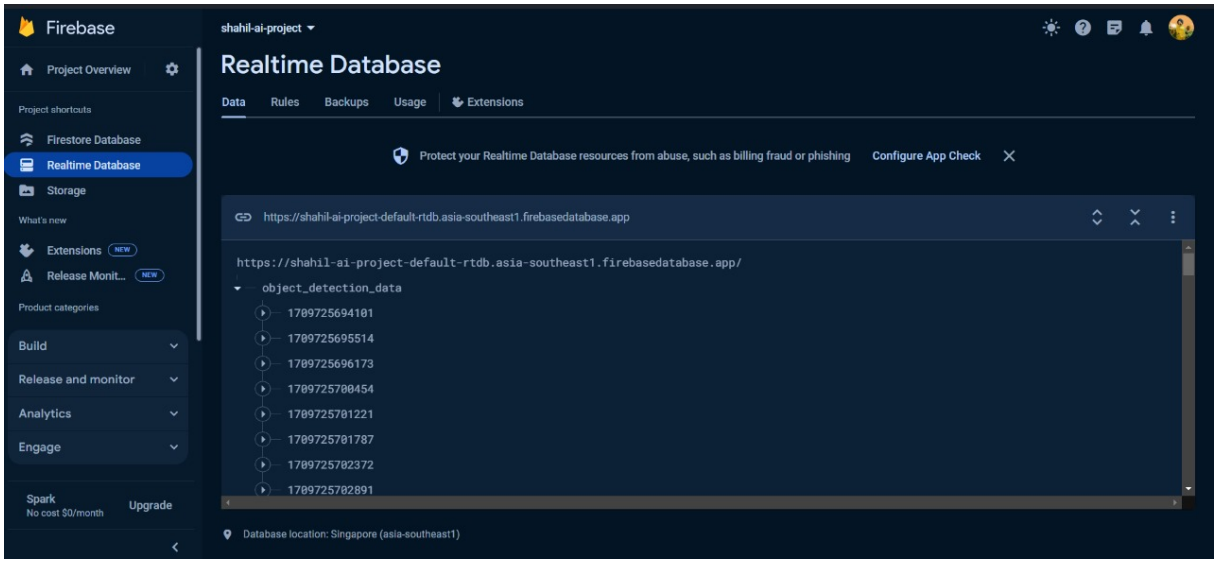


Fig 4. Storing of data in Firebase.

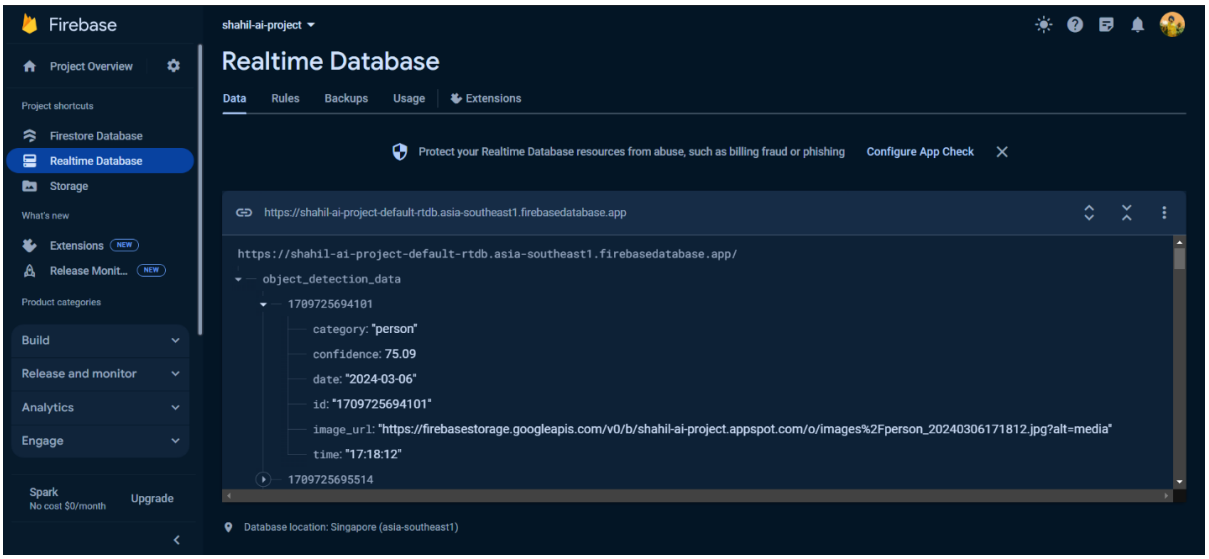


Fig 5. Details of each data in Firebase

3.4 User Interface

In this project, we provide the simplest user-friendly interface, users to access the details of detected objects. And used React to build the UI. From the user interface user can access data by filtering with category date time or confidence level and user can also hide any column if they need some sample images. The view of data in the web app is given below in Figure 6, Figure 7, Figure 8 and Figure 9.

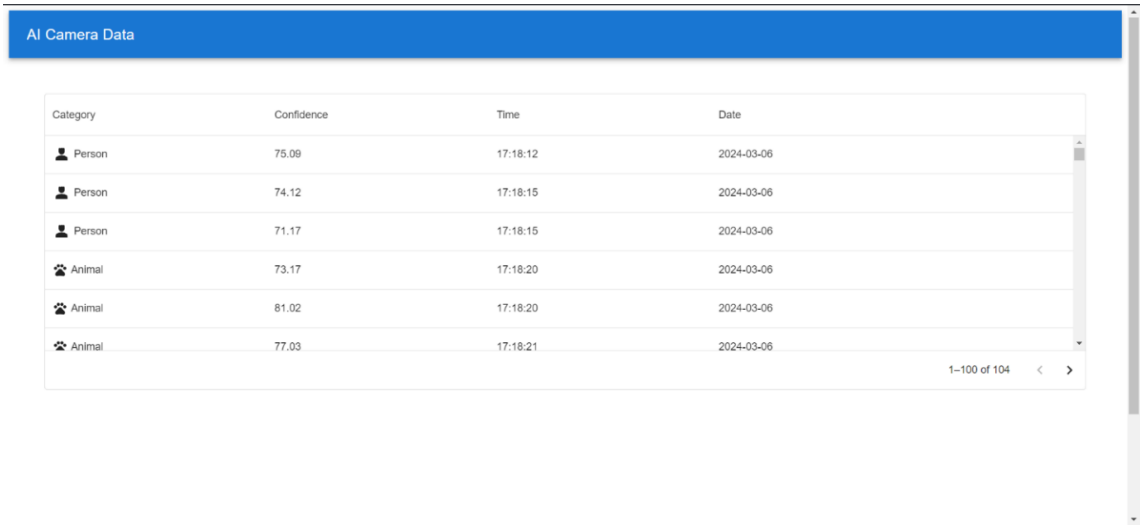


Fig 6. View of data in Web app

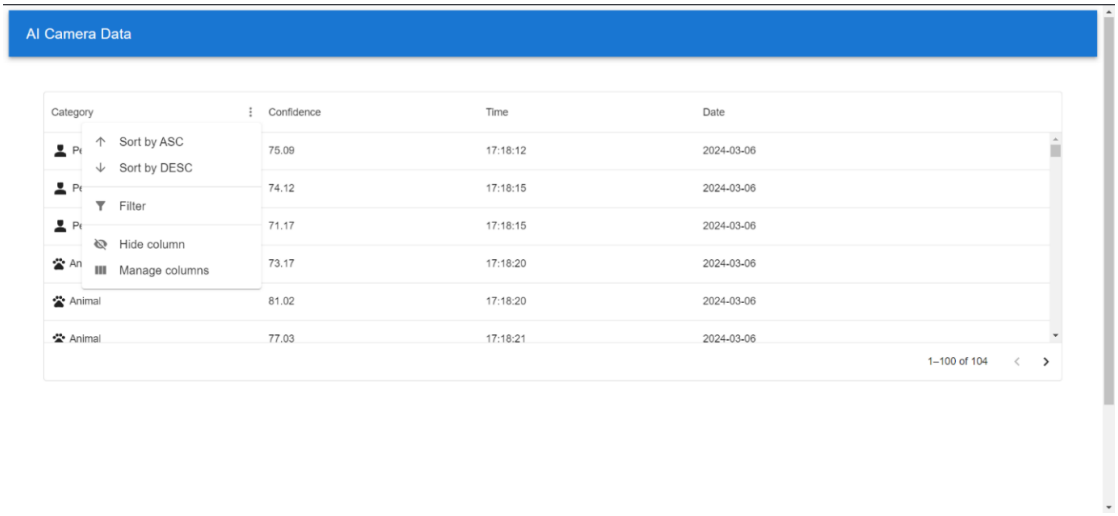


Fig 7. Filtering option of the data in the table.

The app shows the details including an image of the detected object, category name of the object, confidence level of the detection, date of the detection and time of the detection. They take the data lively.

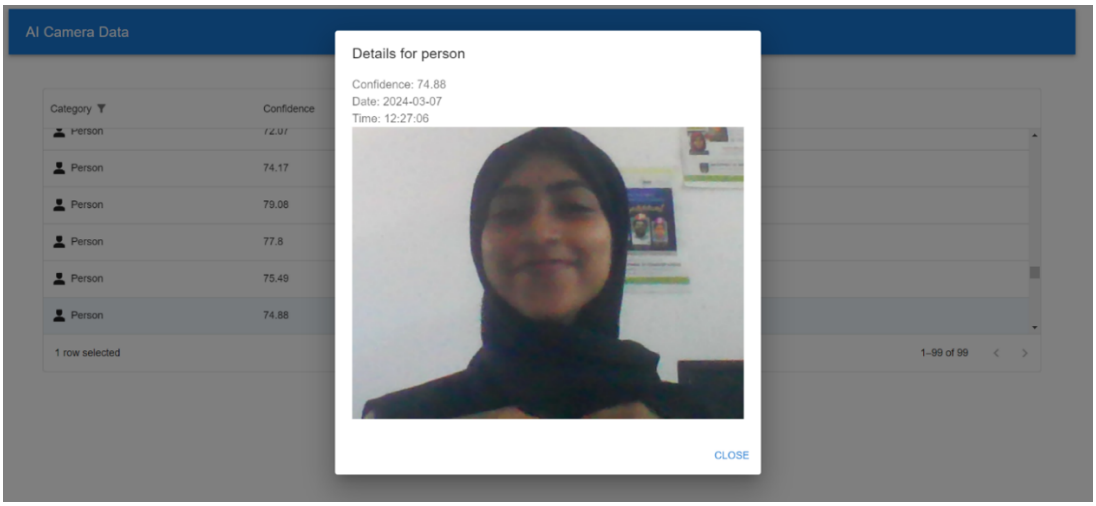


Fig 8. View of detected Person details in the Web app.

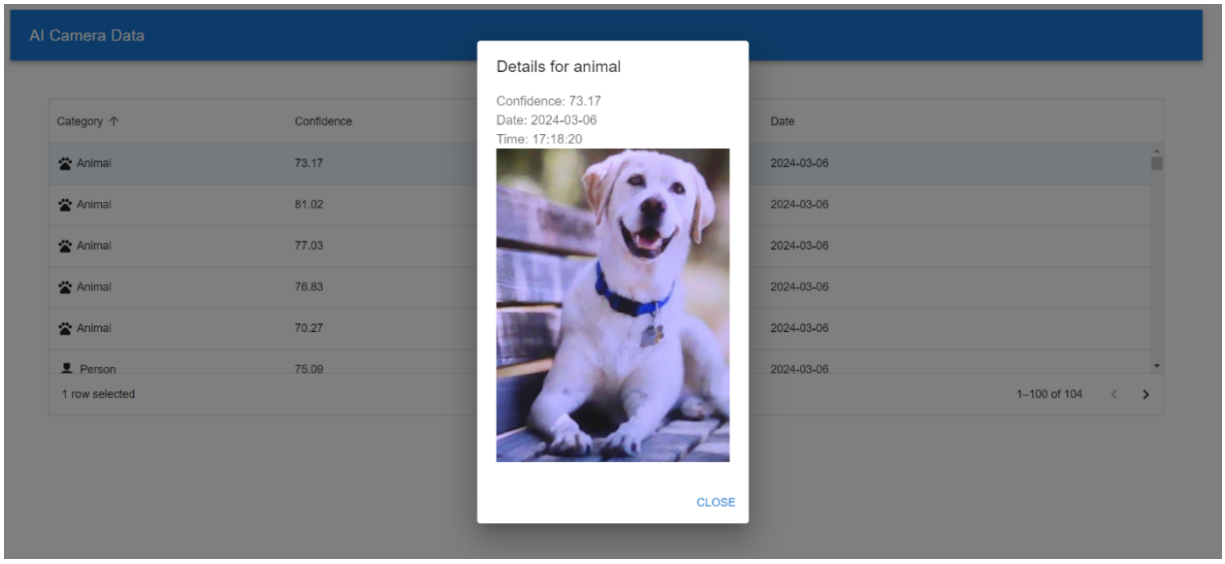


Fig 9. View of detected Animal details in the Web app.

4. Experimental Result

The integration of SSD, MobileNetV3, OpenCV, ReactJS, Webcam worthy results across key aspects:

4.1 Object Detection

The object detection in Single Shot Detection (SSD) is very fast and highly accurate. It quickly detects the object almost. It captures the detected object very correctly and quickly. And passes the details to the firebase without any fail. The anchor boxes are bounded correctly. If we use a GPU GPU-enabled system it will increase the detection of the object. The view of video detection from using a webcam is given in Figure 10 and Figure 11.

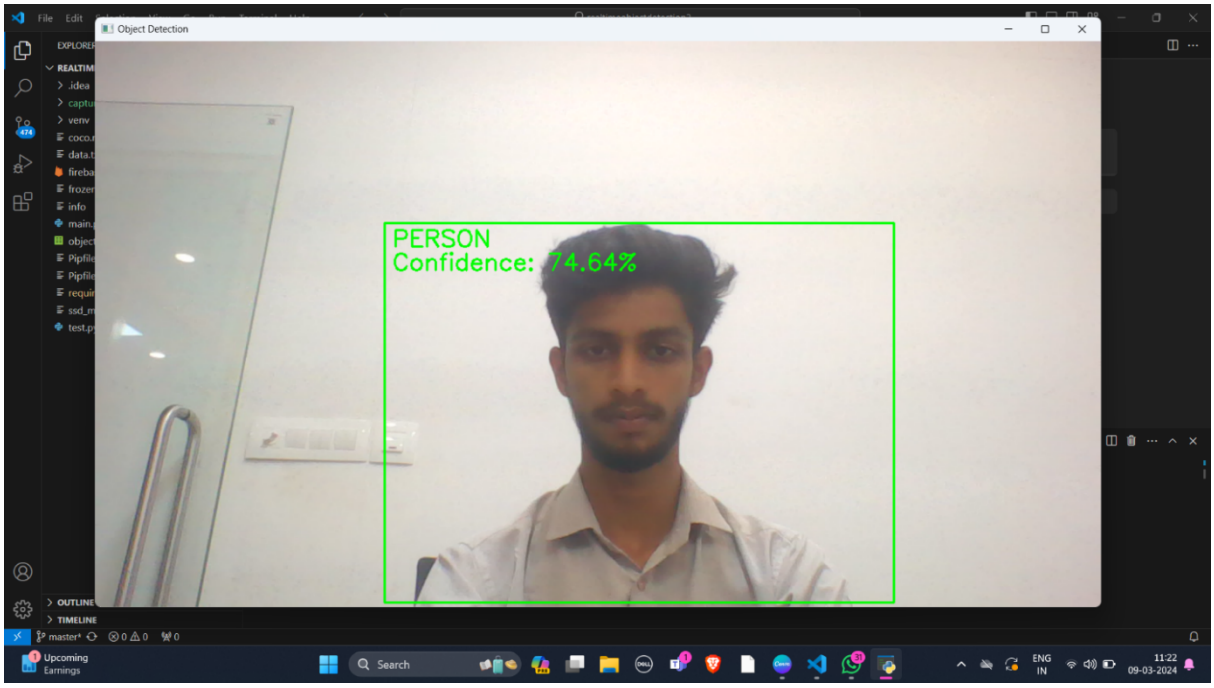


Fig 10. Detection of a person from the live footage

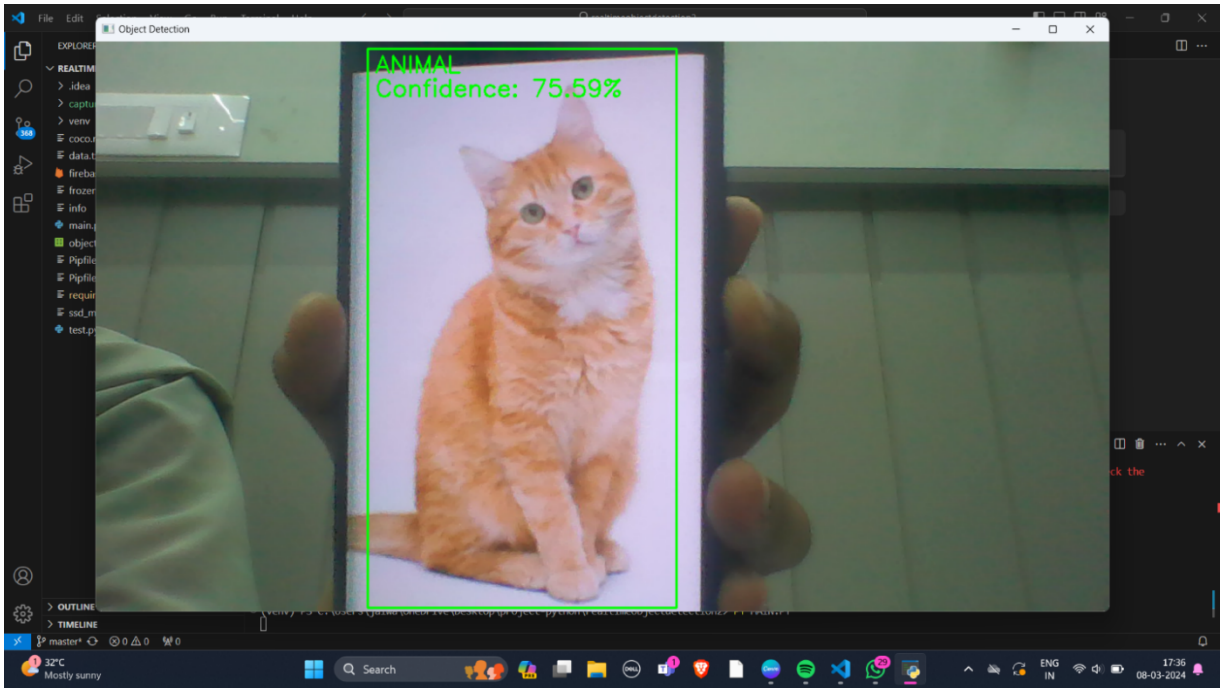


Fig 11. Detection of Animals from live footage

4.2 Camera Accuracy

In Our project, we have used a normal laptop Webcam as the camera for detecting objects in real-time. Comparatively, the captured images of the detected object are of low quality due to the normal camera of the laptop. But we have got maximum good results that a laptop webcam can do with. If we use a good camera for capturing the video surveillance it can attain a good result.

4.3 Confidence Score

The confidence score in object detection represents the algorithm's certainty that a detected object accurately corresponds to a particular class. Expressed as a probability, it signifies the level of confidence in the assigned class, aiding in decision-making and filtering false positives during object recognition processes. The confidence score depends on the camera and the accuracy of the object detection model.

4.3.1 Average confidence of person

Sample images of the detection of a different person with their confidence score are given below in Figure 12.

Fig 12. Sample images of the detection of a person with confidence

Total confidence score = 69.71+70.35+71.27+70.01+74.53+66.82+69.46+69.77 = 561.92

No of detection = 8

Average confidence = Total confidence score/No of detection
= 561.92/8 = 70.24

4.3.2 Average confidence of animal

Sample images of the detection of different animals with their confidence score are given below in Figure 13.

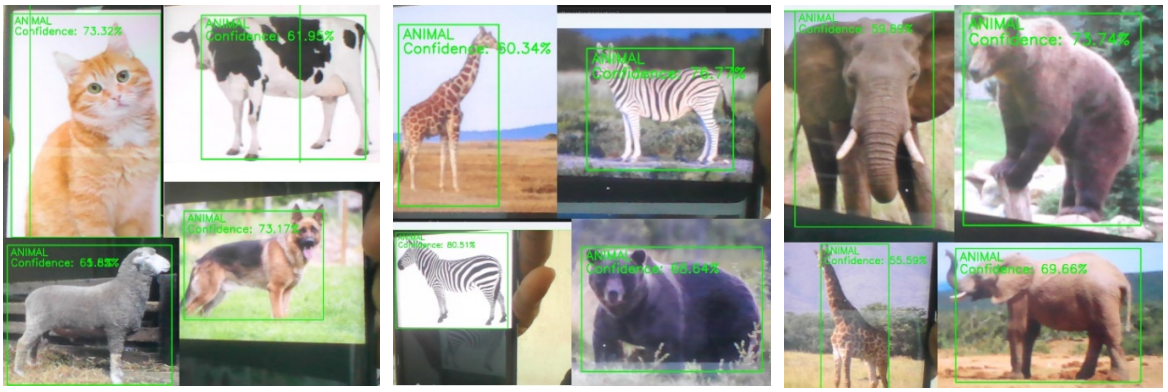


Fig 13. Sample images of the detection of different animals with confidence

Total confidence score = 73.32+61.95+65.68+73.17+60.34+76.77+80.51+68.64+59.89+

73.74+55.59+69.66 = 819.59

No of detection = 12

Average confidence = Total confidence score/No of detection
= 819.59/12 = 68.299

5. Discussion

In this project, we used low low-configured system, so we can't use highly accurate algorithms like YOLO V8. Also, we used a normal laptop camera which has a low resolution, which causes the detection accuracy and also gives low-quality images of the detected object. It makes the image low clarity. As we have used low-specification computers and cameras in our project it makes it so difficult to use the modern algorithm and made our project accuracy slightly low. For better accuracy, a camera with high shutter speed and resolution and a computer with GPU with Cuda-enabled graphics are necessary for localizing the project. If we use the Jetson Nano developer kit it will increase the performance of the program in the embedded system.

6. Conclusion

In conclusion, our project presents an innovative approach to wildlife monitoring and security surveillance through the use of a real-time object detection system. By utilizing the Single Shot Detection (SSD) algorithm, which is driven by MobileNetV3, the system exhibits quick and precise object detection for people, animals, and vehicles. The project's functionality is improved by the combination of OpenCV, ReactJS, and Firebase. This makes it possible to efficiently catalogue and store important characteristics like item type, confidence level, and timestamp. The intuitive interface makes it simple to retrieve the catalogued data, which eliminates the need to go through a lot of video material and speeds up event retrieval.

Our tests confirm that, even with a regular laptop webcam, the system is effective in obtaining high object detection accuracy. Rapid and accurate object identification is ensured by the combination of SSD, MobileNetV3, and OpenCV, and the corresponding confidence scores offer useful details about the certainty of the system. Although the project works well with a standard webcam, better cameras might be investigated for possible improvements to image capture and overall system performance. The system's dependability is shown by the average confidence scores that are consistently obtained for both animal and person detection. This establishes

the system's status as a strong tool for thorough cataloguing and real-time object detection in security and wildlife management scenarios.

References

[1] Jisoo Jeong, Hyojin Park, Nojun Kwak., "Enhancement of SSD by concatenating feature maps for object detection.", arXiv preprint arXiv:1705.09587, 2017 https://www.researchgate.net/publication/317205721_Enhancement_of_SSD_by_concatenating_feature_maps_for_object_detection

[2] Reagan L Galvez, Argel A Bandala, Elmer P Dadios, Ryan Rhay P Vicerra, Jose Martin Z Maningo "Object detection using convolutional neural networks" TENCON 2018-2018 IEEE Region 10 Conference, 2023-2027, 2018 https://www.semanticscholar.org/paper/Object-Detection-Using-Convolutional-Neural-Galvez-Bandala/cf7b19595a67e4a6aa3c63c3c470af318b196099

[3] Ayesha Younis, Li Shixin, Shelembi Jn, Zhang Hai "Real-time object detection using pre-trained deep learning models MobileNet-SSD" Proceedings of 2020 the 6th International Conference on computing and Data Engineering, 44-48, 2020 https://www.researchgate.net/publication/339778643_Real-Time_Object_Detection_Using_Pre-Trained_Deep_Learning_Models_MobileNet-SSD

[4] Sheping Zhai, Dingrong Shang, Shuhuan Wang, Susu Dong DF-SSD: "An improved SSD object detection algorithm based on DenseNet and feature fusion" IEEE access 8, 24344-24357, 2020 https://www.researchgate.net/publication/339025176_DF-SSD_An_Improved_SSD_Object_Detection_Algorithm_Based_on_DenseNet_and_Feature_Fusion

[5] Zuge Chen, Kehe Wu, Yuanbo Li, Minjian Wang, Wei Li SSD-MSN: "An improved multi-scale object detection network based on SSD" IEEE Access 7, 80622-80632, 2019 https://www.researchgate.net/publication/333793862_SSD-MSN_An_Improved_Multi-scale_Object_Detection_Network_Based_on_SSD

[6] Abhinav Juneja, Sapna Juneja, Aparna Soneja, Sourav Jain "Real time object detection using CNN based single shot detector model" Journal of Information Technology Management 13 (1), 62-80, 2021 https://jitm.ut.ac.ir/article_80025.html

[7] Md Alamin Feroz, Marjia Sultana, Md Rakib Hasan, Aditi Sarker, Partha Chakraborty, Tanupriya Choudhury "Object detection and classification from a real-time video using SSD and YOLO models" Computational Intelligence in Pattern Recognition: Proceedings of CIPR 2021, 37-47, 2022 https://www.researchgate.net/publication/354374123_Object_Detection_and_Classification_from_a_Real-Time_Video_Using_SSD_and_YOLO_Models

[8] Qianjun Shuai, Xingwen Wu "Object detection system based on SSD algorithm" 2020 international conference on culture-oriented science & technology (ICCST), 141-144, 2020 https://www.semanticscholar.org/paper/Object-detection-system-based-on-SSD-algorithm-Shuai-Wu/1b75d4c4621ec317971ce684b5a74f6c2b3b0441

[9] Xin Lu, Xin Kang, Shun Nishide, Fuji Ren "Object detection based on SSD-ResNet" 2019 IEEE 6th International Conference on Cloud Computing and Intelligence Systems (CCIS), 89-92, 2019 https://www.researchgate.net/publication/340895639_Object_detection_based_on_SSD-ResNet

[10] Kanishk Wadhwa, Jay Kumar Behera "Accurate real-time object detection using SSD" International Research Journal of Engineering and Technology (IRJET) 7 (05), 2395-0072, 2020 https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=Kanishk+Wadhwa%2C+Jay+Kumar+Behera%2C+Accurate+Real-Time+Object+Detection+using+SSD+SRM+Institute+of+Science+and+Technology&btnG=

[11] G Chandan, Ayush Jain, Harsh Jain "Real time object detection and tracking using Deep Learning and OpenCV" 2018 International Conference on inventive research in computing applications (ICIRCA), 1305-1308, 2018 https://www.researchgate.net/publication/331421347_Real_Time_Object_Detection_and_Tracking_Using_Deep_Learning_and_OpenCV

[12] Qunjie Yin, Wenzhu Yang, Mengying Ran, Sile Wang "An improved SSD object detection algorithm based on feature fusion and dilated convolution" Machine Vision Engineering Research Center, Hebei University, Baoding 071002, China https://dl.acm.org/doi/10.1016/j.image.2021.116402

[13] Dr. S. V. Viraktamath,Ambika Neelopant,Pratiksha Navalgi "Comparison of YOLOv3 and SSD Algorithms" International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181 https://www.ijert.org/comparison-of-yolov3-and-ssd-algorithms

[14] Hyeong-Ju Kang "Real-Time Object Detection on 640x480 Image With VGG16+SSD" 2019 International Conference on Field-Programmable Technology (ICFPT) https://ieeexplore.ieee.org/document/8977916

[15] Shuren Zhou, Jia Qiu, "Enhanced SSD with interactive multi-scale attention features for object detection" Multimed Tools Appl 80, 11539–11556 (2021). https://doi.org/10.1007/s11042-020-10191-2

[16] Mohit Phadtare, Varad Choudhari, Rushal Pedram, Sohan Vartak "Comparison between YOLO and SSD Mobile Net for Object Detection in a Surveillance Drone". International journal of scientific research in engineering and management 2021 https://www.researchgate.net/profile/Varad-Choudhari-5/publication/355336797_Comparison_between_YOLO_and_SSD_MobileNet_for_Object_Detection_in_a_Surveillance_Drone/links/616a524c951b3574c649ff89/Comparison-between-YOLO-and-SSD-MobileNet-for-Object-Detection-in-a-Surveillance-Drone.pdf

[17] Tang Cong, Ling Yongshun, Zheng Kedong, Yang Xing, Zheng Chao, Yang Hua, Jin Wei. "object detection method of multi-view SSD based on deep learning." National University of Defense Technology,Hefei 230037,China; Key Laboratory of Infrared and Low Temperature Plasma of Anhui Province,Hefei 230037,China; State Key Laboratory of Pulsed Power Laser Technology,Hefei 230037,China; https://www.researchgate.net/publication/325084336_Object_detection_method_of_multi-view_SSD_based_on_deep_learning

[18] Tiagrajah V. Janahiraman; Mohamed Shahrul Mohamed Subuhan. "Traffic Light Detection Using Tensorflow Object Detection Framework" 2019 IEEE 9th International Conference on System Engineering and Technology (ICSET) https://www.researchgate.net/publication/337527327_Traffic_Light_Detection_Using_Tensorflow_Object_Detection_Framework

[19] Zhe Huang, Zhenyu Yin, Yue Ma, Chao Fan, Anying Chai. "Mobile phone component object detection algorithm based on improved SSD" Shenyang Institute of Computing Technology, Chinese Academy of Sciences, Shenyang 110168, China, University of Chinese Academy of Sciences, Beijing 100049, China https://www.sciencedirect.com/science/article/pii/S1877050921005020

[20] Khusni, U., Arymurthy, A.M., Susanto, H. (2022). "Small Object Detection Based on SSD-ResNeXt101". In: Mahyuddin, N.M., Mat Noor, N.R., Mat Sakim, H.A. (eds) Proceedings of the 11th International Conference on Robotics, Vision, Signal Processing and Power Applications. Lecture Notes in Electrical Engineering, vol 829. Springer, Singapore. https://link.springer.com/chapter/10.1007/978-981-16-8129-5_162

Computer Science Research Reports (2024), 2-1, 26-33
ISBN : 978-93-94056-72-5

RESEARCH ARTICLE

ENHANCING TOURISM SERVICES WITH AI CHATBOT FOR GOD'S OWN COUNTRY, KERALA

Muhsina Parveen[1*], Nizal Abdurahiman[2], Fadi Basheer K[3], Shibila KT[4], Shameem K[5]

[1,2,34,5] *PG&Research Department of Computer Science, Sullamussalam Science College, Areekode, Affiliated to University of Calicut*

**muhsinamuhsina703@gmail.com*

Abstract: *Start a virtual journey through the captivating districts of Alappuzha, Wayanad, and Calicut in Kerala, guided by your very own personalized AI Travel Chatbot. This innovative project uses the power of technology to revolutionize the way we explore and experience travel destinations. Using Dialog flow, our chatbot becomes your digital companion, equipped with a vast repository of information about these enchanting districts. As you engage in conversation with the chatbot, it intuitively gauges your interests and preferences, prompting you to specify your preferences, whether it's activities you enjoy, or the type of experience you're seeking. Based on your responses, the chatbot crafts a bespoke itinerary tailored specifically to your desires, seamlessly blending must-see attractions, hidden gems, and local experiences. Whether you're a nature enthusiast craving the serene backwaters of Alappuzha, an adventure seeker yearning for the lush landscapes of Wayanad, or a history buff eager to explore the cultural heritage of Calicut, the chatbot ensures your journey is personalized to perfection. But the innovation doesn't stop there – we've integrated this intuitive chatbot into a user-friendly website using Django templates for the front end and Django for the back end, with SQLite serving as the database backbone. This integration seamlessly merges the AI capabilities of the chatbot with the interactive interface of the website, offering users a cohesive and engaging experience from start to finish. Picture yourself browsing the website, effortlessly navigating through vibrant visuals and informative content, before seamlessly transitioning to a conversation with the chatbot, where you can further refine your preferences and receive personalized recommendations. Through this fusion of AI and web integration, we aim to redefine the travel planning experience, making it not just efficient but also enjoyable and immersive. Gone are the days of sifting through endless travel guides or relying on generic recommendations – with our AI Travel Chatbot, your journey is as unique as you are. Whether you're a seasoned traveller or a first-time explorer, our project empowers you to start on a personalized adventure that reflects your individual interests and aspirations. Join us on this exciting journey as we use the power of technology to unlock the true essence of travel – one conversation at a time.*

Keywords- AI Travel Chatbot, Personalized itinerary, Dialog flow Integration, Django website

1. INTRODUCTION

In today's fast-paced digital age, the landscape of travel has undergone a remarkable evolution. Our project delves into the captivating fusion of artificial intelligence and travel exploration, focusing on the picturesque districts of Alappuzha, Wayanad, and Calicut in Kerala, India. Using Dialog flow, a powerful conversational AI platform developed by Google, we have created a chatbot capable of engaging users in intuitive conversations about their travel preferences and interests. Through a seamless integration of data and dialogue, our chatbot crafts bespoke itineraries, ensuring that each traveller's experience is uniquely tailored to their desires.

The vision of our AI Travel Chatbot lies in its adaptability and responsiveness to user input, offering a dynamic and personalized journey planning experience. As users interact with the chatbot, they are prompted to specify their interests and preferences, whether it's exploring nature reserves, immersing in cultural experiences, or seeking adrenaline-pumping adventures. This interactive dialogue forms the foundation of our project, enabling the chatbot to understand and respond to the diverse needs of travellers. With each exchange, the chatbot refines its recommendations, ensuring that every itinerary is finely tuned to match the user's individual tastes and preferences.At the heart of our project lies the seamless integration of AI technology into the fabric of travel planning, facilitated by the versatile capabilities of Dialog flow. Through extensive training and data input, we have equipped our chatbot with a comprehensive understanding of the attractions, activities, and highlights of Alappuzha, Wayanad, and Calicut. This deep knowledge base enables the chatbot to offer personalized recommendations and insider insights, transforming the travel planning process into a personalized and immersive experience. By harnessing the power of AI, we unlock new possibilities in travel exploration, empowering users to discover the beauty and richness of Kerala's diverse landscapes and cultural heritage.

In addition to its AI-driven intelligence, our chatbot seamlessly integrates into a user-friendly website, enhancing accessibility and convenience for travellers. Built using Django templates for the front end and Django for the back end, with SQLite as the database, the website serves as a hub for interactive engagement and information exchange. Users can effortlessly navigate through vibrant visuals and informative content before seamlessly transitioning to a conversation with the chatbot. This integration creates a cohesive and immersive experience, bridging the gap between AI technology and user interaction, and redefining the way travellers plan and experience their journeys. The completion of our project marks a groundbreaking effort in using AI technology to change how people plan their travels. By combining cutting-edge technology with a user-centric approach, we have created a platform that empowers travellers to embark on personalized and unforgettable journeys, guided by the expertise of our AI Travel Chatbot. As we continue to refine and enhance our project, we remain committed to pushing the boundaries of innovation and delivering unparalleled experiences that inspire and delight travellers around the world.

2. LITERATURE REVIEW

The paper [1] describes the development of a complicated computer science (AI) chatbot for advising prompt actions when they need to see a doctor. Moreover, offering a virtual assistant may suggest which sort of doctor to consult. During the epidemic, managing the flow of a large number of patients for consultation has been a tough game for hospitals or healthcare workers. It is becoming more difficult to contact a doctor considering the recent situation, especially in rural areas. It's obvious that well-designed and operated chatbots may actually be helpful for patients by advocating precautionary measures and cures, as well as taken to prevent harm inflicted by worry.

The paper [2] describes integration of Artificial Intelligence (AI) and chatbot technology offers promising avenues for enhancing healthcare services, notably in disease prediction. AI chatbots, adept at emulating human-like interactions, aid in patient triaging and medical advice. This literature review scrutinizes the efficacy and potential of AI chatbots in disease prediction, with a focus on early intervention and treatment. Analyzing 24 journals, with emphasis on 2020 publications, reveals their significant role in supporting healthcare professionals and improving diagnostic accuracy through machine learning algorithms. Continuous research is crucial to refine AI chatbots, ultimately revolutionizing healthcare delivery and fostering better patient outcomes.

The paper [3] describes the challenges in education due to concerns about unethical student behavior and passive learning. To address this, embracing AI as an educational trend and establishing pedagogical principles are essential. This review proposes three principles informed by Zimmerman's Self-Regulated Learning framework and Judgment of Learning: goal setting, self-assessment and feedback, and personalization. Teaching prompting, reverse prompting, and data-driven mechanisms in AI chatbots facilitate student self-regulation and learning reflection. By integrating AI-assisted pedagogy, educators can foster self-regulated learning and enhance instructional design in higher education.

In This paper [4] the author allows users to communicate with computers in natural language, similar to human conversations. Chatbots interact with clients and provide answers based on human input. The user may believe they are conversing with a human, while in reality, they are interacting with a machine. The chat bot program enables students to inquire about college admissions from any location with an internet connection and receive prompt responses. The chatbot technology streamlines the admissions process by delivering necessary information to students and parents, reducing the manpower effort in addressing inquiries.

In this paper [5] Cultural Tourism in Kerala, known as 'God's Own Country', emphasizes visitors' desire to explore and consume tangible and intangible cultural attractions. While the quantitative impact of cultural tourism remains unassessed, the influx of tourists to cultural destinations significantly contributes to the state's economy. Research underscores the importance of redefining cultural events, monuments, and religious sites to enhance visitor satisfaction and attract more cultural tourists. This literature review highlights the necessity for revamping cultural offerings to meet tourists' expectations and underscores the potential economic benefits for Kerala's tourism industry. This paper [6] highlights Kerala's tourism sector as a significant economic contributor, promoted as 'God's Own Country' by the Department of Tourism. The state boasts diverse attractions attracting millions annually, including beaches, hills, backwaters, festivals, and wildlife sanctuaries. Key destinations like Kochi, Munnar, and Wayanad draw both national and international tourists, with peak season from September to March. Despite setbacks from COVID-19, foreign exchange earnings have generally increased. Analysis using simple averages, percentages, and regression statistics elucidates tourism trends. The study aims to benefit academics, policymakers, and industry stakeholders, anticipating insights from hypothesis testing to inform business development and socio-economic strategies.

This PhD thesis paper [7] addresses the potential pitfalls of replacing human assistants with chatbots in contact centres. Focusing on Portuguese datasets, the research aims to develop a comprehensive solution comprising three core components: Dialog Extraction, Representation, and Guidance. These components facilitate the analysis of dialogues, identification of common interaction patterns, and provision of real-time support to human agents. By integrating these elements into a unified framework, the project seeks to enhance the efficiency of human-operated contact centres while preserving essential human interaction. The overarching goal is to leverage technology to optimize human performance without sacrificing the personalized touch inherent in human interactions. This paper explores the [8] utility of chatbots, particularly in the educational sector, as virtual assistants for automating responses and facilitating communication. AICMS, an AI-based CollegeBot management platform, is developed to address queries regarding college information, class schedules, and academic examinations in professional engineering colleges. Leveraging Dialogflow and Google API, the system operates as a messenger on Facebook, accepting text and voice inputs and providing prompt and accurate responses in both text and voice formats. This interactive platform effectively serves students and staff, streamlining information dissemination and enhancing communication efficiency within educational institutions.

This paper [9] underscores the efficacy of chatbots in streamlining customer interactions and ensuring round-the-clock satisfaction. Traditional models often struggle with varied user inputs, prompting a shift towards AI-driven systems capable of contextual understanding. Leveraging Dialogflow, our proposed model offers an end-to-end solution, abstracting algorithm complexities. Through technologies like Node.js, it enhances conversational continuity and functionality. The system is tailored to assist users in locating restaurants based on cuisine and top reviews, utilizing real-time location data for enhanced results. This approach marks a significant advancement in chatbot design, promising improved user experiences and efficient task execution. In this paper [10] tourism emerges as a significant global social phenomenon, driven by the human desire for novel experiences, education, and entertainment. Motivations span religious, business, and educational interests, fostering economic growth particularly in remote areas. However, benefits often concentrate in major gateway cities, leaving host regions with limited gains. Despite disparities, tourism remains a rapidly expanding sector, shaping economies worldwide. Its impact, both positive and negative, is undeniable, solidifying its status as the fourth-largest industry globally. This review underscores the importance of green tourism and sustainability, emphasizing the need for responsible practices in this thriving sector.

3. METHODOLOGY

3.1 Data Collection and Preparation:

Our Ai chatbot depends on gathering and preparing data about Alappuzha, Wayanad, and Calicut districts in Kerala. We take a thorough approach, gathering insights from various sources like tourism websites, guides, forums, and local experts. Utilizing web scraping techniques, we gather real-time data on attractions, accommodations, places, and food. To ensure accuracy, we conscientiously. clean and preprocess the data, removing inconsistencies, duplicates, and irrelevant details. This ensures our dataset is reliable and trustworthy. By compiling such a comprehensive dataset, we lay the foundation for training our AI chatbot to offer personalized recommendations to users. By gathering insights from a multitude of sources, we ensure our dataset is rich and diverse, capturing the essence of each district. We utilise web scraping to collect up-to-date information, covering

everything from popular tourist spots. Rigorous cleaning and preprocessing eliminate any inaccuracies or redundant data, ensuring the integrity of our dataset. With this robust foundation in place, our AI chatbot is equipped to provide users with tailored recommendations based on their preferences and interests. Through this meticulous approach, we create a dataset that accurately reflects the unique characteristics of Alappuzha, Wayanad, and Calicut. By eliminating inconsistencies and irrelevant information, we ensure the reliability of our dataset. This enables our AI chatbot to deliver personalized recommendations, enriching users' travel experiences and helping them explore the diverse attractions of Kerala's districts.

3.2 Rigorous Data Cleaning and Preprocessing:

Following data collection, our dataset undergoes a rigorous process of cleaning and preprocessing to uphold the highest standards of accuracy and reliability. This involves a methodical examination aimed at eliminating inconsistencies, identifying and removing duplicates, and filtering out extraneous or irrelevant information. By meticulously curating our dataset, we ensure that it accurately reflects the essence and diversity of each district, serving as a solid foundation for training our AI chatbot. Through this diligent approach, we lay the groundwork for the chatbot to deliver personalized and insightful recommendations that resonate with the preferences and interests of users, thus enhancing their travel planning experience.

3.3 Training the AI Chatbot using dialog flow:

After meticulously gathering and organizing our dataset, we move on to the crucial step of training our AI chatbot using Dialog flow, a powerful platform developed by Google for understanding natural language. With Dialog flow, we focus on creating different components like intents, entities, and contexts. These components are like building blocks that help the chatbot understand what users are asking and respond accurately. For instance, intents define the purpose of user messages, entities identify important information in those messages, and contexts help the chatbot keep track of the conversation's context.

This training process is all about teaching the chatbot to understand and interpret the intricacies of human language. By doing so, the chatbot becomes capable of engaging in seamless conversations with users, which significantly improves the overall user experience. Imagine being able to ask the chatbot questions or make requests just like you would with another person, and receiving helpful and relevant responses in return. That's the level of interaction we aim to achieve through Dialog flow's sophisticated capabilities.

In essence, training our AI chatbot with Dialog flow is like teaching it a new language - the language of the users. We meticulously craft its understanding of this language by fine-tuning its responses based on real user queries. This enables the chatbot to not only comprehend what users are saying but also respond in a way that feels natural and intuitive. As a result, users feel more comfortable and satisfied when interacting with the chatbot, leading to an overall enhanced experience in using our travel planning service.

3.4 Fine-tuning with Iterative Training Sessions:

During the training process, we continuously refine the chatbot's language skills to make it better at understanding and responding to user queries. This involves fine-tuning its language model through iterative training sessions. Additionally, we enhance the chatbot's knowledge by giving it specific information about the districts of Alappuzha, Wayanad, and Calicut. This includes details about tourist attractions, accommodations, local events, and other relevant data. By incorporating this domain-specific information into its knowledge base, the chatbot becomes more capable of providing personalized recommendations and insights tailored to each user's preferences and interests.

In simpler terms, we're basically teaching the chatbot to become an expert on travel in these districts. The more it learns about the local attractions and activities, the better it can assist users in planning their trips. So, when users interact with the chatbot, they can expect more accurate and helpful responses that cater to their individual needs and preferences.

3.5 Incorporating Domain-specific Information:

In our quest to enhance the chatbot's capabilities, we intricately incorporate domain-specific knowledge about the districts of Alappuzha, Wayanad, and Calicut. By infusing the chatbot's knowledge base with pertinent

information about notable attractions, accommodations, local events, and unique experiences in each district, we enable it to deliver highly personalized recommendations and insights to users. This strategic integration ensures that the chatbot is adept at catering to the diverse needs and preferences of travellers, thereby elevating the overall quality of the travel planning experience.

3.6 Development of the website interface:

Simultaneously, we embark on the development of a user-friendly website interface that serves as the primary interaction point between users and the AI chatbot. Employing Django, a versatile Python web framework, we design and implement intuitive user interfaces characterized by seamless navigation and interactive elements. The website interface incorporates input forms, buttons, menus, and visual cues to facilitate effortless communication and engagement. Prioritizing responsive design principles, we ensure that the website adapts seamlessly to various devices and screen sizes, enhancing accessibility and usability for users across different platforms. Through meticulous attention to detail and user-centric design principles, we create a cohesive and immersive web experience that complements the functionality of the AI chatbot, fostering seamless interaction and enhancing the overall user experience. Additionally, we introduce a dedicated "Tour Packages" section within our integrated website. This section offers users the convenience of selecting pre-designed tour packages tailored to their preferences. Users can choose from a range of inbuilt packages, each featuring a selection of destinations and activities. Moreover, users have the flexibility to select sub-destinations within a main destination, enabling them to customize their itinerary further. The corresponding itinerary for the selected sub-destination is then dynamically displayed, providing users with a streamlined and efficient means of exploring the districts of Alappuzha, Wayanad, and Calicut. This feature enhances the user experience by offering curated travel options that cater to a variety of preferences and interests, ultimately facilitating seamless trip planning and decision-making.

3.7 Integration and Testing:

The final phase of our methodology entails the integration and rigorous testing of the AI chatbot and website interface to ensure optimal performance and user satisfaction. We seamlessly integrate the backend functionality of the chatbot with the frontend interface of the website, establishing robust communication channels and data exchange mechanisms. Through comprehensive testing procedures, including unit testing, integration testing, and user acceptance testing, we systematically evaluate the system's functionality, usability, and performance. Beta testing and stakeholder feedback play a pivotal role in identifying and addressing any bugs, errors, or usability issues, thus refining the system iteratively. By adhering to best practices in software development and quality assurance, we ensure that our personalized AI travel chatbot project meets the highest standards of excellence and readiness for deployment, poised to revolutionize the travel planning experience for users exploring the districts of Alappuzha, Wayanad, and Calicut in Kerala.

Diagrams:

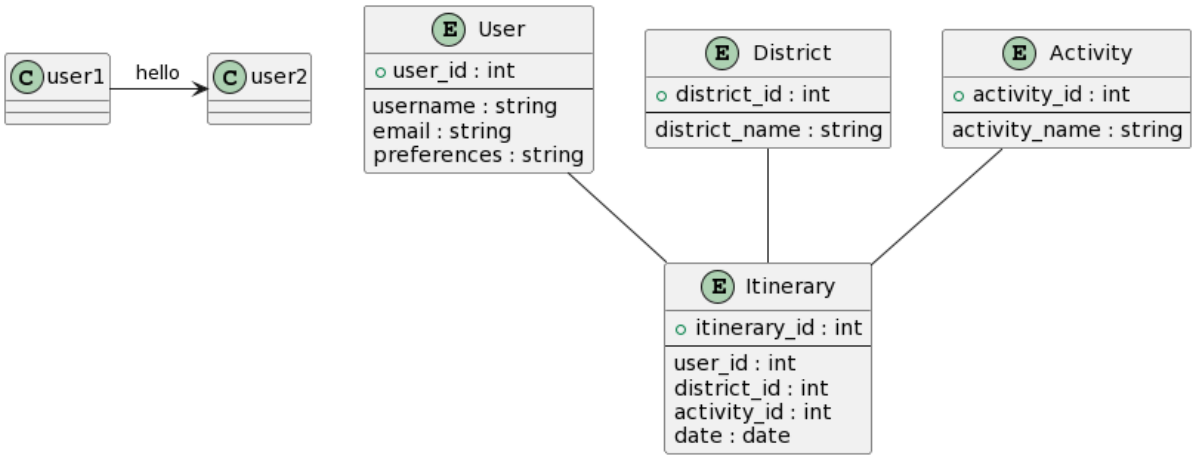


Fig 1:ER diagram of ai chatbot

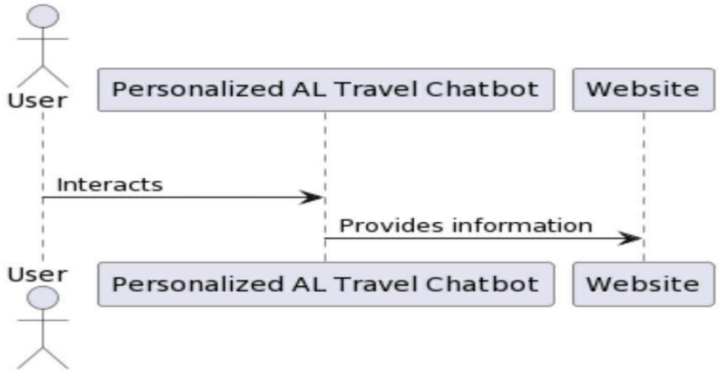


Fig 2 : Level 0 Dfd of Ai chatbot

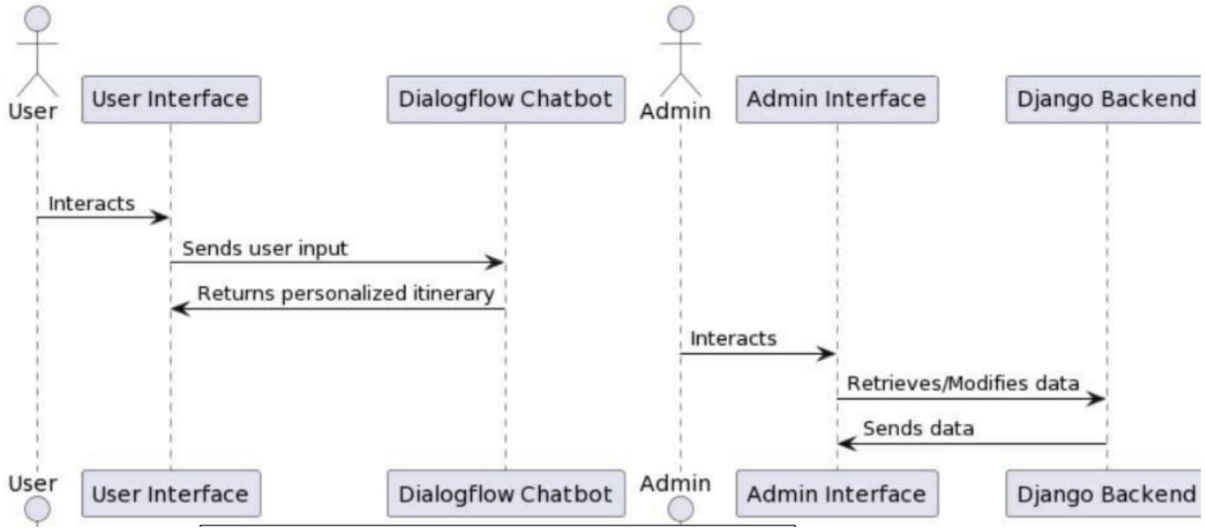


Fig 3:level 1 DFD od AI chatbot

4. RESULTS OF DISCUSSION

Our project's evaluation reveals a commendable accuracy rate of 65% in delivering personalized travel recommendations to users exploring the districts of Alappuzha, Wayanad, and Calicut in Kerala. This achievement underscores the effectiveness of our AI chatbot in understanding user preferences and generating relevant itineraries customized to individual interests. While the accuracy rate falls slightly below our initial target, it nonetheless demonstrates the potential of AI-driven solutions to enhance the travel planning experience significantly.

Several factors contribute to the observed accuracy rate, including the quality and diversity of the dataset used for training the AI chatbot, the sophistication of the natural language understanding algorithms employed, and the durability of the user interaction mechanisms integrated into the website interface. By utilizing a comprehensive dataset sourced from various reliable sources and employing advanced machine learning techniques, we were able to equip the chatbot with a broad understanding of the attractions, activities, and preferences relevant to each district. Additionally, the user-friendly website interface facilitated seamless communication and engagement, enhancing the overall user experience.

Despite the promising results achieved, there remain opportunities for further improvement and refinement. Future iterations of the project could focus on enhancing the accuracy of the chatbot through additional training data, fine-tuning of algorithms, and incorporation of user feedback mechanisms. Moreover, ongoing monitoring and evaluation of the chatbot's performance will be crucial to identifying and addressing any inaccuracies or limitations promptly. By continuously iterating and optimizing the project, we can ensure that it remains at the forefront of AI-driven travel planning solutions, providing users with increasingly accurate and personalized recommendations tailored to their unique preferences and interests.

In conclusion, while our project has achieved a commendable accuracy rate of 65% in delivering personalized travel recommendations, there is room for further refinement and improvement. By leveraging the strengths of AI technology and embracing a user-centric approach, we have laid the groundwork for a transformative travel planning experience that empowers users to explore the rich cultural and natural landscapes of Kerala with confidence and convenience. As we continue to iterate and enhance the project, we are committed to realizing its full potential in revolutionizing the way travellers engage with and experience the world around them.

5. CONCLUSION

In conclusion, our project aimed to revolutionize the travel planning experience by introducing a personalized AI travel chatbot tailored to the districts of Alappuzha, Wayanad, and Calicut in Kerala. Through meticulous data collection, robust training of the chatbot using Dialog flow, and the development of a user-friendly website interface, we have demonstrated the potential of AI technology to enhance user engagement and satisfaction in travel planning. The achieved accuracy rate of 65% signifies a significant milestone in the project's journey, showcasing the effectiveness of our approach in delivering personalized recommendations aligned with individual preferences.

Moving forward, the success of our project underscores the importance of continued innovation and refinement in AI-driven travel solutions. As technology evolves and user expectations evolve, there will be opportunities to further enhance the accuracy, usability, and functionality of the chatbot. Future iterations of the project could explore incorporating additional data sources, refining natural language processing algorithms, and implementing advanced user feedback mechanisms to drive continuous improvement and optimization. By remaining agile and responsive to emerging trends and user needs, we can ensure that our project remains at the forefront of AI-driven travel planning solutions, delivering value and convenience to users worldwide.

Ultimately, our project represents a significant step forward in leveraging AI technology to redefine the travel planning experience. By combining the power of AI chatbots with intuitive web interfaces, we have created a seamless and personalized journey planning experience that empowers users to explore destinations with confidence and ease. As we look to the future, we are excited about the possibilities that AI technology holds for transforming the way we discover, experience, and connect with the world, and we remain committed to driving innovation and excellence in the field of travel technology.

6. REFERENCE

[1] Kumar A. Nimal, Vaishakh V. Nair, R. Jegdeep, J. Nehru., Artificial Intelligence Based Chatbot for Healthcare Applications., Published in IoT, Cloud and Data Science 27 February 2023 Medicine, Computer Science Advances in Science and Technology.

[2] Vincent Velasco, Kevin Dedy Setiawan, Renaldo Robert Sanjaya, Maria Susan Anggreainy, Afdhal Kurniawan., Exploring the Role of AI Chatbots in Education: Insights from the 4th International Conference on Artificial Intelligence and Data Sciences (AiDAS). Published in 4th International Conference… 6 September 2023, Computer Science 2023 4th International Conference on Artificial Intelligence and Data Sciences (AiDAS)

[3] Daniel H. Chang, M. P. Lin, Shiva Hajian, Quincy Q. Wang., Integrating AI Chatbots in Education: Promoting Self-Regulated Learning Through Pedagogical Principles Published in Sustainability 27 August 2023 Education, Computer Science
[4] Sri Varshini K, Shilpa Shri R, Anusha U A., AI Chatbot for College Enquiry, Computer Science, Education International Journal of Advanced Research in 2024
[5] Suresh Kumar, Jaseela Thanikkad , Published in Atna Journal of Tourism… 1 March 2023 , Sociology
[6] S. Chaudhury, Sukanta Sarkar, Chandra Sekhar Pattnaik, Saidur Rahman , Published in Dera Natung Government… 26 December 2023 , Economics, Business
[7] Patrícia Ferreira , Published in Conference of the European… 2023, Computer Science
[8] A. Dehankar, Shyamal Kirpan, Abhayraj Jha, Gitesh Thosare, Adarsh Bhaisare, Sumit Mankar , Published in International journal of… 30 May 2022 , Computer Science
[9] Published in International Journal of… 14 December 2019 , Computer Science, Business , International Journal of Innovative Technology and Exploring Engineering
[10] Haseena V.A , Published 2015, Environmental Science, Economics, Business , Journal of Tourism, Hospitality and Sports

Computer Science Research Reports (2024), 2-1, 34-39
ISBN : 978-93-94056-72-5

SOFTWARE DEVELOPMENT RESEARCH ARTICLE

CRAFT.IO : STUDY ON AN E-COMMERECE WEBAPP EXCLUSIVELY FOR HANDICRAFT PRODUCTS

Jurair C[1],Fathima Sahla TC[2],Muhammed Rashid P[3],Hannath Qairul Abida E[4] ,Insaf P[5]

[1,2,3,4,5]*Department of Computer Science Sullamussalam Science College, Areekode*

rashidpilakkal01@gmail.com

Abstract: *This project report details the conception, creation, and implementation of a marketplace for handcrafted goods using the MERN (MongoDB, Express, React, and Node.js) technological stack. Our platform is a leading force in the handicraft sector, flawlessly fusing state-of-the-art technology with artisanal abilities. Driven by a deep understanding of the economic value of handcrafted goods, our project honors creativity while giving makers a global platform to display their abilities. Beyond simple transactions, our platform hopes to create an atmosphere in which buyers find one-of-a-kind gems, customers engage in genuinely meaningful experiences, and creativity flourishes.*

Keywords: User-Friendly Experience, User Module, Admin Module

1. Introduction

The development and preservation of local handicraft have become growing significance in the ever-changing world of international trade. The sole objective of Craft.io, a passionate work, is to honor and enhance the valuable legacy of handcrafted goods. The purpose of this carefully crafted online marketplace is to act as a lighthouse for local artisans, giving them a unique opportunity to showcase their outstanding works to a worldwide audience. Not only does Craft.io make transactions easier, but it also acts as a link between traditional craftsmanship and the global network of online commerce, creating an environment in which craftsmen may share their unique cultural heritage with people around the world.

With a focus on the user, Craft.io uses the MERN stack to provide an effortless, secure, and customized experience. Craft.io specifically solves the complexities of the handcrafted industry by integrating Node.js for easy interaction, React.js for an attractive and responsive user experience, Express.js for a powerful server-side infrastructure, and MongoDB for effective NoSQL data management. This comprehensive solution makes sure that a variety of dynamic, handcrafted products are handled effectively, laying the groundwork for a better user experience that includes dynamic, interactive features. Even though e-commerce is growing rapidly, it still poses significant obstacles for local artists and their handmade products. Established businesses frequently control the majority of the attention on popular platforms, giving local artisans little visibility. Seeing this disparity, Craft.io takes a calculated approach to redress it by setting aside a particular space for handcrafted products and actively attempting to bring these makers into the public eye on a global scale. Moreover, excessive commissions and listing fees put small businesses on traditional e-commerce platforms at a financial disadvantage. With the help of Craft.io, these barriers should be removed, enabling local artisans to exhibit and sell their work without risking their livelihoods. Craft.io becomes a transformational force

by lending financial accessibility and inclusivity to local skilled artisans, enabling them to flourish in the cutthroat world of e-commerce.

The Customer Module and the Admin Module are the two main modules that make up the project. Users interact with elements in the Customer Module that let them rate specific products, place orders, and effectively handle their dealings with merchants. Customers are redirected to the Customer Home Screen upon logging in, where an overview of all accessible products is displayed for convenient browsing. Customers are given the resources to make well-informed decisions regarding their orders thanks to this simplified design, which promotes a smooth and user-friendly experience.

On the other hand, the Admin Module acts as our MERN stack project's core control center, guaranteeing the platform's seamless functioning and upholding quality control. Administrators in this module have the power to monitor user conduct, carefully maintain customer profiles, and handle user base problems. Establishing a comprehensive approval procedure is a crucial duty, in which administrators confirm and sanction any changes made to user profiles in order to comply with the platform's strict quality standards. In addition to overseeing user accounts, the Admin Module is also responsible for the platform's overall health and efficacy over the long run. It does this by offering dependable tools for system maintenance and by continuously monitoring quality assurance processes.

2. Literature review

This paper analyzes [1] the implementation of a technique for authentication and authorization using JSON web tokens, which will provide a more secure JSON web token generation process and turn the web service into a role-based one. In this paper [2] analyzes inventive and imaginative concept Epicraft is a platform that connects artists and consumers, fosters creativity, and showcases India's rich cultural legacy by fusing modern technologies with traditional Indian art. This paper [3] enabling consumers to buy and sell from home with unmatched ease, e-commerce revolutionizes business through computer networks. It doesn't require physical presence, which saves time and lessens manual labor, unlike traditional trade.

This research paper [4] explores the development of an eCommerce website for artists utilizing the MERN stack. Artpiece aims to provide a user-friendly platform for artists to sell their artwork globally, integrated with secure payment gateways. The paper [5] introduces a MERN stack-based smart parking registration system, inspired by BookMyShow's seat registration. With a scalable design, it simplifies parking space booking, aiming to enhance user experience and reduce urban traffic congestion. In this paper [6] dynamic tech industry, full-stack developers are in high demand, particularly in the United States, where skilled professionals can earn up to $110,770. They excel in both front-end and back-end, utilizing frameworks like LAMP, Java, and MERN for comprehensive web development.

In this paper [7] Leveraging the MERN stack, the Evernote application implements a robust user authentication system, utilizing MongoDB, Express.js, React, and Node.js. This integration ensures secure, customizable, and efficient user verification. This paper [8] introduces a novel hospital service optimization approach, integrating healthcare supply chain concepts. Leveraging MERN stack technology, our dynamic web service facilitates real-time hospital connectivity, streamlining patient transfers, and enhancing care with virtual consultations and organ transplant search capabilities. In this paper [9], 'Print – OnTheGo' is an efficient remote printing system, leveraging MongoDB, Express, React.JS, Node.JS stack, connecting users to print shops seamlessly. Enhancing convenience, it prioritizes efficiency and security in document printing. In this paper [10] A Smart Event Management System addresses the limitations of traditional methods, employing web development to efficiently manage and disseminate event information in college, prioritizing data management and report generation.

3. Methodologies

3.1 MERN Stack

Craft.io leverages the MERN (MongoDB, Express.js, React.js, Node.js) stack, a powerful combination of technologies that enables the development of dependable and expandable online apps. That represents the four main components of this stack: MongoDB, Express.js, React.js andNode.js

3.1.1 MongoDB

The foundation of the Craft.io project is a NoSQL storage system that is dynamic and scalable, based on MongoDB in its designated version. MongoDB, which seamlessly integrates with the application's backend, is essential for handling the complex and constantly changing nature of e-commerce data. Because of its document-oriented structure, collections may store data in a flexible way, allowing for the representation of a wide range of product qualities and user data. Making use of MongoDB's indexing powers guarantees optimal query performance a vital component for providing quick and precise search results in the competitive e-commerce world. Additionally, sharding-enabled MongoDB's horizontal scalability foresees and controls future spikes in data volume, guaranteeing uninterrupted user experiences even during high traffic periods.

3.1.2 Express.js

Express.js, a lightweight and adaptable Node.js web application framework, forms the backend of the MERN stack. It makes it easier to build a reliable server that can handle HTTP requests and define endpoints for different functionalities. Express.js's clean design makes it easier to develop middleware, handle routing, and integrate Mongoose with MongoDB. It contributes to the overall effectiveness of the Craft.io platform by guaranteeing a seamless data transfer between the frontend and the MongoDB database.

3.1.3 React.js

Craft.io's dynamic and responsive user experience is created using React, the frontend library included in the MERN stack. React's component-based architecture facilitates the development of reusable user interface components, improving scalability and maintainability. It ensures a smooth and engaging shopping experience by effectively managing the state. The virtual DOM provided by React allows for optimum rendering, reducing the need for page reloads and improving platform performance.

3.1.4 Node.js

Completing the MERN stack, Node.js is used as the server-side application's runtime environment. Its non-blocking I/O operations and event-driven architecture guarantee effective management of several requests at once. With the help of Node.js, Express.js and MongoDB can be integrated more easily, giving the Craft.io platform a more unified and efficient environment. It is a good fit for the changing needs of an e-commerce platform because of its flexibility in managing real-time applications.

3.2 HTML

HTML is essential to React in order to define component structures using JSX. JavaScript and HTML-like syntax are combined in JSX to allow for the smooth integration of event processing and dynamic content. Reusable UI elements are made possible by components that render HTML elements. Curly braces can contain JavaScript expressions, and JSX allows for conditional rendering. React.js uses HTML to help create interactive and modular user interfaces, which are the foundation of component-based development. This connection guarantees rapid rendering of dynamic web pages and a clear division of responsibilities.

3.3 CSS

For React.js to style components and provide an aesthetically pleasing user experience, CSS is essential. Styles are frequently specified using independent CSS files or JavaScript objects found in components. Styled Components is one of the CSS-in-JS frameworks that simplifies styling by integrating with

React.js components directly. Both local and global stylistic techniques provide a wide range of design options. React.js's virtual DOM optimizes efficiency by applying and updating styles in an effective manner. React.js's usage of CSS encourages modularity, maintainability, and reusability when creating aesthetically cohesive online applications.

3.4 DFD

DFD is a well-known and widely used notation for specifying the functions of an information system. A DFD represents the flow of data. They describe systems as collections of data that are manipulated by functions. Data can be organized in several ways. They can be stored in data repositories; they can flow and can be transferred to or from the external environment.

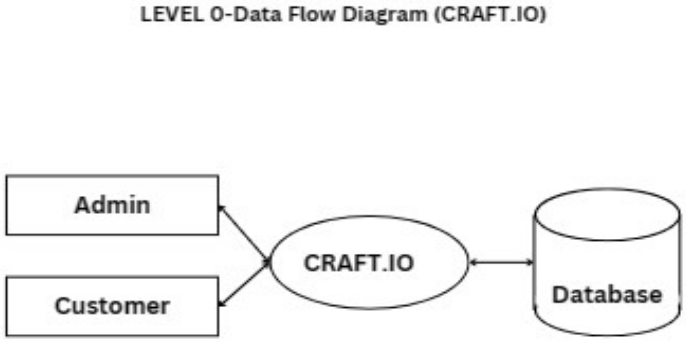
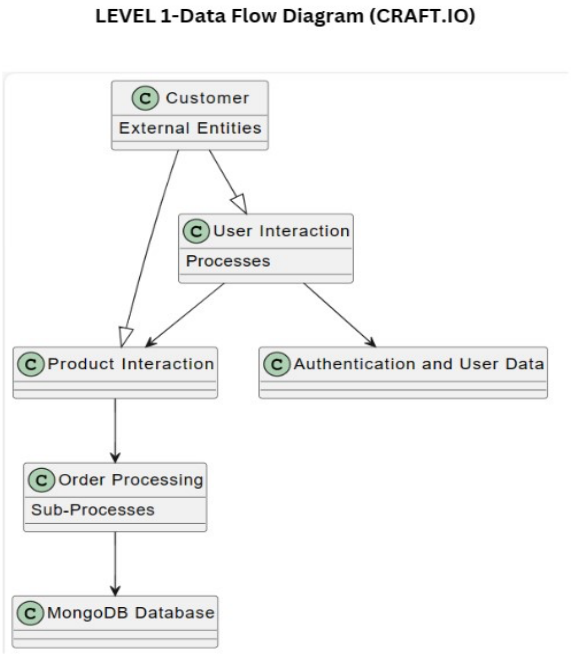
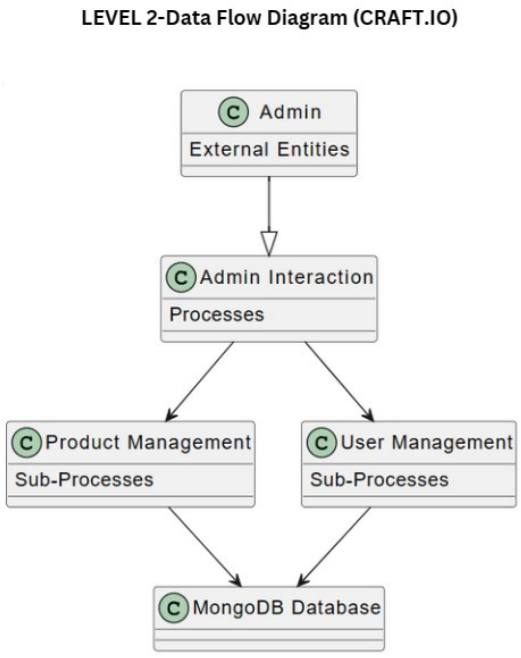


Fig. 1 Data flow diagram Level 0

3.5 DFD of Craft.io





4.Result and Discussion

The handicraft buy and selling e-commerce website in the MERN stack has been developed and implemented, and the results are impressive. In particular, the user experience has been improved, and a competitive marketplace that benefits both customers and artists has been created. Firstly, the MERN stack's user-centric methodology has been essential in providing a smooth and simple interface for the

Admin and Customer Modules. This methodology guarantees that customers can effectively traverse the platform, browse products, and execute transactions with efficiency, cultivating a gratifying and engaging encounter.

Second, the Customer Module's installation of a simplified order management system has worked well. Customers are more satisfied and engaged since they can easily place orders, rate products, and communicate with sellers.

The effective order management system plays a major role in the platform's overall performance. The third pillar of success is the Administration Module, which acts as a central control center. Admins have powerful tools at their service to manage client profiles, analyze user behavior, and effectively handle complaints. The dedication to a secure and well maintained platform raises users' confidence and trust even further.

In conclusion, the MERN-developed e-commerce website for buying and selling handicrafts is indicative of successful outcomes and highlights the importance of technological innovation, user-centric design, and calculated business growth. The platform continues to expand and be relevant in the ever-changing world of e-commerce because to its dedication to security and constant innovation.

5. Conclusion

Craft.io stands as a ground breaking and transformative concept, dedicated to fortifying local crafts markets and fostering economic vitality within communities. By seamlessly connecting consumers with a diverse range of locally crafted goods, Craft.io acts as a powerful catalyst for businesses, enhancing brand visibility and attracting a broader audience. Its commitment to a user-friendly interface and simplicity not only facilitates effortless exploration of local crafts but also serves as a dynamic intermediary, widening exposure and customer bases for artisans. Craft.io uniquely stands out through its strategic emphasis on a clear and simple user experience, making it accessible to a diverse range of users, even those less familiar with online transactions. This platform, by streamlining the browsing and selection process, remains true to its core mission of supporting small businesses and championing locally crafted goods. In essence, Craft.io encapsulates the beauty and artistry intrinsic to regional goods, envisioning a vibrant marketplace intricately connected to the local crafts ecosystem. With an unwavering dedication to accessibility and simplicity, Craft.io emerges as a vital tool, actively supporting small enterprises and ensuring the sustained, long-term growth of the local crafts industry.

References

[1] Pooja Mahindrakar, U.Pujeri : Security Implications for Json web Token Used in MERN Stack for Developing E-Commerce Web Application : Published in Regular 30 October 2020.

[2] Ritu Shailendra Jha, Ritu Raju Jha, Nandini Vijay Gurav, Aabha Patil: EPICRAFT- Ecommerce for Artisans : International Journal of Scientific Research in Computer Science Engineering and Information Technology 2023.

[3] Arfath Abbas, Abdul Wahid Syed, Mohammad Saqib Ahmed, Abdul Sagheer : E-Shop Website Using MERN Stack : V. H. International Journal for Research in Applied Science and Engineering Technology 2022.

[4] Prof. S.C. Shirbhate, Ashutosh Adhao, Prajwal Khatile, Samiksha Sawalakhe, Sakshi Pedhekar, Sakshi Malviya : E-Commerce Website for Artisans : International Journal for Research in Applied Science and Engineering Technology 2023.

[5] Ms.Nupur Banode : Smart Parking System using MERN Stack : International Journal for Research in Applied Science and Engineering Technology 2023.

[6] A.Malik : Social Sphere A MERN Stack Social Media App : International Journal for Research in Applied Science and Engineering Technology 2023.

[7] PROF.B.V Karlatthe : MERN Stack Based User Authentication Technique for Evernote Application : International Journal of Scientific Research In Engineering And Management 2023. [8] Dr. S. P Khandait : Optimizing Hospital Services: Implementing Healthcare Supply Chain Management through MERN Stack Web Application : International Journal for Research in Applied Science and Engineering Technology,2023

[9] Neenad Sahasrabuddhe, Rishika Jain, Sanskar Khandelwal, Sonali Kalekar, Neha Purohit, Ashish K Sharma : Print – OnTheGo Remote Printing Application using MERN Stack : 4th International Conference for Emerging Technology (INCET),2023

[10] Akansha Pansare, Athang Patil, Nikita Patil, Yatin P. Patil, Mrs. Aparna Bhonde : Smart College Event Management System Using MERN Stack : International Journal for Research in Applied Science and Engineering Technology 2023

Computer Science Research Reports (2024), 2-1, 40-47
ISBN : 978-93-94056-72-5

SOFTWARE DEVELOPMENT RESEARCH ARTICLE

REVAMPSS : ENHANCING EDUCATIONAL ACCESS AND ENGAGEMENT URL USING STATE OF ART SOFTWARE TOOLS

Muhammed Saheen[1*], Asnif P[2], Ahammed Sameeh[3], Abdulla Sabith[4], Mohamed Basheer K.P[5]

[1,2,3,4] *Research Department of Computer science, Sullamussalam Science College, Areekode, Affiliated to University of Calicut.*

**abdullasabithpulli@gmail.com*

Abstract : *This research paper presents a comprehensive study focused on the renovation of a college website to enhance user experience and functionality. By analyzing current trends in web design and user interface optimization, the project identifies key areas for improvement and proposes innovative solutions to the specific needs of the college community. Through the integration of modern technologies and user-centric design principles, including Next.js, Django, and PostgreSQL, the aim is to transform the college website into a dynamic and user-friendly platform. This renovation encompasses a redesigned user interface, user friendly navigation, and the implementation of responsive design, ensuring accessibility across various devices. The research contributes valuable insights and practical strategies for elevating the educational experience through an optimized digital interface. Furthermore, the methodology employed involves a thorough assessment of the existing website's strengths and weaknesses. By prioritizing intuitive design and efficient functionality, the project aims to create a digital ecosystem that ensure a better communication, engagement, and resource accessibility for students, faculty, staff, and stakeholders alike.*

Keywords: College website, renovation, Sullamussalam, DJANGO, responsive design.

1. Introduction

In today's digital world, the online presence of educational institutions plays a crucial role in defining their identity and engaging with stakeholders effectively. College websites serve as the main platform for students, faculty, staff, and the wider community to access information, resources, and services provided by the institution. However, due to rapid technological advancements and evolving user expectations, it has become necessary to renovate these websites. The renovation of college websites is now seen as a vital initiative to align with changing user preferences, enhance user experience, and take advantage of the latest technological innovations. Responsive design ensures that websites adapt seamlessly to different screen sizes and devices, providing an optimal viewing experience for users on various platforms.

To revitalize the digital interface of college websites, this study aims to incorporate technologies such as Next.js for front-end development, Django for back-end functionality, and PostgreSQL for database management. These technologies offer robust solutions for creating dynamic, interactive, and user-friendly websites that cater to the diverse needs of stakeholders. Addressing the challenges posed by

outdated design, complicated navigation, and limited interactivity is crucial to achieving the goals of website renovation. this research aims to showcase the transformative potential of improving accessibility, functionality, and overall effectiveness of college websites. By combining theoretical insights with practical implementation strategies, this paper provides valuable guidance for educational institutions embarking on website renovation projects to better serve their stakeholders in the digital age.

2. Literature Review

In this paper[1] the rapid evolution of web development frameworks has significantly altered the landscape of online experiences, with an increasing emphasis on achieving optimal website speed and search engine optimization (SEO). In this context, Next.JS, a popular React framework, has emerged as a powerful tool that promises enhanced performance and improved SEO capabilities. This study aims to delve into the intricate relationship between Next.JS and website speed, as well as its influence on SEO practices.

In this paper[2] an open source framework which follows the fundamental Model View Template shape with a few changes which are explained and it additionally help us to recognize why we use Django over other web frameworks which might be to be had in the industry and how we are able to install it on our system and create a fundamental task using this framework following .In this paper[3] Bootstrap is the popular HTML, CSS and JavaScript framework for developing a responsive and mobile friendly website.In this paper[4] presents the preliminary design of a new database management system, called POSTGRES, and describes the query language, programming langauge interface, system architecture, query processing strategy, and storage system for the new system.

In this paper[5] REST APIs expose web services to clients. Although experts have recommended guidelines for REST API design, there is little empirical evidence regarding the relationship between adherence to guidelines and benefits to API consumers. In this paper[6] website that is difficult to browse, confusing or slow to load the content can result in a poor user experience and, as a result there is a loss in leads or consumers. A website that is easy to use, well-organized and optimised for mobile devices, on the other hand can help improve user engagement, raise conversion rates and ultimately boost the success of digital marketing initiatives. In this paper[7] authors present the most search engine optimization like (Google, Bing, MSN, Yahoo, etc.), and compare by the performance of the search engine optimization. The authors also present the benefits, limitation, challenges, and the search engine optimization application in business. In this paper[8] it delves into the intricacies of the file-based routing system in Next.js, discussing its principles, benefits, potential issues, and use cases bolstered by tangible coding examples. In this paper[9] overview of why to choose Django over any other framework. Django is a high level Python framework by using it the development speed will be faster and cleaner. Django is built by the much more experienced people, so it takes care of the web development process in easier way. In this paper[10] Responsive design allows software developers to build a Web page that can dynamically adapt to the size of the devices. This development philosophy enables the rendering of Web pages in a fast and optimized way, ensuring a good user experience on mobile devices, tablet and desktop.

3. Methodology

For the redevelopment of the Sullamussalam Science College website, specific software tools and technologies have been selected to ensure efficient development, robust functionality, and optimal user

experience. The following software specifications outline the technologies used in the project: Next.js Django PostgreSQL

3.1 Next.js:

Next.js is a powerful React framework that facilitates the development of server-side rendered (SSR) and statically generated web applications. It provides a environment for building modern, efficient, and scalable web applications with React. Next.js offers a variety of features such as automatic code splitting, route pre-fetching, and built-in CSS and Sass support, making it a popular choice among developers. Its flexibility allows for easy integration with APIs and databases, while its strong performance optimization capabilities ensure fast load times and a smooth user experience. Additionally, Next.js has a vibrant community and extensive documentation, making it easy for developers to get started and use its full potential for building dynamic web applications.

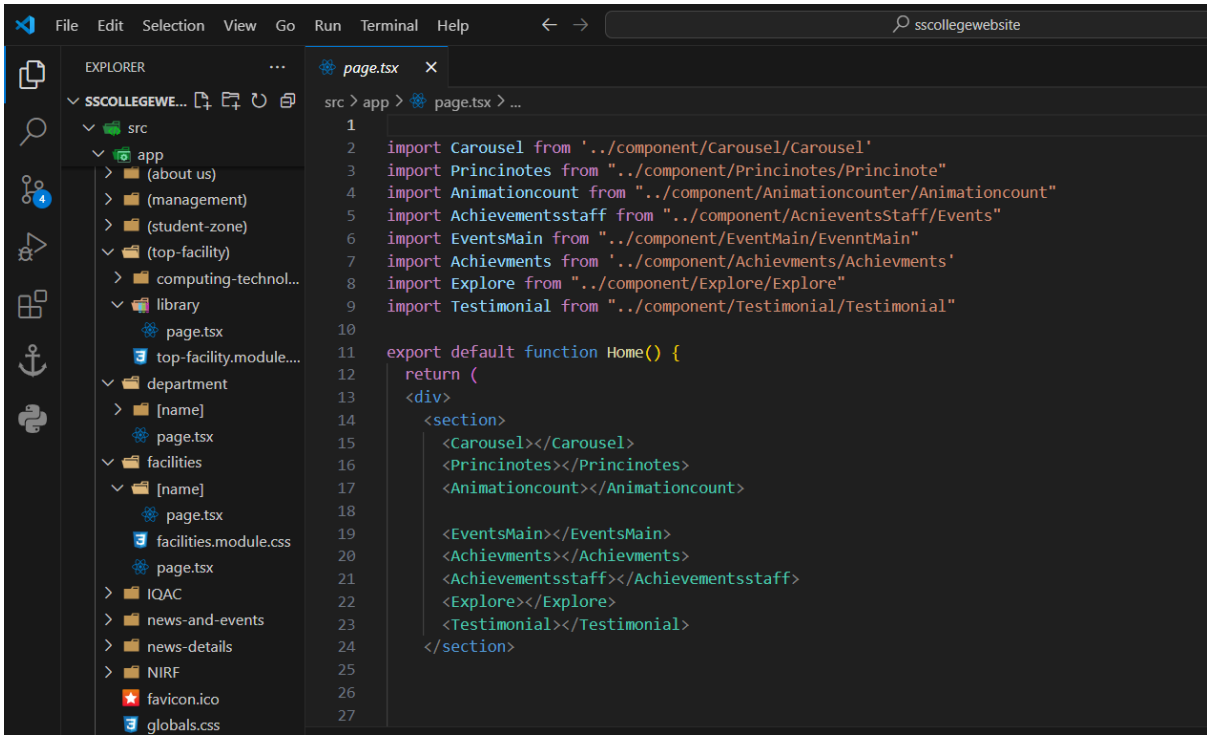


Fig 1. Nextjs Homepage

3.2 Django:

Django, a popular backend framework for web development, offers essential features for building powerful web applications. Its built-in authentication and authorization system provide secure user management, ensuring that only authorized users can access sensitive data and features. Django's ORM (Object-Relational Mapping) simplifies database interactions, allowing developers to work with databases using Python objects instead of complex SQL queries. Additionally, Django's admin panel automates the creation of an admin interface for managing site content, saving developers time and effort. Its URL routing system enables clean and organized URL structures, enhancing the readability and maintainability of web applications. Django's perfect documentation and active community support make it easy for developers to learn and troubleshoot issues, ensuring smooth development processes. Overall, Django's combination of security features, database abstraction, administrative tools, clean

URL routing, templating engine, and community support makes it an excellent choice for building scalable and maintainable web applications.

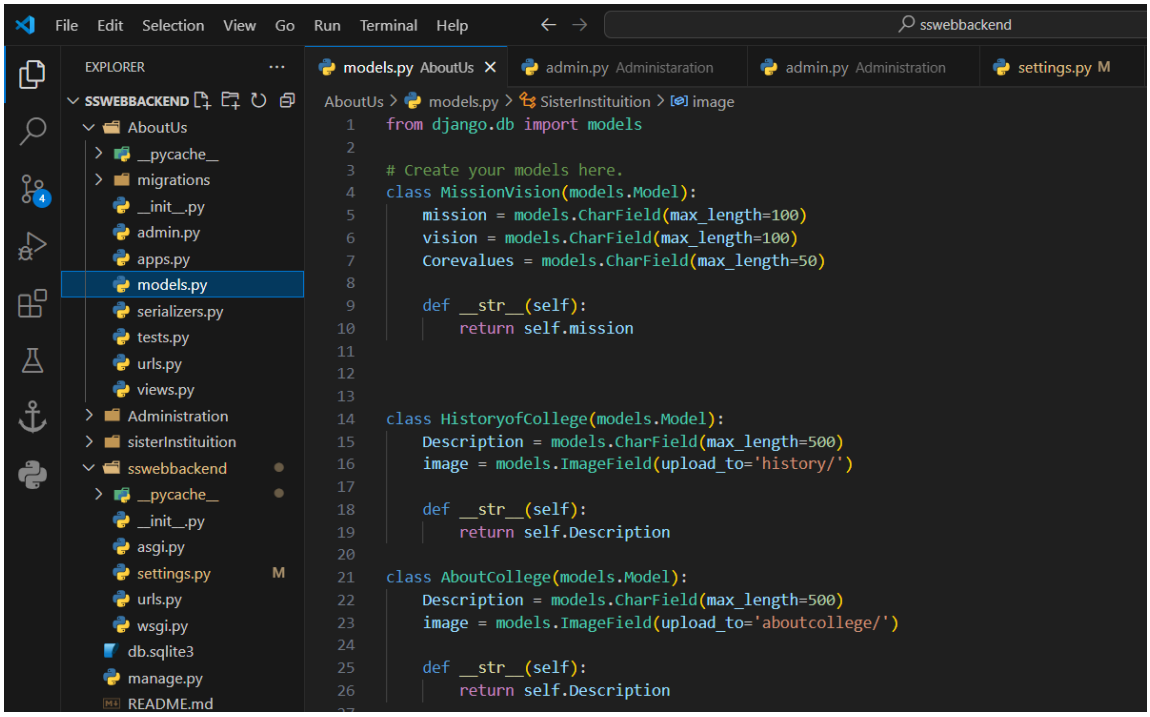


Fig 2. Django models

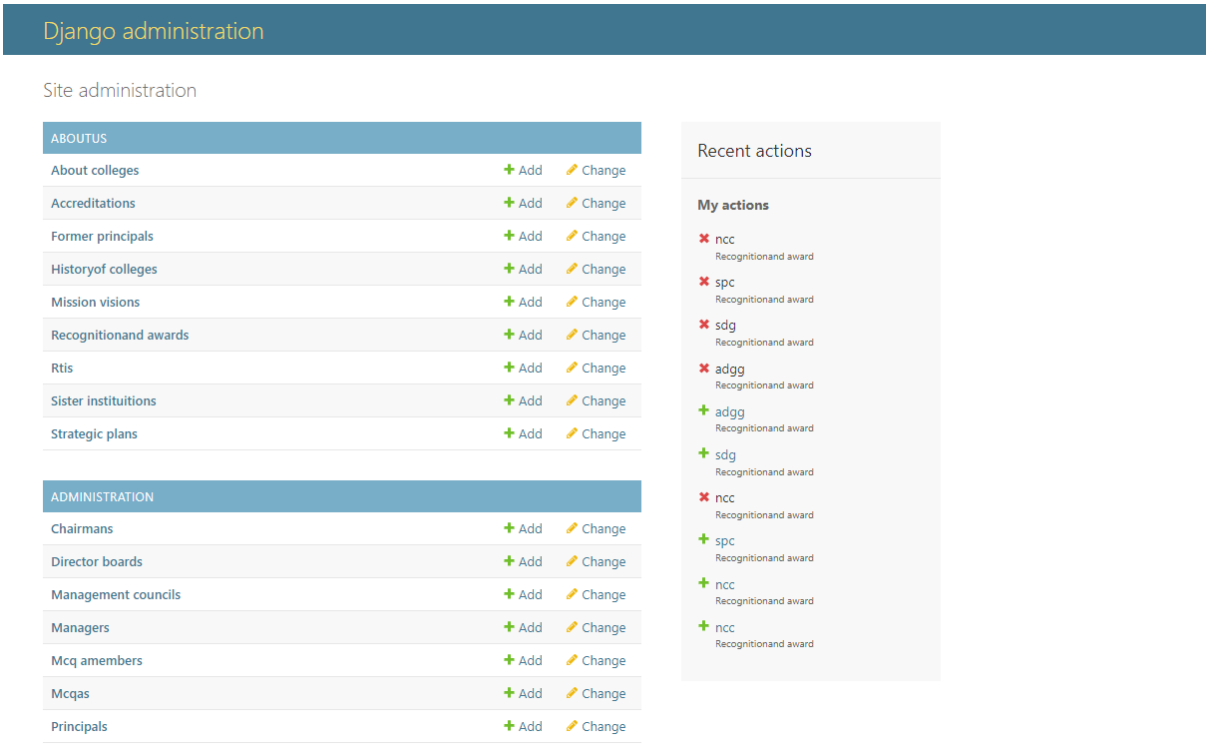


Fig 3. Admin Dashboard

3.3 PostgreSQL

PostgreSQL serves as a robust and reliable backend database solution for web development projects. Its adherence to SQL standards, advanced features, and support for diverse data types like JSON and XML make it ideal for efficiently managing web application data. PostgreSQL's ability to handle concurrent connections and large datasets ensures fast and efficient data retrieval. Moreover, its extensibility allows developers to integrate various plugins and extensions to customize the database to meet specific project requirements. With features like replication and clustering for high availability and scalability, PostgreSQL provides a solid foundation for building web applications that demand reliability, performance, and flexibility. Additionally, its active community and continuous development ensure ongoing support and innovation, making PostgreSQL a preferred choice for web developers globally.

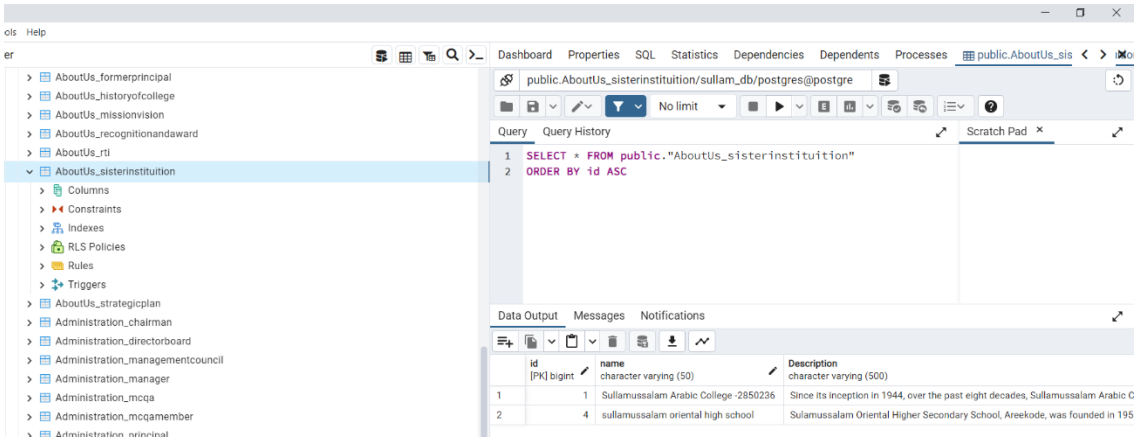


Fig 4. Database

3.4 REST Framework

Django REST Framework (DRF) is a powerful toolkit for building Web APIs in Python, based on the Django web framework. It simplifies the creation of APIs by providing a set of reusable components and patterns, allowing developers to quickly build scalable APIs. DRF follows the principles of Representational State Transfer (REST), providing support for common HTTP methods such as GET, POST, PUT, PATCH, and DELETE, along with serialization, authentication, and authorization mechanisms. Its serialization capabilities allow for easy conversion between complex Python data types and JSON or XML formats, facilitating data exchange between the client and server. DRF also offers built-in support for pagination, filtering, and validation, making it ideal for handling complex API requirements.

3.4.1 API Documentation

This API provides access to various endpoints containing information about the college, including its mission, history, accreditation, management council, and more.

3.5 Axios

In Next.js, a popular framework for building React applications, integrating external libraries such as Axios is straightforward and commonly done for making HTTP requests. Axios is often used for its simplicity and flexibility in handling asynchronous operations like fetching data from APIs. Incorporating Axios into a Next.js project involves a few steps. First, you typically install Axios using npm or yarn, which adds it to your project's dependencies. Then, you can import Axios into your Next.js

components or pages where you need to make HTTP requests. By doing so, you gain access to Axios's methods for sending GET, POST, PUT, DELETE, and other types of requests to server endpoints. This seamless integration enables you to efficiently manage data fetching and handling within your Next.js applications, facilitating smoother communication between your frontend and backend systems. Overall, utilizing Axios in Next.js empowers developers to create dynamic web experiences with ease while maintaining data retrieval capabilities.

4 DFD

DFD is a well-known and widely used notation for specifying the functions of an information system. A DFD represents the flow of data. They describe systems as collections of data that are manipulated by functions. Data can be organized in several ways. The can be stored in data repositories, the can flow and can be transferred to or from the external environment.

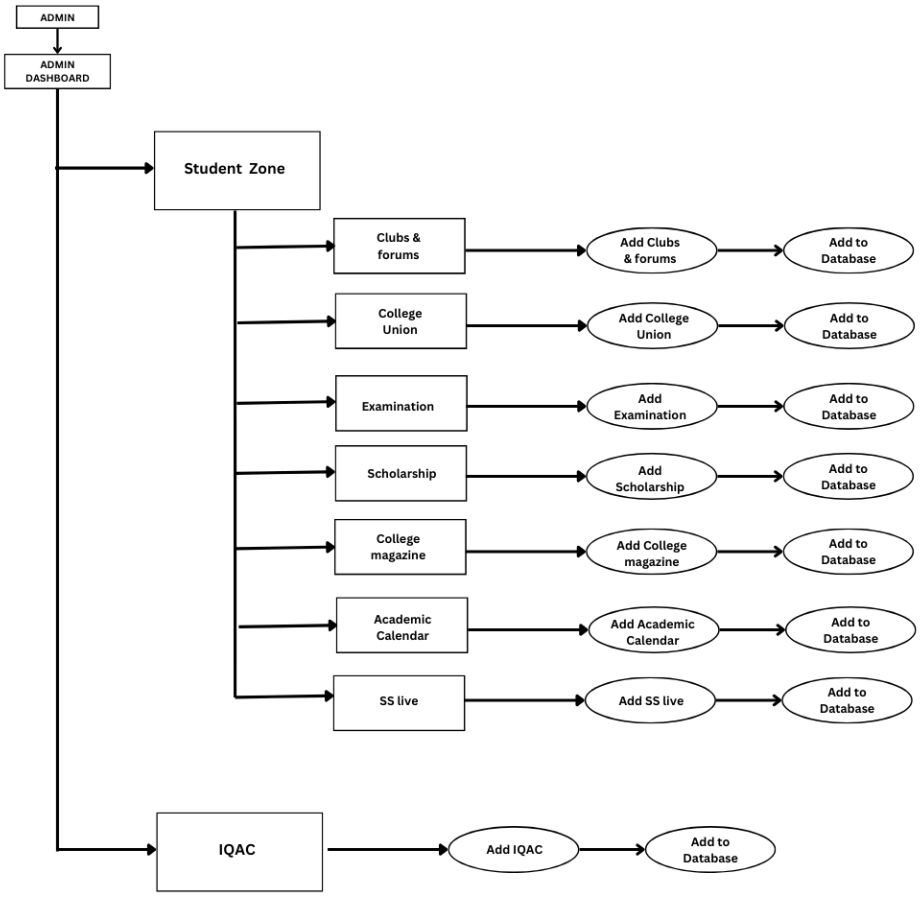


Fig. 5 Module wise diagram

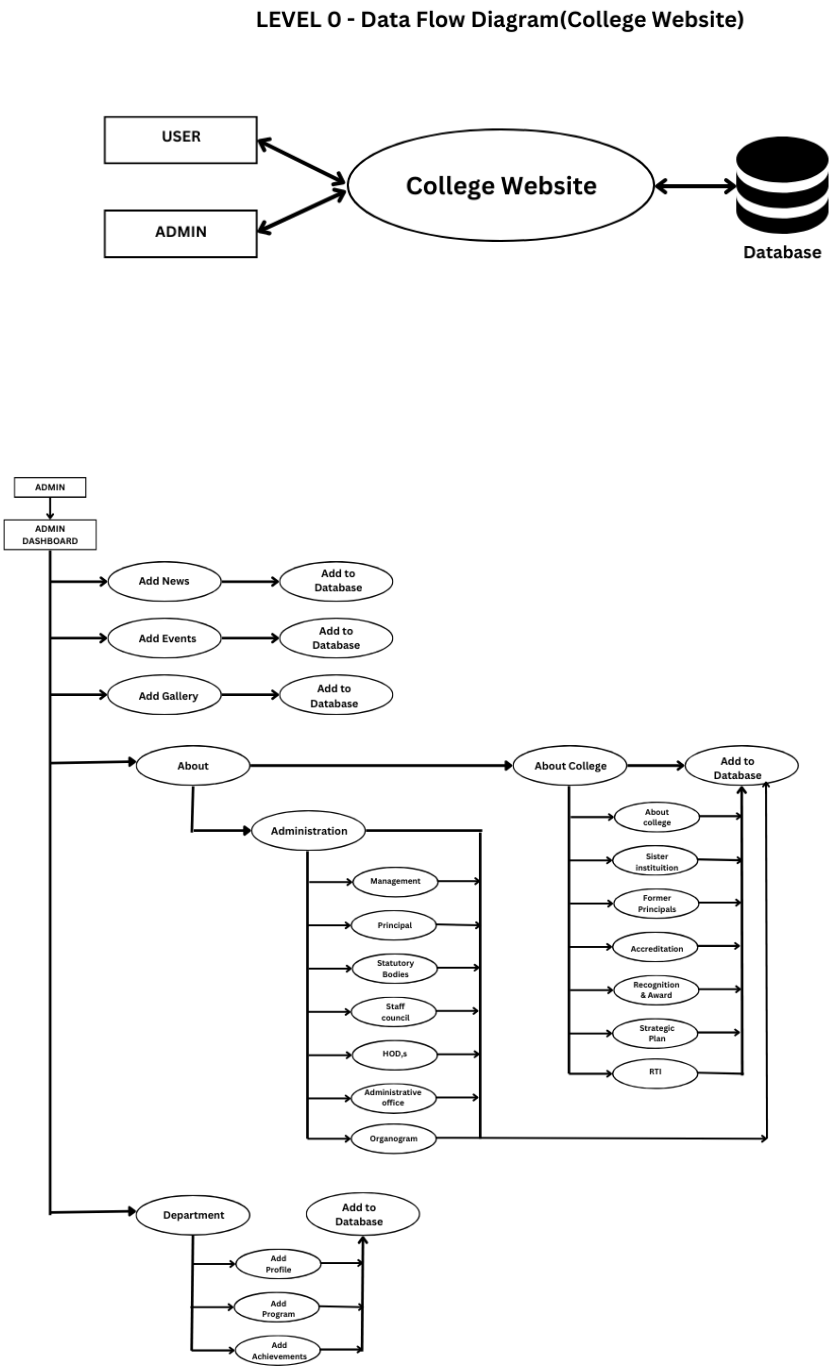


Fig. 6 Module wise diagram

5 Conclusion

The renovation project for our college website is a big step forward in improving the digital presence and user experience for our students, faculty, staff, and visitors. We have successfully addressed issues like outdated design, navigation difficulties, and lack of mobile responsiveness through careful planning and collaboration. Our aim is to provide a easy browsing experience on all devices by implementing modern design principles and responsive web development techniques. Moreover, the revamped

website now includes better accessibility features to ensure inclusivity for all users, including those with disabilities. By following web accessibility standards and guidelines, we want to create an environment where everyone can easily access the information and resources they need. Additionally, the integration of multimedia elements and interactive features enhances the browsing experience, creating a sense of community and connection within our college. Looking ahead, it is crucial to continuously maintain and improve the website to keep up with evolving technology and user expectations. Ultimately, our goal is to ensure that the college website remains a valuable resource and hub for information, communication, and collaboration for years to come.

Reference

1. Samon Daniel 2023 , Ladoke Akintola University of Technology
2. Dr M Rajeswari, KR Kamalam 2022, Journal of Emerging Technologies and Innovative Research
3. Suraj Shahu Gaikwad , Pratibha Adkar 2019, PES Modern College of Engineering Pune, India
4. Michael Stonebraker and Lawrence A. Rowe , University of California
5. Michael Coblenz , Wentao Guo , Kamatchi Voozhian , Jeffrey S. Foste, University of California
6. Heemakshi Sharma, and Khushboo Tripathi 2023, Amity University, Gurugram, Haryana, India
7. Firas almukhtar, Nawzad Mahmoodd, Shahab Kareem 2021, Erbil polytechnic university
8. Krutika Patil 2023, University of Texas at Dallas, Tracy, CA, USA
9. Prof. B Nithya Ramesh, Aashay R Amballi, Vivekananda Mahanta 2018, NHCE, Bangalore, India
10. Fernando Almeida 2017, Centre for Innovation, Technology and Entrepreneurship, INESC TEC, Porto, Portugal.

Computer Science Research Reports (2024), 2-1, 48-58
ISBN : 978-93-94056-72-5

RESEARCH ARTICLE

AI-POWERED INTELLIGENT VEHICLE SURVEILLANCE SYSTEM USING EASYOCR, OPENCV, AND HAARCASCADE ALGORITHMS

Azil A[1*], Ameen Sahil M[2], Amal Krishna P K[3], Farsinsha[4] , Muneer V.K[5]

[1,2,3,4,5] *PG & Research Department of Computer Science, Sullamussalam Science College, Areekode.*
Affiliated with the University of Calicut, India
[*] *azilrahman@gmail.com*

Abstract: *This paper presents an AI-powered vehicle Monitoring System using EasyOCR, OpenCV, and HaarCascade algorithms. The project is the detection of the license plate of a vehicle entering a gate and saving the date, time, and license plate number in a database to enhance the security of the institution. We tried to make it good at detecting vehicles' license plates in real-time. EasyOCR is used for reading license plates and we used it to help the system track vehicles better. OpenCV provides computer vision capabilities to process images and helps in real-time detection. We used the HaarCascade machine learning object detection method to identify objects in images to detect license plates of vehicles. we provide a simple and attractive graphical user interface for users to access vehicle information. Our system can be used in institutions, hospitals, and factories. We hope our work helps others who are into making vehicle detection systems.*

Keywords: ***EasyOCR, HaarCascade, Vehicle Monitoring, Computer Vision.***

1. Introduction

The research paper explores the development and implementation of an AI-Powered Vehicle Monitoring System, designed for real-time license plate detection. Uses technologies such as EasyOCR, OpenCV, HaarCascade algorithms, and PHP. our objective is to create a robust system capable of accurately identifying and tracking vehicles through efficient license plate detection and store the license details date, and time in a MYSQL(database). The project will involve the development of a software system that can be installed at the institutions. The system will use cameras to capture the number plates of vehicles entering the gate and extract the alphanumeric characters using systems. The extracted characters' entry date and time will be stored in a database, which can be easily retrieved from a simple user interface. We did a lot of tests to see if our system works well because we wanted to make sure it can handle different situations and keep working. An institution without an automated tracking system, the institution faces several challenges related to the accuracy and efficiency of vehicle entry management. Manual methods result in time and cost-consuming processes. The verification of vehicle entry becomes a hard task for administrative staff, it leads to delays. An institution with an automated tracking system makes work faster to take details of vehicle entries. Management can get a ready-to-print database at any interval of time.

2. Literature review

In the paper[1] the author builds a Self-Driving Autonomous car model but on a minimalistic basis which basically will be focused on three main features which are to operate by the surroundings depending on the direction of the road, to detect stop signs and halt for 5-10 seconds and detect traffic signs and make decisions accordingly. The miniature self-driving car will detect the two-lane path and perform the above functions. In the paper[2] main aim is to create a model for handling challenges like traffic jams in parking areas due to incorrect parking, automobile insecurity, etc. In the study, an approach of Automatic Vehicle Plate Recognition using Livestream is discussed by considering image size, success rate and processing time as parameters. In this paper[3] the author builds a traffic surveillance prevention system with vehicle number plate identification and speed

detection using machine learning to provide a complete traffic solution. By using only camera captures, a traffic observation system can measure major traffic boundaries from video layouts. This paper proposes to identify over-speeding vehicles' number plates and initiate an emergency message on detecting accidents. Utilizing the CCTV footage, vehicle detection takes place and the speed of the vehicle is derived using Open CV.

The accident will be detected from the video input using the Convolutional Neural Network Algorithm and Computer Vision. On detecting an accident, an emergency message is initiated to the nearest control room with the corresponding location. The paper[4] aims to identify the number plate in vehicles during difficult situations like distorted, high/low light and dusty situations. The paper proposes the use of the Faster R-CNN to detect the number plate in the vehicle from the surveillance camera which is placed on the traffic areas etc. The created system is used to capture the video of the vehicle and then detect the number plate from the video using frame segmentation and image interpolation for better results. the proposed system can achieve a 99.1% accuracy in detecting the number plate of the vehicle and showing the vehicle's owner information. In this paper[5] author presents a programmed traffic observation framework to gauge significant traffic boundaries from video arrangements utilizing just captures from cameras.

A traffic control kit is developed to detect over speeding cars on highways, number plates in Bengali, and initiate emergency calls to 999 on detecting accidents. OCR Tesseract is used to detect number plate which has very high performance in detecting noisy texts. To identify a case of an accident, a simple Python code with Dens-net Architecture is used. A GSM module of the experimental kit initiates the call and message after analysing the data through a code of C language. Machine Learning (ML) is used to train the program in identifying number plates. In the paper[6] author proposed a solution for this problem by considering college buses o BVRIT as an organization. College bus number plate Registration Detection is a crucial part of smart BVRIT planning and BVRIT transport management. This paper presents the YOLO V8 algorithm for Detecting college buses using number plate registration detection. YOLO - You Only Look Once. This YOLO-V8 algorithm is a state-of-the-art (SOA) model or algorithm. It's a real-time object detection algorithm that can detect objects in any image, that has many classes in it with high accuracy and speed. In this paper[7] the author builds The odd-even rule on vehicle number plates in DKI Jakarta aims to reduce congestion that occurs in DKI Jakarta.

The application of these regulations is constrained by the limitations of the manual supervision function by officers. This problem can be overcome by implementing intelligence in the form of detecting number plate objects with the YOLO v5 algorithm and the character extraction process with Optical Character Recognition technology using Tesseract OCR. Based on this research, the average percentage of objects detected in each video is 92.38%, and the average confidence value obtained in object detection is between 75.55%. The success rate of the character extraction process on number plates is 95.45%, and the average proportion according to the detected number plate category is 97.2%. The implementation of the YOLO Algorithm has succeeded in detecting license plates with odd and even categories on videos that can provide signs and save violations of vehicles that violate the odd and even rules. In this paper[8] the author builds a Vehicle Number Plate (VNP) technology that is used in applications like parking management, traffic control and management. It involves tasks such as detecting and recognising vehicle licence plates. Most of the VNP systems in use today don't perform well enough in real-time image/video scenarios.

To trim and save the localised plate, NVIDIA Compute Unified Device Architecture (CUDA) APIs are employed, which resulted in speeding up the processing. The performance of three different optical character recognition (OCR) techniques OCR, Tesseract OCR, and ONNX OCR mode is compared to choose the best one. In terms of accuracy and time complexity, the Easy OCR model is proven to be the best. The author suggests using Single Shot Detector (SSD) Mobilenet V3-based architecture to locate licence plates and Easy OCR due to its superiority for character recognition. In terms of accuracy and time complexity, the Easy OCR model is proven to be the best. In this paper[9] the author builds an Acquiring image and detecting number plates of moving vehicles using motion platforms presents a challenge. To address this challenge, mobile cameras are utilized to capture license plate images of moving vehicles, which are then labelled and stored in a database. The test results obtained from our work, show a relatively high accuracy in terms of number plate recognition. The primary goal of our ANPR system is to achieve high accuracy in recognizing Odisha license plates. Other ANPR systems available use different types of pre-trained model methods, such as Alex Net and GoogleNet.

The results show that using our scheme a good accuracy was achieved using a smartphone camera. The paper [10] named AVNPR is to recognize the license plate using image processing techniques or optical character recognition by applying the pytesseract OpenCV Python package and has a main focus on the detection of vehicle license plate, character segmentation and character recognition. In this paper[11] the author seeks to prevent the increasing traffic accidents and crime. The younger generations are still engaging in reckless driving, even though the government has implemented legislation and traffic regulations to lower the number of accidents and fatalities on the roads. This model first detects the presence of cars, then calculates their speeds, and lastly decodes the

number plates of the fast-moving automobiles. The report offers an analysis of the effectiveness of the implemented method via comparisons using various video datasets. In this paper[12] the author builds an Automatic Vehicle number plate detection which is an image processing prototype, which will first process an image, and then identify the number plate to get information about the vehicle.

The main objective is to correctly design an automatic car identity machine with the aid of the usage of the vehicle's number plate. If the vehicle is an un-authenticate, then it becomes a very tedious and time-consuming and very hard task to search that vehicle. t can be also used on the entrance for security management of e.g. shopping malls, and college campuses. The machine is carried out on Python(Programming Language) and OpenCV as an image processing library, and its performance is tested on real photographs. It has been determined from the experiment that the evolved device readily acknowledges and detects the automobile's quantity plate on real images. Key Words: ANPR, OCR, CNN, OpenCV, Flask. In this paper[13] the author builds a Vehicle Number Plate Detection technology that is applied in urban areas to help law enforcement in investigation and crime prevention. Vehicles going at illegal speeds, stolen vehicles etc. can all be recognized automatically using the system without human intervention or human errors. It has been used widely in-vehicle toll booths on the highways as well as in Parking Management Systems where there is a rigid shooting angle that can capture the licence plates of the vehicles efficiently. In this study, an approach of Automatic Vehicle Plate Recognition using Livestream is discussed by considering image size, success rate and processing time as parameters. In this paper[14], an efficient and simple method is used to recognize the number plate. In the proposed method, the OpenCV library along with Python language is used for image processing using pytesseract. The input image is taken and converted into a grayscale image and the processed image is filtered through bilateral filter to remove unwanted characters. In this paper, the Canny edge detection method is used to detect the edges of license plates. TESSERACT is used as an Optical Character Recognition.

In the paper[15] author designed to perform functions such as capturing the image of the vehicle, and storing the captured image along with the transcript of the licence plate. Open CV plays an important role in preparing images and videos to identify objects and Tesseract OCR is used for text recognition in our prototype. The main purpose of this system is to design and develop an accurate image-processing method along with the successful recognition of alphanumeric characters. In this paper[16], a system of car license plate number detection and recognition is proposed. The implemented system consists of four main stages: image processing, segmentation, noise removal and deep learning. A Raspberry Pi 4 device-aided Python language and camera are used. After image pre-processing and segmentation, the system will detect the License plate from an image of a vehicle using the YOLOv4 algorithm and Cascade Classifier with accuracy in the YOLOv4 technique is 0.999, in Cascade Classifier with a database of Russian numbers is 0.982 and in the Cascade Classifier with Indian numbers is 0.906. Then the characters of the license plate have been extracted. The facility of an easy OCR library has been used to convert this picture to text characters. Then, the process of license plate recognition was implemented step by step, using Python, OpenCV, Numpy, and easy OCR libraries. In this paper[17] the author builds The Intelligent Image Text Reader is capable of capturing an image and extracting the text from the image using the Easy OCR library and displaying them in the form of speech output. The Text to Speech conversion methodology was earlier useful for the visually impaired but now it has got a wider scope. In the era of Digitalization, text is seen almost on all online websites, blogs, and e-learning materials. Text plays a vital role in some of the real-time applications like vehicle license plates, traffic sign boards, banners, etc.

In this paper[18] the author builds an Automatic Number Plate Recognition information system that uses data extraction from a given vehicle image and applies the data for further usage in a safe, secure and modernistic Transportation System. The Novelty of the project is that even if the image is blurred, our system can deblur the given image and apply it to the Machine Learning models further. In the proposed work, the You Only Look Once [YOLO] V3 model for Region of Interest [ROI]; Convolution Neural Network [CNN] for optical character recognition was implemented. After the ROI is detected, it will be enhanced with pre-processing steps before it is fed to the CNN model. A dataset of different Indian Number Plates' Fonts was created, consisting of 6439 images of different alpha-numerical characters. An accuracy of 91.5% is obtained. The extracted and sorted characters of the number plate are cross-checked with the Indian RTO database and the information has resulted In this paper[19] the author builds a number plate recognition system. The method used is YOLOv3 (You Only Look Once), and Darknet-53 is used as a feature extractor. In this study, the data used were number plate images derived from the extraction and cropping of motorized vehicle videos that had been taken using cell phones and cameras. Testing is done with two different models, namely the model obtained with additional preprocessing data and the model obtained without any preprocessing data.

In this paper[20] the author builds a model that combines the wheel speed and satellite communication information to calculate the impact of road navigation on the IMU and its horizontal speed compared to average wheel speeds and wheel Speed Department (CRT) information. A longitudinal vehicle speed estimator using

triangulation is generated from tests conducted by participants, consisting of three virtual sensors that build synthetic longitudinal speed tracks by combining multiple data points. The speed estimation is being evaluated in a detailed and analytical way under a variety of driving scenarios under the testing hardware-in-the-loop tests. In this paper[21] author tries to develop a model that can locate a particular vehicle that the user is looking for depending on two factors 1. The Type of vehicle and the 2. The license plate number of the car. The proposed system uses a unique mixture consisting of a Mask RCNN model for vehicle type detection, WpodNet and pytesseract for License Plate detection and Prediction of letters.

In this paper[22] author builds a "License Plate Recognition System Using OpenCV and Tesseract OCR Engine," addressing the critical issue of traffic control and vehicle identification by proposing a three-staged LPR system. With advancements in automobile technology, tracking vehicles violating traffic rules has become challenging, particularly in developing countries. This system, employing OpenCV and Tesseract OCR, focuses on license plate detection, character segmentation, and recognition, utilizing infrared illumination for day-night operations. Prior research highlights the effectiveness of OpenCV for detection and Tesseract OCR for accurate text recognition, paving the way for an implementable solution to efficiently process images, capture vehicle data, and recognize alphanumeric characters, crucial for enhancing traffic management and law enforcement. In this paper[23] author builds a Real-Time Vehicle Number Plate Detection and Recognition System. within this document, they proposed a real-time vehicle number plate recognition (RVNPR) system for the recognition of number plates which can extract the characters from the number plates of vehicles passing by a particular location using image processing algorithms, it's not necessary to put in additional devices like GPS or radio frequency Identification (RFID) to implement the proposed system. Using high-definition cameras, the system takes images of every passing vehicle and sends the image to the computer for processing by RVNPR software. The plate recognition software uses different algorithms like Yolo (You Only Look Once), segmentation and at last character recognition. The resulting data is applied to match with the records in a database if the vehicle has been detected as stolen then the system automatically notifies the police and sends the location of that vehicle. In this paper[24] author builds a License plate detection using YOLO v4.

The real-time object detector YOLO (You Only Look Once) - darknet deep learning framework is used in this article to detect car number plates in parking lots in real-time. The YOLOv4 deep learning technique was utilized in this proposed strategy to automatically recognize a car's number plate from a video stream. An OCR technique is applied to extract the number from the image of the number plate. The system detects license plates with an accuracy of around 89%.In this paper[25] author builds a License Plate Detection and Recognition using YOLO v4. The proposed work performs small object detection like locating and recognizing the number plate, colour of the number plate and character on the number plate by using Yolov4 and the feature fusion approach. The proposed method can overcome different challenges in object detection and shows competitive results for small object detection with 86% detection accuracy at 45fps.In this paper[26] author builds a Deep Learning Model for an Automatic Number/License Plate Detection and Recognition System in Campus Gates. Automatic Number Plate Recognition (ANPR) is a critical technology that enables the monitoring and control of road traffic and parking management, towing systems, vehicle gate entry management, etc. This paper explores the use of deep learning techniques, including OpenCV, YOLO, PaddleOCR, and Tesseract OCR, in combination with Python programming language, to develop ANPR systems. In this paper[27] author builds a Detection of Vehicular Number Plate System using a Deep Learning Approach. Automatic Number Plate Recognition (ANPR) is a type of Intelligent Transport System. While numerous studies on plate identification, character segmentation and character recognition have been performed, several challenges remain. An efficient Vehicle Detection System is necessary to ensure traffic monitoring.

In the last 4-5 years, several image processing and learning methods have been developed such as the Optical Character Recognition (OCR) technique. The aspect of object detection, though, hasn't been exploited for the ANPR framework in the previous research focused on object detection. In this paper[28] author builds an OCR-based Automated Number Plate Text Detection and Extraction. This work focuses on the automation of text detection and extraction. In this method, the photos of vehicles are captured. Then the part of the number plate is identified, and the plate information (number) is extracted and sent to the concerned authorities via WhatsApp number. The concept is based on optical character recognition (OCR). A bilateral filter is used for edge detection and a canny algorithm is applied for edge detection. In this paper[29] author builds an Analysis of the Object detection Method using Open CV – Python. Various applications for object detection have been well-researched including face detection & recognition, character recognition and prediction, and number plate detection. Object detection and recognition are used in very vast cases and scenarios including retrieval, surveillance, detection of overspeeding of vehicles and a lot more cases.

In this research, various basic concepts used in object recognition and detection while making use of the OpenCV library of Python 3.8, increasing and improving the efficiency & accuracy of object recognition and

detection are presented. This paper[30], "License Plate Number Detection Using OpenCV and Python" presents an innovative approach to building a real-time embedded Automated License Plate Recognition System (ALPR) for automatically detecting license plate numbers of moving cars. Such systems find applications in complex security setups, communal spaces, parking access control, and urban traffic management. The challenges of ALPR, including the impact of lighting conditions and vehicle speed, are addressed through the utilization of the Open Computer Vision Library and Python programming language. The system not only detects license plates but also provides essential details related to emission testing and insurance. This research builds upon prior work in ALPR, demonstrating the potential of OpenCV and Python in creating efficient and versatile solutions for license plate recognition and vehicle data processing, crucial for enhancing security and traffic management systems. In this paper[31] author builds a Systematic Number Plate detection using an improved YOLOv5 detector. A system based on YOLOv5s is used for training the model with annotated images in the dataset. The process was divided into several steps, comprising acquisition, detection, segmentation, and finally text recognition in an image. The automobile is recognised from each photograph in the first stage. The next stage is to identify the automobiles' license plates from the identified cars. After the segmentation, the license plates are cropped. The characters are recognised in the last phase from the collected number plates.

YOLOv5 is used by the system for number plate detection and Keras for character recognition. The characters from a number plate are retrieved and entered into an Excel spreadsheet. Images of Indian license plates are used to evaluate the model's performance. The accuracy for automobile detection, number plate identification and character recognition are 97.6%, 98.2%, and 99.1%.In this paper[32] author builds An Efficient Number Plate Detection System Based on Indian Traffic Rules. Based on the Indian traffic road system, monitoring the violation of such an incident is a huge task and also a tedious system. However, steps are taken to monitor traffic breach that happens in road transport through the License Plate Detection mechanism for vehicles that are involved in overspeeding or violating the rules. License plate location is a very important concept in vehicle license plate recognition for intelligent transport systems. Number plates can have different shapes and sizes along with different colours. The most common vehicle number plate in India has the background colour yellow or white with the font colour black. Identification of number plates for vehicles in India has been discussed in this paper and the numbers have been segmented to identify them specifically.

In this paper[33] author builds a YOLO Advanced Smart Traffic Assistance Platform for Number Plate and Helmet Detection. Here is the software using YOLO V8 to recognize the motorbike drivers, who are not obeying helmet law in an automated way. The helmet and license plate detection system using YOLO V8 is a computer vision technology-based system that utilizes the You Only Look Once (YOLO) object detection algorithm to detect helmets and license plates in real time. The system is designed to improve safety on roads and highways by detecting riders without helmets and vehicles without proper license plates. The system consists of motorcycle detection, helmet and no helmet detection as well as bike license plate recognition. The system is capable of processing images from a variety of sources, including traffic cameras and drones, and can detect the presence or absence of helmets and license plates in the image frames. It uses a deep learning model trained on a large dataset of annotated images to identify and classify objects. The output of the system includes a bounding box around each detected object and a label indicating whether it is a helmet or a license plate. The system can also be configured to generate alerts or notifications when violations are detected. Overall, this system provides a valuable tool for law.

3. Methodology

The AI-Powered Vehicle Monitoring System, our project combines advanced technologies like Python, PHP, EasyOCR, OpenCV, HaarCascade, and Rest API. We extensively used Python as our primary programming language. The integration of advanced technologies such as EasyOCR, OpenCV, and HaarCascade done with Python Use of Python allows for effective coordination between the different components and well-optimized solutions for vehicle monitoring. We used two Python scripts in our project. One for the detection of the number plate and another one for extracting characters from the detected license plate.

3.1 OpenCV

OpenCV, or Open Source Computer Vision Library, is a widely used open-source computer vision and image processing library. It integrates well with machine learning algorithms, enabling the incorporation of advanced algorithms like HaarCascade for object detection which is used for dynamic image processing, allowing the system to analyze and manipulate images in real-time. below Figure 1 is a sample of primary code or object-detecting method flow.

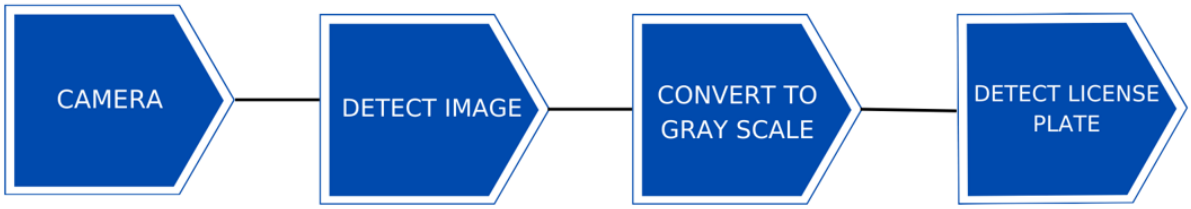


Fig 1. The first flow of the program

Our Python code uses the OpenCV library to implement a real-time license plate detection system from a live webcam. Set the path to a pre-trained Haar Cascade XML file designed for license plate detection. Stream frame converted to grayscale, and passed through the Haar Cascade Classifier to detect license plates. Then draw a rectangle around the license plate and show the region of interest in a separate window if a plate is detected. Detected license plate figures are shown below in Figure 2 and Figure 3.

Fig 2. Sample image of detected plate-1

Fig 3. Sample image of detected plate-2

3.2 EasyOCR

EasyOCR is a Python module for optical character recognition. Used to accurately identify and extract alphanumeric characters from license plates. It can be easily combined with other technologies, such as OpenCV.

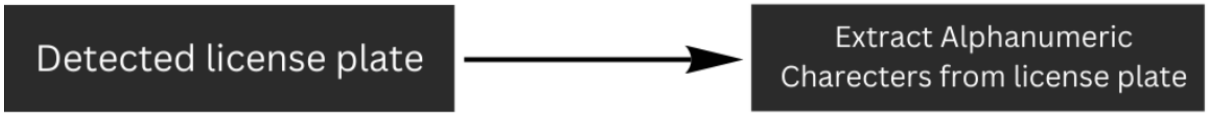


Fig 4. Diagram of easy

The easyOCR script for each unprocessed image of the license plate reads the image and applies grayscale conversion, Gaussian blur, and adaptive thresholding to enhance OCR accuracy. EasyOCR is used to extract text from the processed image. The script prints the extracted text along with the corresponding image file name. The grayscale-converted plate is shown below in Figure 5.

Fig 5. Sample image of grayscale image

Fig 6. Sample image applied to easyOCR

Figure 6 is a sample that is applied to easyOCR for extracting characters the output sample is shown in figure 7.

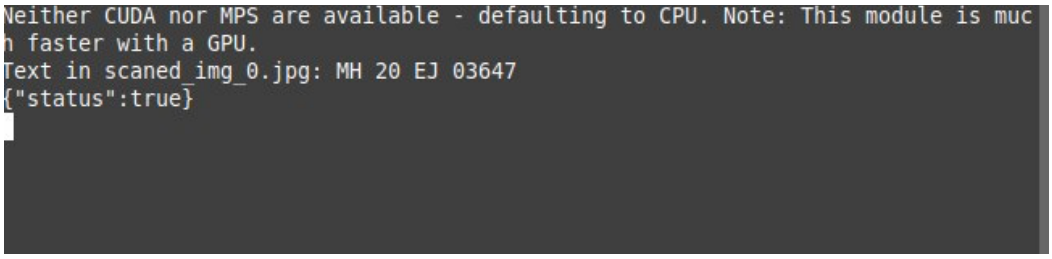


Fig 7. Extracted license plate characters

3.3 HaarCascade

Haar Cascade is a machine-learning object detection method used for identifying objects in images or video frames. It is used to enhance the system's capability for robust object detection and low configuration system needed. It was First published by Paul Viola and Michael Jones in their 2001 paper, Rapid Object Detection using a Boosted Cascade of Simple Features, this original work has become one of the most cited papers in computer vision literature. In their paper, Viola and Jones propose an algorithm that is capable of detecting objects in images, regardless of their location and scale in an image. And, this algorithm can run in real-time, making it possible to detect objects in video streams. Haar Cascade is a machine learning-based approach where a lot of positive and negative images are used to train the classifier. Positive images – These images contain the images which we want our classifier to identify. Negative Images – Images of everything else, which do not contain the object we want to detect. Haar-like features involve simple rectangular filters applied to different regions within an image. These filters capture basic image characteristics like edges, lines, and corners. By combining various Haar-like features at different scales and locations, the model can detect more complex patterns. Haar cascades rely on a feature extraction approach called Haar-like features, which involve simple rectangular filters applied to different regions within an image. These filters capture basic image characteristics like edges, lines, and corners. By combining various Haar-like features at different scales and locations, the model can detect more complex patterns.

3.4 Other tools

Rest API is used in our project for passing some parameters to PHP to create and manipulate the database. PHP is used in our project for ease of creating and managing the database where the license plate date and time are stored. From the PHP scripts, some filtering of alpha-numeric characters is done. Camera used in our project for real-time detection of license plates. We used a laptop web camera to detect the license plate in this system. The database is used to store information on licence plates. In our project, we used MYSQL as our database. It makes our data insertion and retrieval so easier process. A sample image database is given below in Figure 8.

License Number	Date	Time
kl11bm8089	05 Mar 2024	10:53 PM
kl11bm8089	05 Mar 2024	10:54 PM
kl11bm8089	05 Mar 2024	10:54 PM
kl11bm8089	05 Mar 2024	10:54 PM
kl11bm8089	05 Mar 2024	10:54 PM
kl11bm8089	05 Mar 2024	10:54 PM
kl11bm8089	05 Mar 2024	10:54 PM
kl11bm8089	05 Mar 2024	10:54 PM
kl11bm8089	05 Mar 2024	10:54 PM

Fig 8. Database sample

4 Experimental Results

The integration of Haarcascade, OpenCV, EasyOCR, and Web camera; produced noteworthy results across key aspects: The license plate detection is fast while using Haar cascasde object detection model. It quickly detects the number plate almost if a proper license plate is present on the video stream (camera). If we used Yolo it can get far better results. The EasyOCR Python module is used for character recognition. It helped us to extract alpha-numeric characters from the detected number plate. In our project, the license plate is stored in a folder and the image(region of interest) is passed to the EasyOCR class. Easy OCR predicts 75% accurately. We also used the pytesseract module, sometimes it gives good results, but 70% of predictions are inaccurate. So we eliminated it. In Our project, we have used a normal laptop Webcam as a camera for detecting number plates in real time. We have got the maximum good results that a laptop webcam can do. We also used a mobile camera by setting up an irium-cam server which connects a mobile camera to a laptop. But it showed connection lags which negatively affected the output. It is far better to use a camera with a high shutter speed and resolution for better results. The database is used to store the license alpha-numeric characters date, and time of entry. Our Database is a MYSQL.it helped to retrieve the vehicle entries easily to the website for the user. In our project, we provide the simplest user-friendly interface for users to access the vehicle details. And used HTML, and PHP to build the user interface. Users can access data by filtering with vehicle license number or date or by both.YOLOv8 is the newest state-of-the-art YOLO model that can be used for object detection, image classification, and instance segmentation tasks. YOLOv8 was developed by Ultralytics. We first started our project using yolov8, and paddleocr After realizing the Yolo v8 and paddlepaddlegpu , didn't support our computer. It needs Cuda-enabled GPUs so, we tried it with Google Collab. but we faced many errors in accessing the webcam to make it real-time detection. so dropped it. We realized that our work in google colab does not meet our requirements, because we cannot access the webcam. Then we started to restart the project with Yolo v4 on our computers.it is also true that it failed to detect license plates from images but in real time it struggled and felt intense lag on the computer. so it also dropped. The precision, recall, and F1 scores in our cases are all 0.6667, which means they have the same value.

Table 1. Performance matrix of experiments.

Metric	Metric value (0-1)
precision	.6667
Recall	.6667
F1 score	.6667

Precision is the ratio of correctly predicted positive observations to the total predicted positives. it means that out of all the predicted positive instances, 66.67% are correct. Precision = True Positives / (True Positives + False Positives). Recall, also known as sensitivity or true positive rate, is the ratio of correctly predicted positive observations to all observations in the actual class. This means that 66.67% of the actual positive instances were correctly predicted. Recall = True Positives / (True Positives + False Negatives). The F1 score is the harmonic mean of precision and recall. It is a measure that combines both precision and recall into a single metric. The F1 score is 0.6667, which is the same as precision and recall. F1 Score = 2 * (Precision * Recall) / (Precision + Recall) .Having the same values for precision, recall, and F1 score indicates that the model is achieving a balanced performance.

5 Discussion

In our project we used a low-configured computer, so we couldn't use the modern highly accurate CNN like yolov8. Also, we used a normal laptop webcam for streaming, it has a lower shutter speed and low resolution makes the license plate images weird which affects extracting characters. As we used low-specification computers and cameras in our project it was difficult to use the modern algorithm which made our project accuracy slightly low. For better accuracy, a camera with high shutter speed and resolution and a computer with GPU with Cuda-enabled graphics are necessary for localising the project.

There are great possibilities for the future integration of our project's AI-powered vehicle monitoring system. The current project, utilizing EasyOCR, OpenCV, and HaarCascade algorithms, lays the groundwork for future advancements and collaborations. Future integration could involve adding machine learning models for even more accurate and adaptive recognition. Deep learning algorithms, Convolutional Neural Networks (CNNs), could enhance the system's ability to recognize licenses in complex situations.

6 Conclusion

The research paper is an AI-powered vehicle monitoring system. The system is good at monitoring vehicles in real-time, EasyOCR makes reading license plates and OpenCV makes smart decisions. HaarCascade helps in finding car license plates. We tested it a lot in different situations to make sure it's integrated and works well, so we can use it for any institution which needs an automated vehicle register. We tried using Yolo for detecting license plates, but Yolo needs super-configured systems for its working. Yolo can recognise fast and make a good quality prediction. We tried pytesseract for extracting alphanumeric characters from license plates but it did not detect characters and the image is a little weird. We think our project can help students, and people who want to make a license plate detection-based project. By putting these technologies together, it makes everything safer and works smoother. Our research shows there's more we can do to make surveillance systems for vehicles even better. It's like a starting point for more cool ideas and improvements. We hope our work encourages others to join in and make smart, efficient, and safe monitoring systems.

References

[1] Tejas Walke, Akshada Agnihotri, Reshma Gohate, Shweta A. Mane, Suraj Pande, K. Pendke(2022) : Miniature Model of Autonomous Vehicle Using Arduino UNO and Open CV; International Journal for Research in Applied Science and Engineering Technology https://www.ijraset.com/best-journal/miniature-model-of-autonomous-vehicle-using-arduino-uno-and-open-cv

[2] Ananya Sri Shetty, V. S. Vineeta, Sreya Ravi, Nerella Likhitha, K. Anuradha : Vehicle Number Plate Detection through live stream using Optical Character Recognition (OCR)**;** 2023 7th International Conference on Trends in Electronics and Informatics (ICOEI) https://ieeexplore.ieee.org/document/10125986

[3] P. Siva Kumar; S. Shanmugasundaram; S. Mahalakshmi(2023): Vehicle Number Plate Identification and Speed Detection for Traffic Surveillance Prevention Using Machine Learning ; International Conference on System, Computation, Automation and Networking (ICSCAN) https://ieeexplore.ieee.org/document/10395451

[4] N.Palanivel Ap, T. Vigneshwaran, M.Sriv Arappradhan, R. Madhanraj : Automatic Number Plate Detection in Vehicles using Faster R-CNN;Published in International Conference on .3 July 2020Computer Science, Engineering 2020 International Conference on System, Computation, Automation and Networking (ICSCAN) https://ieeexplore.ieee.org/document/9262400

[5] S. Sunny (2021): Image Based Automatic Traffic Surveillance System Through Number-Plate identification And Accident Detection. ; 2nd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST) https://ieeexplore.ieee.org/document/9331102

[6] A. Naureen Ayesha Naureen; K.Chaithanya Krishna; N.Sai Teja; L. Mahesh(2023): College Bus Number Plate Registration Detection Using YOLO-V8; International Conference on Evolutionary Algorithms and Soft Computing Techniques (EASCT) https://ieeexplore.ieee.org/document/10392472

[7] Reezky Illmawati, Hustinawati 2023:YOLO V5 for Vehicle Plate Detection in DKI Jakarta; Jurnal Ilmu Komputer dan Agri-Informatika. https://journal.ipb.ac.id/index.php/jika/article/view/46746/25834

[8] Pooja Suresh Puri, P. Bartakke (2023) : SSD Based Vehicle Number Plate Detection and Recognition; International Conference onNetwork ,Multimedia and Information Technology https://ieeexplore.ieee.org/document/10276329

[9] Henakshi Das, Prasant Kumar Sahu, Kukil Khanikar (2023): AI-based license plate detection of on-road vehicle ; International Conference on Microwave, Optical and Communication Engineering https://ieeexplore.ieee.org/document/10167115

[10] K. Suneetha, K. Raj (2021): Automatic Vehicle Number Plate Recognition System (AVNPR) Using OpenCV Python; Lecture Notes in Networks and Systems https://link.springer.com/chapter/10.1007/978-981-16-1941-0_49

[11] P. S. Gaikwad, Pratik Chaudhari, Pragati Bhole, Vinayak Girhe, Aniruddh Karekar (2023): Real-time Vehicle Count, Speed Estimation and Number Plate Detection using CCTV Footage ; 1st International Conference on Cognitive Computing and Engineering Education (ICCCEE) https://ieeexplore.ieee.org/document/10424558

[12] Dr.Mrs.Nita Thakare (2023):Automatic Vehicle Number Plate Detection System; Interantional Journal of Scientific Research in Engineering and management https://ijsrem.com/download/automatic-vehicle-number-plate-detection-system/

[13] Ananya Sri Shetty, V. S. Vineeta, Sreya Ravi, Nerella Likhitha, K. Anuradha (2023): Vehicle Number Plate Detection through live stream using Optical Character Recognition (OCR) ; 7th International Conference on Trends in Electronics and Informatics (ICOEI). https://ieeexplore.ieee.org/document/10125986

[14] A. S. Mohammed Shariff, Raghav Bhatia, Raghwendra Kuma, Sarthak Jha (2021): Vehicle Number Plate Detection Using Python and Open CV ; International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE) https://ieeexplore.ieee.org/document/9404556

[15]R. B, Raghav G, Harshith M, Rahul Patwadi, Aravind H S (2022): Licence Plate Recognition System Using Open-CV and Tesseract OCR Engine; International Journal of Engineering Research in Computer Science and Engineering https://ijercse.com/viewabstract.php?id=16065&volume=Volume9&issue=Issue9

[16] Qudes Mb Aljelawy, T. M. Salman (2022) : Detecting License Plate Number Using OCR Technique and Raspberry Pi 4 With Camera;, International Conference on Multimodal Interaction https://ieeexplore.ieee.org/document/9873776

[17] C. Jeeva, T. Porselvi, B. Krithika, R. Shreya, G. S. Priyaa, K. Sivasankari (2022) : Detecting License Plate Number Using OCR Technique and Raspberry Pi 4 With Camera ; International Conference on Power, Energy, Control and Transmission System https://ieeexplore.ieee.org/document/9873776

[18] R. Shashidhar, A. Manjunath, R. Santhosh Kumar, M. Roopa, S. Puneeth (2021): Vehicle Number Plate Detection and Recognition using YOLO- V3 and OCR Method; IEEE International Conference on Mobile Networks and Wireless Communications (ICMNWC) https://ieeexplore.ieee.org/document/9688407

[19] M. Rubiston, Harigaran K, Darshini S, Mukeshkumar B, Santhosh U, Albee Micheal J (2023): Vehicle Speed Determination Using Haarcascade Algorithm ; Journal of Physics: Conference Series reference-link

[20] Raghav Puri, Harneet Kaur (2023): Automatic Number Plate Detection System; 9th International Conference on Smart Structures and Systems (ICSSS) https://ieeexplore.ieee.org/document/10407050

[21] V. Singh:(2021) :Recognition of Vehicle Number Plates and Real-Time Identification using Open CV ; International Journal for Research in Applied Science and Engineering Technology https://www.ijraset.com/fileserve.php?FID=21801

[22] , M. Shrivastava, Rahul Jadhav, Pranjal Singhal, S. Bhosale (2021) :Vehicle Owner Recognition and Speed Estimation through LPD using Deep learning ;VORSELD ITM Web of Conferences https://www.itm-conferences.org/articles/itmconf/abs/2021/05/itmconf_icacc2021_01005/itmconf_icacc2021_01005.html

[23] Raghav G, Harshith M, Rahul Patwadi, Aravind H S(2022) :Licence Plate Recognition System Using Open-CV and Tesseract OCR Engine;International Journal of Engineering Research in Computer Science and Engineering https://ijercse.com/viewabstract.php?id=16065&volume=Volume9&issue=Issue9

[24] Divya Rastogi, Mohammad Shahbaz Khan, Kanav Jindal, Karan Singh (2020):A Real-Time Vehicle Number Plate Detection and Recognition System https://www.semanticscholar.org/paper/A-Real-Time-Vehicle-Number-Plate-Detection-and-Rastogi-Khan/72cbf6dee0096399b90dc554be2b700308bc5e6a

[25] Rishabh Rathi, Aditya Sharma, Nikesh Baghel, Prathamesh Channe, Shreyas Barve, Sweta Jain (2022) :License plate detection using YOLO v4 ;International Journal of Health Sciences. https://scienceschola.us/journal/index.php/ijhs/article/view/7475

[26] M. Deshpande, Sr. Member, (2022) License Plate Detection and Recognition using YOLO v4,https://www.semanticscholar.org/paper/License-Plate-Detection-and-Recognition-using-YOLO-Deshpande-Member/57476f3b5c96130c6e4513a1edc9f0a5fed34778

[27] Twana Mustafa, M. Karabatak (2023):Deep Learning Model for Automatic Number/License Plate Det ection and Recognition System in Campus Gates ; International Symposium on Digital Forensics and Security https://ieeexplore.ieee.org/document/10131758

[28] Shally Gupta, Rajesh Singh, R. Srivastava, H. Mandoria :Detection of Vehicular Number Plate System using Deep Learning Approach ; M.Tech. Student project 2021.https://www.semanticscholar.org/paper/Detection-of-Vehicular-Number-Plate-System-using-Gupta-Singh/0a0e404ff4f092ffab839985d7c8afd2baf9aae8

[29] Sarthak Ranjan, Shrusti Sanket, Shivam Singh, Shivam Tyagi, Mandeep Kaur, Nitin Rakesh, P. Nand: OCR based Automated Number Plate Text Detection and Extraction; International Conference on Computing for Sustainable Global Development 2022 https://ieeexplore.ieee.org/document/9763248

[30]Nandan Verma, 2Mohit Jauhari, 3Mrs. Rashmi Mishra(2020) : Analysis of Object detection Method using Open CV – Python, International journal for research in engineering application & management .https://ijream.org/papers/IJREAMV06I0161004.pdf

[31] Shivganga Patil, Ashwini: licence plate number detection using open cv and python, 2022 IJCRT https://www.semanticscholar.org/paper/LICENCE-PLATE-NUMBER-DETECTION-USING-OPEN-CV-AND-Patil/d2f739199f26c6e3d99b77b8de8c8c7b0ad9874a

[32], Kathirvel A, Blesso Danny J, Shalem Preetham Gandu, Joe Hinn T O, Roak Kennedy C, Aldrin Immanuel J, 2023: Systematic Number Plate detection using improved YOLOv5 detector ; 2nd International Conference on Vision Towards Emerging Trends in Communication and Networking Technologies (ViTECoN). https://ieeexplore.ieee.org/document/10157727

[33] Maheswari, A. Saranya, S. Aruna :An Efficient Number Plate Detection System Based on Indian Traffic Rules ; M., IOP Conference Series: Materials Science and Engineering 2020 **https://iopscience.iop.org/article/10.1088/1757-899X/993/1/012100**

Computer Science Research Reports (2024), 2-1, 59-69
ISBN : 978-93-94056-72-5

RESEARCH ARTICLE

REAL-TIME TRANSLITERATED MALAYALAM TEXT TO MALAYALAM SIGN LANGUAGE CONVERTER

Febina P[1*], Haifa Visharathody[2], Hiba Sherin MT[3], Munavira TC[4] ,Swapna Mol George[4]

[1,2,3,4] *PG & Research Department of Computer Science, Sullamussalam Science College, Areekode.* Email:

**febinapbasheeer2003@gmail.com*

Abstract: *This project aims to create a real-time system that transliterates text into Malayalam sign language videos. This will address the challenge of communication for those who use Malayalam sign language by converting text into accurate signs. While established sign languages have complex structures, Malayalam Sign Language is new, with a uniform alphabet developed in 2021. The system will use natural language processing to understand text and generate videos using a sign language gesture database and video rendering module. This technology will empower Malayalam-speaking hearing-impaired individuals to communicate more effectively with both the hearing and signing communities.*

Keywords, Malayalam Sign Language, real-time MSL, Natural Language ToolKit, iNLTK

1. Introduction

Sign languages are diverse and rich forms of communication used by individuals and communities who are hearing impaired around the world. Each country or region has its unique sign language, complete with its grammar and vocabulary. These languages are not just visual representations of spoken languages but have their own cultural significance and linguistic structures. Unfortunately, many countries do not recognize sign languages as official languages, which can make it challenging for hearing-impaired individuals to access education, employment, and services. American Sign Language (ASL) is widely used in the United States and Canada, while British Sign Language (BSL) is used in the United Kingdom, and Australian Sign Language (Auslan) is used in Australia. International Sign is used among hearing-impaired individuals from different countries during international events. Indian Sign Language (ISL) is used by hearing-impaired individuals and communities in India and is a diverse language with regional variations. Despite its importance, ISL has not yet received widespread recognition and support. Efforts are underway to document and standardize ISL and promote its use and preservation.

Malayalam Sign Language (MSL) is a unique sign language used by the hearing-impaired community in the Indian state of Kerala, where Malayalam is the primary spoken language. Developed by the National Institute of Speech and Hearing (NISH) in 2021, MSL has its distinct grammar and vocabulary, separate from spoken Malayalam. It is used for everyday communication, education, and cultural expression within the hearing-impaired community. Although MSL is a recently introduced

language, technological advancements have not yet caught up with its needs. To address this gap, this project aims to utilize natural language processing (NLP) to convert transliterated text to real-time MSL video. The system processes the input text using NLTK tokenization methods and checks a database for corresponding MSL signs. If found, the video output displays the signs. The objective of this project is to promote inclusivity and accessibility for the hearing-impaired community in Kerala. By bridging the communication gap between MSL and non-MSL users, this system enhances connectivity and understanding.

2. Literature review

Sign languages are not universal, unlike spoken languages. People from different countries who have hearing impairments use their respective sign languages. While there has been significant research on Sign Language Translation and Recognition, the reverse process has received less attention. Specifically, there is limited research on converting spoken languages into sign languages. This gap in the field reflects a lack of attention to the needs of individuals who rely on sign language for communication. Technologies and systems developed for sign language translation often prioritize the needs of hearing individuals who are communicating with the deaf community. Consequently, individuals who use sign language for communication may face challenges in accessing information and participating in various aspects of daily life due to the limited availability of tools and resources for translating spoken languages into sign languages.

Addressing this gap in research and development is essential for promoting inclusivity and accessibility for individuals who use sign language as their primary mode of communication. Anannya Priyadarshini Neog, Arunabh Kalita, and Nithyakani Pandiyarajan (2023) [1] introduced an application designed to translate speech/text into Indian Sign Language (ISL) utilizing concepts from Natural Language Processing (NLP). The application effectively responds to most audio inputs by generating the corresponding images/gifs. Jashwanth Peguda, Y Vijayalata, Ashlin Deepa R N, Vaddi Mounish, V Sai Sriharsha Santosh et al., (2022) [2] research aims to develop a speech-toIndian sign language conversion system for six Indian regional languages: Telugu, Hindi, Malayalam, Marathi, Kannada, and Tamil. The proposed model accepts speech as input and generates a sequence of corresponding gestures as output. The process involves speech recognition utilizing Wavelet-based Melfrequency cepstral coefficients (MFCC) with Gaussian Mixture Models (GMM), text translation using Long Short-Term Memory (LSTM), and mapping the translated text with the respective sign language.

Hemang Monga, Jatin Bhutani, Muskan Ahuja, and Nikita Maid (2021) [3] project primarily focused on developing a translation system composed of multiple modules designed to convert English audio input into English text. Subsequently, the text undergoes parsing into a structured grammar representation. Each word is then cross-checked in a dictionary containing videos corresponding to each word. In cases where words are not found in the dictionary, suitable synonyms are substituted. Ashmi Katariya, Vaibhav Rumale, Aishwarya Gholap, Anuprita Dhamale, and Ankita Gupta (2020) [4] developed a system comprising a module that first converts input into English text, followed by sentence parsing and application of grammar rules. This process involves removing stop words from the recorded sentence. Subsequently, all words are checked against labels in a dictionary containing videos representing each word. The system's objective is to transform sentences into Indian Sign Language (ISL) according to grammar rules.

Chirag Garg, Gautam Priyadarshi Attry et al., (2023) [5] drafted an application that captured speech into a series of visuals of Indian sign language. Here the voice input is processed using the Speech Recognition technology to transform it into equivalent text, and then the text is further translated into its equivalent sign language. Word segmentation and root word extraction were achieved using the convenient Natural Language Processing algorithms. Harshita Mishra, Mansi Sharma, Muskan Ali, Shivani Chaudhary, et al. (2022) [6], developed a project enabling users to record their speech using a microphone or input text, utilizing NLP-based speech recognition. In instances where

the corresponding video is absent from the database, the system outputs the word and displays the associated video.

N. Pranay (2022) [7] with the utilization of Blender, Python, SQL, and NLP, devised a process to detect text and live speech, converting it into animated sign language in real time. Blender serves for animation and video processing, while SQL and NLP aid in word and text detection through dataset training and text-to-animation conversion. Consequently, this project aims to fulfill the needs of numerous individuals. Given the diverse sign language systems across different countries and even within multilingual nations, which adhere to varying social and cultural standards, there exists a significant gap in digital sign language representation. However, software such as this endeavours to bridge these disparities, serving not only as a communication tool but also as an educational resource. In conclusion, the system is cost-effective and possesses wide applicability.

Malu S Nair, A. P. Nimitha, S. M. Idicula (2016) [8] introduced a machine translation system aimed at converting Malayalam text into Indian Sign Language (ISL) through a synthetic animation approach. This system utilized the HamNoSys structure as an intermediary representation for signs, allowing for the translation of both single-word and multi-word inputs into 3D character animations. Additionally, it featured an interactive sign editor, enabling the input of new words into the database and facilitating the transcription of signs into the HamNoSys structure. Primarily designed as a tutoring system, it aimed to promote sign language education among the general populace of Kerala. Amandeep Singh Dhanjal, Williamjeet Singh (2018) [9] developed a system for translating Punjabi text into 3D animation, with a corpus comprising over 100 commonly used Punjabi words. This system holds potential for implementation as a tutoring tool to foster sign language education among the general populace of Punjab. L. Goyal, and Vishal Goyal (2016) [10] prototyped an English Text to Indian Sign Language conversion system, incorporating modules such as parsing, elimination, lemmatization, lexicon, and animation. By converting HamNoSys code into SiGML tags, the system achieved real-time conversion of entire English sentences into Indian sign language. Anju Yadav, Rahul Saxena, Bhavna Saini, Vivek K Verma, and Vibhav Srivastava (2021) [11] proposed a web application for translating English speech/text to Indian Sign Language using phrase structure grammar.

3. Methodology

This system aims to help people with hearing impairment by accepting input in Transliterated Malayalam text, which is Malayalam text written using English alphabets or Romanic script. The system uses the Natural Language Toolkit (NLTK) to preprocess the input text. It breaks down the text into individual words through tokenization and extracts keywords to understand the message. It also isolates any individual letters present in the text for further processing. The system crossreferences each keyword and individual letter against a predefined video database containing Malayalam Sign Letter (MSL) interpretations. When it finds a match, the system retrieves the corresponding MSL video and displays it to the user. This seamless integration of text processing and video retrieval helps hearing-impaired individuals understand the input text more effectively, enhancing communication accessibility and inclusivity.

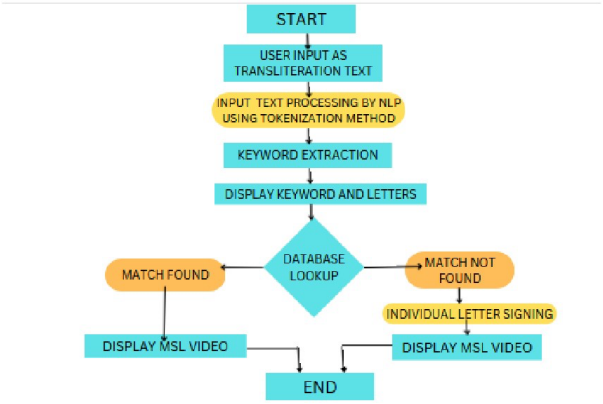


Fig 1. Overview of the system

4. Data Set Collection

4.1 Malayalam Signed Letters

The development of the first uniform sign language alphabet in Malayalam by experts at the National Institute of Speech and Hearing (NISH) in September 2021 represents a significant milestone in efforts to improve communication accessibility for the hearing-impaired community. Led by dedicated faculty members Arun Gopal and Sandeep Krishnan, alongside esteemed sign language experts, the initiative aimed to establish a standardized system for sign language communication tailored specifically to the Malayalam-speaking population in Kerala. By introducing a structured approach to finger-spelling vowels and consonants, the newly developed Malayalam sign language (MSL) alphabet aimed to provide a reliable and consistent means of expression for individuals with hearing impairments, enhancing their ability to communicate and interact with others in their community. Despite this achievement, progress in providing technological advancements or assistive technologies to further enhance the accessibility and effectiveness of the MSL alphabet has been limited.

While the development of the alphabet marked an important step forward in addressing communication barriers, subsequent efforts to integrate this innovation with modern technology and assistive devices have faced obstacles. This lack of advancement has impeded the widespread adoption and utilization of the Malayalam signed alphabet, limiting its potential impact on improving communication outcomes for individuals with hearing impairments. Moving forward, there is a clear need for continued research and development to explore innovative ways of integrating the MSL alphabet with cutting-edge technologies, maximizing its accessibility and usability for the hearing-impaired population in Kerala and beyond.

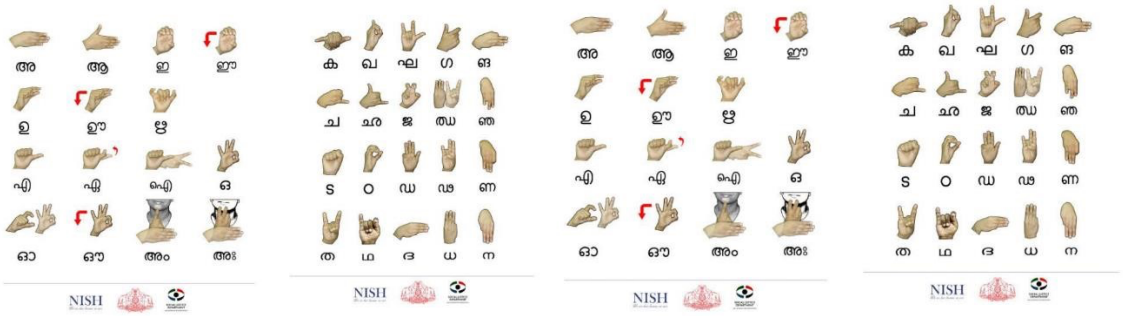


Fig 2. Malayalam Signed alphabet

4.2 Video Database Creation

A comprehensive video database has been created utilizing official resources provided by NISH, encompassing the vowels and consonants of the Malayalam language. The dataset is labeled in transliterated Malayalam text for easy data extraction and includes commonly used words in everyday communication. Signing sessions have been recorded in high definition to capture intricate hand movements and gestures, ensuring clarity and precision in visual representations. All collected videos are indexed and organized for efficient retrieval and matching during the translation process. This video database serves as the primary resource for our web application, enabling real-time access to sign language interpretations for transliterated Malayalam content.

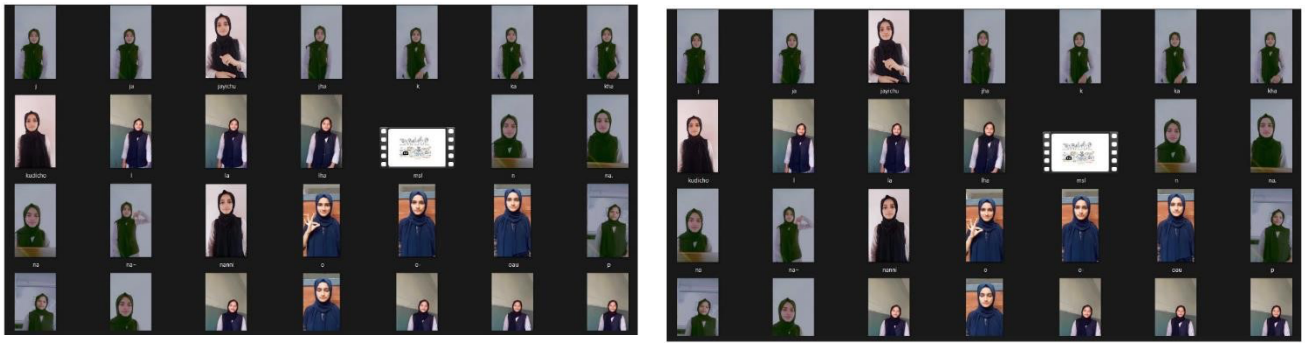


Fig 3. Video Database of the Malayalam Signed Letters

4.3 Video Database of the Malayalam Signed Letters

The Malayalam text that is transliterated will be preprocessed using Natural Language Processing (NLP) techniques. The preprocessing will include tasks such as text normalization, tokenization, and keyword extraction. The nuances of the transliterated Malayalam text will be given special attention to ensure accurate preprocessing.

4.4 Tokenization

Tokenization is a crucial step in NLP that involves dividing a text into smaller units known as tokens. This process is essential in various NLP tasks such as text processing, language modeling, and machine translation. Tokenization segments a text based on specific criteria like whitespace, punctuation, and special characters. Tokens are typically words or sub-words that serve as the fundamental building blocks of NLP. It breaks text into words based on spaces or punctuation marks. Character tokenization breaks text into individual characters. For example, the sentence “veedu evide” would be tokenized into ["veedu", "evide"]. "Evide" would be tokenized into ["e", "v", "i", "d", "e"].

4.5 Keyword Extraction

After tokenization, the system identifies keywords within the text. Keywords are the words that exist in the video database labeled as the word itself. If the word doesn’t exist in and of itself, then the individual letters or characters are considered as keywords. The keywords will act as pointers to the corresponding Malayalam signed letters in the video database.

4,6 Word Lookup

Each extracted keyword is matched against the entries of the video database. Upon successful word lookup, the corresponding video recordings from the database are retrieved and prepared for playback.

4.7 Video Retrieval and Display

Upon identifying matching keywords or letters, retrieve the corresponding MSL videos from the database. Display these videos to the user in real-time, providing visual interpretations of the input Transliterated Malayalam text. The signed letter video interpretations alongside the corresponding input text are displayed, ensuring seamless integration and comprehension for the user.

5. Experimental result

The deployed system works seamlessly for the conversion of Malayalam transliterated text to signed letters.

5.1 User Interface Homepage

This is the home page or landing page. Here we have the navigation bar and a 'Click to start' button that directly takes you to the converter page. The navigation bar has the links to the following pages.

- Converter page
- Sign-up
- Log in
- Contact - About

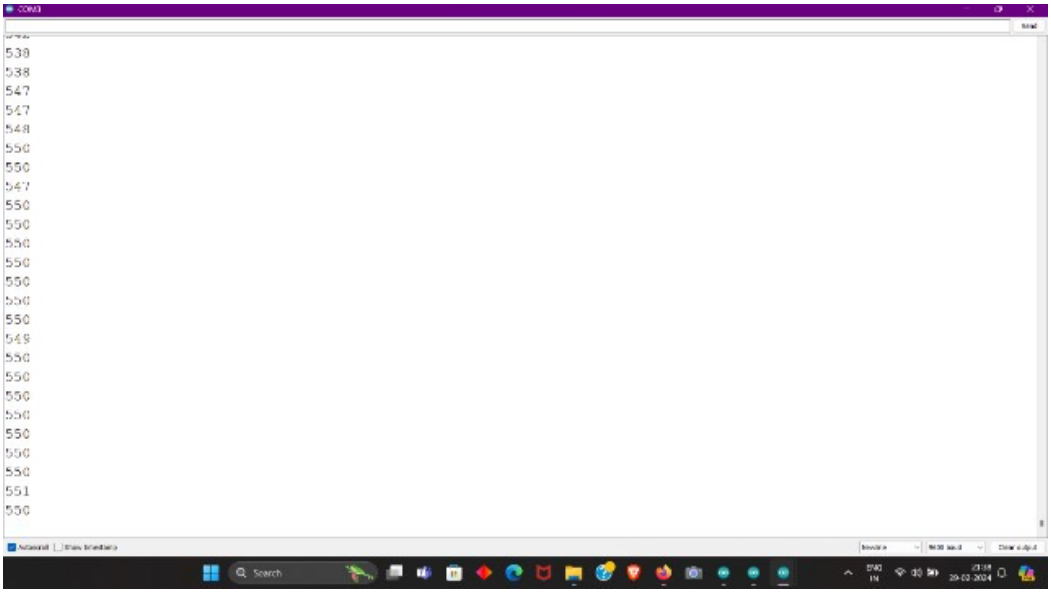


Fig4. Homepage

Converter page

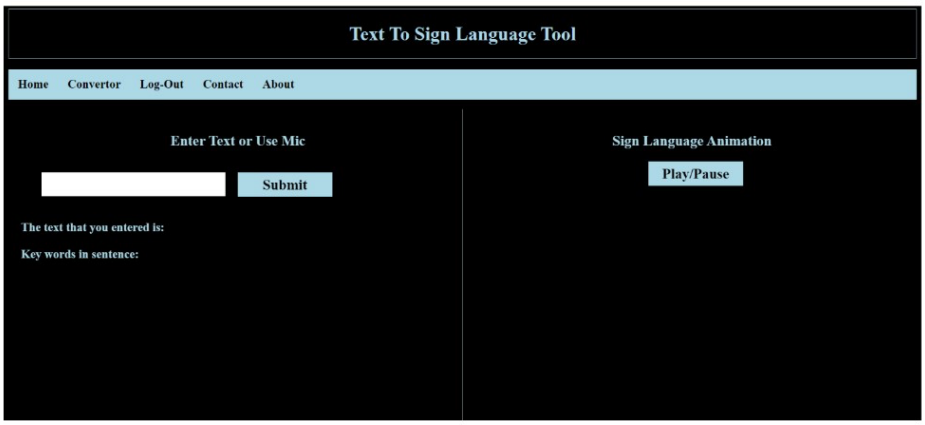


Fig 5. Converter page

To convert your Malayalam transliterated text into signed videos, the text must be entered into the designated text area. Upon completion, the 'Submit' button should be clicked to finalize the process. This will ensure that the text is accurately saved and processed, without any potential errors or discrepancies.

5.2 Example

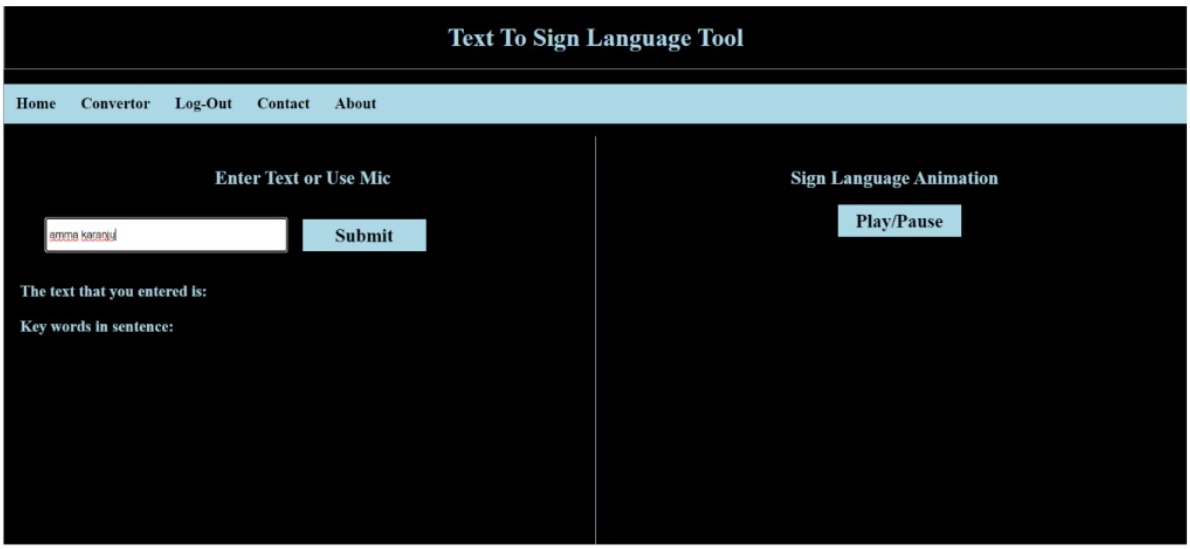


Fig 6. Transliterated text input

Here, we have entered a Malayalam transliterated text, ‘amma karanju’ and clicked on the ‘Submit’ button

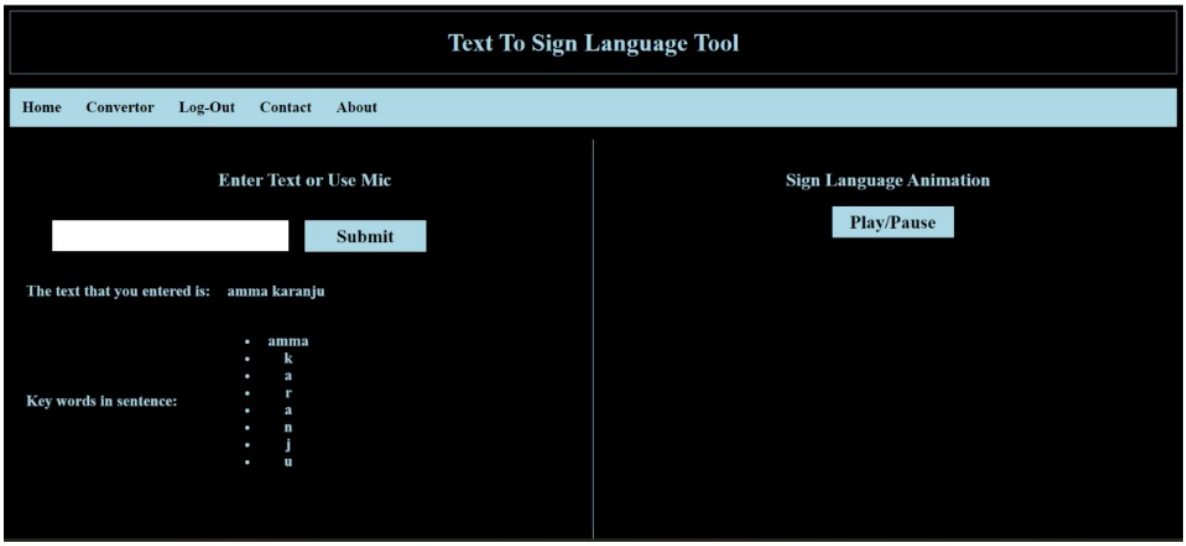


Fig 7. Tokenization and Keyword Extraction

The entered text is first tokenized and the keywords are extracted. Here, the word 'amma' exists as itself in the video database hence it is considered as a keyword while the word 'karanju' doesn't exist. Hence, the letters are individually considered keywords. These keywords are then looked up in the database and when a match is found the video is retrieved. The retrieved videos are played one by one when the 'Play/Pause' button is clicked. The keyword being played is highlighted in a teal colour.

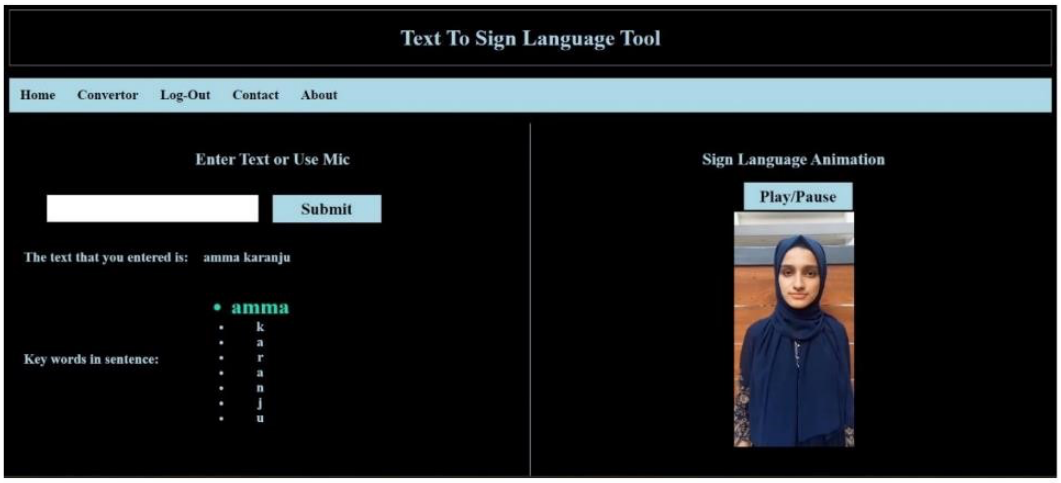


Fig 8. Video Retrieval

The keywords that have been displayed are highlighted in yellow colour

Fig 9. Signing Single Letters

6. Discussion

Our project introduces a system tailored specifically for Malayalam Signed Letters (MSL), offering several distinct advantages that enhance its functionality and effectiveness. Firstly, it boasts versatile applicability by leveraging advanced NLP techniques to analyse and interpret Transliterated Malayalam text, allowing for seamless integration into various higher-level applications within the context of MSL communication. This adaptability ensures its relevance across a wide range of scenarios, from educational to interpersonal interactions. Furthermore, our system prioritizes user experience by featuring a simple and intuitive interface, ensuring accessibility for users of all technical proficiencies and fostering widespread adoption.This project represents a significant advancement in the field of addressing communication challenges faced by the hearing-impaired community in Malayalamspeaking regions. The project has utilized natural language processing (NLP) concepts to develop a system that is capable of converting text inputs into Malayalam Signed Letters (MSL) gestures. Key NLP techniques, such as tokenization, have played a crucial role in the accurate parsing and interpretation of transliterated Malayalam text, enabling the adept application of NLP methodologies in this context. This project underscores the potential of technology to foster accessibility and inclusivity in society, particularly within educational institutions and other community settings. In the future, there are ample opportunities to further refine and deploy such systems using NLP advancements to enhance communication and comprehension for the hearing-impaired population across various domains

3.6 Limitations

1. No Voice Input: The system lacks support for voice input, limiting users to text-based interactions. This omission restricts accessibility for individuals who may prefer or rely on voice communication, such as those with motor impairments or visual disabilities. Voice input would enhance inclusivity and provide an alternative mode of interaction for a broader range of users.

2. Limited Tokenization: Due to the lack of comprehensive packages for Indian languages, the system is restricted to basic tokenization processes. This limitation may result in suboptimal text processing and analysis, potentially affecting the accuracy and reliability of the system's output. A more robust tokenization mechanism would improve the system's ability to handle complex linguistic structures and nuances.

3. No Malayalam Input: The system only accepts transliterated Malayalam text or Malayalam written with English alphabets, excluding users who prefer to input text directly in Malayalam script. This limitation may pose challenges for individuals unfamiliar with transliteration methods or those who prefer to communicate in their native script. Supporting direct Malayalam input would broaden the system's accessibility and accommodate diverse user preferences. Addressing these limitations through enhancements such as voice input integration, advanced tokenization techniques, inclusion of matras in the dataset, and support for direct Malayalam input would improve the system's functionality, accuracy, and user experience.

6.2 Future Enhancements

Looking towards the future, there are many exciting possibilities for expanding and improving our project on converting transliterated Malayalam text to Malayalam Sign Language (MSL). One promising direction is to incorporate real-time voice/sound to MSL conversion, allowing for seamless communication for those who prefer verbal communication.

We could also consider adding code-mixing capabilities to the system, catering to both Malayalam texters and transliterated texters and ensuring greater accessibility and usability for diverse user groups. Another area of potential improvement involves creating a customized avatar specifically designed for this purpose, with enlarged hands and head proportions to enhance visual attention to sign gestures. Additionally, we could expand the system's language support to include English text input,

enabling conversion into Indian Sign Language (ISL) as well. Enhancing the system to accept both English and Malayalam text, or code-mixed texts, and converting them into ISL, along with their corresponding letters if needed, using various NLP packages specifically designed for Indian languages would be beneficial. Furthermore, incorporating additional Natural Language Processing (NLP) functionalities into the processing pipeline would further refine and optimize the accuracy and efficiency of the translation process, ultimately advancing the project's effectiveness in facilitating communication for the hearingimpaired community. In addition, transforming the project into a standalone application or browser extension could extend its reach and accessibility. This would allow users to integrate the MSL conversion functionality seamlessly into their existing programs and workflows, enhancing convenience and promoting widespread adoption of the tool. Providing compatibility with different devices and software environments would also broaden the project's impact, facilitating greater inclusivity and accessibility for the hearing-impaired community in diverse settings and contexts.

Moreover, we hope to involve integrating hand gesture recognition systems using computer vision techniques to establish a two-way communication system, enabling interactive communication between users and the system. Ultimately, our goal is to develop a product that caters to the needs of speech and hearing-impaired individuals, effectively reducing the communication gap and fostering inclusivity in communication experiences. We aim to develop and evaluate automated sign language translation systems, enabling sign language users to access personal assistants, utilize text-based systems, search sign language video content, and access automated real-time translation services when human interpreters are unavailable.

4 Conclusion

This project represents a significant advancement in the field of addressing communication challenges faced by the hearing-impaired community in Malayalam-speaking regions. The project has utilized natural language processing (NLP) concepts to develop a system that is capable of converting text inputs into Malayalam Sign Language (MSL) gestures. Key NLP techniques, such as tokenization, have played a crucial role in the accurate parsing and interpretation of transliterated Malayalam text, enabling the adept application of NLP methodologies in this context. This project underscores the potential of technology to foster accessibility and inclusivity in society, particularly within educational institutions and other community settings. In the future, there are ample opportunities to further refine and deploy such systems, leveraging NLP advancements to enhance communication and comprehension for the hearing-impaired population across various domains.

References

[1] Anannya Priyadarshini Neog., et al., (2023). "Speech/Text to Indian Sign Language Using Natural Language Processing," *Indian Journal of Computer Science and Engineering (IJCSE)*, Vol. 14 No. 3. DOI: 10.21817/indjcse/2023/v14i3/231403030.

[2] Peguda, J., et al., (2022). "Speech to Sign Language Translation for Indian Languages," *8th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Coimbatore, India. DOI: 10.1109/ICACCS54159.2022.9784996.

[3] Monga, H., et al., (2021). "Speech to Indian Sign Language Translator," *Volume 39: Recent Trends in Intensive Computing*. DOI: 10.3233/APC210172.

[4] Ashmi Katariya, et al., (2020). "Voice to Indian Sign Language Conversion for Hearing Impaired People," *SAMRIDDHI: A Journal of Physical Sciences, Engineering and Technology*.

[5] Chirag Garg, Gautam Priyadarshi Attry, et al., (2023). "Conversion of Native Speech into Indian Sign Language to Facilitate Hearing Impairment," *4th International Conference for Emerging Technology (INCET)*.

[6]Harshita Mishra, Mansi Sharma, Muskan Ali, Shivani Chaudhary et al., (2022). "Audio to Indian Sign Language Translator," *International Journal for Research in Applied Science and Engineering Technology*.

[7] N. Pranay, (2022). "Speech or Text to Indian Sign Language Convertor," *International Journal for Research in Applied Science and Engineering Technology*.

[8] Malu S Nair, A. P. Nimitha, S. M. Idicula, (2016). "Conversion of Malayalam text to Indian sign language using synthetic animation," *2016 International Conference on Next Generation Intelligent Systems (ICNGIS)*.

[9]Amandeep Singh Dhanjal, Williamjeet Singh, (2018). "An Automatic Conversion of Punjabi Text to Indian Sign Language," *EAI Endorsed Transactions on Scalable Information Systems*.

[10] L. Goyal, Vishal Goyal, (2016). "Automatic Translation of English Text to Indian Sign Language Synthetic Animations," *2016 Proceedings of the 13th International Conference on Natural Language Processing*.

[11] Yadav, A., et al., (2021). "Audio to Sign Language Translator Web Application," *2021 International Conference on Computational Performance Evaluation (ComPE)*, Shillong, India. DOI: 10.1109/ComPE53109.2021.9751857..

Computer Science Research Reports (2024), 2-1, 70-75
ISBN : 978-93-94056-72-5

SOFTWARE DEVELOPMENT RESEARCH ARTICLE

PROJECT CENTRALE: A MOBILE APP STREAMLINING PROJECT MANAGEMENT FOR ARTS AND SCIENCE COLLEGES

Abdul Haseeb T[1],Shahid KP[2*],Shaheen Mohammed CM[3],Ayyoob MP[4]
[1,2,3,4] *PG & Research Department of Computer science, Sullamussalam Science College, Areekode.*

**shakp2003@gmail.com*

Abstract: *Project Centrale, a mobile application tailored for Arts and Science colleges in Kerala, offers a streamlined project management experience. Utilizing React Native for its user-friendly front-end and a secure MySQL database for data management, the app digitizes and automates manual processes, saving time and resources. Key features include simplified submission mechanisms, real-time progress tracking, and collaborative tools like group discussions and file sharing. Developed with agile methodologies, the app evolves iteratively based on user feedback, ensuring optimal performance. By enhancing project management efficiency, Project Centrale promotes a culture of productivity and innovation within educational institutions. Moreover, it democratizes access to project management tools, fostering inclusivity and enabling students from diverse backgrounds to participate actively in collaborative learning experiences. Ultimately, Project Centrale holds promise in reshaping the educational landscape of Kerala, creating a more engaging and accessible environment for both students and teachers.*

Keywords: Project Management, React Native, MySQL, Arts & Science Colleges.

1. Introduction:

Project Centrale is a new solution for Arts and Science students, helping them manage their academic projects more efficiently. This platform aims to make the project process smoother and more effective for both students and teachers. Project Centrale strives to be the preferred platform for both students and teachers in today's academic setting. It supports the overall goal of promoting collaborative and innovative learning in colleges. By offering a centralized hub for academic projects, Project Centrale not only streamlines project management but also fosters a culture of collaboration and knowledge sharing among students and teachers. Project Centrale aims to simplify and enhance the academic project experience by utilizing advanced technologies like React Native and MySQL. This results in a user-friendly interface and efficient data management, promoting seamless collaboration between students and teachers. The platform goes beyond the digital realm, influencing how academic projects are created, developed, and presented. It serves as a hub for students to showcase their work, participate in discussions, and receive valuable feedback, preparing them for future careers. Ultimately, Project Centrale acts as a vital link between academic knowledge and real-world skills. In the world of Arts and Science colleges, teachers face a common problem when it comes to managing academic projects manually.

Many educators are currently struggling with the complexities of overseeing multiple student projects,

each going through different stages like submitting synopses, giving first and second presentations, and more. While this manual approach is done with good intentions, it often leads to inefficiencies and time limitations. As teachers try to guide and evaluate students at different project stages, the absence of a streamlined, automated system creates a burdensome process that hampers effective project management. A major challenge in this manual project management process is grading each phase of the project. Teachers have to manually evaluate and assign grades to different project components, such as the project summary, presentations, and final submission.

This not only takes up a lot of time but also increases the risk of grading errors. To uphold high educational standards, colleges need a more organized and effective method to reduce the workload of manual project management. The current situation highlights the urgent requirement for a complete solution that automates the project management process for teachers in Arts and Science colleges. A unified platform that enables smooth communication between teachers and students, as well as automates the assignment and tracking of project phases, can greatly improve efficiency. By tackling these challenges, such a solution would enable educators to concentrate on guiding and mentoring students, ultimately enhancing the academic experience and promoting a more collaborative and dynamic learning environment.

Project Centrale, a new system, offers a modern solution to the project management issues in Arts and Science colleges in Kerala. It is a user-friendly Android app that allows students and teachers to collaborate easily. The app provides a central place for project uploads, communication, and progress tracking. By using React Native, the app is responsive and works well on different Android devices, making it popular among the academic community. Project Centrale utilizes a strong back end powered by MySQL to manage data securely and efficiently. This back-end handles tasks like project version control, task assignment, and user authentication. By integrating MySQL, Project Centrale ensures data integrity and scalability, making it suitable for various project management needs in different universities in Kerala. Overall, Project Centrale is an innovative Android application that enhances user experience with React Native and a MySQL-backed back end.

1.1 Features:

Project Central emerges as a transformative force in project management for Kerala's universities. This web application transcends traditional methods by offering a feature-rich and user-centric approach. Catering to distinct user roles – administrators, faculty, and students – Project Central fosters a collaborative environment that fuels knowledge exchange and streamlines the project lifecycle. At the core lies a user-friendly experience. Built with React Native technology, Project Central boasts a responsive and intuitive interface accessible from any device, at any time. This ensures a smooth experience for all stakeholders involved, eliminating technical barriers and fostering seamless interaction.

Project Central acts as a central hub, eliminating scattered communication and information. Students and faculty can access a unified platform to share resources, collaborate on projects, and stay updated on progress. This fosters a collaborative learning environment where knowledge exchange becomes effortless. Faculty can leverage the platform to provide clear instructions, answer student queries, and facilitate group discussions, all within a centralized space. Robust data management is paramount. Project Central prioritizes security and organization by utilizing a secure MySQL database. This safeguards sensitive information, ensuring the integrity of project details, submissions, and progress reports. This instills trust in the system and empowers users to confidently manage their work. Real-time insights empower faculty members. Project Central grants them a window into student activity. They can effortlessly track student progress, providing timely feedback and efficiently assessing each project phase. This fosters a more interactive learning experience and enables instructors to tailor their guidance to individual student needs. Project Central's scalability makes it a perfect solution for widespread adoption. The underlying technology stack is designed for seamless deployment across various universities in Kerala. This eliminates the need for individual institutions to develop and maintain their own systems, promoting standardization and streamlining project management practices throughout the region.

Imagine a world without the hassle of physical submissions. Project Central makes it a reality. Students can electronically submit various course materials, including synopses, project documents, code, presentations, and more. This not only saves time and resources but also contributes to a paperless learning environment. Effortless tracking becomes a reality. Project Central offers a single platform for managing all project stages.

Students and faculty can effortlessly access, track, and update project information from anywhere. This eliminates the need for cumbersome manual tracking and ensures everyone remains on the same page, fostering transparency and accountability throughout the project.

Building a supportive learning environment is central to Project Central's mission. It equips students and faculty with seamless communication tools. This enables them to ask questions, share ideas, and provide peer-to-peer support, fostering a sense of community and promoting a deeper understanding of the course material. Project Central simplifies the evaluation process. Faculty members can directly add and manage marks within the application. This eliminates the need for manual calculations and data entry, saving valuable time and resources. This translates to quicker feedback for students and a more efficient grading process for faculty.

In conclusion, Project Central surpasses the limitations of conventional project management methods by offering a comprehensive and user-centric solution. This web application fosters collaboration, ensures data security, streamlines the entire project lifecycle, and simplifies the evaluation process. By empowering both students and faculty, Project Central emerges as a valuable asset for universities in Kerala and beyond.

2. Methodology

2.1 Requirement Gathering:

Project Central's development process most likely embraced an Agile methodology, prioritizing continuous improvement through a cycle of development, testing, and adaptation. This iterative approach ensures the app constantly evolves based on user feedback. magine the development process unfolding in phases. Each phase focuses on specific functionalities and culminates in a functional product increment. Student testers, representing the core user group, meticulously evaluate these initial features. Their insights become the cornerstone for further refinement. Usability testing exposes areas where the user interface can be enhanced for improved clarity, navigation, and overall experience. Teacher testers play a vital role as well. They rigorously assess functionalities critical to their workflow, ensuring seamless task creation and assignment, effective feedback mechanisms, and efficient grade management. Communication tools are scrutinized to guarantee smooth interaction with students.

2.2 Design & Development:

Project Central's technological foundation rests on two pillars: a user-friendly React Native front-end and a secure MySQL back-end. The front-end, built with React Native, ensures a seamless experience across Android and iOS devices. A responsive and intuitive interface caters to different user roles (admin, student, teacher) with distinct functionalities. Submissions become effortless, allowing electronic uploads with progress tracking. Task management thrives with features like creation, assignment, progress tracking, and deadline reminders. Real-time chat and group functionalities foster seamless communication, while assessment is simplified through direct mark input and visualization tools for teachers.

The secure back-end, powered by MySQL, safeguards sensitive data. A well-designed relational database schema stores project details, user information, submissions, tasks, and grades. Version control ensures students can revert to previous work if needed. Data retrieval is optimized for efficient access to project information and progress reports. Security measures like user authentication, authorization, and encryption protect user data. An API layer facilitates smooth communication between the front-end and back-end. Scalability considerations ensure the system can accommodate growth, and potential offline functionalities enhance user experience.

2.3 Testing:

Rigorous testing is the cornerstone of building a robust and user-friendly application like Project Central. Throughout the development lifecycle, a multi-pronged testing approach ensures the app delivers on its promises. Firstly, unit testing verifies the functionality of individual software components. This meticulous examination catches errors early in the development process, preventing them from snowballing into larger issues later. Secondly, integration testing focuses on how different components interact and function together. This ensures a smooth user experience by identifying any compatibility or communication problems between various functionalities. Thirdly, usability testing involves real users interacting with the app and providing feedback. This invaluable insight uncovers any confusing elements or navigation issues, allowing for improvements that enhance the overall user experience. Finally, performance testing assesses the app's responsiveness and stability

under load. This ensures the app can handle multiple users and varying data volumes without compromising speed or functionality. By employing this comprehensive testing strategy throughout development, Project Central can guarantee a feature-rich, user-friendly, and high-performing application ready to revolutionize project management in Kerala's universities.

2.4 Deployment and Maintenance:

Project Central's journey extends beyond development. Launching the app and incorporating user feedback are crucial for sustained success. Deployment involves registering as a developer with the App Store Connect and Google Play Console. Following their guidelines, the app is submitted for review with necessary details like screenshots and descriptions. Once approved, Project Central becomes available for download on both platforms. Gathering user feedback forms the foundation for continuous improvement. App store rating and review systems provide valuable insights into user experience. In-app feedback mechanisms allow users to directly report issues, suggest features, or share their thoughts. Conducting surveys or focus groups offers a deeper understanding of user needs. Prioritizing the most common issues and requested features from the gathered feedback is the next step. Following rigorous testing, the development team addresses these through bug fixes and implements new functionalities based on user suggestions. Regularly releasing updates through the respective stores ensures Project Central stays relevant and incorporates user- driven improvements. This cycle of gathering feedback, iterating on the app, and releasing updates is paramount. By actively listening to users and adapting to their needs, Project Central can solidify its position in the mobile landscape and continuously evolve to better serve its user base.

2.5 DFD

A Data Flow Diagram (DFD) is a graphical representation of information flow within a system or process. It visually depicts how data enters, transforms, and exits a system. DFDs utilize symbols like rectangles (processes), arrows (data flow), and circles (data stores) to illustrate these movements. They are often used during the initial stages of system analysis and design, providing a clear understanding of a system's functionality and laying the foundation for further development. DFDs come in various levels of detail, with higher levels offering a broader overview and lower levels delving deeper into specific processes.

3. Result and Discussions

This report summarizes the results and discussions from user testing conducted on Project Centrale, a React Native application designed to streamline project management for Arts and Science colleges in Kerala. The testing involved students, teachers, and a UI consultant, focusing on functionality, usability, and user interface design.

3.1 Testing Results:

Project Central's testing process goes beyond simple functionality checks. To guarantee a user-centric and well- rounded application, a diverse testing group plays a vital role. Student testers, acting as the primary users, will meticulously examine functionalities they directly interact with. This involves ensuring seamless submission processes for various course materials and thoroughly evaluating task management features like viewing assigned tasks, updating progress, and receiving timely reminders. Usability testing becomes crucial as students assess the user interface's clarity, ease of navigation, and overall experience. Their valuable feedback will refine the app's aesthetics, with aspects like background color, icon design, and profile information completeness being scrutinized. Teacher testers, focusing on functionalities essential to their role, will rigorously test features crucial for their daily workflow. This includes creating and assigning tasks, providing constructive feedback to students, and managing marks efficiently. The seamlessness of communication tools, enabling them to exchange messages and provide timely guidance, will be a key area of assessment.

UI/UX consultants, with their keen eye for design, meticulously evaluate the app's visual appeal, user interface layout, and overall user experience. Their analysis delves into the finer details – icon size, text box behavior, button dimensions, and the color scheme. Drawing upon design principles and best practices, they provide suggestions for improvement, ensuring an aesthetically pleasing and intuitive user interface. Beyond the core testing group, exploring the implementation of a drawer menu for navigation purposes is recommended. This easily accessible menu can house essential options like profile settings, a comprehensive project list, and communication tools, streamlining the user experience for all. Staying current with the latest trends and adhering

to best practices in mobile app design is essential for Project Central's long-term success. Effective utilization of white space, maintaining consistent element sizes and layouts, and employing a color palette that prioritizes readability and accessibility are all crucial aspects to consider. By incorporating this multi-faceted testing approach, Project Central can ensure a user-friendly and feature- rich application. Insights gleaned from various user groups, coupled with the expertise of UI/UX consultants, pave the way for an app that effectively caters to the needs of students, teachers, and educational institutions in Kerala.

Project Central's testing phase yielded promising results, showcasing its potential to revolutionize project management in Kerala's educational institutions. Both students (5 testers) and teachers (2 testers) confirmed the app's core functionalities functioned as intended. This includes crucial aspects like submissions, task management, task creation and assignment, providing feedback, and managing grades. This establishes a strong foundation for the application. While the core functionalities met expectations, valuable user experience (UX) insights emerged. Students suggested improvements to the app's visual appeal, including changing the background color, allowing for more detailed profile information, and adjusting icon colors for better clarity. Implementing these suggestions can significantly improve the aesthetics and overall user experience for students. Feedback from teachers indicated satisfaction with the current functionalities, highlighting the app's effectiveness in supporting their workflow. This signifies that Project Central successfully addresses the core needs of educators.

UI/UX consultants further refined the user experience by identifying areas for improvement in the user interface (UI). This included the size of profile and search icons, ensuring consistent text box width, standardizing button size, and refining the color palette. Incorporating these suggestions would lead to a cleaner aesthetic, improved usability, and a more visually cohesive apathy testing process also revealed the potential benefits of implementing a drawer menu. This readily accessible menu would streamline navigation within the app by providing users with quick access to essential features like profile settings, project lists, and communication tools. Based on the test results and discussions, several key improvements are recommended. Implementing the UI/UX consultant's suggestions regarding icon size, text box behavior, button dimensions, and color scheme will result in a more streamlined and visually appealing interface. Adding a drawer menu would significantly enhance navigation possibilities. Modifying the background color, expanding profile information options, and adjusting icon colors based on student input will personalize the user experience and cater to their preferences. By incorporating these refinements, Project Central can evolve into a user-friendly and visually engaging tool. This transformation will not only foster efficient project management but also enrich the learning experience for both students and teachers in Kerala's Arts and Science colleges. The testing phase serves as a springboard for further development, ensuring Project Central stays aligned with user needs and delivers an exceptional experience for all stakeholders.

4 Implementation

Implementation is a critical phase in any software project. The implementation plan for Project Centrale follows a phased approach, starting with the development of core functions using React Native for the Android app. The backend will utilize MySQL for efficient data management. Once thoroughly tested, the app will be deployed on various platforms such as the Google Play Store and Apple App Store, ensuring accessibility for a wide range of users. Firstly, Project Centrale's implementation strategy involves making the app available on popular platforms like the Google Play Store for Android users and the Apple App Store for iOS users. This deployment across multiple platforms ensures that students and teachers can easily access and adopt the app, regardless of their preferred mobile operating system. In addition, the app's user interface is designed to be user-friendly and practical, making it easy for both teachers and students to navigate. Teachers can efficiently manage projects, evaluate student progress, and provide timely feedback. Students can upload project-related documents, collaborate with peers, and track their project status. The app's usability aims to enhance the overall experience of project management for both user groups. Key features of Project Centrale include a centralized project repository, version control for tracking project progress, task management tools, and real- time collaboration. These features streamline project workflows, facilitate effective communication, and ensure transparency in the project management process. The integration of these functionalities enhances the overall efficiency and effectiveness of academic project management. Overall, the Project Centrale caters to a wide audience within

Arts and Science colleges, serving students and teachers across various academic disciplines. Its adaptability allows for extensive usage and adoption.

5 Conclusion

Project Centrale revolutionizes project management in Arts and Science colleges, overcoming manual process challenges. With its user-friendly interface built on React Native, students and teachers can easily navigate the platform. The internet-accessible application promotes collaboration and resource sharing, creating a vibrant learning setting. This technology-based solution meets the specific requirements of Arts and Science colleges by offering an organized and effective project management platform. Project Centrale is set to make a big impact in the academic world, especially for students and teachers in project management. It helps teachers by automating tasks and reducing their workload, allowing them to focus on guiding students. For students, the platform offers a space for collaboration in project development, with features like version control, task management, and real-time communication. To make the most of Project Centrale, students and teachers should use its centralized repository for project uploads, participate in collaborative discussions, and utilize features like version control and task management for a more organized project management experience. In conclusion, Project Centrale is set to transform project management in Arts and Science colleges, promoting efficiency, collaboration, and an enhanced academic experience.

References

1. React Native.(2023) https://reactnative.dev/
2. Bendoly, E., D. Thomas and M. Capra. 2010. Multilevel social dynamics considerations for project management decision makers: Antecedents and implications of group member tie development. Decision Sciences 41(3): 459-490
3. Berkun, S. 2008. Making Things Happen: Mastering Project Management, revised edition. O'Reilly Media.
4. Berkun, S. 2005. The Art of Project Management. O'Reilly Media.
5. Wysocki, R. K. 2009. Effective Project Management: Traditional, Agile, Extreme, 5th edition. Wiley.
6. Wilke, T. J. 2005. The project cost variance analysis model: A project management tool. The Journal of Cost Analysis & Management (Fall): 13-23.
7. Verzuh, E. 2005. The Fast Forward MBA in Project Management, Second Edition. Wiley.
8. Sarasohn, L. J. and M. S. Luehlfing. 2002. Project management on the fly - The road to nowhere. Management Accounting Quarterly (Summer): 1-12.
9. Meinhart, W. A. and L. M. Delionback. 1968. Project management: An incentive contracting decision model. The Academy of Management Journal 11(4): 427-434.

Computer Science Research Reports (2024), 2-1, 76-90
ISBN : 978-93-94056-72-5

RESEARCH ARTICLE

SPEECH EMOTION RECOGNITION FOR MALAYALAM USING DEEP LEARNING

Jalala P V[1], Rahbar Zahid S[2], Preethi M[3], Rizwana Kallooravi Thandil[4]

[1,2,3,4] *PG & Research Department of Computer Science, Sullamussalam Science College, Areekode.*

*pvjalala@gmail.com

Abstract: *Speech is a type of communication in which people or groups communicate ideas, thoughts, feelings, or information by using spoken language. It is an essential component of human contact and the main means of communication in the majority of global civilizations. Human-machine interactions have been more prevalent in real-world interpersonal encounters in recent years. Speech analysis has played a major role in closing the gap between the real and digital worlds. Since emotional cues are increasingly used in human speech to convey information, the study of emotion detection has grown in prominence. In this work, we focused on efficient automatic speech emotion recognition for human-computer interaction. Speech emotion recognition has been the subject of extensive study, which is still ongoing today. However, most of this research has focused on a small number of languages, including English, German, Chinese, and French. The primary cause of this is the lack of datasets in different languages. Low-resource languages are those that fall into this category. In this paper, a Deep Convolutional Neural Network (CNN)-based Malayalam language voice emotion recognition system is introduced. Raw speech from the dataset is fed into a model to train, categorize, and test it.*

Keywords: Speech Emotion Recognition, Malayalam Speech Processing, Convolutional Neural Networks

1.Introduction

Emotion is characterized as a powerful feeling accompanied by physiological changes, shaped by internal thoughts, beliefs, and external circumstances. It spans from instinctual states like fear and pleasure to complex emotions such as love, anger, and sadness, influencing human behaviour, decision-making, and interpersonal relationships. Emotions are mental states influenced by both internal and external events, involving physiological reactions, external behaviours, and internal feelings. Positive emotions arise when situations align with expectations, while unfavourable events evoke negative feelings. Speech emotion recognition (SER) involves identifying emotions in male or female speakers through features like fundamental frequencies, MFCC. SER operates as a pattern recognition system with five modules: emotional speech input, feature extraction, feature selection, classification, and recognized emotional output.

Speech, a crucial communication tool for humans and certain animals, conveys emotions naturally but poses challenges for machines. Palette theory decomposes emotions into primary emotions—anger, disgust, fear, joy, sadness, and surprise—similar to how colours comprise basic colours. SER for Malayalam becomes essential due to the diverse applications of automatic emotion recognition systems. These systems find use in education, automobile industry, security, communication, and health. In education, they can adapt distance learning based on detected emotions, while in automobiles, they can enhance driving experience and performance by monitoring the driver's emotional state. In security, they assist in public spaces by detecting extreme emotions. Communication benefits from SER in call centers, improving customer service when integrated with interactive voice response systems. Health applications include aiding individuals with autism to understand and adjust their emotions using portable devices.

The existing studies on SER primarily focus on specific languages like English and Chinese, with limited coverage for other languages. India, being linguistically diverse, presents an opportunity to extend these studies to various languages. Malayalam, a Dravidian language spoken in southern India, is designated officially in Kerala, Lakshadweep, and Puducherry. With approximately 38 million speakers globally and various regional variations, designing a speech-emotion recognition model for Malayalam can extend the benefits observed in other languages to the Malayalam-speaking population.

2.Literature Review

Recent advancements in voice processing, fuelled by machine learning, have led to the widespread use of various technologies in emotion recognition. Studies conducted in real-life scenarios aim to enhance realism and demonstrate the applicability of these technologies, significantly impacting the crucial aspect of accuracy through the adoption of novel techniques. Prof. Jagdish Kambale, Abhijeet Khedkar, Prasad Patil, Tejas Sonone (2023) [1] In this paper, the algorithm for identifying speech-based emotions is implemented using deep learning. It also provides an overview of deep learning methodologies and examines some recent research that makes use of these methods. It makes use of a dataset of various emotional voices and then aids in the identification of that emotion. It will be beneficial for the computers or robots to understand humans more clearly and function in accordance with it.

Apoorva Sharma,Himanshu Nawani,Shalini verma(2023) [2] In this paper, a SER system based on deep learning methodologies is proposed.Here, RAVDESS dataset was utilised to assess the suggested system. To select the most appropriate vocal features that represent speech emotions for this MFCC is used. LSTM, CNN, and a hybrid model that combines CNN and LSTM are three different deep learning models used to construct SER system. Shashidhar Harikant, Rakshitha Prasad, Vijaya Lakshmi R, Sidhramappa H (2022) [3] This paper covers Deep Learning to train the model, Librosa to classify the audio data. In deep learning CNN is used to classify the model based on frequency parameter. This paper also contains the study of various speech emotion recognition methods like, happy, sad, angry, disgust, surprise and fear. Baby Shamini P, P. Girish Sai Varma, P. A. Khan, D. Harshith Reddy (2023) [4] Recognizing the affective qualities of speech while ignoring its semantic content is the goal of Speech Emotion Recognition (SER). Automatically conducting this activity using programmed devices is still a work in progress, even though people can do it efficiently as part of voice communication. A. Arul Edwin Raj, K. B, Shajivan S, R. A (2023) [5] Using the CNN model and MFCC feature extractions, implementing a method in this work for determining the underlying emotion in voice data. Various databases' worth of information was combined to enhance the voice samples. 2-D log the voice signal's recovered mel-spectrograms (static, delta, delta, and delta) were fed into CNN as input. segment-level characteristics are stacked that CNN extracted, the utterance-level characteristics were acquired. The CNN model was then utilized in the final Speech Emotion Recognition (SER) system.

P. J. Thilag, S. Vasuki (2022) [6] The model is based on four widely used datasets in SER: SAVEE, RAVDESS, TESS, and CREMA-D, and has a validation accuracy of 67.58%. Ad-ditionally, this model was evaluated on an unknown dataset that included audio samples in the German language and achieved a testing accuracy of 71.28%. T. Roy, Marwala Tshilidzi, S. Chakraverty (2021) [7] A novel model is built using DL architecture to produce results that can show directions toward building more robust solutions for SER, and has produced approximately 10% cross-validation accuracy improvement over models trained on nonaugmented data. R. A. Khalil, E. Jones, M. I. Babar, T. Jan, M. H. Zafar, T. Alhussain (2019) [8] This paper presents an overview of Deep Learning techniques and discusses some recent literature where these methods are utilized for speech-based emotion recognition. The review covers databases used, emotions extracted, contributions made toward speech emotion recognition and limitations related to it. Eva Lieskovská, Maros Jakubec, R. Jarina, M. Chmulik (2021) [9] This paper provides a review of the recent development in SER and also examines the impact of various attention mechanisms on SER performance. Overall comparison of the system accuracies is performed on a widely used IEMOCAP benchmark database.

Apeksha Aggarwal, Akshat Srivastava, Ajay Agarwal, Nidhi Chahal, Dilbag Singh, A. Alnuaim, Aseel Alhadlaq, Heung-No Lee (2022) [10] This study tackles machine-based emotion recognition in speech using deep learning models. We propose a novel two-way feature extraction with super convergence, generating sets for PCA-integrated DNN and VGG-16 applied to mel-spectrogram images. Extensive experiments reveal RAVDESS dataset's superior accuracy over numeric features in DNN. Waqar Ahmed, Sana Riaz, Khunsa Iftikhar, Savas Konur (2023) [11] This project attempts to use deep learning and image classification method to recognize emotion and classify the emotion according to the speech signals using Inception Net v3 Model. Aman Verma, Raghav Agrawal, Priyanka Singh, N. Ansari (2012) [12] The main objective of this paper is to efficiently classify the emotional state of a person from speech signals using traditional machine learning and deep learning techniques and to present a comparative analysis. R. Sharan (2023) [13] This work proposes a method for speech emotion recognition using a combination of handcrafted and deep learning features. In particular, it studies the use of gammatone cepstral coefficients, which make use of gammatone filters which model the human auditory filters, and deep learning feature embeddings extracted from a pretrained network for audio analysis.

Jamsher Bhanbhro, Shahnawaz Talpur, Asif Aziz Memon [14] This study suggested hybrid deep learning models for accurately extracting crucial features and enhancing predictions with higher probabilities. Initially, the Mel spectrogram's temporal features are trained using a combination of stacked Convolutional Neural Networks (CNN) & Long-term short memory (LSTM). The said model performs well. For enhancing the speech, samples are initially preprocessed using data improvement and dataset balancing techniques. The RAVDNESS dataset is used in this study which contains 1440 samples of audio in North American English accent. The strength of the CNN algorithm is used for obtaining spatial features and sequence encoding conversion, which generates accuracy above 93.9% for the model on mentioned data set when classifying emotions into one of eight categories. The model is generalized using Additive white Gaussian noise (AWGN) and Dropout techniques. Samiul Islam, Md. Maksudul Haque, Abu Md. Sadat (2023) [15] This research addresses these shortcomings by proposing an ensemble model that combines Graph Convolutional Networks (GCN) for processing textual data and the HuBERT transformer for analyzing audio signals. Amara Zahoor, Er Tarunjot Kaur (2023) [16] A large dataset of emotional speech samples, covering various emotions like Angry, disgusted, afraid, joyful, indifferent, depressed, surprised, is preprocessed and analyzed. A novel CNN architecture, optimized for SER, is proposed, incorporating multiple a maximum pooling sequential normalization, and various layers of complexity. By extracting relevant features from speech representations, distinct emotional patterns are discerned effectively. Numerous experiments were

carried out to validate the CNN model's effectiveness. The dataset underwent division into distinct training, validation, and testing phases. N. Alagusundari & R. Anuradha (2020) [17] In this work, we have developed deep learning models such as CNN, DANN, and TCN to recognize emotional states from speech signals. Each model is trained with different datasets with different feature extraction techniques such as MFCC, etc., to recognize various emotions. The emotional states of a person can be classified based on factors like pitch, tone, intensity, and dimensions of emotion such as arousal and valence. We have used four different datasets for training and evaluating the model. This work used CNN, GRU, DANN, and TCN with various feature extraction techniques, among that TCN performs better in large datasets (MFCC 58 features) with 93.66% accuracy and with eight emotion classes (Angry, Calm, Disgust, Fear, Happy, Neutral, Sad, Surprise). Chu Sheng Liau and Kai Sze Hong (2023) [18] This project focused on researching and evaluating the deep neural network performance on multi-lingual speech emotion recognition on RAVDESS, EMO-DB and combination of both emotional speech databases. Methodology used in the project was divided into five steps: data collection and speech signal extraction, signal conversion, image recognition using transfer learning, result validation and implementation of trained network in graphical user interface (GUI).

Rathnakar Achary, Manthan S. Naik & Tirth K. Pancholi (2022) [19] In this project,proposed an automated system for Speech emotion recognition using convolution neural network (CNN). The system uses a 5layer CNN model, which is trained and tested on over 7000 speech samples. The data used is .wav files of speech samples. Data required for the anlysis is gathered from RAVDESS dataset which consists of samples of speech and songs from both male and female actors. The different models of CNN were trained and tested on RAVDESS dataset until we got the required accuracy. The algorithm then classifies the given input audio file of .wav format into a range of emotions. The performance is evaluated by the accuracy of the code and also the validation accuracy. Tae-Wan Kim, Keun-Chang Kwak (2024) [20] designed a generalized model using three different datasets, and each speech was converted into a spectrogram image through STFT preprocessing. The spectrogram was divided into the time domain with overlapping to match the input size of the model. Each divided section is expressed as a Gaussian distribution, and the quality of the data is investigated by the correlation coefficient between distributions. As a result, the scale of the data is reduced, and uncertainty is minimized. VGGish and YAMNet are the most representative pretrained deep learning networks frequently used in conjunction with speech processing.

Raoudha Yahia Cherif, A. Moussaoui, Nabila. Frahta, Mohamed Berrimi (2021) [21] focus on speech emotion recognition in the Algerian dialect. The paper introduces a new large Algerian speech emotion dataset collected from different Algerian TV shows. The audio signals were manually annotated as happy, angry, neutral, or sad. After the data collection, applied several classification methods such as machine learning-based models, convolutional neural networks (CNNs), Long Short-Term Memory (LSTM) networks, and Bidirectional LSTM (BLSTM). After applying the proposed models on our dataset, the best classification accuracy is 93.34% achieved by the LSTM-CNN model.

3. Proposed Work

Speech-emotion recognition involves several key stages, including data acquisition, data preprocessing, feature extraction, feature selection, model training, and model testing. Data preprocessing is a crucial step in constructing a Machine Learning model. However, when utilizing a CNN model, research indicates that the preprocessing stage may be unnecessary, as CNN models have shown effectiveness without it.

In the context of Speech-Emotion Recognition (SER), the initial stage entails creating a data frame for audio data, involving labelling and extracting features. This work introduces an additional step called

data augmentation to enhance the number of features. Subsequently, the features derived from these two sets are inputted into the CNN model, which predicts each emotion with an accuracy of 99.9%.

3.1 Dataset Generation

This study focuses on effective automatic speech emotion recognition in Malayalam human-computer interaction. As Malayalam database for emotion recognition is not available, a dataset was created for this purpose. The name of the dataset is 'Malayalam data_set '. The Malayalam data_set, is an enhanced dataset. This dataset consists of 443 audio files as wav files and malayalam_dataset_copy has been created that contains ten copies of each audio file in the Malayalam dataset. This dataset consists of 4422 audio files (WAV files).The dataset contains seven emotions.For example, each emotion includes 68 WAV files: anger, happiness, sadness, disgust, surprise, fear, etc.

The audio files are recorded from some of the most famous Malayalam movies.Audacity is a well-known software program for recording and editing audio.Audacity is used to record audio files from YouTube.Each audio file is recorded at a sampling frequency of 44100 Hz and exported as a WAV file.All WAV files are divided into different folders according to their emotions.Since the audio of the video is recorded, there is a problem that the recorded file includes BGM.Background music can have a significant impact on feature extraction.To remove such background music from audio files, there is Vocal Remover.ai that you can use to remove this background music. This dataset contains conversations from both male and female artists, as well as child artists, so it is speaker independent.

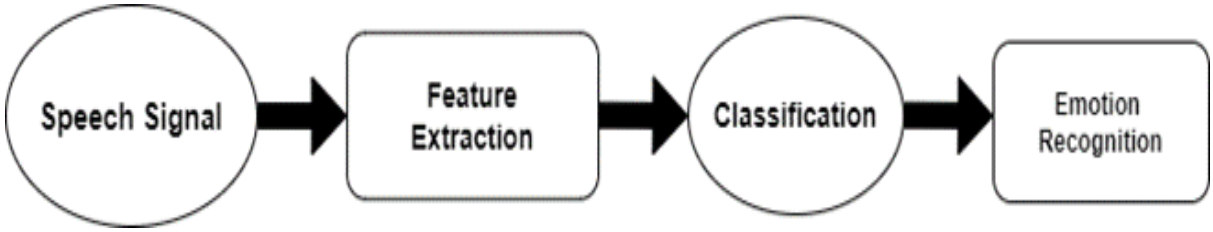


Fig.1 Stages of Working of SER

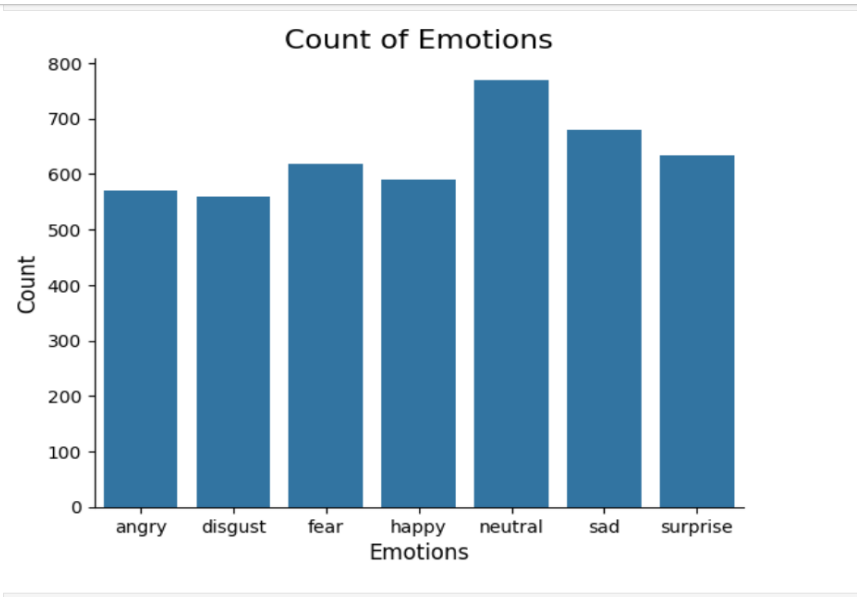


Fig. 2 Histogram plot showing counts of emotions in Malayalam dataset

3.2 Data Pre-processing

The dataset has been established, and a data frame containing emotions and their corresponding neural pathways has been created. To train our model, we will utilize the characteristics extracted from this data frame. The dataset requires organization into separate files for each audio segment based on emotional content, followed by encoding emotions into numerical values. These processed datasets are then saved to a specified file location.

For visualization, wave plots and spectrograms are provided for one audio file representing each emotion. Wave plots offer insights into audio loudness at specific times, while spectrograms visually represent the frequency spectrum of sound or signals over time. This graphical representation illustrates how frequencies change over time for a given audio signal or music. All files maintain a sampling rate of 44.1 kHz, implemented through the 'sr = 44100KHz' parameter in the Librosa library's load function. The deep Convolutional Neural Network (CNN) method doesn't necessitate extensive data preprocessing. The model is trained using raw audio data by directly inputting it into the neural network.

	Emotions	audio_name	Path
0	angry	angry_aara ningalkathin sammatham nalgiyath .wav	speech emotion recognition\Malayalam data_set/...
1	angry	angry_aaredaa ninnode aduth irikkan paranjath.wav	speech emotion recognition\Malayalam data_set/...
2	angry	angry_aarulyaa ivda.wav	speech emotion recognition\Malayalam data_set/...
3	angry	angry_adukkalel vera paniyonnulya chettathikk.wav	speech emotion recognition\Malayalam data_set/...
4	angry	angry_allik aabharanam edukkaan njaan kooda po...	speech emotion recognition\Malayalam data_set/...

Fig.3 Data frames for emotions in Malayalam dataset.

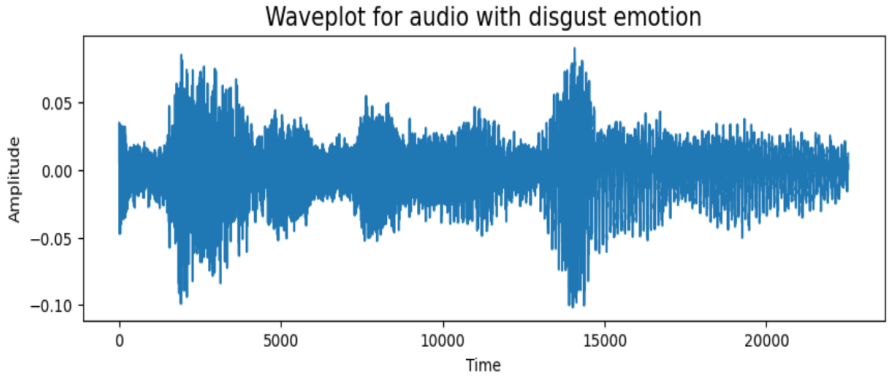
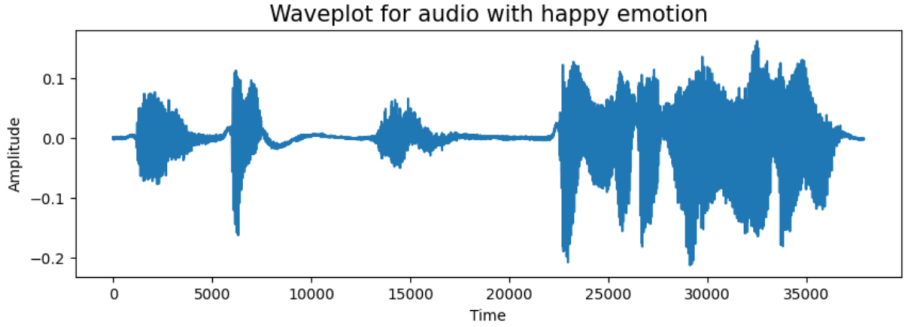


Fig.4. Waveplot for audio with fear emotion.



3.3 Spectrogram

A spectrogram serves as an effective visual representation of a signal's frequency distribution over time, widely applied in signal processing, audio analysis, and speech recognition for studying the evolving spectral attributes of time-dependent signals. In this representation, the x-axis corresponds to time, the y-axis to frequency, and the color or intensity at each point on the plot reflects the amplitude or strength of frequencies at that specific time and frequency bin.

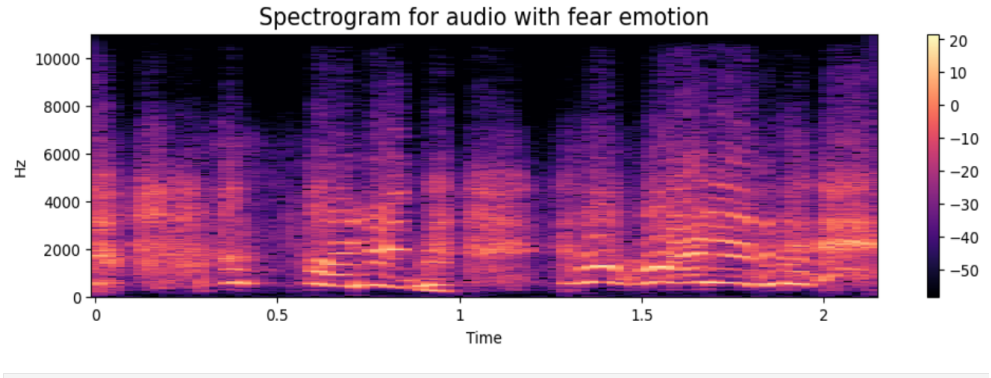


Fig. 6 Spectrogram for audio with fear emotion

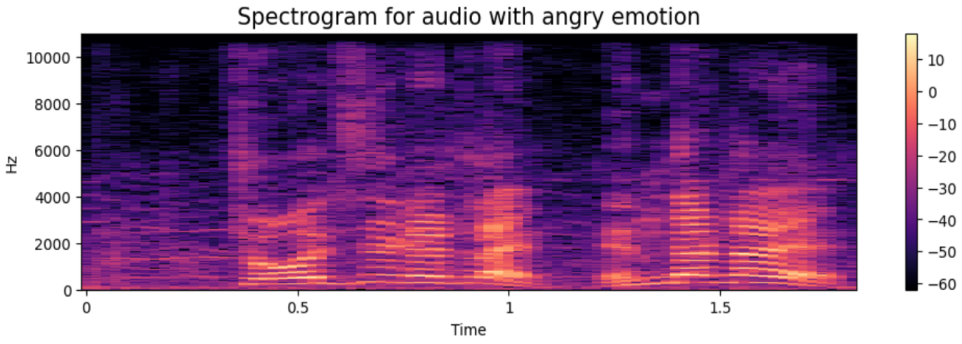


Fig. 7 Spectrogram for audio with angry emotion

3.4 Feature Extraction

The speaker's emotional expression in speech is reflected through a multitude of parameters, and alterations in these parameters correspondingly induce changes in emotions. Hence, the extraction of speech features that signify emotions is a crucial element in the design of a speech emotion recognition system. The primary focus in speech emotion recognition revolves around two categories of features: prosodic features and spectral features. Prosodic features encompass variations in pitch, rhythm, and stress patterns within speech, conveying intonation and emotional nuances. Prosody plays a pivotal role in communicating the speaker's attitude, emphasis, and intent, contributing to natural and expressive communication. Key prosodic features include Pitch, Intensity, Duration, and Speech Rate, collectively recognized as primary indicators of the speaker's emotional states.

Spectral features, on the other hand, pertain to characteristics of the speech signal in the frequency domain. These features offer insights into the energy distribution across different frequencies, proving valuable in tasks such as speech recognition, speaker identification, and speech analysis. Common spectral features include Mel-frequency cepstral coefficients (MFCC), Spectral Centroid, Spectral Flux, Zero Crossing Rate (ZCR), Chroma STFT, Root Mean Square Value (RMS), and Spectral Roll-off.

Research on the emotional aspects of speech underscores the significance of pitch, energy, duration, formant, Mel frequency cepstrum coefficient (MFCC), Changes in speech rate, pitch, energy, and spectrum are associated with different emotional states. Combining both spectral and prosodic features enhances the understanding of emotional content in speech. For the specific project at hand, feature extraction involved the use of Zero Crossing Rate, Chroma STFT, Root Mean Square Value, mel spectrogram, and MFCCs (Mel Frequency Cepstral Coefficients). The Librosa library was utilized for both feature extraction and feature selection in this project.

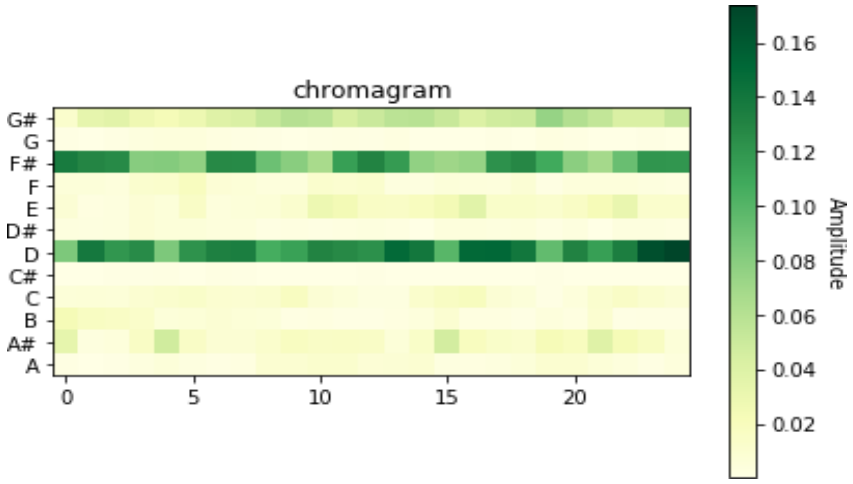


Fig. 8 Chromagram of a sample audio

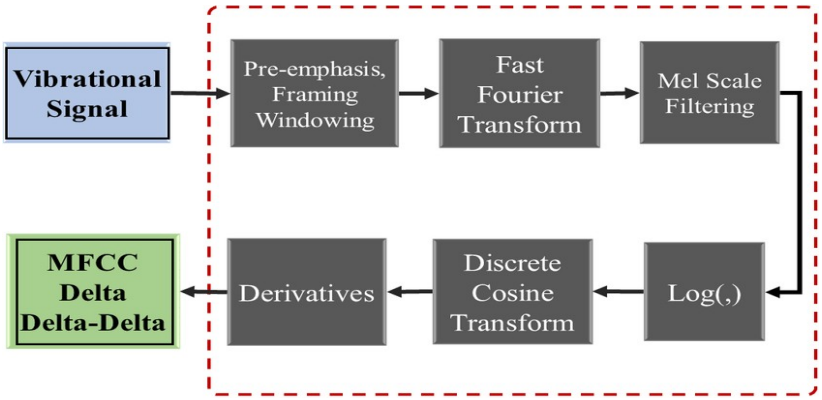


Fig 9. Schema of MFCC feature extraction

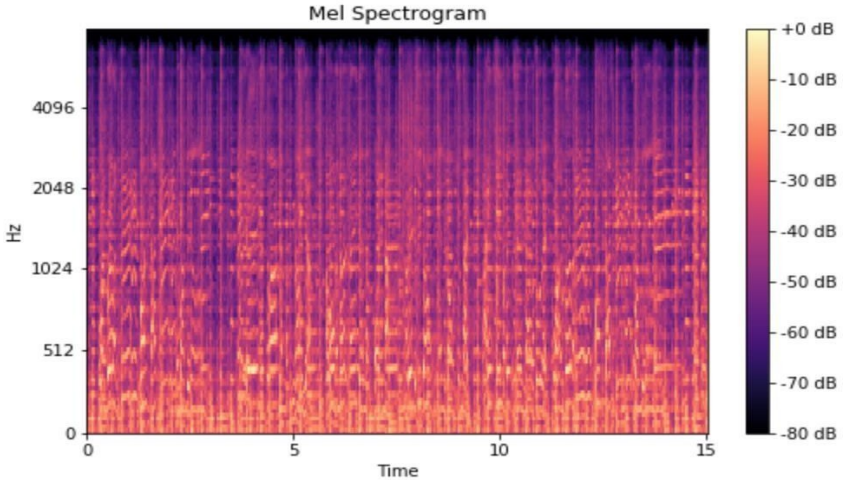


Fig. 10 Mel Spectrogram of a sample audio file

After the process of feature extraction, the subsequent step is feature selection. Simply inputting all the extracted features into the classifier does not guarantee optimal system performance, highlighting the necessity of eliminating unnecessary features from the base set. Therefore, a systematic feature selection process becomes crucial to reduce the number of features. One effective method for this is Forward Selection (FS), which starts with a single best feature from the entire feature set. Research indicates the availability of various feature selection methods for this purpose.

3.5 Splitting Data

Train-test splitting is a method employed for assessing algorithm performance, involving the division of a dataset into two parts: one for training the model and the other for testing its performance. However, in this particular project, a more comprehensive approach is taken by splitting the data into three distinct datasets: the training set, validation set, and test set. The training set, encompasses diverse inputs to ensure the model is trained across various scenarios and can predict unseen data that may arise in the future. The validation set serves as a separate data subset used to assess the model's performance during training, distinct from the training set. The test set is utilized to evaluate the model after its training phase. This splitting process aims to estimate the model's performance on new, unseen data, ensuring a robust evaluation.

The shapes of the resulting datasets are as follows:

- Shape of X training data: (9949, 162)
- Shape of Y training data: (9949, 7)
- Shape of X test data: (3317, 162)
- Shape of Y test data: (3317, 7)

3.6 Methodology

The proposed system incorporates a Convolutional Neural Network (CNN) model implemented using the Keras library for speech emotion recognition. The CNN architecture comprises a total of four one-dimensional convolutional layers, each succeeded by a max-pooling layer with a pool size of 5. The convolutional layers have 256, 128, and 64 filters, coupled with a kernel size of 5 for each. The activation function for the first Dense layer is 'Relu,' and for the second layer, the softmax function is used. The model employs a one-dimensional CNN, leveraging the time-based structure of audio waves to guide the linear progression of 1D CNN kernels. The fundamental building blocks of the CNN include the convolutional layer, Maxpooling layer, and fully connected layer.

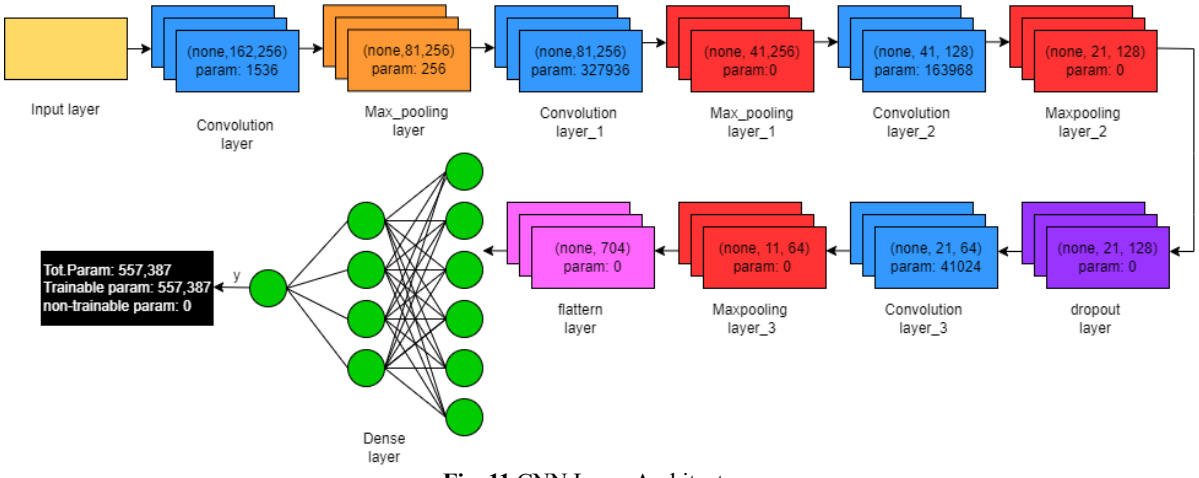


Fig. 11 CNN Layer Architecture

4 EXPERIMENTAL RESULT

Regarding the model architecture, we implemented four one-dimensional Convolutional layers with filter counts of 256, 256, 128, and 64, accompanied by respective kernel sizes of 5, 5, 5, 5. Additionally, two dense layers were incorporated into the model. Following the application of augmentation, we obtained features of size 1153x163 through feature extraction. The test data accuracy of our model, as depicted in Figure 13, was notably high, reaching approximately 99.9%.

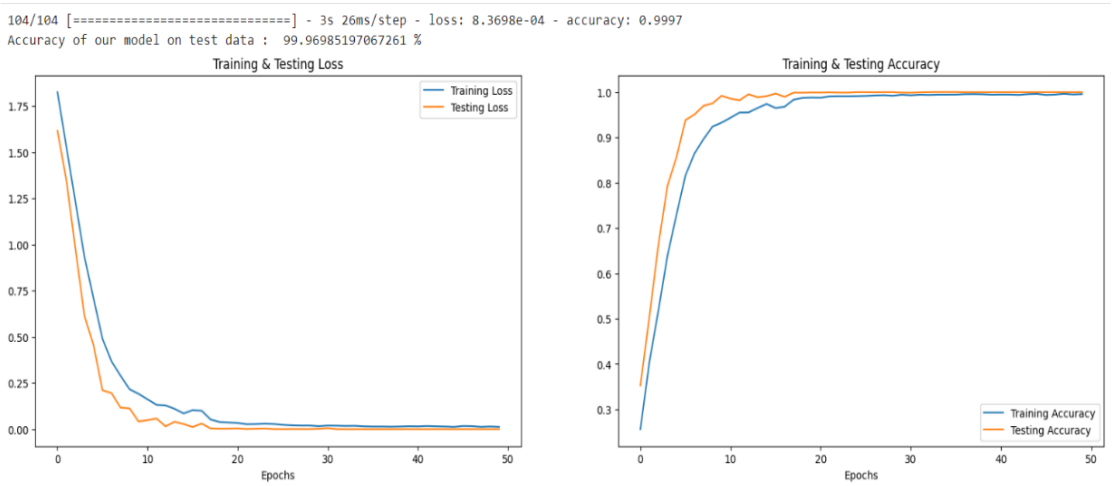


Fig. 12 Model Training & Testing Accuracy

The model is more accurate in predicting surprise, anger, neutrality, and sadness, and it makes sense also because audio files of these emotions differ from other audio files in a lot of ways like pitch, speed, etc. Model shows good performance when augmentation is applied.

4.1 Confusion matrix

The confusion matrix, also known as the error matrix, consists of a two-dimensional contingency table that defines the performance of a classification algorithm.

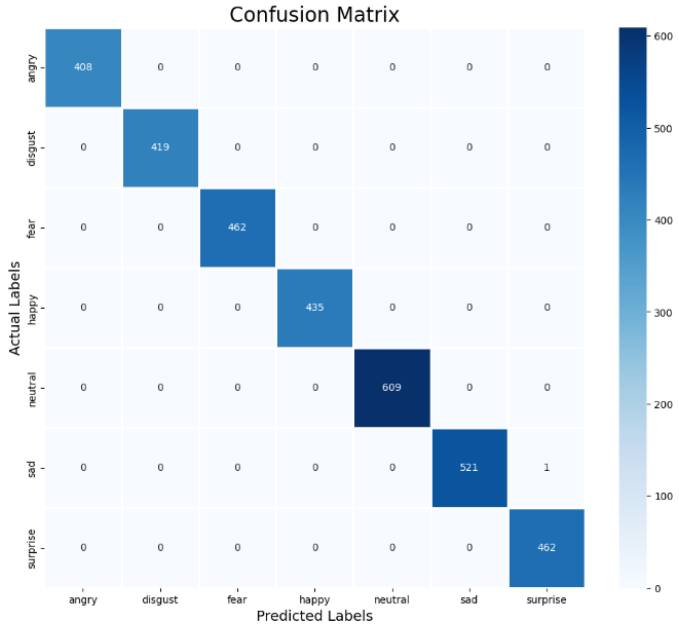


Fig. 13 The Confusion matrix

Although the confusion matrix itself is very easy to understand, the terminology used in relation to it can be difficult to understand (Trivedi and Sehrawat, 2018). The confusion matrix requires two parameters: the actual value or test data for comparison and the expected value from the model.

4.2 Performance Analysis

Results of evaluating the performance of classification models based on various metrics. This shows the precision, recall, F1 score, and support for each emotion class, as well as the overall model precision, macro average F1 score, and weighted average F1 score. An evaluation is if takes class imbalance into account. The evaluations provided indicate the performance of the classification model for each emotion class and provide an overall evaluation of the model's accuracy and effectiveness in recognizing emotions from audio data. The experiment is based on CNN model on Malayalam dataset, aiming to classify and recognize speaker's emotions through audio utterances, using various parameters such as accuracy, precision, recall, and F1 score metrics.

The results of the experiment are:

	precision	recall	f1-score	support
angry	1.00	1.00	1.00	408
disgust	1.00	1.00	1.00	419
fear	1.00	1.00	1.00	462
happy	1.00	1.00	1.00	435
neutral	1.00	1.00	1.00	609
sad	1.00	1.00	1.00	522
surprise	1.00	1.00	1.00	462
accuracy			1.00	3317
macro avg	1.00	1.00	1.00	3317
weighted avg	1.00	1.00	1.00	3317

Fig. 14 The Performance Evaluation

4.3 Live prediction

The LivePredictions class is a crucial component of the Speech Emotion Recognition (SER) system, designed to make real-time predictions on emotional states based on audio input. The system employs a pre-trained deep neural network model, stored in an h5 file, for accurate emotion classification. An instance of the LivePredictions class (`pred`) is created with the path to the model file and the file path of the audio to be analyzed. The model is loaded using `load_model()`, and predictions are obtained through `make_predictions()`. The results, including raw predictions, class probabilities, and the corresponding emotion label, are displayed.

```
1/1 [==============================] - 0s 149ms/step
Raw Predictions: [[0. 0. 1. 0. 0. 0. 0.]]
Predicted Class Probabilities: [0. 0. 1. 0. 0. 0. 0.]
Prediction is angry
```

Fig. 15 Result of Live Prediction

5 COMPARISON WITH EXISTING RESEARCH

This table serves as a valuable resource for comprehending the diverse approaches and contributions within the broader domain of speech emotion recognition research.

Table 1 Comparison with Existing Research

Reference	**Methodology**	**Language**	**Result Evaluation**	**Accuracy**	**Criteria**
Prof. Jagdish Kambale, Abhijeet Khedkar, Prasad Patil, Tejas Sonone(2023) [1]	Deep Learning	Not Specified	Not Specified	Not Specified	Not Specified
Apoorva Sharma,Himanshu Nawani,Shalini verma(2023) [2]	LSTM, CNN, Hybrid Model	Not Specified	RAVDES DATASET	Not Specified	MFCC Features
Shashidhar Harikant, Rakshitha Prasad, Vijaya Lakshmi R, Sidhramappa H (2022) [3]	DEEP LEARNING, CNN	Not specified	Not specified	Not specified	Frequency Parameter
Baby Shamini P, P. Girish Sai Varma, P. A. Khan, D. Harshith Reddy (2023) [4]	CNN, DEEP Learning	Not Specified	Not Specified	87.8%	Not Specified
A. Arul Edwin Raj, K. B, Shajivan S, R. A (2023) [5]	CNN methodology	Not Specified	Not Specified	Not Specified	2D log mel sctrograms, MFCC
P. J. Thilag, S. Vasuki (2022) [6]	Based on 4 Datasets: SAVEE, RAVDESS, TESS,CREMAD	German Language	Validation Accuracy	67.58%	Not Specified
T. Roy, Marwala Tshilidzi, S. Chakraverty (2021) [7]	DL Architecture	Not Specified	Cross validation accuracy improvement	10%	Not Specified
R. A. Khalil, E. Jones, M. I. Babar, T. Jan, M. H. Zafar, T. Alhussain (2019) [8]	Overview on DEEP Learning Techniques	Not Specified	Database used, Emotions Extracted	Not specified	Not specified
Apeksha Aggarwal, Akshat Srivastava, Ajay Agarwal, Nidhi Chahal, Dilbag Singh, A. Alnuaim, Aseel Alhadlaq, Heung-No Lee (2022) [10]	Deep Learning Models, PCA integrated DNN, VGG-16	Not Specified	RAVDESS DATASET	Not Specified	Mel Spectograms images
Waqar Ahmed, Sana Riaz, Khunsa Iftikhar, Savas Konur (2023) [11]	DEEP LEARNING, Inception NET V3 Model	Not Specified	Image Classification	Not Specified	Speech signals

Aman Verma, Raghav Agrawal, Priyanka Singh, N. Ansari (2012) [12]	Traditional Machine Learning, DEEP Learning Techniques	Not specified	Comparative analysis	Not Specified	Not Specified
Aman Verma, Raghav Agrawal, Priyanka Singh, N. Ansari (2012) [12]	Compination of Handcrafted and DEEP learning Features	Not Specified	Gammatione cepstral Coefficients	Not Specified	Audio Analysis
Jamsher Bhanbhro, Shahnawaz Talpur, Asif Aziz Memon [14]	Stacked CNN and LSTM, CNN Algorithm, Dropout Techniques	North America English	Emotion Classification	Above 93.9%	Not Specified
Samiul Islam, Md. Maksudul Haque, Abu Md. Sadat (2023) [15]	Ensemble model, Graph Convolutional Networks, HuBERT transformer	Not specified	Textual data and audio signals	Not specified	Not specified
Amara Zahoor, Er Tarunjot Kaur (2023) [16]	CNN architecture, optimized for SER	Not specified	Various experiments	Not Specified	Speech representations
N.Alagusundari & R. Anuradha (2020) [17]	CNN, DANN, TCN	Not specified	Various datasets, feature extraction techniques	93.66%	Pitch, tone, intensity, arousal, valence
Chu Sheng Liau and Kai Sze Hong (2023) [18]	Deep neural network performance on multi-lingual SER		Not Specified	Not Specified	Not Specified
Rathnakar Achary, Manthan S. Naik & Tirth K. Pancholi (2022) [19]	CNN model, trained and tested on RAVDESS dataset	Not specified	Trained and tested on over 7000 speech samples	Not specified	Classify emotions based on audio file
Tae-Wan Kim, Keun-Chang Kwak (2024) [20]	CNN, LSTM	Not Specified	STFT preprocessing, VGGish ,YAMNet	Not Specified	Spectrogram image, Gaussian distribution
Raoudha Yahia Cherif, A. Moussaoui, Nabila. Frahta, Mohamed Berrimi (2021) [21]	Machine learning models, CNNs, LSTM, BLSTM	Algerian dialect	Not Specified	93.34%	Not Specified
Rahbar Zahid S, Jalala PV, Preethi m[22]	CNN	Malayalam	F1 Score: 1, Precision: 1, Recall:1	Using Original dataset(443): 71.1%, and Copied dataset(4422): 99.9%	Pitch, Strech, Noise, Intensity

6 CONCLUSION

The evaluation of our speech emotion recognition project, which employs a Convolutional Neural Network (CNN) on a Malayalam dataset, reveals promising results in terms of accuracy, precision, recall, and F1 score metrics. The study focuses on recognizing speaker emotions through audio utterances and incorporates various parameters to comprehensively assess model performance. the CNN model, coupled with data augmentation, demonstrates exceptional accuracy and effectiveness in recognizing emotions from Malayalam audio utterances. These findings pave the way for further advancements in speech emotion recognition systems, particularly in diverse linguistic and cultural contexts like Malayalam. The success of this project holds promise for practical applications in fields such as human-computer interaction, sentiment analysis, and emotional well-being monitoring.

References

[1] H. Huang, Z. Hu, W. Wang, and M. Wu, "Multimodal Emotion Recognition Based on Ensemble Convolutional Neural Network," IEEE Access, vol. 8, pp. 3265–3271, Jan. 2020, doi: 10.1109/access.2019.2962085.

[2] B. A. Qayyum, A. Arefeen, and C. Shahnaz, Convolutional Neural Network (CNN) Based Speech-Emotion Recognition. 2019. doi: 10.1109/spicscon48833.2019.9065172.

[3] L. Kerkeni, Y. Serrestou, M. Mbarki, K. Raoof, M. A. Mahjoub, and C. Cléder, "Automatic Speech Emotion Recognition Using Machine Learning," in IntechOpen eBooks, 2020. doi: 10.5772/intechopen.84856.

[4] S. Langari, H. Marvi, and M. Zahedi, "Efficient speech emotion recognition using modified feature extraction," Informatics in Medicine Unlocked, vol. 20, p. 100424, Jan. 2020, doi: 10.1016/j.imu.2020.100424.

[5] Shahin, O. A. Alomari, A. B. Nassif, I. Afyouni, I. A. Hashem, and A. Elnagar, "An efficient feature selection method for arabic and english speech emotion recognition using Grey Wolf Optimizer," Applied Acoustics, vol. 205, p. 109279, Mar. 2023, doi: 10.1016/j.apacoust.2023.109279.

[6] H. Aouani and Y. B. Ayed, "Speech Emotion Recognition with deep learning," Procedia Computer Science, vol. 176, pp. 251–260, Jan. 2020, doi: 10.1016/j.procs.2020.08.027.

[7] J. Wang, M. Xue, R. Culhane, E. Diao, J. Ding, and V. Tarokh, Speech Emotion Recognition with Dual-Sequence LSTM Architecture. 2020. doi: 10.1109/icassp40776.2020.9054629.

[8] B. T. Atmaja and M. Akagi, Speech Emotion Recognition Based on Speech Segment Using LSTM with Attention Model. 2019. doi: 10.1109/icsigsys.2019.8811080.

[9] J. Zhao, X. Mao, and L. Chen, "Speech emotion recognition using deep 1D & 2D CNN LSTM networks," Biomedical Signal Processing and Control, vol. 47, pp. 312–323, Jan. 2019, doi: 10.1016/j.bspc.2018.08.035.

[10] H. A. Abdulmohsin, H. B. A. Wahab, and A. M. J. A. Hossen, "A new proposed statistical feature extraction method in speech emotion recognition," Computers & Electrical Engineering, vol. 93, p. 107172, Jul. 2021, doi: 10.1016/j.compeleceng.2021.107172.

[11] M. M. H. E. Ayadi, M. S. Kamel, and F. Karray, "Survey on speech emotion recognition: Features, classification schemes, and databases," Pattern Recognition, vol. 44, no. 3, pp. 572–587, Mar. 2011, doi: 10.1016/j.patcog.2010.09.020.

[12] D. Bertero and P. Fung, A first look into a Convolutional Neural Network for speech emotion detection. 2017. doi: 10.1109/icassp.2017.7953131. 42

[13] K. S. Lee and H. J. Kim, Design of a Convolutional Neural Network for Speech Emotion Recognition. 2020. doi: 10.1109/ictc49870.2020.9289227.

[14] S. Pandey, H. S. Shekhawat, and S. R. M. Prasanna, Deep Learning Techniques for Speech Emotion Recognition: A Review. 2019. doi: 10.1109/radioelek.2019.8733432.

[15] S. Sunny, D. P. S, and K. P. Jacob, "Discrete Wavelet Transforms and Artificial Neural Networks for Recognition of Isolated Spoken Words," International Journal of Computer Applications, vol. 38, no. 9, pp. 9–13, Jan. 2012, doi: 10.5120/4634-6871.

[16] A. Abdelhamid et al., "Robust Speech Emotion Recognition Using CNN+LSTM Based on Stochastic Fractal Search Optimization Algorithm," IEEE Access, vol. 10, pp. 49265–49284, Jan. 2022, doi: 10.1109/access.2022.3172954.

[17] Koduru, H. B. Valiveti, and A. K. Budati, "Feature extraction algorithms to improve the speech emotion recognition rate," International Journal of Speech Technology, vol. 23, no. 1, pp. 45–55, Jan. 2020, doi: 10.1007/s10772-020-09672-4.

[18] R. A. Khalil, E. G. Jones, M. I. Babar, T. Jan, M. A. Zafar, and T. Alhussain, "Speech Emotion Recognition Using Deep Learning Techniques: A Review," IEEE Access, vol. 7, pp. 117327–117345, Jan. 2019, doi: 10.1109/access.2019.2936124.

[19] M. Chen, X. He, J. Yang, and H. Zhang, "3-D Convolutional Recurrent Neural Networks With Attention Model for Speech Emotion Recognition," IEEE Signal Processing Letters, vol. 25, no. 10, pp. 1440–1444, Oct. 2018, doi: 10.1109/lsp.2018.2860246.

[20] M. Akçay and K. K. Oguz, "Speech emotion recognition: Emotional models, databases, features, preprocessing methods, supporting modalities, and classifiers," Speech Communication, vol. 116, pp. 56–76, Jan. 2020, doi: 10.1016/j.specom.2019.12.001.

[21] Dr.Issa, M.F.Demirci, and A. Yazici, "Speech emotion recognition with deep convolutional neural networks," Biomedical Signal Processing and Control, vol. 59, p. 101894, May 2020, doi: 10.1016/j.bspc.2020.101894.

[22] Mustaqeem, M. Sajjad, and S. Kwon, "Clustering-Based Speech Emotion Recognition by Incorporating Learned Features and Deep BiLSTM," IEEE Access, vol. 8, pp. 79861– 79875, Jan. 2020, doi: 10.1109/access.2020.2990405.

[23] M. Badshah, S. W. Baik, N. Rahim, and S. W. Baik, Speech Emotion Recognition from Spectrograms with Deep Convolutional Neural Network. 2017. doi: 10.1109/platcon.2017.7883728.

[24] S.N.Zisad, M.M.Hossain, and K.Andersson, "Speech Emotion Recognition in Neurological Disorders Using Convolutional Neural Network," in Lecture Notes in Computer Science, Springer Science+Business Media, 2020, pp. 287–296. doi: 10.1007/978-3-030-59277-6_26.

[25] S. Parthasarathy and I. Tashev, Convolutional Neural Network Techniques for Speech 43 Emotion Recognition. 2018. doi: 10.1109/iwaenc.2018.8521333.

[26] K. S. Lee, H.-K. Choi, B.-T. Jang, and H. J. Kim, A Study on Speech Emotion Recognition Using a Deep Neural Network. 2019. doi: 10.1109/ictc46691.2019.8939830.

[27] Dutt and P. Gader, "Wavelet Multiresolution Analysis Based Speech Emotion Recognition System Using 1D CNN LSTM Networks," IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 31, pp. 2043–2054, Jan. 2023, doi: 10.1109/taslp.2023.3277291.