

AEGISCARE: A MULTIMODAL AI ASSISTANT FOR INTELLIGENT ELDERLY SAFETY AND HOME MONITORING

Inara Basheer¹, Minisiya Mol.M², Niba Naufal³, Shamsiyya P⁴, Ms Sruthi A⁵

¹Inara Basheer, Bachelor of Technology in CSE, NCERC

²Minisiya Mol.M, Bachelor of Technology in CSE, NCERC

³Niba Naufal, Bachelor of Technology in CSE, NCERC

⁴Shamsiyya P, Bachelor of Technology in CSE, NCERC

⁵Sruthi A, Professor in CSE, NCERC

Abstract - AegisCare Elderly Assistance and Monitoring System is an advanced technological platform designed to enhance elderly care by shifting the focus from reactive supervision to proactive safety monitoring, addressing the challenges faced by elderly individuals in independent living environments. The system operates by continuously capturing real-time visual data through a camera, along with optional voice inputs, and processes this information locally to maintain privacy and reduce dependency on cloud-based solutions. The captured data is analyzed using advanced computer vision and Artificial Intelligence techniques to detect critical events such as falls, prolonged inactivity, and object presence in the environment. The system integrates multiple functional modules including fall detection, inactivity monitoring, object detection, and voice interaction to provide a comprehensive assistance framework. A key component of the system is the intelligent alert mechanism, which generates immediate notifications through sound alerts, voice messages, and email communication to caregivers or family members. By combining real-time monitoring, intelligent analysis, and user-friendly interaction, the system ensures improved safety, timely response, and enhanced quality of life for elderly individuals living independently.

Key Words: elderly care, computer vision, fall detection, monitoring system, AI

1.INTRODUCTION

The AegisCare Elderly Assistance and Monitoring System represents a significant advancement in modern healthcare technology, addressing the limitations of traditional elderly care methods, which are largely reactive and dependent on manual supervision. Conventional systems are effective only after an incident has occurred, but they are insufficient in handling real-time risks such as falls, prolonged inactivity, and delayed emergency response—issues that are increasingly common among elderly individuals living independently. These challenges are further intensified by the lack of continuous monitoring and timely intervention, which can lead to serious health consequences.

The core objective of AegisCare is to bridge the gap between periodic human supervision and the need for continuous, real-time monitoring of elderly individuals. This is achieved through an automated, intelligent, and integrated framework that combines computer vision, Artificial Intelligence (AI), and voice-based interaction to provide a comprehensive assistance system.

By continuously analyzing visual input and user interaction, the system is capable of detecting critical events such as falls, inactivity, and object-related assistance requirements, and responding immediately with appropriate alerts.

The system leverages advanced computational technologies including computer vision techniques for image processing, machine learning concepts for intelligent decision-making, and real-time data analysis to ensure accurate detection and response. The integration of multiple modules allows the system to function as both a monitoring and assistance tool, thereby improving safety, independence, and quality of life for elderly individuals. Through its intelligent design and real-time capabilities, AegisCare provides a reliable and scalable solution for modern elderly care challenges.

2.LITERATURE SURVEY

2.1 Domain Overview and Motivation

The field of intelligent elderly care systems has experienced significant growth in recent years, driven by the increasing need for continuous, efficient, and accessible healthcare solutions for aging populations. Traditional elderly care approaches largely depend on reactive measures, where assistance is provided only after an incident occurs, such as a fall or medical emergency. This model has proven to be inadequate in ensuring timely intervention and continuous monitoring, especially for individuals living alone. As a result, there is a growing demand for automated systems that can provide real-time monitoring and proactive assistance.

Recent research highlights the importance of integrating Artificial Intelligence (AI), computer vision, and smart sensing technologies to develop systems capable of detecting critical events and responding instantly. These systems aim to reduce dependency on manual supervision while improving safety and independence for elderly individuals. Consequently, modern elderly care solutions focus on creating intelligent frameworks that combine real-time monitoring, automated detection, and immediate alert mechanisms to enhance overall quality of life.

2.2 Data Sources and Multimodal Inputs

The effectiveness of intelligent elderly monitoring systems largely depends on the ability to collect and process continuous and reliable data from multiple sources.

Existing research demonstrates that modern systems utilize various input modalities, including camera-based monitoring, wearable sensors, and voice inputs, to capture comprehensive information about the user's condition and environment. Among these, vision-based systems have gained significant attention due to their non-intrusive nature and ability to monitor activities without requiring users to wear devices.

Camera-based inputs provide continuous visual data that can be used to analyze posture, movement, and object interactions in real time. Additionally, voice inputs enable user interaction, allowing individuals to communicate with the system and request assistance. By combining these multimodal inputs, the system can achieve a more accurate and dynamic understanding of the user's activities and surroundings. This integrated approach enhances the reliability of detection and improves the system's ability to respond effectively to various situations.

2.3 Data Processing and Analysis Techniques

Despite the availability of large volumes of input data, raw data collected from cameras and sensors is often noisy, unstructured, and unsuitable for direct analysis. Therefore, effective preprocessing and analysis techniques are essential to ensure accurate and reliable system performance. In vision-based systems, preprocessing involves converting images into suitable formats, reducing noise, and extracting relevant features such as edges, shapes, and motion patterns.

Advanced computer vision algorithms and machine learning techniques are then applied to analyze this processed data. These techniques enable the system to detect patterns such as sudden posture changes for fall detection, absence of movement for inactivity detection, and object recognition for assistance tasks. Additionally, real-time processing plays a crucial role in ensuring immediate response to critical situations. By combining preprocessing, feature extraction, and intelligent analysis, the system can achieve high accuracy and reliability in detecting events and generating appropriate alerts.

3. MODULE DESCRIPTION

Fall Detection Module –

This module detects falls using the YOLOv8 pose estimation model. The system captures live video and extracts human body keypoints such as head, shoulders, and legs. By analyzing sudden changes in posture and vertical-to-horizontal transitions, the system identifies fall events accurately. Upon detection, it triggers real-time alerts and records short video clips for evidence. Email notifications are also sent to caregivers to ensure immediate response.

Inactivity Detection Module –

This module detects prolonged inactivity using frame difference and motion analysis. It compares consecutive video frames to calculate motion levels. If no significant movement is detected for predefined durations (20 seconds and 30 seconds), the system generates warning and alert messages.

It also sends email notifications to caregivers when inactivity crosses critical thresholds.

Object Detection Module –

This module uses AI-based vision processing to detect objects in the environment. Frames captured from the camera are converted into Base64 format and sent to a Groq API (LLaMA model). The system receives object details such as name, position, size, and nearby objects, and provides contextual voice feedback to assist users in locating items.

Speech Interaction Module –

This module enables voice-based interaction between the user and the system. It captures user input (object queries) and processes them to guide the object detection module. This allows elderly users to interact with the system hands-free.

Text-to-Speech Module –

This module converts system outputs into voice responses. It provides real-time spoken feedback such as object location, alerts, and warnings, improving accessibility for elderly users.

Alert and Notification Module –

This module is responsible for generating alerts in emergency situations. It provides:

- Voice alerts
- Visual alerts on screen
- Email notifications to caregivers

Email alerts are triggered for both fall detection and prolonged inactivity, ensuring remote monitoring and timely intervention.

Vision Processing Module –

This module handles video acquisition and preprocessing. It captures frames from the camera, applies transformations such as grayscale conversion and noise reduction, and prepares data for motion detection, fall detection, and object recognition.

Web Integration Module (Django Framework) –

This module manages the web interface and system control. It handles:

- User sessions
- Video streaming (real-time camera feed)
- API communication
- Interaction between frontend and backend

It ensures smooth operation and real-time communication between all modules.

4. TECHNOLOGIES USED

Computer Vision (OpenCV) – Used for capturing and processing real-time video input from the camera.

YOLOv8 Pose Estimation – Used for detecting human posture and identifying fall events by analyzing body keypoints.

Groq API (LLaMA Model) – Used for AI-based object detection and understanding objects from camera input.

Django Framework – Used for backend development, handling system logic, APIs, and user management.

REST APIs – Enable communication between frontend and backend components.

Email Notification System – Sends alerts to caregivers during emergency situations like falls and inactivity.

Text-to-Speech (TTS) – Converts system outputs into voice alerts for better accessibility.

Speech Recognition – Allows voice-based interaction for user commands.

Database Management System (SQLite) – Stores user data, activity logs, and alert records efficiently.

4.1 SOFTWARE REQUIREMENTS

Operating System – Windows, Linux, or macOS for system development and execution.

Programming Language – Python for backend development, AI integration, and system logic implementation.

Framework – Django for developing the backend, handling APIs, and managing user interactions.

Computer Vision Library – OpenCV for real-time video capture and image processing.

AI Model – YOLOv8 (Ultralytics) for pose estimation and fall detection.

API Integration – Groq API (LLaMA model) for AI-based object detection and analysis.

Speech Processing Libraries – SpeechRecognition for voice input and pyttsx3 for text-to-speech output.

Database Management System – SQLite for storing user data, alerts, and activity logs.

API Communication – Requests library for handling communication with external services.

Development Tools – Visual Studio Code or PyCharm for coding and debugging.

Web Browser – Google Chrome or Microsoft Edge for running the web interface.

4.2 HARDWARE REQUIREMENTS

RAM – Minimum 4 GB (8 GB recommended) Required for smooth execution of real-time video processing and backend operations.

Storage – Minimum 20 GB free space Used for storing project files, recorded video clips, database, and system logs.

Camera – for capturing real-time video input for fall detection, inactivity monitoring, and object detection.

Microphone – Used for capturing voice input for speech interaction and user commands.

Speakers – Required for voice alerts and system responses using text-to-speech.

Internet – Needed for API communication (Groq API) and sending email notifications to caregivers.

5.SYSTEM DESIGN



Fig 5.1 Home Page

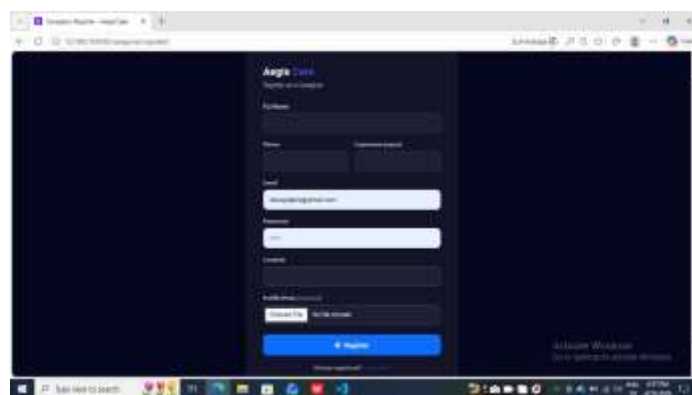


Fig 5.2 Caregiver Registration

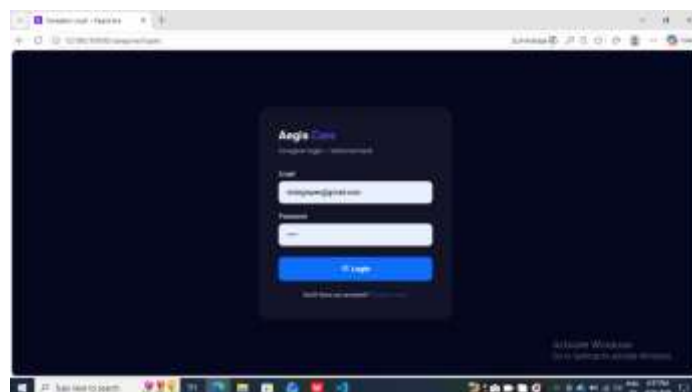


Fig 5.3 Caregiver Login



Fig 6.1 Fall Detection

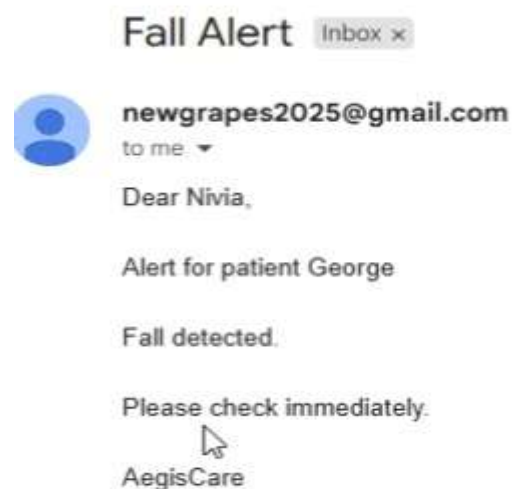
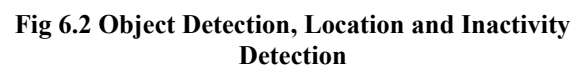
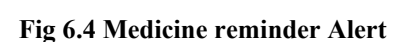


Fig 6.3 Fall Alert notification through gmail



7. DISCUSSION AND ANALYSIS

7.1 ADVANTAGES

Real-Time Fall Detection

The system uses YOLOv8 pose estimation to detect falls instantly by analyzing body posture and movement. This enables quick identification of emergency situations and ensures immediate response.

Inactivity Monitoring and Early Alerts

The system continuously monitors user movement and detects prolonged inactivity. Alerts are generated at different thresholds, helping caregivers respond early to potential health risks.

Integrated Multi-Function System

AegisCare combines multiple functionalities such as fall detection, inactivity monitoring, object detection, and assistance features into a single platform, reducing the need for multiple applications.

Non-Intrusive Monitoring

The system uses camera-based monitoring instead of wearable devices, eliminating discomfort and ensuring ease of use for elderly individuals.

Voice-Based Interaction

Speech recognition and text-to-speech features allow users to interact with the system easily and receive guidance through voice, improving accessibility.

AI-Based Object Detection Assistance

The system helps users locate misplaced objects using AI, enhancing convenience and supporting independent living.

Automated Alert System with Email Notifications

The system generates real-time alerts through voice, sound, and email notifications to caregivers, ensuring timely intervention during emergencies.

Privacy and Local Processing

Most processing is performed locally, reducing dependency on cloud services and improving data privacy and system performance.

7.2 FUTURE WORKS

The future scope of the AegisCare system focuses on enhancing its intelligence and expanding its capabilities to provide more advanced elderly care. One important improvement is the use of predictive AI models to analyze user behavior and detect potential health risks at an early stage, enabling proactive intervention.

The system can be further enhanced by integrating wearable devices and IoT sensors to monitor physiological parameters such as heart rate and oxygen levels. Combining this data with vision-based monitoring will improve accuracy and reliability. Future developments may also include advanced voice interaction using Natural Language Processing, mobile applications for remote monitoring, and cloud-based dashboards for caregivers.

Additionally, integration with healthcare systems and telemedicine services can enable direct communication with medical professionals.

Overall, these improvements will make AegisCare a more intelligent, scalable, and comprehensive solution for elderly monitoring and assistance.

8. CONCLUSION

The existing elderly care systems are often limited by fragmented functionalities and a reactive approach, where assistance is provided only after an incident occurs. This lack of continuous monitoring and integrated support can lead to delayed responses during critical situations such as falls or prolonged inactivity, ultimately affecting the safety and well-being of elderly individuals. Additionally, many systems rely on wearable devices or manual supervision, which may not always be practical or reliable.

The AegisCare system addresses these limitations by providing an integrated and intelligent solution for real-time monitoring and assistance. By combining technologies such as computer vision, YOLOv8-based fall detection, inactivity monitoring, and AI-based object detection, the system ensures continuous observation and timely response to emergencies. The inclusion of voice interaction and automated alert mechanisms, including email notifications, further enhances usability and caregiver support.

Overall, AegisCare demonstrates the effectiveness of integrating multiple functionalities into a single platform to improve elderly safety, independence, and quality of life.

The system provides a scalable and efficient solution that can be further extended to support advanced healthcare applications in the future.

9. ACKNOWLEDGEMENT

We would like to express our sincere gratitude to our Mentor, Ms. Sruthi A for her invaluable guidance, encouragement, and continuous support throughout the development of this project. Her expertise and insightful feedback played a crucial role in enhancing our understanding and improving the project's quality.

We are also thankful to Nehru College of Engineering and Research Centre, Thiruvilwamala, for providing the necessary resources and a conducive learning environment, enabling us to explore and apply our technical knowledge effectively.

Finally, we deeply appreciate the unwavering support and encouragement from our family and friends, whose motivation kept us inspired throughout this journey. This project has been a significant learning experience, and we are truly grateful to everyone who contributed to its successful completion.

REFERENCES

- [1] Rieke, N., et al. (2020). The future of digital health with federated learning. *NPJ Digital Medicine*, 3, 119.
- [2] Sittig, D. F., et al. (2020). Precision medicine, AI, and the future of personalized health care. *Yearbook of Medical Informatics*, 29(1), 10–18.
- [3] Cao, Z., Hidalgo, G., Simon, T., Wei, S. E., & Sheikh, Y. (2021). OpenPose: Realtime multi-person 2D pose estimation using part affinity fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
- [4] Jocher, G., et al. (2023). Ultralytics YOLOv8: State-of-the-art object detection and pose estimation. Available: <https://docs.ultralytics.com>
- [5] Wang, L., et al. (2023). Blockchain and IoT enabled privacy-preserving medical data sharing for smart healthcare. *Computers & Electrical Engineering*, 107, 108605.
- [6] Thaker, K., et al. (2024). Enhancing health recommendations through patient metadata integration: A persona-based evaluation approach. *HealthRecSys 2024*.
- [7] Ultralytics. (2024). YOLOv8 Pose Estimation Documentation. Available: <https://docs.ultralytics.com/tasks/pose/>
- [8] OpenCV. (2024). Open Source Computer Vision Library Documentation. Available: <https://opencv.org>