Seizure Detection in Dark Environments

PROJECT PHASE 2
ADD416

Project Phase 2 Report submitted in partial fulfillment for the degree
of
B.Tech in Artificial Intelligence and Data Science

Submitted by

AAKASH JASON KAVALAKKAT (Reg No: MUT21AD001)

AKASH R (Reg No: MUT21AD004)

BIJIL BOBY (Reg No: MUT21AD019)

BRAYAN DEEP (Reg No: MUT21AD021)

Muthoot Institute of Technology & Science

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

(Affiliated to APJ Abdul Kalam Technological University)

March 2025

DECLARATION

We hereby declare that this submission is our own work and that, to the best of our knowledge and belief, it contains no material previously or written by another person nor material which has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgment has been made in the text.

AAKASH JASON KAVALAKKAT (MUT21AD001)

AKASH R (MUT21AD004)

BIJIL BOBY (MUT21AD019)

BRAYAN DEEP (MUT21AD020)

Place:

Date:

Muthoot
Institute of Technology & Science

Varikoli P.O., Puthencruz-682308

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

CERTIFICATE

This is to certify that the Project Phase 2 report entitled **"Seizure Detection in Dark Environments"** submitted by **AAKASH JASON KAVALAKKAT (Reg No: MUT21AD001), AKASH R (Reg No: MUT21AD004), BIJIL BOBY (Reg No: MUT21AD019), BRAYAN DEEP (Reg No: MUT21AD021)** of Semester VIII is a bonafide account of the work done by them under our supervision during the academic year 2024-25.

Guide	**Coordinator**	**Head of the Department**
Parvathy E.N	Sheena Y	Dr. Sumam Mary Idicula
Asst. Professor	Asst. Professor	Professor
Dept. of AI&DS	Dept. of AI&DS	Dept. of AI&DS

ACKNOWLEDGEMENT

We are grateful to Almighty who has blessed us with good health, committed and continuous interest throughout the project work.

We express our sincere thanks to our project guide, **Ms. Parvathy E.N**, Assistant Professor, Department of Artificial Intelligence and Data Science, Muthoot Institute of Technology and Science for her guidance and support which were instrumental in all the stages of the project work and without whom the project could not have been accomplished.

We are grateful to our project coordinator **Ms. Sheena Y**, Assistant Professor, Department of Artificial Intelligence and Data Science, Muthoot Institute of Technology and Science, for her guidance and support.

We also wish to express our sincere appreciation to **Dr. Sumam Mary Idicula**, Professor & Head of the Department, Muthoot Institute of Technology and Science, who was willing to spend her precious time to give some ideas and suggestion towards this project.

We would like to thank **Dr. Neelakantan P.C.**, Principal, Muthoot Institute of Technology and Science, Varikoli for providing us all the necessary facilities.

We would like to express our sincere gratitude to all the teaching and non-teaching staff of Artificial Intelligence and Data Science Department for providing a good environment.

In our daily work we have been blessed with a friendly and cheerful group of fellow students. We are very much thankful to all the members of S8 AI&DS 2021-2025 batch.

Besides this, several people have knowingly and unknowingly helped us in the successful completion of this project. We express our sincere gratitude to all of them.

Abstract

Seizures occurring in hospitalized patients, particularly during nighttime, pose a significant risk if left undetected. This study presents a real-time Seizure Detection and Alert System designed to operate effectively in low-light hospital environments. The system processes video footage from surveillance cameras by enhancing visibility using Gamma Intensity Correction (GIC) and Contrast Limited Adaptive Histogram Equalization (CLAHE), followed by noise reduction with FastDVDnet to improve clarity and consistency across frames. For seizure detection, a deep learning-based skeleton tracking model (VSViG: Video-based Seizure Detection via Skeleton-based Spatiotemporal ViG) is used to extract patient pose information and identify seizure-related movements. The model generates a seizure probability score in real time, and if the probability exceeds a predefined threshold, an automated alert is sent to the nurse station for immediate intervention. By integrating computer vision techniques, deep learning-based action recognition, and real-time alerting, our system provides an effective, automated seizure monitoring solution to enhance patient safety and ensure timely medical response.

Contents

List of Figures

List of Abbrevations

GIC	Gamma Intensity Correction
CLAHE	Contrast Limited Adaptive Histogram Equalization
FastDVDnet	Fast Deep Video Denoising Network
VSViG	Video-based Seizure Detection via Skeleton-based Spatiotemporal Vision Graph
EEG	Electroencephalogram
EMG	Electromyography
VEM	Video EEG Monitoring

Chapter 1

Introduction

1.1 Motivation

Seizures are a critical medical condition that can occur at any time, often without warning. In hospitals, especially during nighttime or in low-light environments, it becomes challenging for healthcare staff to continuously monitor patients, increasing the risk of delayed response during a seizure. Immediate attention during a seizure can significantly reduce the chances of severe injury or complications. However, relying solely on human observation can lead to missed events, especially when patients are in isolated rooms or under dim lighting conditions. There is an urgent need for a reliable system that can monitor patients in real-time and instantly alert caregivers when a seizure occurs, ensuring timely medical assistance and improving patient safety.

In real-world scenarios, patients in intensive care units, nursing homes, or home-based care often require constant monitoring, especially those with epilepsy or similar conditions. Caregivers may struggle to detect sudden seizures in low-light conditions, potentially leading to fatal consequences. An automated system capable of detecting abnormal body movements and alerting caregivers could significantly reduce response time and prevent serious harm. Developing a system that works effectively in low-light conditions is crucial, as most seizures occur during sleep or in poorly lit environments. This project aims to bridge that gap by introducing a reliable, real-time monitoring solution that prioritizes patient safety.

1.2 Problem Statement

Patients suffering from epilepsy or other seizure disorders often experience seizures during sleep or in low-light environments, where continuous human monitoring becomes challenging. In hospitals, nursing homes, or home care settings, delayed detection of seizures can lead to severe injuries, complications, or even fatalities. The absence of real-time seizure detection, especially in dark environments, poses a significant risk to patient safety. Therefore, there is a critical need for an automated system that can monitor patients in real-time, detect seizure-like body movements in low-light conditions, and immediately alert caregivers, ensuring timely medical intervention and reducing potential harm to patients.

Chapter 2

Literature Review

2.1 DarkLight Networks for Action Recognition in the Dark

The paper "DarkLight Networks for Action Recognition in the Dark" by Rui Chen et al. addresses the challenge of human action recognition in low-light conditions, which is essential for applications such as night surveillance and autonomous driving. Traditional action recognition models trained on well-lit datasets struggle to perform effectively in dark environments due to the lack of sufficient training data and the inefficiency of conventional video enhancement techniques. To overcome this, the authors introduce DarkLight Networks, a novel neural network architecture that incorporates a dual-pathway structure. This framework processes both the original dark videos and their brightened counterparts using Gamma Intensity Correction (GIC), allowing the extraction of complementary features. A self-attention mechanism is then employed to fuse and refine the information from these two pathways, improving the model's ability to recognize actions in low-light conditions.

The DarkLight Networks achieve state-of-the-art results on the ARID dataset, significantly outperforming existing 3D convolutional neural networks (CNNs) and two-stream architectures. The experimental results demonstrate that the proposed method not only enhances action recognition accuracy in dark videos but also eliminates the need for optical flow estimation, which is often computationally expensive and unreliable in low-light scenarios. By leveraging the dual-pathway structure and self-attention mechanism, DarkLight Networks effectively capture spatio-temporal features, ensuring more robust recognition performance. The study highlights the importance of multi-modal data fusion and self-attention for improving video analysis in challenging visual conditions.

Overall, the paper contributes to the advancement of action recognition research by providing an effective solution for low-light video analysis. The proposed DarkLight Networks offer practical implications for real-world applications where lighting conditions are unpredictable. Future work could explore further optimizations and extensions to other challenging environments, such as extreme weather conditions or thermal imaging scenarios, to enhance the generalizability of the approach.

2.2 Low Light Image Enhancement via Structural Modeling and Guidance

The paper introduces a framework that enhances low-light images by simultaneously modeling both appearance and structure. The core innovation lies in its structure-aware feature extraction through a modified generative model, coupled with a *structure-guided enhancement module* that uses detected edges to guide image enhancement. The combination of structural and appearance modeling enables this method to achieve superior performance in terms of both PSNR and SSIM, making it highly effective for producing sharp, realistic images in low-light conditions. One significant aspect of the proposed framework is the use of Structure-Aware Feature Extractor (SAFE), which extracts structural details through adaptive spatial computations. This approach differs from conventional enhancement methods, which typically focus only on improving appearance, by explicitly considering edges and other structural information. Furthermore, the use of a Generative Adversarial Network (GAN) loss ensures that structural maps are robust, minimizing noise artifacts that often degrade image quality in dark environments.

The paper's experimental results demonstrate that this approach outperforms many state-of-the-art low-light enhancement models across various datasets, both in the sRGB and RAW domains. Its architecture also generalizes well to other tasks beyond low-light enhancement, showcasing the framework's potential for broader applications in image restoration and enhancement under challenging lighting conditions.

2.3 Smart Frame Selection for Action Recognition

The paper presents a method aimed at improving the computational efficiency of action recognition by selecting the most informative frames from video data. This approach reduces the computational cost of analyzing entire video sequences while also enhancing accuracy by discarding frames that are less relevant. The proposed method, known as SMART (Sampling through Multi-frame Attention and Relations in Time), leverages attention and relational networks to evaluate the importance of frames collectively, in contrast to traditional methods that treat frames individually. This approach is inspired by challenges in action recognition, particularly for short, trimmed videos where relevant frames often appear closely together. By jointly considering frames, the method improves the diversity and coverage of key action moments in the video. SMART is tested on multiple datasets, such as UCF101, HMDB51, and ActivityNet, and shows superior accuracy while reducing the computational load by up to 10 times compared to other frame selection strategies. This demonstrates its efficiency, even outperforming state-of-the-art methods like AdaFrame and FastForward on benchmarks.

SMART's innovation lies in combining a lightweight MobileNet-based visual feature extractor with a self-attention mechanism to score frame relevance, followed by relational and temporal attention to capture dependencies between frames. This efficient frame selection approach can also be used as a pre-processing step for more complex action recognition models, improving their performance while maintaining computational efficiency. The framework is versatile, demonstrating strong results across different video types, including both trimmed and untrimmed scenarios.

2.4 ALEN: A Dual-Approach for Uniform and Non-Uniform Low Light image enhancement

The paper introduces ALEN (Adaptive Light Enhancement Network), a novel deep learning framework designed to enhance images captured in low-light conditions. ALEN employs a dual-approach to handle both uniform and non-uniform lighting scenarios, which are common challenges in low-light environments. The system uses a Light Classification Network (LCNet) to classify an image into global or local illumination categories. Based on this classification, ALEN applies two distinct enhancement networks: SCNet (Single-Channel Network) for global illumination to boost luminance, and MCNet (Multi-Channel Network) for enhancing color fidelity by independently processing the RGB channels.

This dual strategy ensures that both brightness and color are effectively restored without introducing noise or artifacts. The model is trained on diverse datasets, enabling it to generalize well across various low-light conditions. ALEN's approach is shown to improve not only the visual quality of images but also the performance of high-level vision tasks such as object detection and semantic segmentation in low-light scenarios. However, the computational complexity of ALEN, involving three networks, poses challenges for real-time applications on resource-limited devices.

2.5 Unsupervised Face Detection in the Dark

The paper presents a novel approach to detect faces in low-light conditions, tackling challenges related to nighttime surveillance, autonomous driving, and other real-world applications where lighting is poor. Instead of relying on extensive low-light annotations, the paper introduces a High-Low Adaptation (HLA) framework that bridges the gap between normal-light and low-light image domains. The framework employs both pixel-level and feature-level adaptation, allowing the model to detect faces in extreme low-light settings by leveraging normal-light training data. This enables face detection models to perform well even in dark environments without requiring annotated low-light datasets.

Key techniques include **bidirectional low-level adaptation**, which enhances low-light images and degrades normal-light images to create intermediate domains that help the model adapt across lighting conditions. Additionally, **multitask self-supervised learning** is used to adapt the model's features from normal to low-light, eliminating the need for annotated low-light data. A **lightweight curve-based enhancement model** is also proposed to improve computational efficiency while maintaining strong detection performance in dark environments. However, the model still relies on large datasets of low-light images for training and faces limitations in extremely dark settings, with potential issues like false positives when distinguishing between bright objects and faces.

This framework aligns with challenges in night vision action recognition, as it addresses low visibility and pixel-level inconsistencies, making it valuable for broader applications beyond face detection. However, careful consideration of its limitations, such as false positives and difficulties in extreme darkness, is crucial for generalizing the model to action recognition tasks.

2.6 Seeing Motion in the Dark

The paper addresses the challenge of processing extreme low-light videos, specifically to enhance visual content from raw sensor data where traditional camera outputs would fail. It proposes a deep learning-based approach, using a Siamese network structure to process raw video data and achieve high-quality sRGB outputs with minimal noise, artifacts, and flickering. The system is designed to enhance both static and dynamic video sequences captured in extremely low-light environments, aiming to improve visibility and consistency in video frames.

Advantages of the paper include a Deep Learning Framework. The paper introduces a learning-based video processing pipeline that uses a deep Siamese network structure for low-light video enhancement. This is crucial for preserving color consistency and minimizing temporal artifacts (flickering). Raw Video Processing is done Rather than relying on JPEG outputs, the system operates on 14-bit raw sensor readings, maintaining more data and detail from the original frame, essential for low-light environments. Noise Reduction, that is Temporal and spatial denoising are crucial components, handled through preprocessing using VBM4D, a state-of-the-art method for video denoising. Siamese Network and Self-Consistency Loss, that is The network is trained using both recovery and self-consistency losses, ensuring that consecutive video frames maintain temporal consistency, a critical aspect for stable video output in dynamic sequences. Temporal Consistency is maintained. The method addresses temporal instability (e.g., flickering) which can be significant in dark environments. By incorporating self-consistency loss, the system produces more stable video outputs over time. ResUnet Structure is used. The network incorporates a ResUnet with 16 residual blocks, which allows for powerful feature extraction and reconstruction, making it suitable for the challenging task of low-light video restoration. Evaluation Metrics used here are Quantitative evaluation which is performed using standard image quality measures such as PSNR (Peak Signal to Noise Ratio) and SSIM (Structural Similarity Index), while perceptual quality is evaluated through user studies.

Limitations of the paper include failure in Extreme Low-Light Conditions. Although the method shows significant improvements in low-light video enhancement, it struggles with moon-light levels (0.01-0.03 lux). There is No Dynamic Range Preservation, Due to preprocessing steps, the system does not preserve high dynamic range (HDR), resulting in overexposed areas around strong light sources (e.g., candle-lit scenes). This may be a limitation for capturing accurate motion details in the night vision action recognition system. Always Single Frame Processing is done in Test Time, While the network generalizes well to dynamic videos, it only processes single frames at test time, which might limit the direct application of temporal information, a key element in action recognition. Problem of Heavy Computation is present. The proposed method requires significant computational resources, which could be a limitation in real-time action recognition systems in dark environments, particularly when embedded systems or low-power devices are involved.

2.7 Smart Image Enhancement Using CLAHE Based on an F-Shift Transformation during Decompression

The paper ”Smart Image Enhancement Using CLAHE Based on an F-Shift Transformation during Decompression” proposes a novel image enhancement technique that integrates Contrast-Limited Adaptive Histogram Equalization (CLAHE) with a two-dimensional F-shift (TDFS) transformation and a non-standard two-dimensional wavelet transform (NSTW) during decompression. The method first adjusts the global brightness of the image using the wavelet synopsis coefficient. CLAHE is then applied twice: first during decompression at the second-to-last level to enhance low-frequency components, and again after complete reconstruction to refine details. This approach preserves compression properties while improving contrast and detail visibility. Experimental results demonstrate that the proposed method outperforms existing techniques by significantly enhancing contrast, information entropy, and average gradient. The study suggests that this method can be particularly beneficial for images with poor quality, making it more effective for data mining and further processing.

2.8 Learning to See Moving Objects in the Dark

The paper introduces a novel system for capturing paired low-light and well-lit videos of the same scene. It focuses on enhancing low-light video frames using a fully convolutional network that combines 3D and 2D operations to improve temporal consistency. The system is designed to work with dynamic scenes, such as street views with moving vehicles and pedestrians, addressing the challenge of low signal-to-noise ratios (SNR) in extreme low-light conditions.

Advantages of the paper include a Dual-Camera System for Dataset Creation, a specialized optical system using two synchronized cameras—one with a neutral density (ND) filter to capture dark scenes and another for well-lit videos. This setup enabled the creation of the SMOID dataset (See Moving Objects in the Dark), comprising 179 video pairs. 3D U-Net is used for Video Enhancement, core innovation of the paper is a modified 3D U-Net architecture that uses a combination of 2D and 3D convolutions to enhance low-light videos. The 3D convolutions capture temporal dependencies across frames, while the 2D convolutions process spatial features. The inclusion of 3D convolutions allows the network to preserve temporal consistency, reducing flickering that often occurs when low-light enhancement is applied frame-by-frame. The pipeline also includes a subpixel depth-to-space layer, which restores the resolution of the output frames after processing. This step is crucial for preserving the spatial quality of the frames.

Limitations of the paper include High Complexity and Real-Time Limitations, The system's complexity, particularly with the inclusion of 3D convolutions and subpixel restoration, makes it computationally intensive. While the network produces high-quality results, it may not be optimal for real-time applications such as action recognition systems that require immediate feedback on movements. It suffers from limited dataset and generalization, Although the SMOID dataset is a significant contribution, it is limited in scope primarily covering street scenes with moving vehicles and pedestrians. This limited variety might affect the network's ability to generalize to other low-light scenarios, such as indoor environments or different types of actions.

2.9 Low-light aware framework for human activity recognition via optimized dual stream parallel network

The paper focuses on enhancing Human Activity Recognition (HAR) in low-light environments, addressing limitations in current models. HAR is vital for applications in smart cities, health monitoring, and security, but existing deep learning models face challenges with low-light, cluttered backgrounds, and high computational costs.

It proposes a two-tier architecture combining edge and cloud computing. First, a lightweight CNN enhances low-light video frames, while a human detection algorithm processes selective frames to optimize resource use. These refined frames are transmitted to the cloud, where a dual-stream CNN and transformer network extract both short- and long-range spatiotemporal features. This is followed by an Optimized Parallel Sequential Temporal Network (OPSTN) which employs squeeze and excitation attention mechanisms to improve HARincomplex environments.

Extensive experiments on three HAR datasets (HMDB51, UCF50, and YouTube Action) show that the framework outperforms state-of-the-art methods in recognizing complex activities under low-lighting conditions. The model achieves competitive accuracy, with significant improvements in both normal and low-light scenarios, demonstrating its robustness and efficiency in real-world surveillance applications.

2.10 Unlimited Knowledge Distillation for Action Recognition in the Dark

The paper addresses the challenge of recognizing human actions in dark environments, where essential visual information is often lost. Traditional methods for action recognition struggle under low-light conditions, leading to decreased performance. To overcome this, it propose Unlimited Knowledge Distillation (UKD), a novel approach that enhances the learning process by distilling knowledge from multiple teacher models into a student model without requiring large computational resources.Unlike conventional knowledge distillation methods, which require substantial GPU memory, UKD assembles knowledge from multiple models post-training, allowing for unlimited teacher models and improving action recognition accuracy. The authors introduce a Preferred Knowledge Distribution (PKD) to assess the importance of different teacher model outputs, ensuring that more valuable information is retained.

Extensive experiments were conducted on the ARID dataset, designed for action recognition in dark environments, showing that UKD significantly outperforms other methods, even surpassing two-stream models with a single-stream approach. The proposed method is computationally efficient and enhances the action recognition capabilities in low-light videos, demonstrating its potential for real-world surveillance and security applications

2.11 Dark Transformer: A Video Transformer for Action Recognition in the Dark

The paper addresses the challenge of recognizing human actions in low-light environments. This is critical for applications such as surveillance and nighttime driving. The existing methods typically handle action recognition and video enhancement separately, but this approach introduces Dark Transformer, a video transformer-based model that integrates both tasks.Dark Transformer leverages spatiotemporal self-attention mechanisms and unsupervised domain adaptation to recognize actions in dark environments effectively. The model extends the Transformer architecture and uses knowledge distillation, making it the first domain-invariant video transformer designed for dark settings. It demonstrates state-of-the-art performance on benchmark datasets such as InFAR, XD145, and ARID.

By employing a weight-sharing triple-branch framework, Dark Transformer processes both source (daytime) and target (nighttime) videos, learning domain-invariant representations that improve cross-domain action recognition. The experimental results show a significant performance increase, validating the model's efficacy in adverse lighting conditions. This work sets a new benchmark for action recognition in dark environments, paving the way for future advancements.

2.12 Movement-based Seizure Detection

The paper "Movement-based Seizure Detection" by Johan B. A. M. Arends explores the use of movement sensors in detecting epileptic seizures, particularly tonic–clonic seizures. It discusses how rhythmic movement components, such as those observed in the clonic phase, play a crucial role in seizure detection. Among the various movement sensor types, accelerometric sensors are the most commonly used due to their ability to capture subtle seizure-related movements. The study reviews multiple clinical trials conducted in video-electroencephalographic (EEG) monitoring units and real-world environments, highlighting variations in sensitivity (31%-95%) and positive predictive values (4%-60%). Field studies indicate that while bed sensors can detect seizures effectively in controlled environments, their accuracy declines in real-world settings, such as residential care for adults with intellectual disabilities. The paper emphasizes that while movement-based detection is valuable, it often requires additional modalities such as heart rate and electromyography (EMG) to enhance accuracy.

The research further explores the potential of advanced accelerometric techniques in understanding seizure evolution and evaluating the effects of antiepileptic drugs on specific seizure phases. Experimental studies using multiple accelerometers and video-EEG recordings reveal that different motor seizure components, such as myoclonic, tonic, and hyperkinetic movements, exhibit varying detection performance. While hyperkinetic movements are the easiest to detect, myoclonic seizures present challenges due to their low positive predictive value. The study suggests that multimodal detection, integrating multiple physiological signals, will be key to improving seizure detection accuracy in the future. Additionally, movement analysis could aid in developing neuromodulation techniques and personalized seizure management strategies, potentially improving patient outcomes.

2.13 Privacy-Preserving Early Detection of Epileptic Seizures in Videos

The paper "Privacy-Preserving Early Detection of Epileptic Seizures in Videos" presents a novel framework (SETR-PKD) designed for early seizure detection while maintaining patient privacy. Traditional video-based seizure detection relies on RGB video inputs, which can compromise privacy and hinder large-scale data collection. The authors propose an alternative approach that extracts optical flow features from video recordings, capturing movement patterns without revealing identifiable details. Additionally, their method employs progressive knowledge distillation, where knowledge from models trained on longer video samples is gradually transferred to models operating on shorter clips. This enables real-time seizure detection while reducing the need for full video processing. The framework achieves an impressive accuracy of 83.9% in detecting tonic-clonic seizures when only half of the seizure progression has occurred, making it a viable solution for real-time monitoring.

The study addresses key limitations of existing seizure detection techniques, such as the need for hospital-based Video EEG Monitoring (VEM), which is expensive and time-consuming. By utilizing optical flow analysis and transformer-based models, SETR-PKD effectively detects seizures without direct patient identification, ensuring compliance with data privacy concerns. The framework was validated on both an in-house hospital dataset and the publicly available GESTURES dataset, demonstrating its adaptability to different seizure types. The paper highlights the importance of privacy in medical AI applications and sets a foundation for future real-time, privacy-aware seizure monitoring systems, potentially improving patient safety and early intervention strategies.

2.14 Dancing in the Dark: A Benchmark towards General Low-light Video Enhancement

The paper introduces a novel benchmark and method for improving video quality in low-light conditions. It addresses the challenges of low-light video enhancement, which are crucial for applications like surveillance and autonomous driving, where visibility in poor lighting is essential. The authors present the "Dancing in the Dark" (DID) dataset, a high-quality video dataset captured using multiple cameras under various lighting conditions. This dataset features pronounced camera motion and precise spatial alignment, making it a valuable resource for training models in low-light video enhancement.

The authors propose a Light Adjustable Network (LAN), a Retinex-based method that iteratively refines illumination to enhance video frames. LAN adjusts to different lighting conditions, producing visually appealing results across diverse real-world scenarios. Extensive experiments demonstrate the superior performance of LAN, both on the DID dataset and in comparison with state-of-the-art methods. The DID dataset provides a more realistic and diverse set of low-light video scenes, and LAN achieves more natural and robust results without overexposure or underexposure, significantly advancing the field of low-light video enhancement

2.15 FastDVDnet: Towards Real-Time Deep Video Denoising Without Flow Estimation

The paper "FastDVDnet: Towards Real-Time Deep Video Denoising Without Flow Estimation" presents a novel deep learning-based approach for video denoising that eliminates the need for explicit motion estimation, making it significantly faster than existing methods. Traditional video denoising techniques rely on motion estimation to align frames before processing, which introduces additional computational costs and potential artifacts due to inaccurate flow calculations. FastDVDnet overcomes this limitation by using a multi-scale, U-Net-based convolutional neural network (CNN) architecture designed to implicitly handle motion. The model processes five consecutive frames at a time and denoises the central frame using two cascaded denoising steps. This architecture allows it to maintain temporal coherence while reducing flickering and preserving fine details. The authors demonstrate that FastDVDnet achieves comparable or superior denoising quality to state-of-the-art methods while being significantly faster—running up to 4000 times faster than patch-based methods and more than 80 times faster than previous deep learning-based approaches.

One of the key advantages of FastDVDnet is its ability to handle various noise levels with a single trained model, making it highly efficient and adaptable. The authors train the model using additive white Gaussian noise (AWGN) and validate its performance on benchmark datasets such as DAVIS-test and Set8. Quantitative evaluations show that FastDVDnet performs well across different noise levels, maintaining high peak signal-to-noise ratio (PSNR) scores and ensuring minimal visual artifacts. The study also highlights the superiority of the proposed multi-scale denoising blocks, the importance of end-to-end training, and the benefits of eliminating motion compensation. Overall, FastDVDnet represents a major step forward in real-time video denoising, offering a practical and high-quality solution suitable for real-world applications like mobile imaging and video enhancement.

Chapter 3

Proposed System

3.1 Objectives

Traditional action recognition systems struggle in low-light environments due to inadequate training datasets and the absence of effective enhancement techniques. This limitation makes it challenging to detect critical actions, such as seizures, in hospital settings where patients often require continuous monitoring. Existing monitoring systems primarily rely on visible-light recordings, which fail to capture essential details in the dark, leading to missed or delayed detection of medical emergencies.

This project aims to develop a deep-learning-based seizure detection system that enhances low-light video footage, extracts key motion patterns, and classifies seizure events in real time. The pipeline includes video enhancement using contrast adjustment techniques, skeletal keypoint extraction for movement tracking, and a deep learning model for accurate classification. Additionally, it provides real-time alerts to caregivers when a seizure is detected.

A key feature of this system is its ability to store and analyze seizure clips for future review, ensuring that critical medical events are documented. By integrating AI-driven motion analysis with low-light enhancement, this project offers a reliable and efficient solution for real-time patient monitoring, improving response times and patient safety in low-visibility conditions. Future extensions could include cloud-based monitoring and multimodal integration with EEG data for improved accuracy.

3.2 Proposed Solution

The proposed solution introduces a real-time seizure detection system optimized for dark environments using a multi-stage video processing pipeline. The system enhances low-light videos through Gamma Intense Correction (GIC) and Contrast Limited Adaptive Histogram Equalization (CLAHE) to improve visibility. FastDVDnet is applied to denoise noisy frames while preserving motion integrity.OpenPose extracts keypoints representing human skeletal motion, reducing background noise. Small 32×32 pixel patches are extracted around keypoints to focus on critical body movements. A deep learning model (VSViG) then analyzes these patches and keypoint sequences to classify seizures based on spatiotemporal motion patterns. If seizure probability exceeds a defined threshold, the system triggers real-time alerts via sound, desktop notifications, and a flashing GUI. The method ensures non-invasive, real-time, and reliable seizure detection, making it suitable for clinical and home-based monitoring, where traditional EEG-based systems may not be practical.

3.2.1 Video Enhancement Using GIC and CLAHE

In low-light environments, video frames often suffer from poor visibility and contrast, making it difficult to extract meaningful motion patterns. To address this, the frames undergo a two-step enhancement process. First, Contrast Limited Adaptive Histogram Equalization (CLAHE) is applied to improve local contrast. CLAHE divides each frame into small regions, or tiles, and applies histogram equalization independently to each region, ensuring that dark regions become more distinguishable without over-amplifying noise.

After CLAHE, Gamma Intense Correction (GIC) is used to adjust the brightness and contrast of the entire frame. GIC applies a non-linear transformation to pixel intensity values, brightening darker regions while preventing overexposure in already bright areas. This transformation is particularly useful for nighttime videos where illumination varies significantly across the frame. Once the enhancement process is complete, the frames are stored for further processing.

3.2.2 Denoising Using FastDVDnet

FastDVDnet is a deep-learning-based video denoising model that processes frames in small groups to remove noise while preserving motion details. It operates on a **five-frame sliding window**, where the central frame is denoised using information from neighboring frames. The network consists of **multiple denoising blocks** that extract spatiotemporal features, enabling it to handle complex noise patterns. By leveraging temporal redundancy, FastDVDnet achieves high-quality denoising without requiring optical flow estimation, making it efficient for real-time applications.

3.2.3 Keypoint Extraction Using OpenPose

Instead of analyzing full-frame movement, which can be computationally expensive and noisy, the system focuses on human skeletal motion. This is achieved through OpenPose, a deep learning-based framework that detects and tracks human body keypoints. Each frame is processed by OpenPose, which identifies 18 keypoints corresponding to joints such as the nose, shoulders, elbows, wrists, hips, knees, and ankles. These keypoints are extracted and stored as numerical coordinates in a JSON file, which maps each frame to its corresponding keypoint positions.

By focusing on skeletal motion rather than pixel-based movement, the model effectively reduces background noise and irrelevant motion artifacts. This makes it particularly effective in scenarios where lighting conditions are poor or where traditional motion detection methods would struggle.

3.2.4 Patch Extraction for Seizure Detection

Once the keypoints are extracted, the system isolates small image patches around the most important keypoints. Instead of analyzing the entire human skeleton, which may introduce unnecessary data, the system selects 15 keypoints that are most relevant for detecting seizure-related movement patterns. Around each selected keypoint, a 32×32 pixel patch is cropped from the enhanced video frame. These patches contain localized movement information, allowing the model to focus on specific body regions rather than processing the entire frame.

The extracted patches are converted into NumPy arrays (.npy format) for efficient processing. These patches serve as the primary input for the seizure classification model, ensuring that only the most relevant regions of the frame are analyzed.

3.2.5 Seizure Classification Using VSViG

To determine whether a seizure is occurring, the system employs VSViG (Vision Spatio-Temporal Vision Graph Model), a deep learning framework designed for skeleton-based motion analysis. Unlike traditional convolutional neural networks (CNNs), which analyze entire images, VSViG focuses on the temporal relationships between keypoints over a sequence of frames.

The classification model takes two primary inputs: the patch sequences extracted around keypoints and the sequential movement of keypoints over time. By analyzing how keypoints change over consecutive frames, VSViG identifies abnormal, repetitive, and involuntary movement patterns that indicate a seizure.

The model outputs a probability score for each frame, ranging from 0 to 1, where higher values indicate a higher likelihood of seizure activity. The final seizure classification is determined by aggregating frame-wise scores and applying a thresholding mechanism. If the computed seizure probability exceeds a predefined threshold, the system flags the event as a seizure.

3.2.6 Alert System and Graphical User Interface

To ensure that caregivers or medical professionals are immediately notified of seizure events, the system includes a real-time alert mechanism. If a seizure is detected, an alert is triggered based on the probability scores computed by the classification model. The alert conditions are designed to minimize false alarms while ensuring that genuine seizures are not missed. If a frame receives a probability score above 0.75, or if multiple consecutive frames register a probability of 1.0, an alert is activated.

The system generates three types of alerts:

- **Sound Alert:** A loud siren is played using the `winsound` module, ensuring that nearby individuals are notified.
- **Desktop Notification:** A notification appears on the screen using the `plyer.notification` module, providing instant feedback to users.
- **Flashing Screen:** The graphical user interface (GUI) flashes a warning message using `tkinter`, providing a visual indicator of the seizure event.

Chapter 4

Project Design

4.1 System Architecture

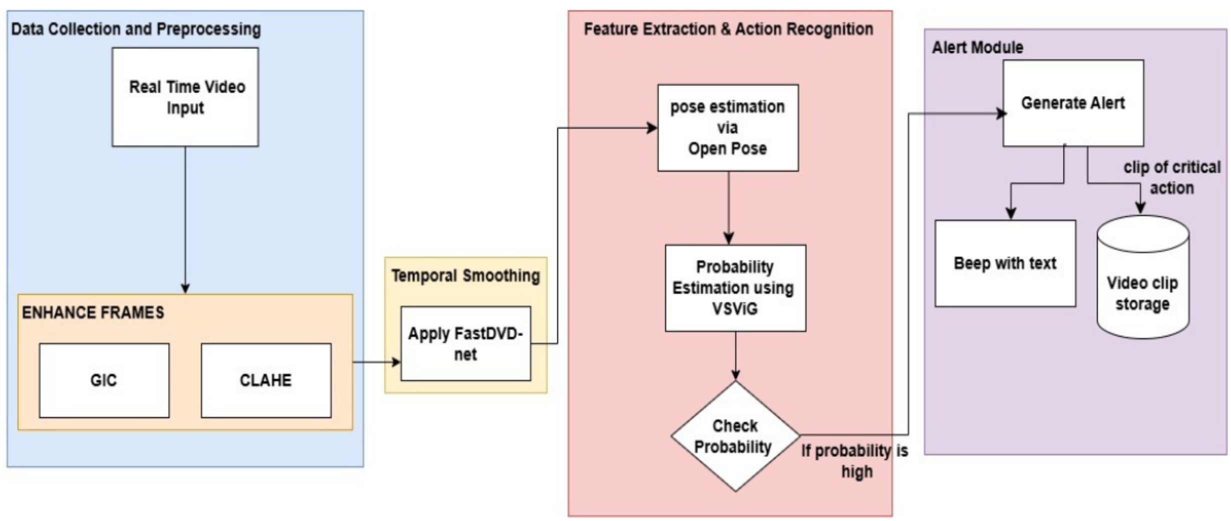

Figure 4.1: Architecture diagram

This system operates in real-time, processing video inputs from dark environments and providing timely alerts if critical actions are detected. The architecture is organized into distinct modules that each play a crucial role in ensuring high accuracy and responsiveness.

The Data Collection and Preprocessing Module is the initial stage, where raw video frames are captured from low-light settings, such as patient rooms during the night. Due to the poor visibility inherent in these environments, preprocessing is essential to enhance the quality of the frames for effective action recognition. This module applies Gamma Intense Correction (GIC) and Contrast Limited Adaptive Histogram Equalization(CLAHE) techniques to brighten and improve the contrast of the frames. Next, the Temporal Smoothing Module is introduced to ensure temporal consistency across frames, which is vital for effective action recognition.This module employs a FastDVDNet to smooth out these frame-level differences, capturing a coherent motion trajectory and minimizing jitter.

The Feature Extraction and Action Recognition Module is the core of the system, where seizure recognition and probability generation take place. The enhanced and denoised frames from the previous module is given to the openpose model, which is a state of the art pose estimation system developed by researchers at Carnegie Mellon University (CMU) that can detect and track the human body in real-time and accurately determine its pose in 3D space. The openpose overlays a skeleton structure on each frame which denotes the positon of person in the frame. There are 18 points that openpose detects, that is sholder, elbow, wrist, etc. We take 15 points out of these 18 for our detection purpose. Based on the location of these points we generate patches for each frame, these patches are then fused with the frame, which highlights the part of the frame that the model can focus on. We take both the spatial and temporal features of the points in the frame. We train the model to recognize these patterns that patients may make during the seizure, which will generate a probability that denotes how likely a seizure may occur.

Finally, the Alert Module is responsible for handling alerts and managing video storage. When a critical action is detected by the healthcare monitoring feature, this module triggers an alert that is sent to healthcare staff. Additionally, it stores a 10-second video clip from the frames leading up to the critical action in a secure database.

4.2 Modules

4.2.1 Data Collection and Preprocessing

This module captures raw video frames from a low-light environment and preprocesses them for improved visibility. The frames are enhanced using GIC and CLAHE, which brighten the frames and enhance contrast, making it easier for the action recognition module to detect movements.

Output from this module are enhanced frames ready for further processing.

4.2.2 Temporal Smoothing

To ensure smooth action recognition, this module uses FastDVDNet or optical flow to maintain temporal consistency across frames. This helps in reducing frame-level noise and jitter, which can interfere with accurate recognition.

Output from this module are smoothed frames with consistent temporal flow, preserving movement across the sequence.

4.2.3 Feature Extraction and Action Recognition

The core of the system is the Feature Extraction and Action Recognition Module, which handles seizure recognition and probability generation. Enhanced and denoised frames are processed by OpenPose, a state-of-the-art pose estimation system developed by Carnegie Mellon University (CMU). OpenPose detects and tracks human body poses in real-time, overlaying a skeleton structure on each frame with 18 key points (e.g., shoulder, elbow, wrist). For seizure detection, 15 points are utilized.

From these points, patches are generated for each frame and fused with the frame to highlight areas of interest. Both spatial and temporal features of these points are analyzed. The model is trained to recognize specific patterns associated with seizures, generating probabilities that indicate the likelihood of seizure occurrence.

4.2.4 Alert Module

This module handles alerting and clip storage. If a critical action is detected, it stores a 10-second video clip preceding the action in a secure database and sends an alert to healthcare staff as a beep notification with text for context.

4.3 Data Flow Diagrams

4.3.1 Level 0

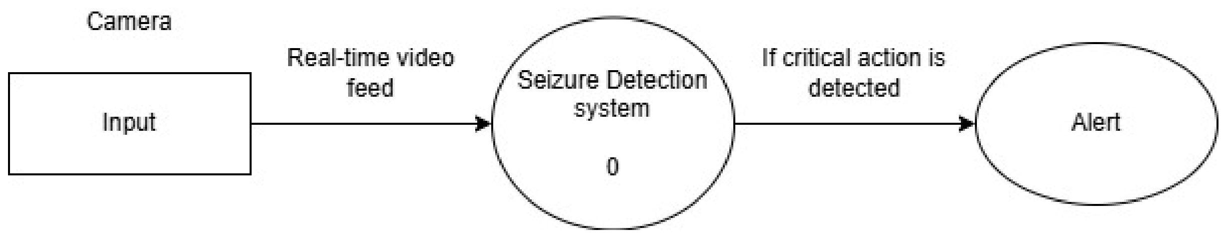

Figure 4.2: Level 0

Level 0 of the DFD gives an overview of the entire system as a single, high-level process. This process receives raw video frames from a low-light environment, enhances them, processes them to recognize actions, and finally issues alerts for critical actions.

4.3.2 Level 1

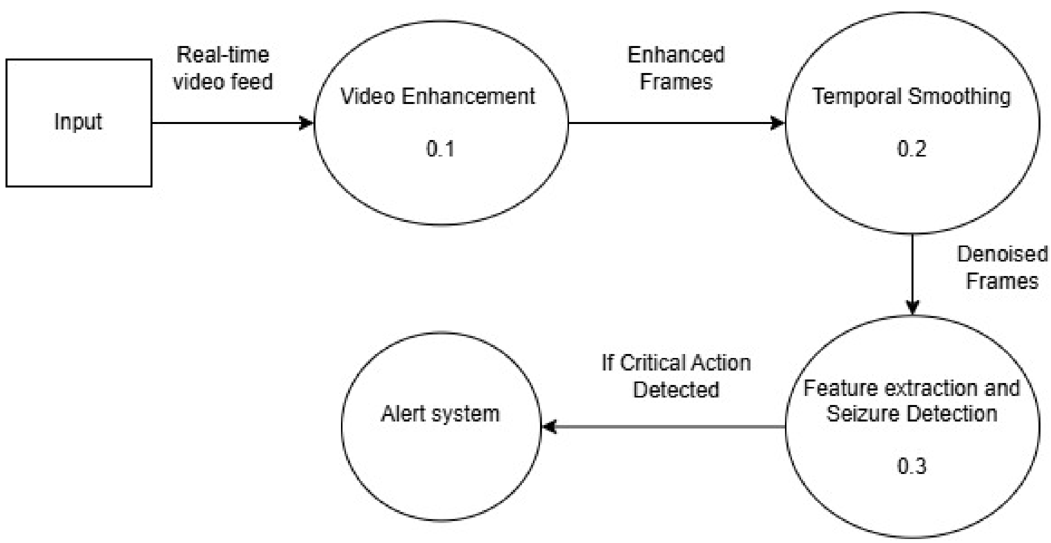

Figure 4.3: level 1

Level 1 of the DFD breaks down the main processes of the system into distinct components: Data Collection and Preprocessing, Action Recognition, and Alert System. Here, each module's role becomes more defined.

The Data Collection and Preprocessing module handles the acquisition of video frames and enhances them for better visibility using techniques like GIC and CLAHE. This enhanced video data is then passed to the next module.Here the temporal smoothing is done by using FastDVDnet.

The Feature Extraction and Action Recognition module, encompassing pose detection and probability generation of likelihood of a patient affected by seizure.

The Alert System module receives the output from action recognition, checks if a seizure has occurred based on probaility, and issues alerts accordingly. If a critical event is detected, this module also manages the storage of a 10-second video clip of the incident.

4.3.3 Level 2

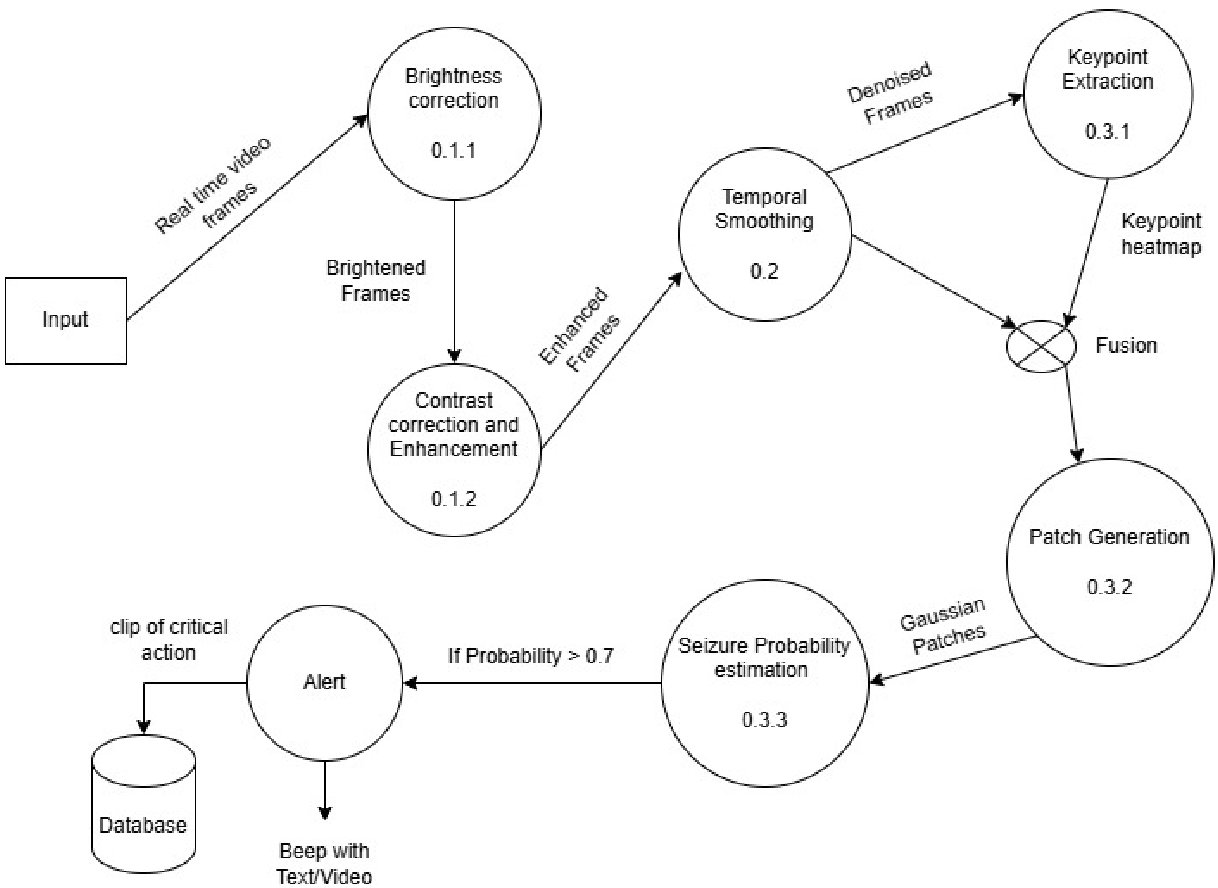

Figure 4.4: Level 2

This process starts with video capture, where raw frames are gathered from the environment. The frames are then processed with Gamma Intense Correction to improve brightness, followed by CLAHE to balance over exposure. This prepares the data for accurate action recognition.

Temporal Smoothing within the action recognition pipeline, temporal smoothing uses FastDVDnet or optical flow to stabilize motion across frames. This ensures that actions are coherent across time and reduces noise.

The system's core is the Feature Extraction and Action Recognition Module, which uses Open-

Pose for pose estimation to detect and track human body movements. It processes enhanced frames, focusing on 15 key points for seizure detection, analyzes spatial and temporal features, and generates probabilities for seizure likelihood.

Alert Module checks if a Seizure is detected, the alert module triggers a notification to healthcare staff. Additionally, this module stores a 10-second video clip leading up to the critical event in a secure database and includes this clip in the alert. This ensures that healthcare staff have both real-time notifications and contextual information, enabling better response and patient care.

4.4 System Requirements

4.4.1 Hardware

Hard Disk: 512 GB or above
CPU: i7/Ryzen 7 (8 core) or above
GPU: NVIDIA RTX 3080 or above
RAM: Minimum 16 GB
High-resolution cameras with low-light capabilities
Sound System capable of producing beep sounds

4.4.2 Software

Windows version 10 and above
Technologies used: Python, TensorFlow or PyTorch etc.
OpenCV for image processing
IDE: Visual Studio
Dataset: Custom seizure datasets

Chapter 5

Implementation

5.1 Screenshots

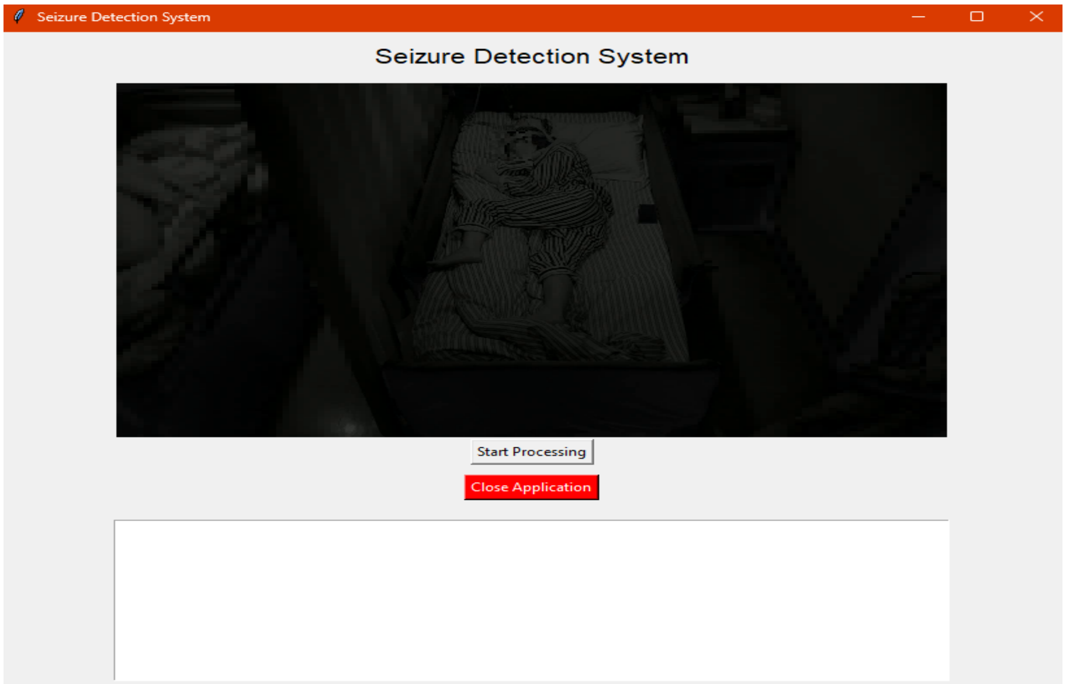

Figure 5.1: Initial Interface

This is the initial interface that is shown when the program starts running. It takes a video as input and is ready to start the processing phase.

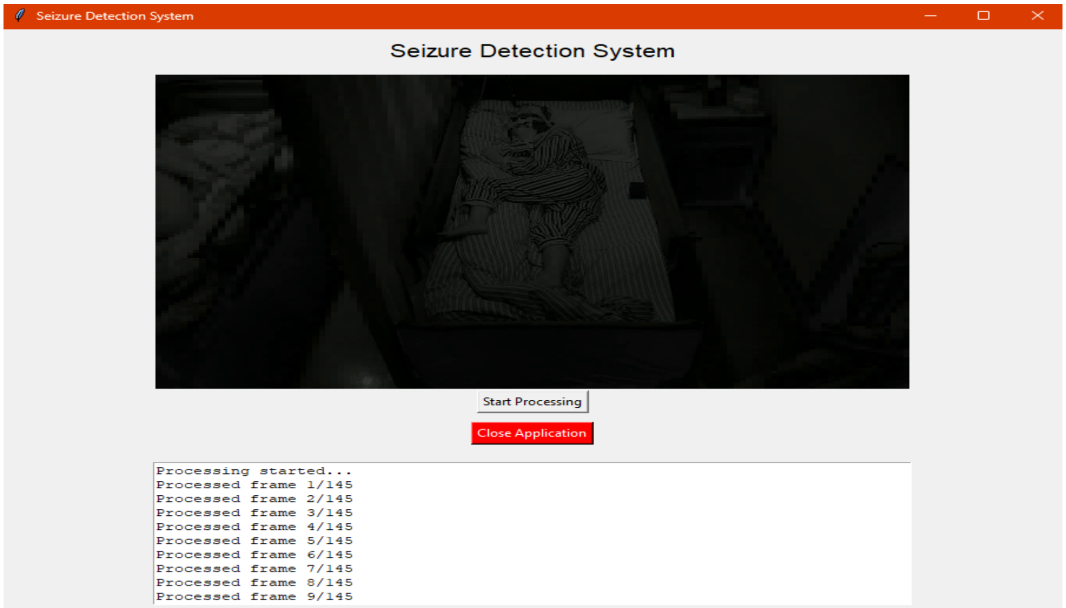

Figure 5.2: Live Frame Processing

As soon as processing starts, each frame of the video is extracted using TensorFlow and is given to the preprocessing module.

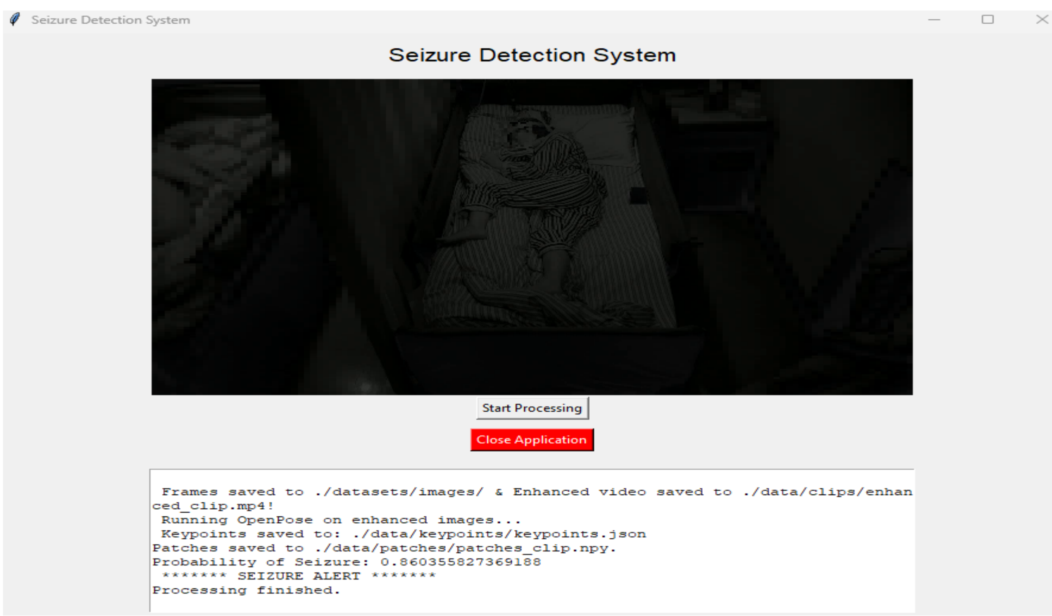

Figure 5.3: Seizure Prediction

After the frame passes through all the modules, the system provides the final probability of seizure occurrence. If this probability exceeds the threshold, an alert is activated.

Chapter 6

Results and Discussions

We have tested the model on a test video from the dataset. The medical data related to the video shows the EEG onset and Clinical onset of the seizure. The data states that the EEG onset happens at the 4th second and the clinical onset happens at the 53rd second. There is a transition period of 49 seconds between the EEG and clinical onset during which our model aims to detect the seizure by calculating the the probability how likely there is a seizure happening.

The proposed system employs a two-step enhancement process as to overcome all the light issues in videos as shown in Fig.6.1.

Figure 6.1: Input Video

First, Contrast Limited Adaptive Histogram Equalization (CLAHE) is applied to enhance local contrast without over-amplifying noise. This method ensures that fine details in dark video frames become more distinguishable. Next, Gamma Intense Correction (GIC) is used to dynamically adjust the brightness of frames, making movements more perceptible. By sequentially applying these enhancement techniques, the system ensures optimal feature extraction for downstream processing as shown in Fig.6.2.

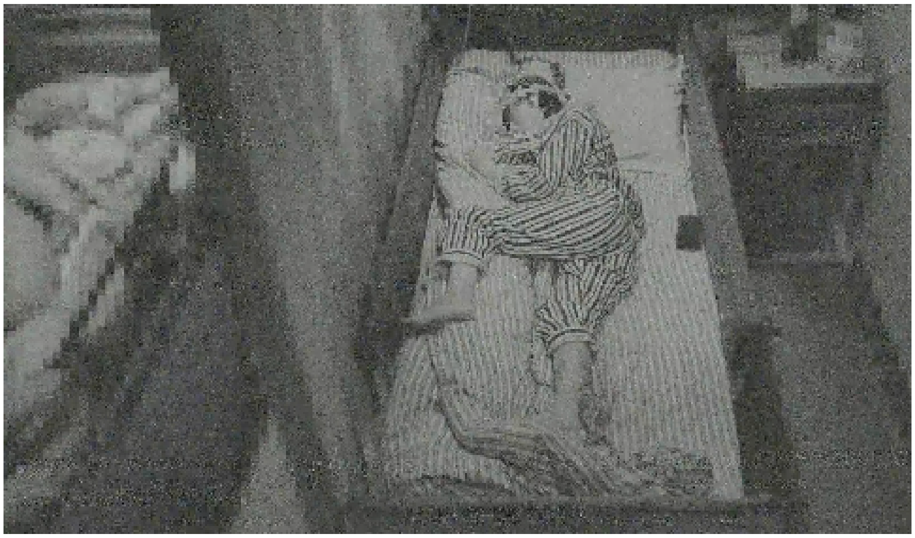

Figure 6.2: Enhanced Output using GIC and CLAHE

Dark environment videos often suffer from noise due to low exposure and sensor limitations. In order to improve the accuracy of seizure detection, a temporal smoothing module based on FastDVDnet is introduced. After applying the technique we get a clear an denoised output as shown in Fig.6.3.

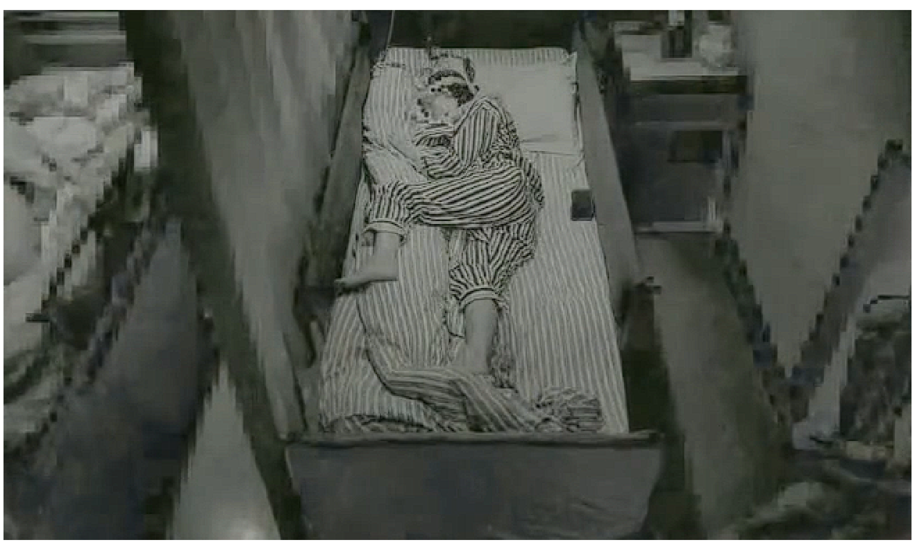

Figure 6.3: Temporal Smoothing using FastDVDnet

And there are differences in the graphs of new input video and denoised video as shown in Fig.6.4.

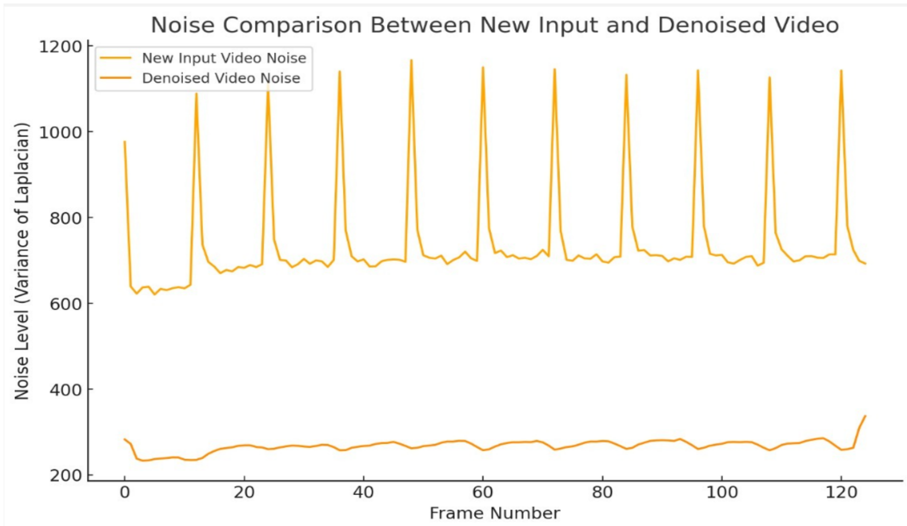

Figure 6.4: Comparison between New input and Denoised Video

And the next phase is the Feature Extraction where we have introduced an customized pose estimation mechanisms. As in seizure specific motion patterns require additional refinement. OpenPose ensures that it can reliably detect and map seizure-related movement variations and denotes the keypoints as shown in Fig.6.4. To analyze seizure-related movements over time, a Vision Spatio-

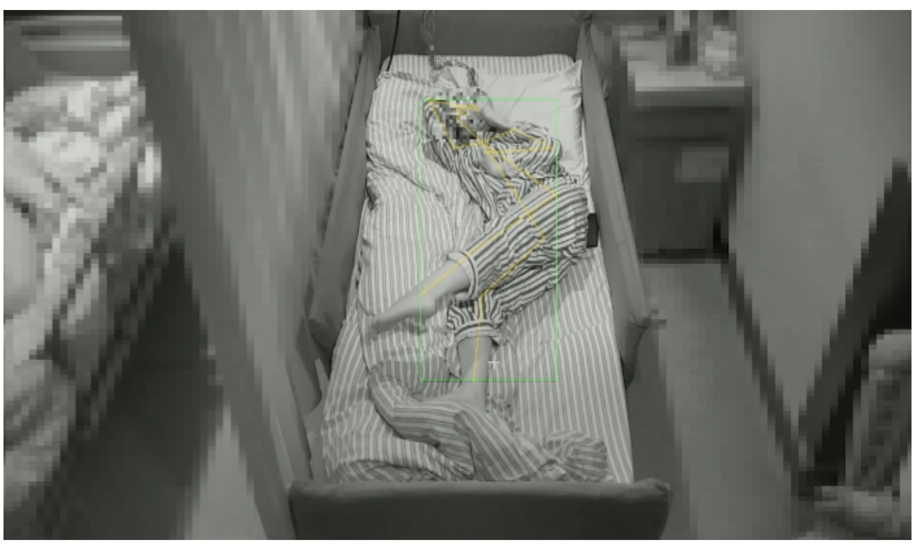

Figure 6.5: Keypoints identification using Openpose

Temporal Vision Graph Model (VSViG) is built. This model processes extracted patches with a combination of graph-based feature learning and temporal motion analysis.Graph-based analysis allows the system to represent keypoint relationships efficiently, identifying patterns in movement coordination. As it generates seizure probability over time as shown in Fig.6.6. if the generated probability goes above our selected threshold, it will trigger the alert.

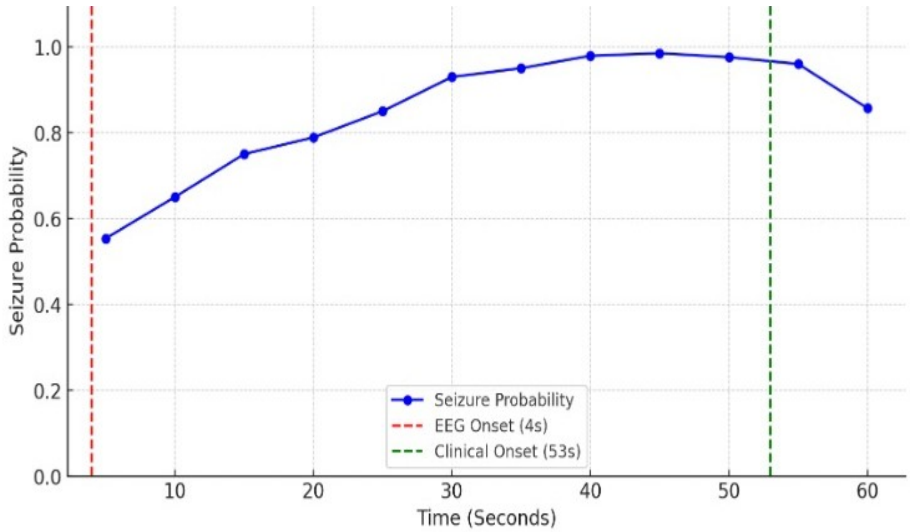

Figure 6.6: Seizure Probability over Time

The proposed seizure detection system achieves high accuracy, sensitivity, and specificity, even in low-light conditions. The integration of video enhancement, temporal smoothing, and OpenPose-based feature extraction significantly improves seizure detection performance compared to baseline models. The alerting system ensures timely intervention, making this approach suitable for clinical and home-based applications. Future work will focus on optimizing computational efficiency and expanding dataset diversity for broader applicability.

6.1 Comparison with Existing System

The existing system of VSViG can only work with well lit environment, it is unable to work in a dark environment with poor lighting conditions, our proposed project is able to work in a dim light environment by using enhancement methods such as GIC and CLAHE. Based on the videos that we used for testing which contains both ones having seizure and ones without it, the model was able to give an optimal accuracy of 72%, most of the accuracy we lost are on false positives as the system may classify some movements as ones with high seizure probability, there are no cases where the system had false negatives.

Chapter 7

Conclusion

The seizure detection system developed in this project effectively integrates multiple deep learning-based modules to identify seizures in low-light environments. By leveraging video enhancement techniques (GIC + CLAHE), the system enhances visibility, ensuring that motion details are preserved for accurate analysis. The OpenPose model extracts skeletal keypoints from video frames, creating a structured representation of movement patterns. Patch extraction around keypoints further refines the data by isolating motion-specific features, which are then analyzed using VSViG, a spatiotemporal deep learning model. This pipeline enables accurate seizure classification, ensuring timely detection through an automated workflow.

The system's modular approach allows for flexibility and future improvements. While denoising using FastDVDnet was initially omitted due to computational constraints, its potential integration could further enhance the clarity of low-quality videos. Additionally, optimizing patch extraction could improve spatial resolution, leading to better classification accuracy. The real-time seizure probability computation ensures prompt detection, making the system a viable solution for continuous monitoring. Implementing detection logs for medical review could further enhance its applicability in clinical settings.

Overall, this project presents a robust AI-driven approach for seizure detection, demonstrating the potential of computer vision and deep learning in medical diagnostics. By combining low-light image enhancement, skeletal keypoint analysis, and deep learning-based classification, the system provides a reliable method for seizure detection in challenging environments. Future enhancements, such as cloud-based scalability, real-time optimization, and improved model efficiency, can further extend its impact, making it a valuable tool for both hospital and home-based monitoring solutions.

Chapter 8

Future Scope

The seizure detection system has significant potential for further advancements to enhance its accuracy, efficiency, and real-world applicability. One major area for improvement is integrating real-time processing to reduce latency, ensuring near-instantaneous detection and alert generation. Optimizing patch extraction and keypoint tracking can refine feature selection, leading to better seizure classification accuracy. Additionally, incorporating FastDVDnet or other deep-learning-based denoising techniques can improve video quality, making keypoint detection more robust in extreme low-light conditions.

Another promising direction is cloud-based deployment, enabling remote monitoring for hospital and home-care environments. A lightweight, mobile-compatible version could be developed to facilitate continuous monitoring via smartphones, making it accessible for caregivers. Furthermore, integrating the system with EEG data could enhance detection accuracy by combining visual and neurological signals for a multimodal approach. Personalized models trained on individual patient data may also improve performance by adapting to variations in movement patterns.

Expanding the system's capabilities to log and analyze seizure events over time could provide valuable insights for medical professionals. This feature would allow for better tracking of seizure frequency, severity, and progression, aiding in treatment adjustments. With these advancements, the system could evolve into a comprehensive, AI-driven healthcare tool for real-time seizure detection and monitoring.

Chapter 9

Bibliography

[1] Chen, R., Chen, J., Liang, Z., Gao, H., Lin, S. (2021). DarkLight Networks for Action Recognition in the Dark. CVPRW.

[2] VSViG: Real-time Video-based Seizure Detection via Skeleton-based Spatiotemporal ViG Yankun Xu, Junzhe Wang, Yun-Hsuan Chen, Jie Yang, Wenjie Ming, Shuang Wang, and Mohamad Sawan, July 2024.

[3] Gedas Bertasius, Heng Wang, and Lorenzo Torresani. Is space-time attention all you need for video understanding? arXiv preprint arXiv:2102.05095, 2021.

[4] Deepti Ghadiyaram, Du Tran, and Dhruv Mahajan. Largescale weakly-supervised pre-training for video action recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 12046–12055, 2019.

[5] Low-Light Image Enhancement via Structure Modeling and Guidance Xiaogang Xu, Ruixing Wang, Jiangbo Lu, Zhejiang Lab, Honor Device Co. Ltd., SmartMore Corporation, 2022.

[6] Chen Chen UIUC, Qifeng Chen HKUST, Minh N. Do UIUC. Vladlen Koltun Intel Labs, Seeing Motion in the Dark (2019), IEEE.

[7] Haiyang Jiang, Yinqiang Zheng. Learning to See Moving Objects in the Dark, IEEE (2019).

[8] R. Fan, X. Li, S. Lee, T. Li, and H. L. Zhang, "Smart image enhancement using CLAHE based on an F-shift transformation during decompression," Electronics, vol. 9, no. 9, p. 1374, Aug. 2020. DOI: 10.3390/electronics9091374..

[9] Altaf Hussain, Samee Ullah Khan, Noman Khan, Imad Rida, Meshal Alharbi, Sung Wook Baik. Low-light aware framework for human activity recognition via optimized dual stream parallel network, Sejong University, Seoul 143-747, Republic of Korea, 2023.

[10] Ruibing Jin, Guosheng Lin, Min Wu, Jie Lin, Zhengguo Li, Xiaoli Li, Zhenghua Chen. Unlimited Knowledge Distillation for Action Recognition in the Dark, Institute for Infocomm Research, Singapore, 2023.

[11] Anwaar Ulhaq. Dark Transformer: A video Transformer for action recognition in dark, School of Engineering and Technology Sydney campus, Australia, 2024.

[12] SMART Frame Selection for Action Recognition Shreyank N Gowda, Marcus Rohrbach, Laura Sevilla-Lara, University of Edinburgh, Facebook AI Research, 2020.

[13] ALEN: A Dual-Approach for Uniform and 0, Jalisco, Mexico, 2024.

[14] Unsupervised Face Detection in the Dark, Wenjing Wang, Graduate Student Member, IEEE, Xinhao Wang, Wenhan Yang, Member, IEEE, and Jiaying Liu, Senior Member, IEEE, 2023.

[15] Fu, Huiyuan, et al. "Dancing in the dark: A benchmark towards general low-light video enhancement." Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023.

[16] Mehta, D., Sivathamboo, S., Simpson, H., Kwan, P., O'Brien, T., Ge, Z.: Privacy-preserving early detection of epileptic seizures in videos. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp.210–219. Springer (2023).

[17] Yue, R., Tian, Z., Du, S.: Action recognition based on rgb and skeleton data sets: A survey. Neurocomputing (2022).

[18] Shoeb, A.H., Guttag, J.V.: Application of machine learning to epileptic seizure detection. In: ICML. pp. 975–982 (2010).

[19] Lee, J., Lee, M., Lee, D., Lee, S.: Hierarchically decomposed graph convolutional networks for skeleton-based action recognition. In: ICCV. pp. 10444–10453 (October 2023).

[20] Hussain, A., Khan, S. U., Khan, N., Rida, I., Alharbi, M., Baik, S. W. (2023). Low-light aware framework for human activity recognition via optimized dual stream parallel network. Alexandria Engineering Journal, 74, 569-583.

[21] J. B. A. M. Arends, "Movement-based seizure detection," Epilepsia, vol. 59, no. S1, pp. 30–35, 2018. doi: 10.1111/epi.14053.

[22] Tassano, M., Delon, J., & Veit, T. (2020). FastDVDnet: Towards real-time deep video denoising without flow estimation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)

Chapter 10

Publications

10.1 Paper Status

The paper titled " Seizure Detection in Dark Environments " has been submitted at the IEEE CCET SPC'25 Conference organized by IEEE CCET SB & Department of Electrical and Electronics Engineering, Carmel College of Engineering and Technology, Punnapara, Alappuzha . The paper is waiting for acceptance.

Seizure Detection in Dark Environments

Aakash Jason Kavalakkat
Department of Artificial Intelligence and Data Science
Muthoot Institute of Technology and Science
Ernakulam, India
aakash.jk2003@gmail.com

Akash R
Department of Artificial Intelligence and Data Science
Muthoot Institute of Technology and Science
Ernakulam, India
akash.rajesh060@gmail.com

Bijil Boby
Department of Artificial Intelligence and Data Science
Muthoot Institute of Technology and Science
Ernakulam, India
bijilbobyofficial@gmail.com

Brayan Deep
Department of Artificial Intelligence and Data Science
Muthoot Institute of Technology and Science
Ernakulam, India
brayandeep.eclipse@gmail.com

Parvathy E N
Asst.Professor
Department of Artificial Intelligence and Data Science
Muthoot Institute of Technology and Science
Ernakulam, India
parvathyen@mgits.ac.in

***Abstract*—Seizures occurring in hospitalized patients, particularly during nighttime, pose a significant risk if left undetected. This study presents a real-time Seizure Detection and Alert System designed to operate effectively in low-light hospital environments. The system processes video footage from surveillance cameras by enhancing visibility using Gamma Intensity Correction (GIC) and Contrast Limited Adaptive Histogram Equalization (CLAHE), followed by noise reduction with FastDVDnet to improve clarity and consistency across frames. For seizure detection, a deep learning-based skeleton tracking model (VSViG: Video-based Seizure Detection via Skeleton-based Spatiotemporal ViG) is used to extract patient pose information and identify seizure-related movements. The model generates a seizure probability score in real time, and if the probability exceeds a predefined threshold, an automated alert is sent to the nurse station for immediate intervention. By integrating computer vision techniques, deep learning-based action recognition, and real-time alerting, our system provides an effective, automated seizure monitoring solution to enhance patient safety and ensure timely medical response.**

***Index Terms*—Seizure Detection, real-time monitoring, low-light video enhancement.**

I. Introduction

Seizures, It is a sudden uncontrolled electrical disturbances in the brain that which can cause physical convulsions, loss of consciousness, and other serious symptoms. As they often require immediate measures to prevent injury or complications. Traditionally, seizure detection has done through electroencephalography (EEG) monitoring, As it provides an high accuracy but is intrusive and impractical for continuous use outside of clinical settings. Another approach involves video-based monitoring; however, this method becomes ineffective in low-light conditions where visibility is significantly been reduced.

Detecting seizures in dark environments poses significant challenges due to the poor lightings, loss of visual details, and increased noises in the video recordings. The lack of sufficient light can obscure critical motion features, making it difficult to identify characteristic seizure patterns accurately. These limitations necessitate the development of advanced video processing techniques that enhance visibility while preserving essential motion cues.

The proposed research presents a real-time seizure detection system designed specifically for dark environments. The system incorporates video enhancement methodologies such as Gamma Intense Correction (GIC) and Contrast Limited Adaptive Histogram Equalization (CLAHE) to improve visibility and brightness. Additionally, a temporal smoothing module based on FastDVDnet is integrated to reduce noise and stabilize motion details. These enhanced frames are processed using OpenPose, a deep learning-based human pose estimation model, to extract skeletal keypoints that represent human motion. By analyzing the extracted keypoints over time, the system classifies seizure events using the Vision Spatio-Temporal Vision Graph Model (VSViG), a deep learning model optimized for spatiotemporal motion analysis.

To ensure immediate response to detect seizures, the system includes an real-time alert mechanism that notifies caregivers through an visual notifications, audio alerts, or through desktop notifications. The entire pipeline is designed to operate in real-time, making it more suitable for both clinical and home-based seizure monitoring applications. By integrating advanced computer vision and deep learning techniques, this research aims to bridge the gap in non-

Figure 10.1: First page of the research paper

Appendix A

Code Snippets

```python
###  Video Enhancement & Save Frames for OpenPose
def enhance_video_and_save_frames(input_video_path, image_output_dir, output_clip_path, gamma=0.5):
    """
    Enhances video frame-by-frame using GIC + CLAHE and saves frames for OpenPose.
    Also saves the enhanced video for patch extraction.
    """

    cap = cv2.VideoCapture(input_video_path)
    frame_width = int(cap.get(cv2.CAP_PROP_FRAME_WIDTH))
    frame_height = int(cap.get(cv2.CAP_PROP_FRAME_HEIGHT))
    fps = int(cap.get(cv2.CAP_PROP_FPS))
    total_frames = int(cap.get(cv2.CAP_PROP_FRAME_COUNT))

    fourcc = cv2.VideoWriter_fourcc(*'mp4v')
    out = cv2.VideoWriter(output_clip_path, fourcc, fps, (frame_width, frame_height))

    frame_count = 0
    while cap.isOpened():
        ret, frame = cap.read()
        if not ret:
            break

        # Enhance frame
        enhanced_frame = apply_gic_with_clahe(frame, gamma)
        #enhanced_frame = frame


        # Save frame as image for OpenPose
        frame_filename = os.path.join(image_output_dir, f"frame_{frame_count:04d}.jpg")
        cv2.imwrite(frame_filename, enhanced_frame)

        # Save frame to output video for patch extraction
        out.write(enhanced_frame)

        frame_count += 1
        print(f"Processed frame {frame_count}/{total_frames}", end="\r")

    cap.release()
    out.release()
    print(f"\n Frames saved to {image_output_dir} & Enhanced video saved to {output_clip_path}!")
```

Figure A.1: Image Enhancement via GIC and CLAHE

```
def test_fastdvdnet(**args):
    start_time = time.time()
    # If save_path does not exist, create it
    if not os.path.exists(args['save_path']):
        os.makedirs(args['save_path'])
    logger = init_logger_test(args['save_path'])
    # Sets data type according to CPU or GPU modes
    if args['cuda']:
        device = torch.device('cuda')
    else:
        device = torch.device('cpu')
    # Create models
    print('Loading models ...')
    model_temp = FastDVDnet(num_input_frames=NUM_IN_FR_EXT)
    # Load saved weights
    state_temp_dict = torch.load(args['model_file'], map_location=device)
    if args['cuda']:
        device_ids = [0]
        model_temp = nn.DataParallel(model_temp, device_ids=device_ids).cuda()
    else:
        # CPU mode: remove the DataParallel wrapper
        state_temp_dict = remove_dataparallel_wrapper(state_temp_dict)
    model_temp.load_state_dict(state_temp_dict)
    # Sets the model in evaluation mode (e.g. it removes BN)
    model_temp.eval()
    with torch.no_grad():
        # process data
        seq, _, _ = open_sequence(args['test_path'],\
                                  args['gray'],\
                                  expand_if_needed=False,\
                                  max_num_fr=args['max_num_fr_per_seq'])
        seq = torch.from_numpy(seq).to(device)
        seq_time = time.time()
        # Add noise
        noise = torch.empty_like(seq).normal_(mean=0, std=args['noise_sigma']).to(device)
        seqn = seq + noise
        noisestd = torch.FloatTensor([args['noise_sigma']]).to(device)

        denframes = denoise_seq_fastdvdnet(seq=seqn,\
                                           noise_std=noisestd,\
                                           temp_psz=NUM_IN_FR_EXT,\
                                           model_temporal=model_temp)
    # Compute PSNR and log it
    stop_time = time.time()
    psnr = batch_psnr(denframes, seq, 1.)
    psnr_noisy = batch_psnr(seqn.squeeze(), seq, 1.)
    loadtime = (seq_time - start_time)
    runtime = (stop_time - seq_time)
    seq_length = seq.size()[0]
    logger.info("Finished denoising {}".format(args['test_path']))
    logger.info("\tDenoised {} frames in {:.3f}s, loaded seq in {:.3f}s".\
                format(seq_length, runtime, loadtime))
    logger.info("\tPSNR noisy {:.4f}dB, PSNR result {:.4f}dB".format(psnr_noisy, psnr))
    # Save outputs
    if not args['dont_save_results']:
        # Save sequence
        save_out_seq(seqn, denframes, args['save_path'], \
                     int(args['noise_sigma']*255), args['suffix'], args['save_noisy'])
```

Figure A.2: Denoising using FastDVDNet

```python
def run_demo(net, image_provider, height_size, cpu, track, smooth):
    net = net.eval()
    if not cpu:
        net = net.cuda()
    stride = 8
    upsample_ratio = 4
    num_keypoints = Pose.num_kpts
    previous_poses = []
    delay = 1
    all_keypoints_data = []  # To store keypoints from all images/frames
    for img in image_provider:
        orig_img = img.copy()
        heatmaps, pafs, scale, pad = infer_fast(net, img, height_size, stride, upsample_ratio, cpu)
        total_keypoints_num = 0
        all_keypoints_by_type = []
        for kpt_idx in range(num_keypoints):  # 19th for bg
            total_keypoints_num += extract_keypoints(heatmaps[:, :, kpt_idx], all_keypoints_by_type, total_keypoints_num)
        pose_entries, all_keypoints = group_keypoints(all_keypoints_by_type, pafs)
        for kpt_id in range(all_keypoints.shape[0]):
            all_keypoints[kpt_id, 0] = (all_keypoints[kpt_id, 0] * stride / upsample_ratio - pad[1]) / scale
            all_keypoints[kpt_id, 1] = (all_keypoints[kpt_id, 1] * stride / upsample_ratio - pad[0]) / scale
        current_poses = []
        frame_keypoints = []  # Keypoints for the current frame/image
        for n in range(len(pose_entries)):
            if len(pose_entries[n]) == 0:
                continue
            pose_keypoints = np.ones((num_keypoints, 2), dtype=np.int32) * -1
            for kpt_id in range(num_keypoints):
                if pose_entries[n][kpt_id] != -1.0:  # keypoint was found
                    pose_keypoints[kpt_id, 0] = int(all_keypoints[int(pose_entries[n][kpt_id]), 0])
                    pose_keypoints[kpt_id, 1] = int(all_keypoints[int(pose_entries[n][kpt_id]), 1])
            frame_keypoints.append(pose_keypoints.tolist())
            pose = Pose(pose_keypoints, pose_entries[n][18])
            current_poses.append(pose)
        all_keypoints_data.append(frame_keypoints)
        if track:
            track_poses(previous_poses, current_poses, smooth=smooth)
            previous_poses = current_poses
        for pose in current_poses:
            pose.draw(img)
        img = cv2.addWeighted(orig_img, 0.6, img, 0.4, 0)
        for pose in current_poses:
            cv2.rectangle(img, (pose.bbox[0], pose.bbox[1]),
                          (pose.bbox[0] + pose.bbox[2], pose.bbox[1] + pose.bbox[3]), (0, 255, 0))
            if track:
                cv2.putText(img, 'id: {}'.format(pose.id), (pose.bbox[0], pose.bbox[1] - 16),
                            cv2.FONT_HERSHEY_COMPLEX, 0.5, (0, 0, 255))
        cv2.imshow('Lightweight Human Pose Estimation Python Demo', img)
        key = cv2.waitKey(delay)
        if key == 27:  # esc
            return
        elif key == 112:  # 'p'
            if delay == 1:
                delay = 0
            else:
                delay = 1
    save_keypoints_to_json(all_keypoints_data, args.output_file)
```

Figure A.3: OpenPose for Pose Estimation

```python
def generate_patches(clip_name, keypoints_file):
    video_path = os.path.join(CLIPS_DIR, clip_name)
    patches_path = os.path.join(PATCHES_DIR, f"patches_{clip_name.split('.')[0].split('_')[1]}.npy")

    # Load keypoints
    with open(keypoints_file, "r") as f:
        keypoints_data = json.load(f)

    # Validate keypoints
    validated_keypoints = []
    for frame_data in keypoints_data:
        if isinstance(frame_data, list) and len(frame_data) == 1 and len(frame_data[0]) == 18:
            validated_keypoints.append(frame_data[0])  # Extract keypoints
        else:
            validated_keypoints.append([[0, 0]] * 18)  # Placeholder for missing frames

    keypoints = np.array(validated_keypoints).reshape(-1, 18, 2)  # Shape: (frames, 18, 2)

    # Open video file
    cap = cv2.VideoCapture(video_path)
    frame_idx = 0
    patches_list = []

    while cap.isOpened():
        ret, frame = cap.read()
        if not ret:
            break

        frame_keypoints = keypoints[frame_idx, :15, :]  # Take first 15 keypoints
        patches = extract_patches(frame, frame_keypoints)
        patches_list.append(patches)

        frame_idx += 1

    cap.release()

    # Save patches as .npy file
    patches_list = np.array(patches_list)  # Shape: (frames, 15, patch_height, patch_width, 3)
    np.save(patches_path, patches_list)
    print(f"Patches saved to {patches_path}.")

    return patches_path
```

Figure A.4: Patch Generation

```python
###  Run Seizure Detection Model
def detect_seizure(patches_path, keypoints_path, model):
    """ Run the seizure detection model """
    patches = np.load(patches_path).astype(np.float32)
    patches = torch.tensor(patches).permute(0, 1, 4, 2, 3).float() # (Frames, 15, 3, 32, 32)

    # Load and preprocess keypoints
    with open(keypoints_path, "r") as f:
        keypoints_data = json.load(f)

    keypoints = []
    for frame_kpts in keypoints_data:
        if isinstance(frame_kpts, list) and len(frame_kpts) == 1 and len(frame_kpts[0]) == 18:
            keypoints.append(frame_kpts[0])  # Use valid keypoints
        else:
            print("Warning: Invalid keypoints detected. Adding placeholder keypoints.")
            keypoints.append([[0, 0]] * 18)  # Placeholder for missing keypoints

    keypoints = np.array(keypoints, dtype=np.float32)  # Convert to float32
    keypoints = keypoints.reshape(-1, 15, 2)  # Ensure shape is (Frames, 15, 2)
    keypoints = torch.tensor(keypoints).float()  # Convert to PyTorch tensor

    # FIX: Add a dummy channel of zeros
    dummy_channel = torch.zeros((keypoints.shape[0], keypoints.shape[1], 1), dtype=torch.float32)
    keypoints = torch.cat([keypoints, dummy_channel], dim=-1)  # Shape: (Frames, 15, 3)


    # Match keypoints to patch frames
    #keypoints = keypoints[:patches.shape[0]]
    num_frames = min(len(patches), len(keypoints))
    patches = patches[:num_frames]
    keypoints = keypoints[:num_frames]

    with torch.no_grad():
        seizure_probabilities = model(patches.unsqueeze(0), keypoints.unsqueeze(0))

    return seizure_probabilities.squeeze().tolist()
```

Figure A.5: Seizure Detection

```python
def adaptive_smoothing(probs):
    window_size = 3 if np.mean(probs) < 0.5 else 5  # Dynamic window selection
    return uniform_filter1d(probs, size=window_size)


p=cumulative_prob = np.mean(seizure_probs)
print(cumulative_prob)


# Step 6: Trigger Alert
ALERT_THRESHOLD = 0.75
if cumulative_prob > ALERT_THRESHOLD:
        trigger_alert()
```

alert.py > ...

```python
import winsound
import time
import plyer
from plyer import notification
import tkinter as tk


def loud_siren():
    for _ in range(5):  # Play the sound 5 times
        winsound.Beep(1000, 700)  # Higher-pitched beep
        time.sleep(0.3)  # Short delay
        winsound.Beep(1500, 700)  # Even higher-pitched beep
        time.sleep(0.3)

def show_alert():
    notification.notify(
        title="🚨 SEIZURE ALERT!",
        message="Seizure probability exceeded threshold!",
        timeout=10,  # Display for 10 seconds
    )

def flash_screen():
    root = tk.Tk()
    root.attributes('-fullscreen', True)  # Fullscreen mode
    root.configure(bg='red')  # Red screen
    root.after(2000, root.destroy)  # Close after 2 seconds
    root.mainloop()
```

Figure A.6: Alert

Appendix B

Plagiarism Report

Seizure Detection in Dark Environments

ORIGINALITY REPORT

24%	17%	15%	14%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to APJ Abdul Kalam Technological University, Thiruvananthapuram Student Paper	8%
2	arxiv.org Internet Source	2%
3	www.ecva.net Internet Source	1%
4	doaj.org Internet Source	<1%
5	openaccess.thecvf.com Internet Source	<1%
6	www.catalyzex.com Internet Source	<1%
7	thesai.org Internet Source	<1%
8	moments.csail.mit.edu Internet Source	<1%
9	iieta.org Internet Source	<1%
10	eccv.ecva.net Internet Source	<1%
11	Submitted to Nottingham Trent University Student Paper	<1%
12	Wenjing Wang, Xinhao Wang, Wenhan Yang, Jiaying Liu. "Unsupervised Face Detection in	<1%

	the Dark", IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023 Publication	
13	deepai.org Internet Source	<1 %
14	Johan B. A. M. Arends. "Movement-based seizure detection", Epilepsia, 2018 Publication	<1 %
15	papers.miccai.org Internet Source	<1 %
16	link.springer.com Internet Source	<1 %
17	web.archive.org Internet Source	<1 %
18	www.researchgate.net Internet Source	<1 %
19	www.mdpi.com Internet Source	<1 %
20	export.arxiv.org Internet Source	<1 %
21	Submitted to American University in Cairo Student Paper	<1 %
22	dblp.uni-trier.de Internet Source	<1 %
23	research.monash.edu Internet Source	<1 %
24	www.grafiati.com Internet Source	<1 %
25	www.grin.com Internet Source	<1 %
26	Avinash Paliwal, Libing Zeng, Nima Khademi Kalantari. "Multi-Stage Raw Video Denoising	<1 %

	with Adversarial Loss and Gradient Mask", 2021 IEEE International Conference on Computational Photography (ICCP), 2021 Publication	
27	Submitted to Nanyang Technological University Student Paper	<1 %
28	Submitted to Technical University of Cluj-Napoca Student Paper	<1 %
29	tudr.thapar.edu:8080 Internet Source	<1 %
30	Submitted to Gitam University Student Paper	<1 %
31	cg.postech.ac.kr Internet Source	<1 %
32	iasir.net Internet Source	<1 %
33	eprints.utm.my Internet Source	<1 %
34	Kaixuan Wei, Ying Fu, Jiaolong Yang, Hua Huang. "A Physics-Based Noise Formation Model for Extreme Low-Light Raw Denoising", 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020 Publication	<1 %
35	Uzair Aslam Bhatti, Jingbing Li, Mengxing Huang, Sibghat Ullah Bazai, Muhammad Aamir. "Deep Learning for Multimedia Processing Applications - Volume Two: Signal Processing and Pattern Recognition", CRC Press, 2024 Publication	<1 %
36	blog.roboflow.com Internet Source	

		<1 %
37	www.vcl.fer.hr Internet Source	<1 %
38	Nguyen, Cindy My Anh. "Learning Domain-Specific Cameras", Stanford University, 2024 Publication	<1 %
39	Rundong Luo, Wenjing Wang, Wenhan Yang, Jiaying Liu. "Similarity Min-Max: Zero-Shot Day-Night Domain Adaptation", 2023 IEEE/CVF International Conference on Computer Vision (ICCV), 2023 Publication	<1 %
40	eurchembull.com Internet Source	<1 %
41	mafiadoc.com Internet Source	<1 %
42	scholar.xjtlu.edu.cn Internet Source	<1 %
43	mbcet.ac.in Internet Source	<1 %
44	Leal, João Gonçalo Nunes. "Sistema de Deteção de Quedas Automático Baseado em Vídeo", Instituto Politecnico do Porto (Portugal), 2024 Publication	<1 %
45	paperswithcode.com Internet Source	<1 %
46	escholarship.org Internet Source	<1 %
47	vdoc.pub Internet Source	<1 %

48	"Medical Image Computing and Computer Assisted Intervention – MICCAI 2023", Springer Science and Business Media LLC, 2023 Publication	<1 %
49	"Pattern Recognition and Computer Vision", Springer Science and Business Media LLC, 2024 Publication	<1 %
50	Zhou, Guanglin. "Navigating Distribution Shifts in ML through Causality and Foundation Models.", University of New South Wales (Australia) Publication	<1 %
51	www.ijesr.org Internet Source	<1 %
52	"Achieving Sustainable Business through AI, Technology Education and Computer Science", Springer Science and Business Media LLC, 2025 Publication	<1 %
53	Himanshu Singh, Saurabh Suman, Badri Narayan Subudhi, Vinit Jakhetiya, Ashish Ghosh. "Action Recognition in Dark Videos using Spatio-temporal Features and Bidirectional Encoder Representations from Transformers", IEEE Transactions on Artificial Intelligence, 2022 Publication	<1 %
54	Submitted to Liverpool John Moores University Student Paper	<1 %
55	Philippe Ryvlin, Sándor Beniczky. "Seizure detection and mobile health devices in	<1 %

	epilepsy: Update and future developments", Epilepsia, 2018 Publication	
56	Protter, Matan, Michael Elad, Vivek K. Goyal, and Manos Papadakis. "", Wavelets XII, 2007. Publication	<1 %
57	Sameera Palipana, David Rojas, Piyush Agrawal, Dirk Pesch. "FallDeFi", Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2018 Publication	<1 %
58	Submitted to The Robert Gordon University Student Paper	<1 %
59	Wenjing Wang, Wenhan Yang, Jiaying Liu. "HLA-Face: Joint High-Low Adaptation for Low Light Face Detection", 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021 Publication	<1 %
60	ir.lib.uwo.ca Internet Source	<1 %
61	www.csauthors.net Internet Source	<1 %
62	"Pattern Recognition", Springer Science and Business Media LLC, 2025 Publication	<1 %
63	Advances in Intelligent Systems and Computing, 2016. Publication	<1 %
64	Guanzhi Ding, Peng Gao, Haojie Wu. "A Domain Confusion Adaptation Method for Semi-supervised Action Recognition in Dark Environments", 2023 4th International Symposium on Computer Engineering and Intelligent Communications (ISCEIC), 2023	<1 %

Publication

65	Ikram Kourbane, Panagiotis Papadakis, Mihai Andries. "Optimized assessment of physical rehabilitation exercises using spatiotemporal, sequential graph-convolutional networks", Computers in Biology and Medicine, 2025 Publication	<1 %
66	Mingfang Zhang, Yifei Huang, Ruicong Liu, Yoichi Sato. "Chapter 18 Masked Video andBody-Worn IMU Autoencoder forEgocentric Action Recognition", Springer Science and Business Media LLC, 2025 Publication	<1 %
67	Wenhao Li, Guangyang Wu, Wenyi Wang, Peiran Ren, Xiaohong Liu. "FastLLVE: Real-Time Low-Light Video Enhancement with Intensity-Aware Look-Up Table", Proceedings of the 31st ACM International Conference on Multimedia, 2023 Publication	<1 %
68	Xiaoyang shen, Haibin Li, Yaqian Li, Wenming Zhang. "Low-Light Image Enhancement Guided by Multi-Domain Features for Detail and Texture Enhancement", Digital Signal Processing, 2024 Publication	<1 %
69	core.ac.uk Internet Source	<1 %
70	www.arxiv-vanity.com Internet Source	<1 %
71	"Proceedings of the International Conference on Computer Science, Electronics and Industrial Engineering (CSEI 2023)", Springer Science and Business Media LLC, 2024 Publication	<1 %

72	Altaf Hussain, Samee Ullah Khan, Noman Khan, Imad Rida, Meshal Alharbi, Sung Wook Baik. "Low-light aware framework for human activity recognition via optimized dual stream parallel network", Alexandria Engineering Journal, 2023 Publication	<1 %
73	Minse Ha, Wan-Gi Bae, Geunyoung Bae, Jong Taek Lee. "Chapter 2 ELLAR: An Action Recognition Dataset forExtremely Low-Light Conditions withDual Gamma Adaptive Modulation", Springer Science and Business Media LLC, 2025 Publication	<1 %
74	Umut Kuran, Emre Can Kuran. "Parameter Selection for CLAHE using Multi-objective Cuckoo Search Algorithm for Image Contrast Enhancement", Intelligent Systems with Applications, 2021 Publication	<1 %
75	Wenhan Yang, Shiqi Wang, Yuming Fang, Yue Wanga, Jiaying Liu. "Band Representation-Based Semi-Supervised Low-Light Image Enhancement: Bridging the Gap Between Signal Fidelity and Perceptual Quality", IEEE Transactions on Image Processing, 2021 Publication	<1 %
76	Zhigang Tu, Yuanzhong Liu, Yan Zhang, Qizi Mu, Junsong Yuan. "DTCM: Joint Optimization of Dark Enhancement and Action Recognition in Videos", IEEE Transactions on Image Processing, 2023 Publication	<1 %
77	Chen Chen, Qifeng Chen, Minh Do, Vladlen Koltun. "Seeing Motion in the Dark", 2019	<1 %

	IEEE/CVF International Conference on Computer Vision (ICCV), 2019 Publication	
78	Gayathri T, Mamatha HR. "How to Improve Video Analytics with Action Recognition: A Survey", ACM Computing Surveys, 2024 Publication	<1 %
79	Xiaoke Shang, Gehui Li, Zhiying Jiang, Shaomin Zhang, Nai Ding, Jinyuan Liu. "Holistic Dynamic Frequency Transformer for image fusion and exposure correction", Information Fusion, 2023 Publication	<1 %
80	Yuecong Xu, Haozhi Cao, Jianxiong Yin, Zhenghua Chen, Xiaoli Li, Zhengguo Li, Qianwen Xu, Jianfei Yang. "Going Deeper into Recognizing Actions in Dark Environments: A Comprehensive Benchmark Study", International Journal of Computer Vision, 2023 Publication	<1 %

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	Off		