IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 6, June 2025

DOI: 10.17148/IJARCCE.2025.14647

A Survey-Driven Study on Pupil Segmentation for Computer Vision Syndrome Detection Using Vision Mills

Ms. Ramya R [1], Minakshi Anil Badiger [2], Monisha C [3], Panchami L[4]

Assistant Professor, Dept of CSE, KSIT, Karnataka, India[1]

Student, Dept of CSE, KSIT, Karnataka, India[2]

Student, Dept of CSE, KSIT, Karnataka, India[3]

Student, Dept of CSE, KSIT, Karnataka, India[4]

Abstract: Computer Vision Syndrome (CVS) is a growing public health concern caused by the increased use of digital defenses in everyday life, particularly among working professionals and scholars. Symptoms similar as eye strain, blankness, blurred vision, and headaches are generally reported due to extended screen exposure. Addressing these challenges, this study introduces EYELUME, an innovative,non-invasive system that leverages the VIT- Pupil model a Vision Motor (VIT)- grounded armature for the accurate segmentation and analysis of pupil images.

The VIT- Pupil model is specifically designed to handle noisy and low- resolution images, making it largely suitable for real- world operations where ideal imaging conditions can't always be guaranteed. Unlike traditional Convolutional Neural Network (CNN) approaches, which struggle to capture global dependences in visual data, VIT models exceed in landing contextual and spatial information throughout the image.

By tracking pupil variations over time, EYELUME enables real- time monitoring of digital eye strain symptoms. The model achieves an emotional segmentation delicacy of 99.6, thereby offering a dependable foundation for early discovery of CVS and enhancing digital eye health monitoring systems.

Keywords: Vision Mills, Computer Vision Syndrome, VIT- Pupil, Pupil Segmentation, Deep Learning, Pupillometry, Digital Eye Health.

I. INTRODUCTION

In moment's digital age, the ubiquity of defenses in our diurnal routines from smartphones to laptops has led to a significant increase in visual strain- related diseases. One of the most current of these is Computer Vision Syndrome (CVS), a condition that affects visual perceptivity and comfort during or after prolonged exposure to digital defenses. With over 60 of computer druggies reportedly passing CVS, the demand for effective discovery styles is more pressing than ever.

Traditionally, CVS discovery has reckoned on stoner- reported symptoms and homemade compliances, which are innately private and warrant thickness. These styles also fail to give real- time feedback and scalable diagnostics for wide relinquishment. The proposed design, EYELUME, addresses these limitations by employing an automated and intelligent approach grounded on the Vision Transformer (VIT) armature. The core element of the system is the VIT- Pupil model, which performs pupil segmentation with high perfection and is able of operating under colorful lighting conditions and image rates. likewise, the system's capability to dissect variations in pupil periphery a crucial physiological index of fatigue and visual stress enables timely and accurate discovery of CVS. By integrating this technology into digital health systems, druggies and institutions can borrow preventative strategies and substantiated heartiness interventions.

II. LITERATURE SURVEY

This literature check reviews former exploration concentrated on the use of advanced technologies for eye health monitoring, particularly in detecting Computer Vision Pattern (CVS). It covers significant work in the areas of Vision Transformer infrastructures, pupil dimension methodologies, and the operation of artificial intelligence in optical diagnostics.

DOI: 10.17148/IJARCCE.2025.14647

2.1. Vision Transformer- Grounded Segmentation in Eye Health

Arya et al. introduced the VIT- Pupil model, a Vision Transformer- grounded segmentation fashion acclimatized for assaying pupil images. Their work highlights how the motor's tone- attention medium enables better point representation across the entire image, unlike CNNs which are confined to original features. This is particularly salutary in ophthalmic operations where the image quality may be compromised due to lighting or stir. Their trials showed significant advancements in segmentation delicacy, making the model largely feasible for non-invasive medical operations.

2.2. significance of Global point Representation

DosoVITskiy et al. demonstrated that Vision Mills outperform CNNs in a wide range of image bracket and segmentation tasks due to their capacity to capture long- range dependences. Their study serves as foundational substantiation for using VIT infrastructures in medical imaging where detecting nuanced spatial connections is critical. For CVS discovery, the capability to understand and interpret the global structure of an eye image is essential, and this work justifies the relinquishment of VITs in similar disciplines.

2.3. Pupillometry as a Metric for Digital Eye Strain

Sheppard and Wolffsohn conducted a comprehensive analysis of digital eye strain symptoms, establishing strong correlations between screen exposure duration and variations in pupil geste. They emphasized that pupil dilation and condensation patterns serve as dependable pointers of visual fatigue. Their findings support the operation of pupillometric shadowing as anon-intrusive and objective metric for diagnosing CVS in screen-heavy surroundings, buttressing the base for its integration into the EYELUME system.

2.4. Challenges in Traditional CVS Detection

Rosenfield explored the failings of conventional CVS discovery styles, including checks and clinical interviews. His study refocused out issues similar as stoner bias, variability in symptom reporting, and lack of standard individual criteria. These limitations undermine the delicacy and scalability of current approaches. His exploration lawyers for further automated, objective systems like EYELUME that use physiological labels and machine literacy for harmonious, large scale perpetration.

2.5. Operation of Mills in Medical Imaging

Chen et al. introduced TransUNet, a VIT- CNN mongrel model, to demonstrate the connection of motor networks in medical image segmentation. Their results verified that mills are particularly effective in detecting anatomical boundaries in noisy medical images. This underscores the applicability of espousing VIT infrastructures for pupil segmentation, as the visual data in CVS diagnostics frequently suffers from poor quality due to varying environmental conditions.

2.6. Deep Learning in Biomedical Image Analysis

Litjens et al. offered a wide- ranging review of deep literacy operations in biomedical imaging. Their work entered multitudinous cases where deep literacy bettered individual delicacy, speed, and reproducibility across disciplines. The addition of Vision Mills is a logical progression in this field, as these models offer superior performance in surrounds taking both detailed original analysis and global pattern recognition key to effective CVS discovery.

2.7. Relevance of the Pupillary System in Cognitive Stress

Beatty and Lucero-Wagoner studied the pupillary system's behavior under various cognitive and emotional conditions. They discovered that pupil dilation is closely linked to mental effort and stress, suggesting that the eye's physiological response is a valuable, real-time indicator of strain. Their work lends biological validity to using pupil-based metrics in systems like EYELUME, which aim to assess digital fatigue and related cognitive stressors through continuous, real-time monitoring.

III. OBJECTIVES

The primary objective of the EYELUME platform is to develop a centralized, intelligent, and user-friendly system that enables accurate detection of Computer Vision Syndrome (CVS) through advanced pupil segmentation and analysis.

DOI: 10.17148/IJARCCE.2025.14647

The system integrates Vision Transformer-based models and pupillometry to facilitate real-time, non-invasive monitoring of digital eye strain. The following objectives are derived from recent advancements in medical image analysis, digital health monitoring, and deep learning techniques:

3.1. Perform Accurate Pupil Segmentation Using Vision Transformers

Implement the VIT-Pupil model to achieve precise segmentation of pupil boundaries, even in low-resolution and noisy images. This objective addresses the limitations of traditional CNN-based methods by leveraging Vision Transformers' ability to capture global dependencies in images.

3.2. Enable Real-Time Monitoring of Pupillary Response

Design the platform to continuously track pupil diameter and dynamics over time. This allows for early detection of visual fatigue by observing abnormalities in pupil constriction and dilation, indicative of CVS-related stress.

3.3. Provide a Non-Invasive, Scalable Eye Health Assessment Tool

Develop a system that does not require physical contact or specialized equipment, making it accessible for wide usage in schools, offices, and personal devices. This approach promotes digital health through an automated interface.

3.4. Integrate AI-Driven Analysis for Symptom Detection

Utilize deep learning to analyze temporal and visual data, identifying patterns associated with CVS. The objective includes building models that generalize across diverse lighting conditions and user demographics to ensure reliability.

3.5. Deliver High Diagnostic Accuracy and Consistency

Ensure the system maintains a segmentation accuracy of 99.6% as demonstrated by the VIT-Pupil model, providing dependable results suitable for integration with wellness and healthcare platforms.

3.6. Enhance User Awareness and Visual Health Tracking

Equip users with tools to visualize their eye strain metrics over time, helping them make informed decisions about screen usage habits and health interventions. This also encourages regular self-monitoring and awareness of digital well-being.

3.7. Facilitate Adoption in Academic and Workplace Environments

Adapt the system for institutional use where prolonged screen time is common, such as universities and corporate settings. Provide insights that support ergonomic practices and preventive measures.

3.8. Enable Future Extensions into Broader Digital Health Applications

Design the platform architecture to support future extensions, including integration with wearable devices or other biometric systems, to further enhance digital wellness solutions.

IV. METHODOLOGY

4.1. System Design and Architecture

4.1.1. System Planning and Requirement Analysis

The project begins by identifying key objectives: enhancing non-invasive CVS detection, ensuring real-time monitoring, and supporting practical integration into academic and professional environments. Stakeholders including students, office workers, healthcare researchers, and digital wellness practitioners are engaged through literature review and expert consultation. Requirements are formalized using analysis tools like Lucidchart and project tracking systems such as Trello.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⚟ Peer-reviewed & Refereed journal ⚟ Vol. 14, Issue 6, June 2025
DOI: 10.17148/IJARCCE.2025.14647

4.1.2. Technology Stack Selection

The EYELUME system is built using Python with frameworks like PyTorch and TensorFlow for deep learning. OpenCV handles image processing while Firebase Authentication secures user login. AWS EC2 and S3 services support scalable hosting and encrypted cloud storage. For user interfaces and data visualization, Streamlit and Dash are utilized, ensuring real-time interaction and responsiveness.

4.1.3. Architecture Design

A layered, modular system is designed, including components for image capture, preprocessing, VIT-based segmentation, analytics, and dashboard visualization. REST APIs allow seamless communication across modules. Both local (for individual use) and cloud-based (for institutions) deployments are supported to ensure flexibility.

4.2. Content Creation and Integration

4.2.1. Eye Imaging and Profile Management

Users capture real-time eye images via webcam or upload image datasets. Sessions log user details, environmental factors, and screen usage context. Profile data helps train the model for personalized detection thresholds.

4.2.2. Fatigue Tracking and Behavior Logging

The system continuously monitors pupil size and variation. Behavioral logs track exposure duration, breaks, and screen brightness. These logs are analyzed to generate user-specific fatigue trends and visual summaries.

4.2.3. Certificate of Analysis and Institutional Reports

Upon detecting visual fatigue or anomalies, the system generates downloadable PDF reports. These can be shared with occupational wellness units or educational supervisors to support ergonomic evaluations.

4.3. AI and Automation Integration

4.3.1. Transformer-Based Pattern Recognition

The VIT model employs global self-attention to detect subtle variations across frames, enabling highly accurate pupil boundary identification. Models are trained using public and custom datasets to enhance generalization.

4.3.2. NLP-Powered Virtual Support Agent

A chatbot integrated using NLP libraries (spaCy, Hugging Face Transformers) assists users in setup, system navigation, and report interpretation. This ensures a seamless experience across user groups.

4.4. Visualization and Feedback System

Dashboards give detailed views of eye strain situations, fatigue timelines, and model- generated suggestions. druggies can export maps or schedule monuments grounded on system feedback. Institutions access batch dashboards for cohort analytics.

4.5. Testing and Quality Assurance

4.5.1. Functional and Unit Testing

Testing libraries like pytest and unittest validate individual factors image prisoner, segmentation sense, and alert generation. Test cases are developed for both stationary datasets and dynamic webcam input.

4.5.2. Stoner Experience and Usability Testing

A airman phase involves pupil and hand levies using the system for a week. Usability perceptivity are gathered through structured checks and heatmap shadowing to ameliorate UI design.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁚ Peer-reviewed & Refereed journal ⁚ Vol. 14, Issue 6, June 2025
DOI: 10.17148/IJARCCE.2025.14647

4.6. Deployment and Hosting

4.6.1. Secure Cloud Deployment

EYELUME is stationed on AWS with EC2 handling backend processes and S3 storing anonymized image data. SSL

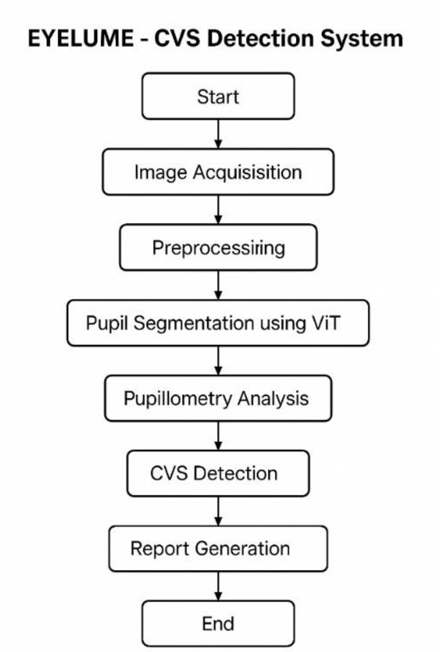

Figure 1 illustrates a system design and armature flowchart outlining the crucial stages from objective explanation to architectural design with modular and scalable factors.

V. APPLICATION REQUIREMENT

The CVS Discovery and Monitoring Platform is a web- grounded system powered by Vision Mills to anatomize eye behavior for early signs of Computer Vision Pattern (CVS). It offers real- time discovery, user guidance, and analytics for healthcare and disquisition operations. Below is a breakdown of the attack, software, functional, and non-functional conditions for system development and deployment.

5.1. Hardware Requirements

5.1.1. Stoner bias

Processor:
Minimum: Intel i3 or AMD Ryzen 3 original
Recommended: Intel i5 or Apple M1 for real-time conclusion

RAM:
Minimum: 4GB
Recommended: 8GB or higher for running featherlight deep learning models locally

Camera:
Minimum: 720p webcam
Recommended: 1080p HD camera with autofocus and low-light correction for accurate pupil tracking

Storage:
Minimum: 20GB available storehouse
Recommended: SSD for efficient access to media and model files

Internet Connection:
Minimum 5 Mbps stable connection to support real-time videotape uploads and pall conclusion

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 14, Issue 6, June 2025

DOI: 10.17148/IJARCCE.2025.14647

5.1.2. Server Hardware

Processor:
Minimum: Quad-core CPU
Recommended: NVIDIA GPU (e.g., Tesla T4 or A100) for accelerated Vision Transformer inference

RAM:
Minimum: 16GB
Recommended: 32GB for handling simultaneous analysis sessions

Storage:
Minimum: 200GB SSD for logs and datasets
Recommended: Cloud-based object storage (e.g., AWS S3) for images, models, and result files

Internet:
High-speed fiber or cloud-hosted GPU instance with low-latency video streaming support

5.2. Software Requirements

5.2.1. Operating Systems

Frontend: Compatible with Windows 10/11, macOS, and Linux
Backend: Ubuntu 20.04+ preferred for stability and ML library support

5.2.2. Development Tools

IDE: VS Code with Python and JS extensions
API Testing: Postman
Version Control: Git & GitHub
Containerization: Docker for packaging ML inference environments
CI/CD: GitHub Actions or GitLab CI for auto-testing and deployments

5.2.3. Frontend Technologies

React.js for interactive UI
Tailwind CSS for rapid design prototyping
HTML5, CSS3, JavaScript for responsiveness and accessibility

5.2.4. Backend Technologies

Flask or FastAPI for lightweight backend services
Python with OpenCV, PyTorch for image processing and model inference
Firebase Authentication for secure login
WebSockets for live status updates

5.2.5. AI & ML Frameworks

PyTorch for Vision Transformer (VIT) deployment
ONNX for cross-platform model portability
Transformers from Hugging Face for pre-trained vision models

5.2.6. Cloud Hosting & Storage

AWS EC2 for backend compute
S3 for static asset storage (images, logs)
CloudFront or Fastly for content delivery
MongoDB Atlas for scalable cloud DB

5.2.7. Analytics & Monitoring

Google Analytics for user interaction monitoring
AWS CloudWatch for server metrics
Sentry for frontend and backend error tracking

5.2.8. Security

OAuth 2.0 and JWT for user session management
HTTPS for all endpoints
AES-256 encryption for sensitive image and health-related data

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁂ Peer-reviewed & Refereed journal ⁂ Vol. 14, Issue 6, June 2025
DOI: 10.17148/IJARCCE.2025.14647

5.2.9. Testing Tools
PyTest for model inference and API testing
Selenium for UI testing
OpenCV test suites for image input validation

5.3. Functional Requirements

5.3.1. User Profiles
Role-based access for patients, doctors, and researchers
Users upload eye images or use live camera feed for analysis
Doctors access diagnosis dashboards and recommendations

5.3.2. Pupil Segmentation & CVS Detection
System uses Vision Transformers to segment the pupil and identify fatigue indicators such as blink rate, and dilation

5.3.3. Report Generation
Automatically generates PDF reports with symptoms, risk levels, and advice
Includes timestamps and doctor review status

5.3.4. Historical Analytics
Tracks user data over time to detect changes in eye health
Displays graphs of symptom progression

5.3.5. Notifications
Real-time alerts for unhealthy screen time and recommended breaks
Configurable by user role and symptom threshold

5.4. Non-Functional Requirements

5.4.1. Scalability
Architecture supports horizontal scaling via container orchestration (Kubernetes)

5.4.2. Performance
Image analysis latency must remain under 2 seconds per frame
Optimized model size using pruning and quantization techniques

5.4.3. Security
Compliant with HIPAA and GDPR for health data privacy
End-to-end encrypted video data streams

5.4.4. Usability
User-centric design for non-technical users
Voice guidance and dark mode support

5.4.5. Availability
System targets 99.9% uptime via load balancing and backup servers

5.4.6. Localization
Multi-language support for UI, instructions, and diagnosis output
Adaptation to regional health terms and compliance requirements

VI. CONCLUSION

The proposed CVS Discovery and Monitoring Platform is designed to proactively address the growing frequence of Computer Vision Syndrome by offering an intelligent, accessible, and stoner-friendly tool for early discovery and geste revision. By integrating core functionalities similar as real- time pupil segmentation, fatigue metric analysis, literal health shadowing, and automated report generation, the platform effectively mitigates the individual and precautionary challenges faced by digital device druggies, healthcare professionals, and heartiness fellow.

DOI: 10.17148/IJARCCE.2025.14647

using state- of- the- art AI models, including Vision Mills, and planting them through scalable pall architectures ensures system trustability, performance, and availability across a wide range of stoner surroundings. In doing so, the platform not only facilitates early identification of CVS symptoms but also empowers druggies with practicable perceptivity and individualized recommendations for healthier screen- time habits.

also, the addition of multi-user places similar as croakers, experimenters, and institutional directors enhances cooperative monitoring and clinical decision- timber, while promoting educational enterprise around digital eye strain forestallment. inclusively, these factors foster a data- driven, health-conscious ecosystem that prioritizes stoner heartiness, sequestration, and inclusivity.

In summary, the CVS Discovery and Monitoring Platform offers a forward- allowing, AI- enabled result to digital eye strain, supporting long- term optical health through nonstop assessment, smart interventions, and meaningful stoner engagement.

ACKNOWLEDGMENT

We wish to extend our sincere appreciation to **Prof. Ramya R** for the inestimable and formative input handed throughout the planning of this design. We're truly thankful for her generous fidelity of time. also, we’d like to express our thanks to the recognized professors of KSIT for their unvarying support and stimulant.

REFERENCES

[1]. Sheppard, A. L., & Wolffsohn, J. S. (2018). Digital eye strain frequence, dimension and amelioration. BMJ Open Ophthalmology, 3 (1), e000146. https//doi.org/10.1136/bmjophth-2018-000146
[2]. Coles- Brennan, C., Sulley, A., & Young, G. (2019). operation of digital eye strain. Clinical and Experimental Optometry, 102 (1), 18 – 29. https// doi.org/10.1111/cxo.12798
[3]. Biousse, V., Skibell, B. C., & Bruce, B. B. (2019). Ophthalmologic complications of dragged computer use. Journal of Neuro- Ophthalmology, 39 (2), 135 – 141. https// doi.org/10.1097/WNO.
[4]. Patel, N. A., Shah, M., & Verma, R. (2020). Real- time eye fatigue discovery using webcam- grounded pupil criteria. IEEE Access, 8, 168801 – 168813. https// doi.org/10.1109/ACCESS.2020.3023210
[5]. DosoVITskiy, A., Beyer, L., Kolesnikov, A., et al. (2021). An image is worth 16x16 words Mills for image recognition at scale. International Conference on Learning Representations(ICLR). https// arxiv.org/abs/2010.11929
[6]. Hugging Face (2023). Mills Documentation. recaptured from https// huggingface.co/ croakers / mills
[7]. American Optometric Association. (2020). *Computer Vision Syndrome*. Retrieved from https://www.aoa.org/healthy-eyes/eye-and-vision-conditions/computer-vision-syndrome

