International Journal of Science, Strategic Management and Technology
Volume **02** Issue **03** March-**2026** | ISSN: **3108-1762 (Online)** | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

Multimodal Human Digital Twin AI for Real Time Productivity and Fatigue Intelligence

Ratchita B[1], Dr. S.V. Anandhi[2]
[1] Student, Dept. of Artificial Intelligence and Data Science, Ramco Institute of Technology Rajapalayam ,India
ratchitab21@gmail.com
[2] Associate Professor, Dept. of Artificial Intelligence and Data Science,Ramco Institute of Technology, Rajapalayam,India,
svanandhi2020@gmail.com

https://doi.org/10.55041/ijsmt.v2i3.206

Cite this Article: B, R. (2026). Multimodal Human Digital Twin AI for Real Time Productivity and Fatigue Intelligence. International Journal of Science, Strategic Management and Technology, *02*(03). https://doi.org/10.55041/ijsmt.v2i3.206

ABSTRACT— This project presents a Multi-modal Human Digital Twin System designed to continuously model human cognitive, emotional, and physical states through integrated vision, audio, and behavioral data. The system captures real-time facial expressions, voice characteristics, and interaction patterns to construct a comprehensive and dynamic digital representation of an individual. Deep learning techniques are employed to extract meaningful features from each modality, while advanced multi-modal fusion methods combine these features to capture complex inter-dependencies across human behaviors. Transformer-based and sequence modeling architectures analyze temporal patterns to estimate productivity, fatigue, and overall well-being. The framework supports both short-term monitoring and long- term trend analysis, enabling predictive assessment of performance degradation and health risks. Real-time insights, fatigue alerts, and personalized recommendations are generated to support proactive interventions. An interactive visualization dashboard presents interpret-able analytic, facilitating informed decision-making for individuals and organizations. By enabling continuous, accurate human state modeling, the proposed system aims to enhance productivity, improve well-being, and promote sustainable work practices through intelligent, data-driven support.
Keywords— Affective Computing, Multi-modal Fusion, Deep Learning, Transformer Networks, Temporal Sequence Modeling, Emotion Recognition, Fatigue Detection, Productivity Analysis, Behavioral Analysis, Predictive Modeling, Artificial Intelligence

I. INTRODUCTION

In recent years, rapid advancements in artificial intelligence, sensor technologies, and data analytics have transformed the way human behavior and performance are understood and monitored. Modern work environments increasingly demand sustained cognitive effort, adaptability, and high productivity, often leading to mental fatigue, stress, and long-term health challenges. Traditional methods for assessing human performance and well-being rely heavily on self-reports, periodic evaluations, or isolated physiological measurements, which are often subjective, intrusive, or insufficient for continuous monitoring. As a result, there is a growing need for intelligent systems capable of continuously observing, analyzing, and interpreting human states in a non-invasive and data-driven manner.

The concept of a Human Digital Twin has emerged as a promising solution to address these challenges. A Human Digital Twin refers to a virtual representation of an individual that mirrors their physical, cognitive, and emotional states in real time. By leveraging multi-modal data sources such as visual cues, audio signals, and behavioral interactions, a digital twin can provide a holistic understanding of human conditions and performance. Unlike static models, a dynamic digital twin continuously evolves based on incoming data, enabling accurate monitoring, prediction, and personalized feedback. This paradigm has gained significant attention in fields such as healthcare, human–computer interaction, occupational safety, and smart work-spaces.

Vision-based data plays a crucial role in understanding human emotions and physical states. Facial expressions, eye movements, head posture, and micro-expressions offer valuable insights into fatigue, stress, engagement, and emotional well-being. Computer vision techniques combined with deep learning models have demonstrated strong capabilities in recognizing facial action units, detecting drowsiness, and estimating attention levels. When continuously captured through cameras.

Behavioral data, including keyboard usage, mouse movements, interaction frequency, task-switching behavior, and activity patterns, provides an additional layer of insight into productivity and work habits. These behavioral cues reflect how individuals interact with digital systems and how efficiently tasks are performed over time. Changes in these interaction patterns may indicate fatigue, reduced concentration, or declining performance. By continuously monitoring behavioral interactions, the system can identify long-term trends and detect deviations from normal performance baselines. This information helps in understanding user behavior and enables the system to provide timely insights and recommendations for improving productivity and maintaining optimal performance.

Audio data serves as another vital modality in human state analysis. Speech patterns, tone, pitch, energy, and speaking rate are closely linked to emotional and mental states. Variations in voice characteristics can indicate stress, fatigue, cognitive load, or reduced alertness. Advances in speech processing and deep learning have enabled robust extraction of acoustic and prosodic features from voice signals, even in real- world environments. Incorporating audio data enhances the system's ability to detect subtle changes that may not be visible through vision alone.

While each modality offers valuable information independently, human states are inherently complex and cannot be accurately inferred from a single data source. Multimodal fusion techniques address this limitation by integrating vision, audio, and behavioral features into a unified representation. Deep learning–based fusion models are capable of capturing interdependencies and correlations across modalities, leading to more robust and reliable predictions. This integrated approach significantly improves the accuracy of cognitive and fatigue estimation compared to unimodal systems.

Temporal dynamics are also essential for understanding human behavior. Human states such as fatigue and productivity evolve over time and are influenced by both short-term conditions and long-term patterns. Transformer-based architectures and sequence modeling techniques enable the system to analyze temporal dependencies across multimodal data streams. These models support both real-time state estimation and predictive analysis, allowing early detection of performance degradation and potential health risk

The proposed Multimodal Human Digital twin System aims to combine these technological advancements into the unified framework for continuous human state modeling .The system processes real-time multimodal data to generate actionable insights, including fatigue alerts, productivity, assesments and personalised recommendations.An interactive visualizations dashboard presents these insights in an intuitive and interpretable manner, enabling individuals and organisations to take proactive measures.

In summary, the integration of multimodal sensing, deep learning, and digital twin technology offers a powerful approach to understanding and supporting human performance. This project focuses on designing a comprehensive and intelligent system that accurately represents human cognitive, emotional, and physical states. By enabling continuous monitoring, predictive assessment, and personalized feedback, the proposed framework contributes to the development of human-centered intelligent systems that prioritize both productivity and well-being.

International Journal of Science, Strategic Management and Technology
Volume **02** Issue **03** March-**2026** | ISSN: **3108-1762 (Online)** | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

II. LITERATURE SURVEY

A Vision-Enabled Fatigue-Sensitive Human Digital Twin (2024)

1) Authors: Anonymous (Journal of Manufacturing Systems, 2024)

2) Methodology: This study introduces a Human Digital Twin framework designed for human-centric human-robot collaboration. The system integrates vision algorithms to segment and classify repetitive assembly actions from live video. A non-contact physical fatigue assessment is achieved by combining vision-based tracking with personalized muscle fatigue profiles. Cognitive fatigue is monitored using smartwatch physiological signals, enabling real-time fatigue evaluation in dynamic manufacturing settings. The framework was extensively tested on a real 26-task assembly process, demonstrating effective tracking of both physical and cognitive fatigue metrics. By integrating multiple sensing modalities, the system supports contextual task allocation and workload optimization. The overall aim was to create a fatigue-aware digital twin that informs human–robot interaction for improved safety and productivity.

B. Construction of Human Digital Twin Model Based on Multimodal Data (2023)

1) Authors: Ruirui Zhong, Bingtao Hu, Yixiong Feng, et al.

2) Methodology: This research proposes a multimodal Human Digital Twin model emphasizing real-time human state acquisition in intelligent manufacturing. A multimodal data acquisition system gathers diverse signals (e.g., IMU, depth camera, EMG, plantar pressure), which are dynamically processed and digitized. The corefusion network combines Bidirectional Long Short-Term Memory (BiLSTM) and Convolutional Neural Networks (CNN) to extract spatiotemporal features from heterogeneous sources. The system's application was demonstrated through human locomotion mode identification, comparing results with traditional models. Experiments showed superior performance for multimodal deep learning in identifying locomotion states and modeling human behavior within a cyber-physical production system. This framework emphasizes dynamic updating of human digital representations for process optimization.

C. Building a Human Digital Twin Using Large Language Models (2024)

1) Authors: (JMIR Formative Research, 2024)

2) Methodology: This work presents HDTwin, a framework leveraging Large Language Models (LLMs) to integrate heterogeneous digital data for modeling human cognitive states, particularly for mild cognitive impairment diagnosis. Diverse data sources — including speech, digital behavior markers, self-reports, and clinical assessments — are fused by the LLM to construct cohesive individual profiles. The system processes time-stamped and multimodal inputs to generate diagnostic predictions comparable or superior to traditional machine learning ensemble models. The framework demonstrates the potential of LLMs for multimodal data retrieval, fusion, and interpretability, making digital twins useful for clinical cognitive diagnostics.

D. Human Digital Twin Framework and Perspectives in Human Factors (2024)

1) Authors: Q. He, L. Li, D. Li, et al.

2) Methodology: This article explores the transition from traditional digital human modeling to dynamic Human Digital Twin systems, focusing on human factors engineering. It proposes a conceptual framework emphasizing realistic human behavior, performance, and interaction influences on system design. The study discusses foundational technologies such as AI, sensor networks, real-time feedback mechanisms, and cognitive state modeling. It outlines perspectives on integrating human factors into digital twin systems for safety, ergonomics, and user experience. Applications span domains including manufacturing, healthcare, and smart environments, with a strong focus on improving human-machine symbiosis through continuous state representation.

E. Human Digital Twins: Systematic Review for Industry 5.0 (2024)

1) Authors: (ScienceDirect article, 2024)

2) Methodology: This systematic literature review provides a cross-domain analysis of Human Digital Twin developments, particularly for Industry 5.0. It clarifies conceptual distinctions between digital models, shadows, and full human digital twins. The paper summarizes enabling technologies — including IoT, edge computing, AI, and data fusion — and highlights core challenges impeding adoption. The review also categorizes HDT approaches and evaluates trends across healthcare, industry, and daily life applications, drawing a general definition and architecture for human digital

twins. The work emphasizes the need to harmonize human behavior, cognition, and social factors within digital twin systems.*F. Secure Storage Allocation Using Fuzzy Based Heuristic Algorithm (Cloud Security)*

1) Authors: M. Sivam, M. Kaliappan, S. J. Shobana, V. Prakash, V. Porkodi

2) Methodology: This research proposes a secure storage allocation scheme for cloud environments using a fuzzy-based heuristic algorithm. The system aims to improve storage allocation efficiency while maintaining security and reliability in distributed cloud infrastructures. The fuzzy logic mechanism evaluates multiple parameters such as storage availability, access frequency, and system load to determine optimal data placement. By applying heuristic optimization, the approach enhances resource utilization and minimizes latency in cloud storage operations. The model also improves decision-making in dynamic cloud environments where storage demand frequently changes.

G. Load Balanced Clustering Technique in MANET using Genetic Algorithms

1) Authors: M. Kaliappan, E. Mariappan, M. V. Prakash, B. Paramasivan

2) Methodology**:**This research introduces a load-balanced clustering technique for Mobile Ad-hoc Networks (MANET) using Genetic Algorithms. The proposed method optimizes cluster head selection to ensure balanced network load distribution among nodes. Genetic algorithms are applied to search for optimal cluster configurations by evaluating parameters such as node mobility, residual energy, and communication cost. The algorithm improves network stability, reduces packet loss, and enhances routing efficiency in dynamic MANET environments where nodes frequently change position.

H. Glaucoma Progression Detection Using K-Means and GLCM Algorithm

1) Authors: S. Vimal, Y. H. Robinson, M. Kaliappan, K. Vijayalakshmi, S. Seo

2) Methodology**:**This study proposes a smart medical prediction system to detect the progression of glaucoma using image processing and machine learning techniques. The approach combines K-Means clustering for segmenting retinal images with the Gray Level Co-occurrence Matrix (GLCM) algorithm for extracting texture features. These extracted features are then analyzed to identify glaucoma patterns and disease progression. The system improves early diagnosis and supports medical professionals in detecting glaucoma at earlier stages, potentially preventing vision loss.

I. Analyzing Public Sentiment on Demonetization Using SVM

1)Authors: M. Kaliappan, B. Guruprakash, J. Rajalakshmi, T. Blessing Karunya, E. Mariappan, M. Ramnath, R. Angel Hepzibah

2)Methodology:This research focuses on analyzing public sentiment regarding demonetization using machine learning techniques. The study collects textual data from online sources such as social media and news platforms. Natural Language Processing (NLP) techniques are applied to preprocess the text, including tokenization, stop-word removal, and feature extraction. A Support Vector Machine (SVM) classifier is used to categorize sentiments into positive, negative, or neutral opinions. The system helps analyze public reaction to economic policies and provides insights into societal perception of financial reforms.

III. SYSTEM PROPOSAL

A. EXISTING SYSTEM

In existing human state monitoring and productivity assessment systems, traditional machine learning techniques such as Support Vector Machines (SVM) are commonly used for classification and prediction tasks. These systems typically rely on handcrafted features extracted separately from vision, audio, or behavioral data, rather than a unified multimodal approach. SVM-based models perform adequately for small and well-defined datasets but struggle to capture complex nonlinear relationships present in real-world human behavior. Moreover, they lack the ability to model temporal dependencies and adapt to continuous, real-time data streams. As a result, existing systems provide limited accuracy, poor scalability, and insufficient insight into dynamic cognitive and fatigue states, making them unsuitable for comprehensive human digital twin applications.

International Journal of Science, Strategic Management and Technology
Volume 02 Issue 03 March-2026 | ISSN: 3108-1762 (Online) | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

1) DISADVANTAGES:

- SVM models depend on manually engineered features, which fail to capture high- level patterns in complex multimodal human data such as emotions, fatigue, and behavioral variations.
- SVM-based systems do not scale well with large, continuous datasets and require frequent retraining, making them inefficient for real-time and long-term monitoring.
- Existing systems lack sequence modeling capabilities, preventing accurate analysis of time-dependent changes in productivity, fatigue, and well-being.

B. PROPOSED SYSTEM

The proposed system introduces a Multimodal Human Digital Twin framework designed to continuously monitor and model human cognitive, emotional, and physical states in real time. The system integrates vision, audio, and behavioral data collected through cameras, microphones, and interaction logs to form a comprehensive representation of an individual. Deep learning models automatically extract high-level features from facial expressions, voice signals, and user interaction patterns, eliminating the need for manual feature engineering. Advanced multimodal fusion techniques combine these features to capture interdependencies across modalities. Transformer-based and sequence learning models analyze temporal patterns to assess productivity, fatigue, and well-being over time. The system supports both real-time monitoring and long-term trend analysis. Predictive modeling enables early detection of fatigue and performance degradation. Real-time alerts and personalized recommendations are generated to support proactive interventions. An interactive visualization dashboard presents intuitive analytics and state summaries. The system adapts dynamically to individual behavior patterns. This unified framework enhances accuracy, scalability, and interpretability. The proposed system aims to improve performance, well-being, and sustainable work practices through intelligent human-centered design.

1) ADVANTAGES:

- The integration of vision, audio, and behavioral data with deep learning improves the reliability and precision of fatigue and productivity estimation.
- Transformer-based temporal analysis enables continuous tracking and early prediction of performance degradation and health risks.
- The system adapts to individual behavior patterns and supports long-term monitoring, making it suitable for diverse real-world applications

C. PROPOSED ARCHITECTURE

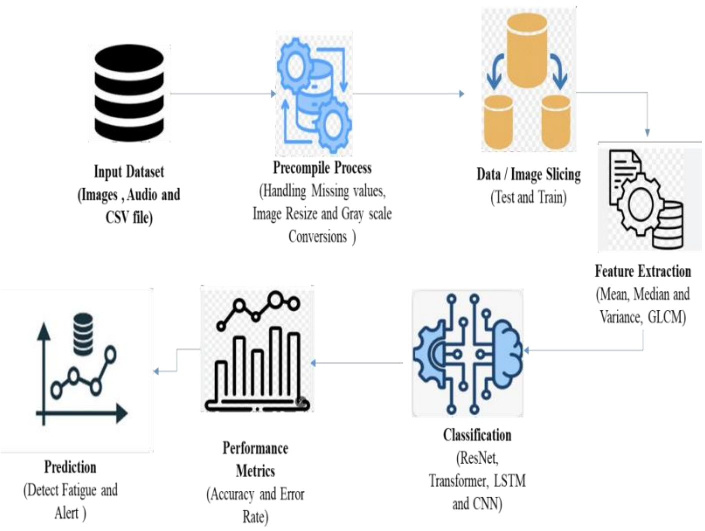

Fig. 1 System architecture

The proposed system follows a modular end-to-end pipeline designed to transform raw multimodal sensor data into actionable fatigue alerts. The architecture is divided into seven critical stages.

IV. IMPLEMENTATION

A. Data Acquisition and Pre-processing

This module is responsible for collecting raw multimodal data required for building the Human Digital Twin. It acquires image data through cameras, audio data through microphones, and behavioral data from system interaction logs. The collected data may contain noise, inconsistencies, and irrelevant information.Therefore, preprocessing operations such as noise removal, resizing, format conversion, and data cleaning are applied. Data splitting is performed to segment continuous streams into manageable units for further processing. This module ensures synchronization across all modalities. Proper preprocessing improves data quality and enhances model accuracy. The output of this module serves as standardized input for subsequent feature extraction stages.

B. Image Processing and Feature Extraction

This module focuses on analyzing visual data to extract meaningful features related to human facial expressions and physical states. Preprocessed images are divided into segments and passed through texture analysis using the Gray Level Co-occurrence Matrix (GLCM). GLCM extracts spatial texture features such as contrast, correlation, and homogeneity.These features are useful for identifying facial cues associated with fatigue and emotional changes.The extracted features are then fed into transformer-based deep learning models. The transformer model captures spatial and temporal dependencies in visual data. This module plays a critical role in visual fatigue and emotion recognition.

C. Audio Processing and Feature Extraction

The audio processing module handles voice signals to assess cognitive and emotional states. Audio data is first preprocessed to remove background noise and normalize amplitude levels. The audio stream is then split into short frames for efficient analysis. Mel Frequency Cepstral Coefficients (MFCC) are extracted to represent important acoustic characteristics such as pitch and tone. These features are processed using Convolutional Neural Networks (CNNs) to identify stress and fatigue-related voice patterns.The model learns subtle variations in speech that indicate cognitive load. This module enhances fatigue detection accuracy by capturing non-visual indicators.

D. Behavioral Data Processing

This module analyzes user interaction behavior to evaluate productivity and work patterns. Behavioral data such as typing speed, mouse movements, activity duration, and interaction frequency are collected.Preprocessing includes normalization and segmentation to ensure consistency across users. Long Short-Term Memory (LSTM) networks are used to model sequential behavioral patterns.LSTM effectively captures long-term dependencies and behavioral trends over time. Changes in interaction behavior often indicate fatigue or reduced focus. This module contributes long-term insights into productivity and work sustainability.

E. Multimodal Fusion

The multimodal fusion module integrates features extracted from image, audio, and behavioral modules. It combines heterogeneous feature representations into a unified feature space. Fusion techniques capture interdependencies across modalities that cannot be identified individually.This integrated representation improves robustness and reliability of human state estimation. The fusion process supports real-time and continuous analysis.By leveraging complementary information, the module reduces false detections. It forms the core of the Human Digital Twin system. The fused output is passed to decision-making components for final analysis.

F. Fatigue Detection and Alert Generation

This module performs final decision- making based on fused multimodal features. Deep learning models analyze the combined data to detect fatigue levels and productivity states. When fatigue or abnormal patterns are identified, the system generates real-time alerts. Notifications and recommendations are delivered to users through the dashboard.This enables proactive interventions to prevent performance degradation. The module supports continuous monitoring and adaptive responses.It also stores results for long-term trend analysis. This final module ensures practical usability and real-world impact of the system.

V. RESULTS AND DISCUSSION

A. Data Collection and Pre-processing Input:

Simulated raw images, audio recordings, and user interaction logs.Initiate data acquisition for all three modalities.Apply preprocessing steps including noise removal, normalization, and splitting.Verify data synchronization across modalities.

B. Image Feature Extraction Module Input:

Set of preprocessed facial images showing various expressions. Input images into the image processing pipeline.Extract GLCM texture features from images.Feed features into the transformer model for representation. GLCM features correctly represent texture patterns in images.Transformer model generates meaningful feature embeddings capturing facial cues. No errors or significant delays during processing.

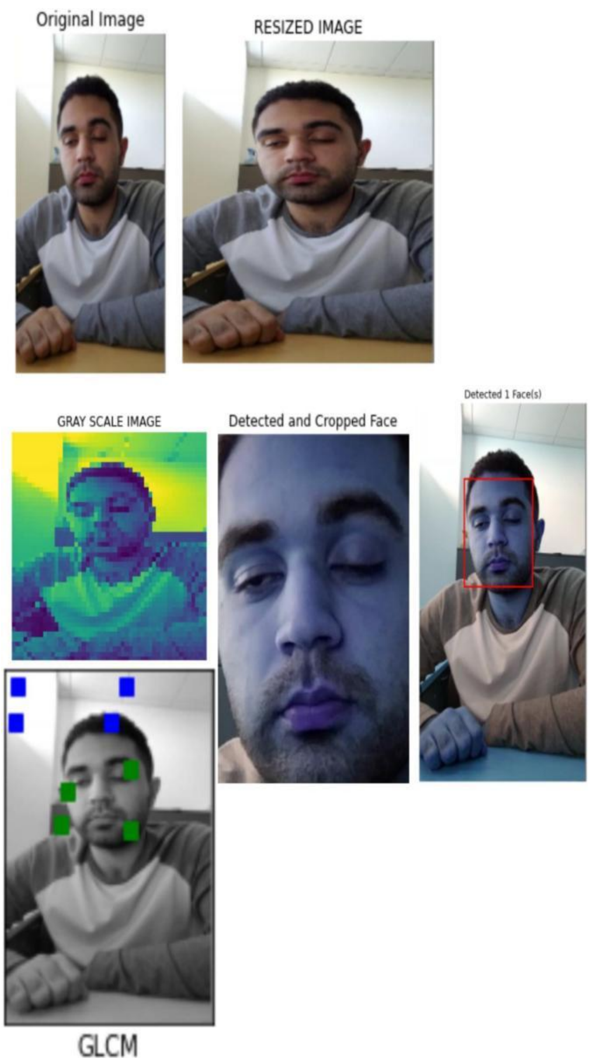

Fig. 2 Image Feature Extraction

C. Audio Feature Extraction Module Input:

Audio clips of normal speech and fatigued speech. Preprocess audio clips (noise filtering, normalization). Extract MFCC features.Classify features with CNN to identify fatigue indicators.

D. Behavioral Data Processing and LSTM Modeling Input:

Simulated user interaction data with varying activity levels. Normalize behavioral data such as keystroke speed and mouse movement. Segment data into time windows.Input into LSTM model to identify fatigue or drop in productivity.

E. Multimodal Fusion and Alert Generation Input:

Synchronized feature sets from image, audio, and behavioral modules indicating fatigue.

Fuse multimodal features using the fusion algorithm. Analyze fused features to detect fatigue threshold breaches. Trigger alert and notification systems.

Fig. 3 System User Interface

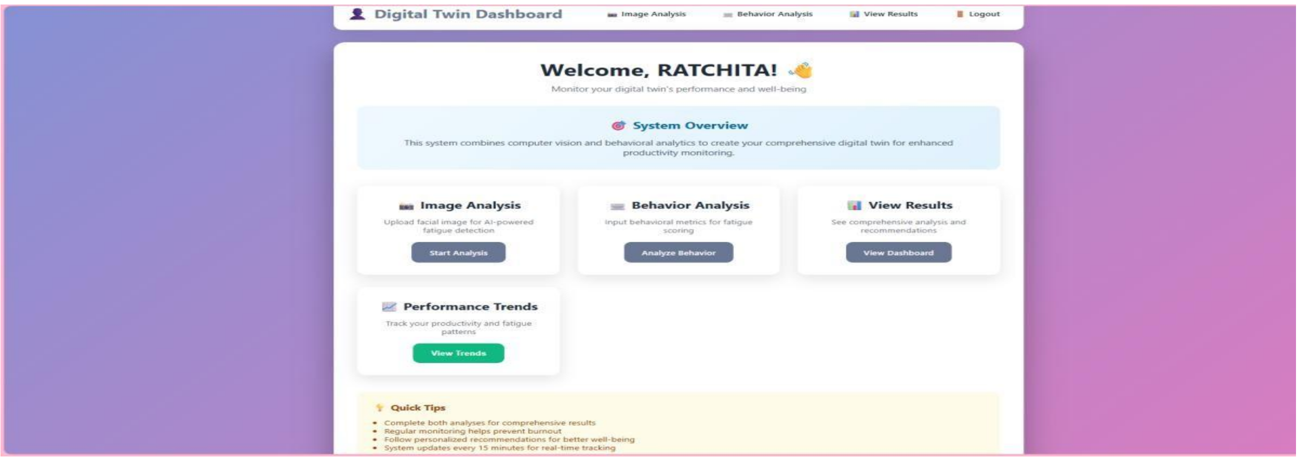

Fig. 4 Digital Twin Dashboard and Functional Navigation

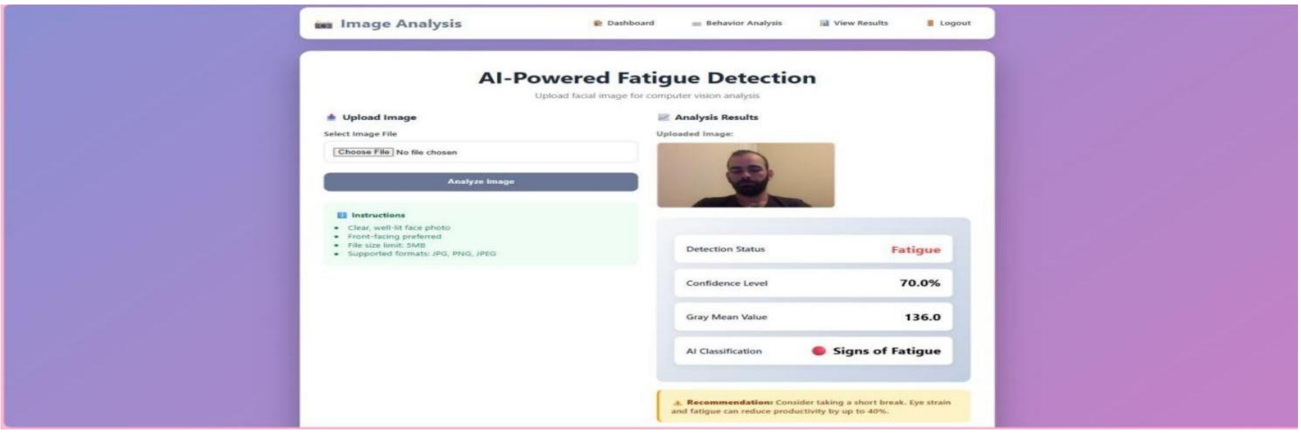

Fig. 5 AI-Powered Facial Fatigue Detection Interface

International Journal of Science, Strategic Management and Technology
Volume **02** Issue **03** March-**2026** | ISSN: **3108-1762 (Online)** | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

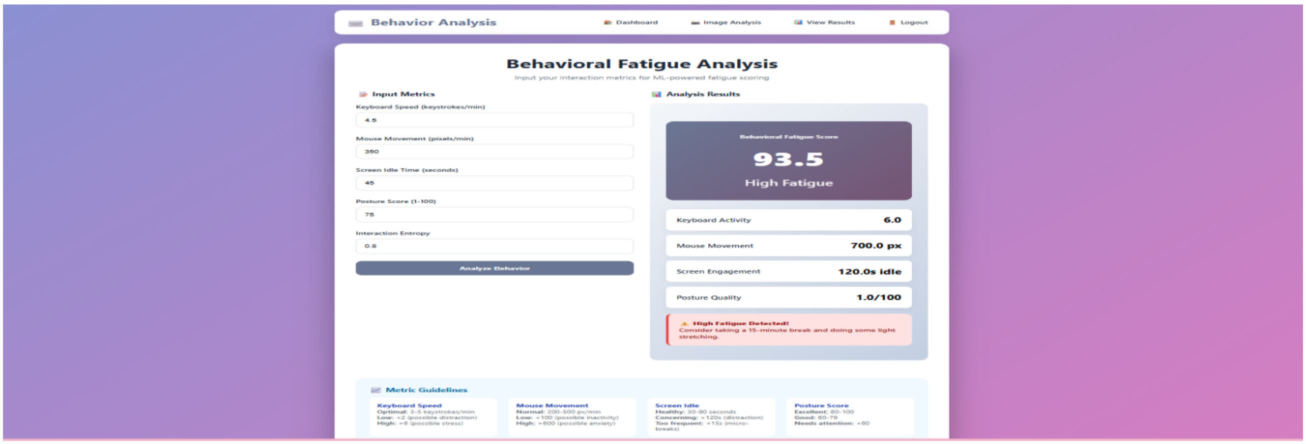

Fig. 6 AI-Powered Behavioral Fatigue Analysis Interface

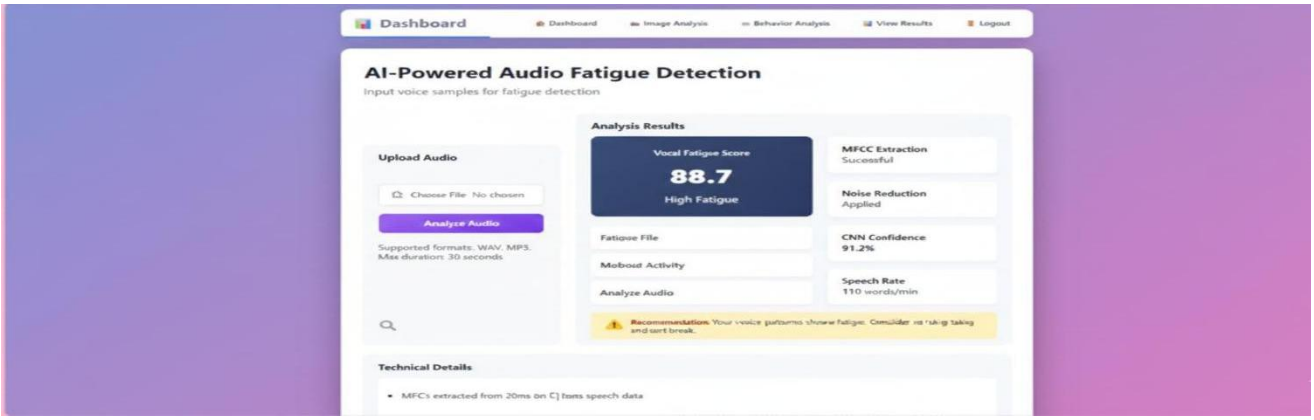

Fig. 7 AI-Powered Audio Fatigue Detection

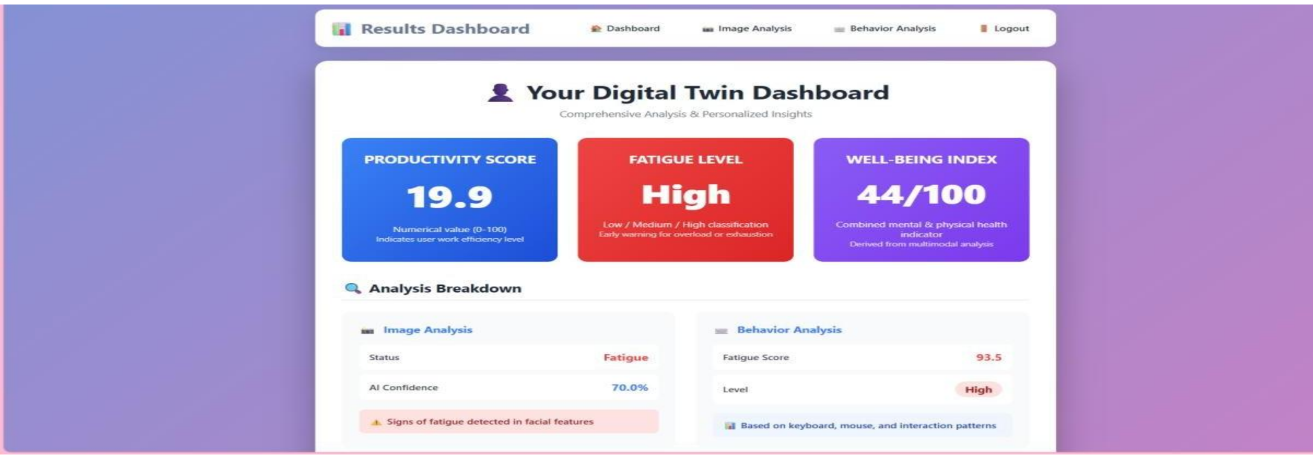

Fig. 8 Comprehensive Multimodal Results and Wellbeing Analysis

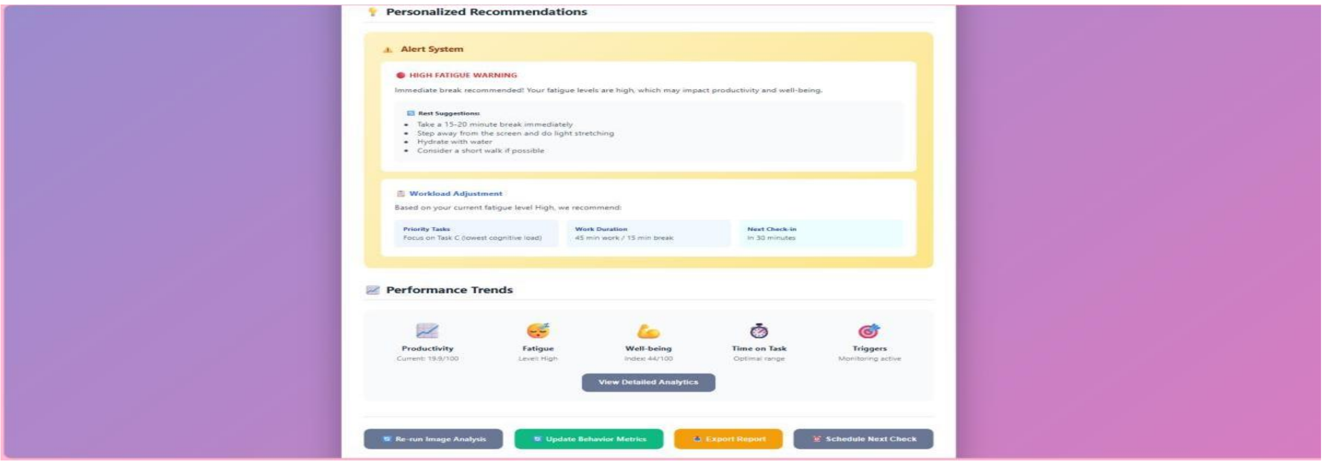

Fig. 9 Personalized Intervention and Performance Trend Analysis

1.	Accuracy	**92.48%**
2.	Error rate	**7.52%**
3.	Precision	**91.30%**
4.	Recall	**93.10%**
5.	F1-Score	**92.20%**

TABLE I

VI CONCLUSION

This project successfully presents a Multimodal Human Digital Twin System for continuous monitoring and analysis of human cognitive, emotional, and physical states. By integrating vision, audio, and behavioral data, the system provides a comprehensive and dynamic representation of human conditions. Deep learning models enable automatic feature extraction, eliminating the limitations of manual methods. Multimodal fusion enhances accuracy by capturing interdependencies across multiple data sources. Transformer and sequence-based models effectively analyze temporal patterns to detect fatigue and productivity changes. The system supports both real-time monitoring and long-term trend analysis. Real-time alerts and notifications help in early identification of fatigue. The interactive dashboard improves interpretability and user engagement. Overall, the proposed framework enhances performance awareness and well-being. The project demonstrates the feasibility of intelligent human-centered monitoring systems. It contributes to sustainable work practices and proactive health management. The system can be applied across various domains such as offices, healthcare, and smart environment

REFERENCES

[1] Y. You and Z. Ji, “Human Digital Twin for Real-Time Physical Fatigue Estimation in Human-Robot Collaboration,” Proc. 25th IEEE Int.Conf. on Industrial Technology (ICIT), Bristol, UK, Mar. 2024, pp. 1–6.
[2] J. Su, Z. Liao, Q. Mao and Z. Sheng, “Deep Learning-based Digital Twin for Human Activity Recognition,” 2023 IEEE 29th Int. Conf. on Parallel and Distributed Systems (ICPADS), Dec. 2023, DOI: 10.1109/ICPADS60453.2023.00024.
[3] R. Zhong, B. Hu, Y. Feng et al., “Construction of Human Digital Twin Model Based on Multimodal Data and Its Application in Locomotion Mode Identification,” Chinese Journal of Mechanical Engineering, vol. 36, art. 126, Oct. 2023.
[4] Y. Lin, L. Chen, A. Ali et al., “Human Digital Twin: A Survey,” Journal of Cloud Computing, vol. 13, art. 131, Aug. 2024.
[5] M. R. Islam, M. Subramaniam and P. C. Huang, “Image-based Deep Learning for Smart Digital Twins: A Review,” Artificial Intelligence Review, Feb. 2025.
[6] “Human-Machine Symbiosis Using Multimodal Trace Data for the Design, Development, Testing and Implementation of Human Digital Twins,” IEEE SMCS Workshop, Feb. 2024.
[7] Q. He, L. Li, D. Li et al., “Human Digital Twin Framework and Perspectives in Human Factors,” Chinese Journal of Mechanical Engineering, Feb. 2024.
[8] “Human Digital Twins: A Systematic Literature Review and Concept Disambiguation for Industry 5.0,” Computer Industry, 2024.
[9] S. Chand, H. Zheng and Y. Lu, “A Vision-Enabled Fatigue-Sensitive Human Digital Twin Towards Human-Centric Human-Robot Collaboration,” Journal of Manufacturing Systems, vol. 77, Dec. 2024.
[10] X. You, C. Duan, J. Gong et al., “A Human Digital Twin Approach for Fatigue- Aware Task Planning in Human-Robot Collaborative Assembly,” Computers & Industrial Engineering, 2025.
[11] T. Shaik, X. Tao, L. Li, H. Xie and J. D. Velásquez, “A Survey of Multimodal Information Fusion for Smart Healthcare,” arXiv:2306.11963, 2023.
[12] J. Li, L. Yao, B. Li and C. Sammut, “Distilled Mid-Fusion Transformer Networks for Multi-Modal Human Activity Recognition,” arXiv:2305.03810, 2023.
[13] S. Xaviar, X. Yang and O. Ardakanian, “Robust Multimodal Fusion for Human Activity Recognition,” arXiv:2303.04636, 2023.
[14] IEEE Standard or Workshop on Digital Twin technologies, IEEE CASE 2024 Special Sessions, 2024.
[15] A. Alkhateeb, S. Jiang and G. Charan, “Real-Time Digital Twins: Vision and Research Directions for 6G and Beyond,” arXiv:2301.11283, 2023.
[16] “Towards Intelligent AIoT: A Comprehensive Survey on Digital Twin and Multimodal Generative AI Integration,” MDPI Mathematics, 2023–2024.
[17] “Deep-Learning for Change Detection Using Multi-Modal Fusion of Remote Sensing Images: A Review,” Remote Sens., vol. 16, no. 20, Oct. 2024.
[18] X. Wang, J. Yu and S. Tang, “Fusionbench: A Comprehensive Benchmark of Deep Model Fusion,” arXiv preprint, 2024.
[19] K. Bagadi et al., “Advanced Self-Driving Vehicle Model Using Integrated Image Processing and Sensor Fusion,” IEEE Access, 2024
[20] “Multimodal Data Storage and Retrieval for Embodied AI: A Survey,” arXiv preprint, 2025
[21] M Kaliappan, E Mariappan, MV Prakash, B Paramasivan, Load Balanced Clustering Technique in MANET using Genetic Algorithms.. Defence Science Journal 66 (3), 251-258.
[22] M Sivam, M Kaliappan, S J Shobana, Prakash, V Porkodi Secure storage allocation scheme using fuzzy based heuristic algorithm for cloud Journal of Ambient Intelligence and Humanized Computing, pp.1-9
[23] Vimal, S., Robinson, Y. H., Kaliappan, M., Vijayalakshmi, K., &Seo, S. (2021). A method of progression detection for glaucoma using K-means and the GLCM algorithm toward smart medical prediction. The Journal of Supercomputing, 77(1), 1–17. https://doi.org/10.1007/s11227-020-03268-0
[24] Kaliappan M, Guruprakash B, Rajalakshmi, J. Blessing Karunya T, Mariappan E, Ramnath M and Angel Hepzibah R, Analyzing Public Sentiment on Demonetization Using SVM: A Machine Learning Approach, Journal of Computer Science 2025, 2482-2487, Published: 8 December 2025.