International Journal of Advanced Engineering Research and Science (IJAERS)
Peer-Reviewed Journal
ISSN: 2349-6495(P) | 2456-1908(O)
Vol-12, Issue-9; Sep, 2025
Journal Home Page Available: *https://ijaers.com/*
Article DOI: *https://dx.doi.org/10.22161/ijaers.129.7*

Enhancing Employee Attention and Task Efficiency Using Advanced Digital Distraction Blocking Mechanisms

Dung Tran Tuan

Vinschool Smartcity, Vietnam
dung083337@stu.vinschool.edu.vn

Received: 13 Aug 2025,

Received in revised form: 14 Sep 2025,

Accepted: 18 Sep 2025,

Available online: 23 Sep 2025

©2025 The Author(s). Published by AI Publication. This is an open-access article under the CC BY license

***Keywords*— Employee attention, digital distraction, Decision Tree Tuned Scalable Spiking Network (DT-S2Net), distraction patterns.**

***Abstract*— *Employee attention is critical for successful task completion, learning, and productivity. However, modern distractions such as notifications, applications, and online information regularly disrupt focus and reduce work effectiveness. Research aims to improve employee attention and task efficiency by creating an advanced digital distraction blocking system that detects distracted behaviors and intervenes to maintain focus. Video recordings of employees executing tasks in online and digital settings were gathered to capture natural attention and distraction patterns. Frames were normalized, faces were identified, and irrelevant background noise was eliminated to ensure consistent data quality. Frame blocking is utilized for preprocessing, and the VGG16-CNN was used to extract deep visual features, including facial expressions, gaze direction, head orientation, and micro-movements that indicate attentiveness or distraction. A Decision Tree Tuned Scalable Spiking Network (DT-S2Net), which combines Decision Trees (DT) for interpretable classification with Scalable Spiking Neural Networks (S2NN) for temporal pattern detection of attention and distraction behaviors. The DT-S2Net method outperformed the existing methods, achieving higher accuracy (99.01%), precision (99.57%), recall (98.77%) and F1-score (99.0%) in recognizing distracted states, allowing for timely digital interventions such as limiting notifications or irrelevant programs, resulting in faster task completion and sustained focus.*

I. INTRODUCTION

The digital transformation of work, also known as work digitalization, has altered how tasks are completed, where they occur, and who completes them. It provides flexibility and different practices by combining computing, persistent communication, and cloud technology. However, it blurs the line between work and personal life, increasing concerns about digital distractions and their effects on productivity and well-being [1]. Distractions in the workplace or during working hours impede effective contributions and, in the long run, reduce labour productivity. Work productivity is a measure of the real economic output achieved per work hour, as well as the potential output feasible within that time. Finally, it reflects how productive a worker can be when totally alert, concentrated, and free of interruptions [2]. The internet is currently utilized for a variety of virtual activities, including social contact, information retrieval, entertainment, and education. In today's office, working without computers and internet connection is practically impossible, emphasizing the importance of understanding digital procrastination, a prevalent behaviour in which employees spend extended periods of time online during work hours [3]. Employee performance is typically assessed using several criteria: quality of work, as evidenced by task completion and skills; quantity, as evidenced by output or targets met; timeliness in completing assigned tasks; efficiency in utilizing resources effectively; and independence, as demonstrated by responsibility, accountability, and a capacity to work independently [4]. The aim of this research is to improve employee attention and task efficiency by creating an advanced distraction blocking system that detects distracted behaviours in real time and provides timely interventions to maintain focus and increase workplace productivity.

The remaining part of the research is organized as follows: Part 2 described the related works, Part 3 represents the methodology, Part 4 is the result and discussion and Part 5 concludes the research.

II. RELATED WORKS

An Electroencephalogram (EEG) was utilized to examine the association between distraction and neural activity in mining [5]. After extracting and choosing certain EEG data, a Random Forest (RF) classifier was built using these EEG features to automatically detect distraction. The RF classifier successfully detected distraction, recognizing discrete EEG signals that distinguished between focused and distracted states. Performance assessment systems were utilized to benefit employees by providing information for promotion decisions, staff training, and disciplinary punishments [6]. Managers use performance evaluation data as motivational and performance-improvement measures. The performance evaluation system, which was based on employee performance targets, has proven effective in terms of system design, administrative practices, and long-term system support. Self-Deep Spiking Network (Self-DSNet) utilizing Self-Organizing Neural Network (Self-ONNs) was used to identify distractions while driving [7]. The Self-DSNet model improved distraction detection by using multimodal data and increasing complicated pattern recognition with Self-ONN layers. RF, Decision Trees (DT), and Support Vector Machines (SVM) Machine Learning (ML) algorithms were used to classify motorist mental distractions [8]. The methods identified features from eye tracking, behavioural, and automobile kinematic data, which were used to differentiate diverted and unfocused drivers in multiple circumstances. Bi-directional feature pyramid networks (BiFPN) were utilized for identifying driver distractions [9]. The BiFPN module was added to the neck structure for improved multi-scale feature fusion without excessive computation. This technique offered excellent recognition accuracy, fast detection speed, and minimal model memory usage, making it ideal for industrial implementation.

Research gap: Existing approaches, such as EEG-based RF classifiers for distraction detection [5], performance evaluation systems for workplace productivity [6], SSL-RL frameworks for workstation activity prediction [7], and ML or BiFPN methods for driver distraction [8,9], offer useful insights but have limitations. These methods frequently use single-modal data, struggle with real-time flexibility, and lack explainability in decision-making. Thus, there is still a research gap in building advanced, multimodal, and interpretable distraction-blocking technologies that effectively improve attention and task efficiency in dynamic office settings.

III. METHODOLOGY

The methodology section includes frame blocking for data pre-processing, VGG16 for feature extraction, and the proposed Decision Tree Tuned Scalable Spiking Network (DT-S2Net) method, which combines Decision Trees (DT) for interpretable classification with Scalable Spiking Neural Networks (S2NN) for temporal pattern detection of attention and distraction behaviors. Figure 1 depicts the overall flow of the research.

3.1 Dataset

The employee attention & task efficiency dataset utilized in the research was obtained from Kaggle (https://www.kaggle.com/datasets/zara2099/employee-attention-and-task-efficiency-data/data). This dataset includes data gathered from 1000 employees executing jobs in digital settings. The key features include: Employee identification (Employee ID) and video recordings (Video Path). Visual attention markers include gaze direction, head orientation, facial emotions, and micro movements. Task-related metrics include task completion time, interrupted notifications, and the number of applications opened. The target variable is the attention status (Focused or Distracted) as determined by observable behaviors. The dataset is intended to investigate attention and distraction patterns to enable efforts that improve work efficiency.

3.2 Frame blocking for Pre-processing

Frame blocking is used in speech recognition to split voice signals into short, overlapping frames for employee attention &task efficiency. This allows for smooth transitions between frames and maintains signal continuity. Employee distraction behaviours (eye gaze shifts, micro-movements, and head orientation) share temporal sequence features with voice and physiological signals; hence, the frame blocking approach is used on video frames to achieve consistent input duration. Each video is separated into fixed-length, overlapping frame pieces, which allows the model to detect both local distraction cues and long-term attention patterns. The frame blocking method is applied to the behavioral signal of employees. The length of each frame is calculated by Equation (1).

$$G_m = G_t + g_p \quad (1)$$

Where G_m is the length of each frame, $G_t$is the frame shift (time lag of the frame from the start of the recording), $and g_p$ is the overlapping part between adjacent frames. For a raw recording with a total length of t_m, given the number of frames $G_o$and frame length G_m, the framing equation of employee attention &task efficiency can be represented in Equation (2).

$$G_t = {(t_m - G_m)}\Big/{(G_o - 1)} \quad (2)$$

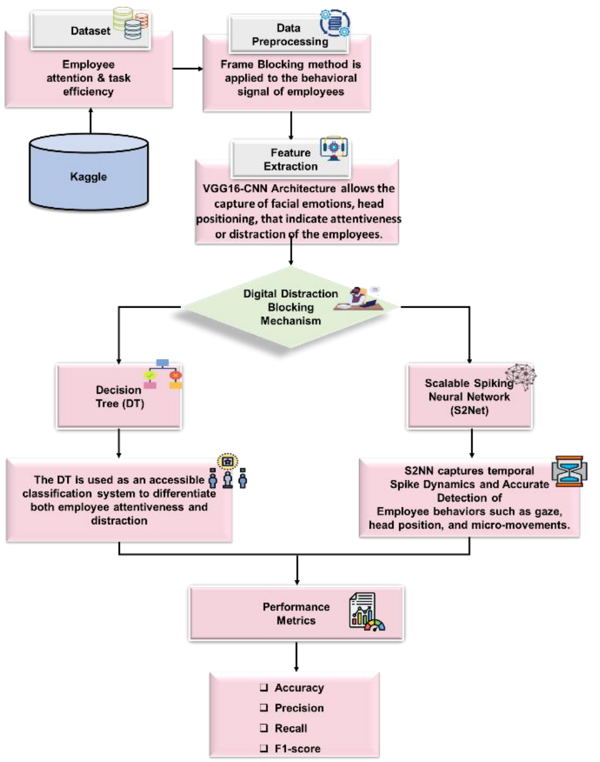


Fig.1: Overall flow

3.3 VGG16 for feature extraction

VGG16 is a Convolutional Neural Network (CNN) architecture widely utilized in applications such as object detection, pattern detection, segmenting pictures, and image classification for employee attention and task efficiency. To extract deep visual information from individual video frames, the VGG16-CNN is used. Its architecture allows the capture of facial emotions, gaze movement, head positioning, and minute movements that indicate attentiveness or distraction of the employees.

Convolutional Layer (CL): The CL uses many kernels or layers to extract features from images for employee attention &task efficiency. A non-linear conversion from the input data is represented in Equation (3).

$$Y = {B_n}^o U \quad (3)$$

Where B_n represents the input image and Y is the productivity of the resultant characteristic employee from the input image.

Dense layer (DL): The DL links all neurons in the preceding layer to those in the current layer. It is the CL capture features, which are then minimized by pooling layers of employee attention & task efficiency. The DL can be mathematically represented in Equation (4).

$$E = \sigma(X \times G \times c) \quad (4)$$

Where E is the DL outcome, $\sigma$is the stimulation coefficient, X symbolizes a weight, E is the flattening layer input, and c is the coefficient of bias of employee attention & task efficiency. The CSV data format converts raw employee attention and task records into structured characteristics, allowing for efficient extraction and model training for distraction detection and efficiency analysis.

3.4 Decision Tree Tuned Scalable Spiking Network (DT-S2Net) for digital distraction blocking mechanisms

DT-S2Net was created by DT for interpretable categorization with S2NN for temporal pattern detection of attention and distraction behaviours. The DT-S2Net approach excelled in detecting distracted states, allowing for prompt digital interventions such as limiting alerts or unnecessary programs, resulting in quicker task completion and continued concentration of the employees.

Decision tree (DT): The DT is used as an accessible classification system to differentiate both employee attentiveness and distraction. A DT is a sequential strategy for dividing a data subset, with each stage selecting a variable and a splitting border. The dataset is then repeatedly split into pure subsets using the impurity measure. The DT defines goodness of split as the distinction between impurity levels before and after splitting. As a result, more accuracy in the separated outputs indicates that the split is more beneficial. As a result, the data set is divided along the division boundary$R$, with the highest goodness of split defined in Equation (5).

$$H(U,S) = J(U) - J(U|S) \quad (5)$$

Where U is a set of training examples, $H(U,S)$represents the goodness of division when the instructional set U is divided by S, whereas $J(U)$ and $J(U|S)$ measure the contaminants prior to and after division, respectively, based on the border. The best split point for lowering contamination is found using the DT method, which makes use of impurity measurements. It chooses key visual cues such as head orientation, gaze direction, facial expressions, and micro-movements. While distance indicators employ length to equalize impurity data, the Gini index evaluates variance between desired attributes.

Improved scalable spiking neural network (S2Net): The spiking neuron is the core computing unit of SNNs. The membrane's energy and intrinsic voltage alter according to the input signals by linked synapses. In Artificial Neural Networks (ANNs), the spiking neuron generates a binary spike signal to represent information and analyze it with great precision. It is transmitted to the next connected one when the membrane potential exceeds the firing threshold, consistent with biological neuron firing principles. By leveraging biologically inspired spiking dynamics, S2Net effectively detects distraction states and ensures timely interventions, promoting sustained focus and faster task completion in digital work environments. The iterative Leaky-Integrate-and-Fire (LIF) model captures temporal dynamics of employee actions (gaze, head position, micro-movements), which are necessary for distraction detection and employee attention & task efficiency represented in Equation (6) and the spike firing condition is represented in Equation (7).

$$T\frac{dv}{dt} = -v + SI, v < U_{th} \tag{6}$$

$$v = v_{rest} \& fireaspike, v \geq U_{th} \tag{7}$$

Where v represents the neural cell potential, and v_{rest} denotes the beginning potential at rest of a neuron, which is typically assumed to be 0, calculated as the product of input current (I) and resistivity. S Represents the shifting voltage of employee attention and task efficiency towards prior to synaptic inputs and $U_{th}$is the firing threshold. Equation (6) represents the membrane potential accumulation of the employee's current attention state below U_{th}. Equation (7) indicates that when u is up to when a neuron emits a spike (U_{th}), the v is reset to its resting potential (v_{rest}). The iterative LIF method is demonstrated in Equation (8) and surrogate gradient approximation is represented in Equation (9) that captures temporal patterns of employee behaviors such as gaze, head position, and micro-movements.

$$v^u = \lambda_s v^{u-1}(1 - \delta^{u-1}) + \sum_k x_k p_k^u \tag{8}$$

$$\delta^{u-1} = \begin{cases} 1, & if v^{u-1} \geq U_{th} \\ 0, & otherwise \end{cases} \tag{9}$$

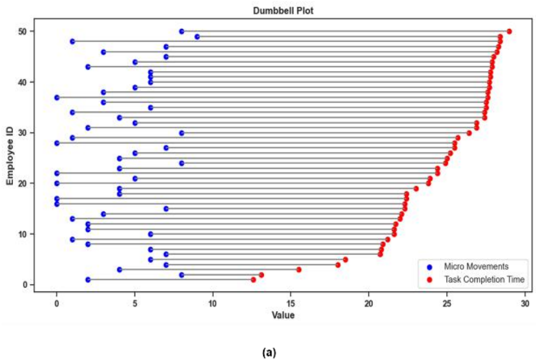
Equation (8) shows that the current membrane potential (v^u) is influenced by the spike (p_k^u). From the pre-neuron and the spike emission δ^{u-1}, each time step adds to the current time step u. The DT-S2Net combines interpretable DT with scalable S2NN to capture spatial features and temporal dynamics of employee distraction, allowing for precise detection, adaptive digital interventions, and increased task efficiency in workplace environments.

IV. RESULT AND DISCUSSION

The results section describes the experimental setup, research findings, and evaluation metrics, followed by a comparison analysis that shows the proposed DT-S2Net outperforms the existing ResNet [10] versions. The discussion focuses on benefits and limitations, while the conclusion emphasizes effectiveness, restraints, and future research areas for workplace distraction control. The experiments were carried out in Python 3.10, with NumPy, Pandas, TensorFlow, and PyTorch for model construction and Scikit-learn for decision tree integration.

4.1 Research Outcomes

Figure 2 demonstrates the link between attention, micro-movements, and task efficiency. (a) Indicates that increased micro-movements result in longer task completion durations. (b) Shows that distracted employees constantly demand more time to complete tasks than focused individuals.



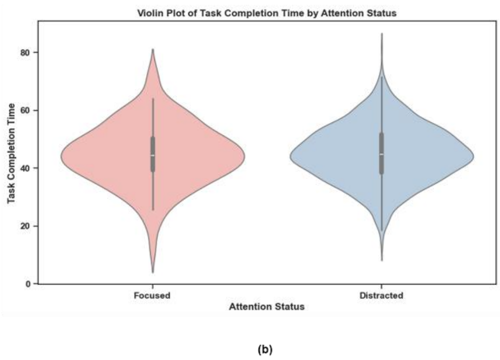


Fig.2: Employee attention in (a) Dumbbell plot and (b) violin plot of task completion by attention status

The numbers show how employees' attention is distributed and how well they complete tasks. Figure 3(a) depicts a pie chart with the majority of employees (91.2%) distracted, and only 8.8% remaining focused. Figure 3(b) shows a histogram of task completion time with a normal-like distribution, indicating that most employees finish tasks within 35-55 time units.

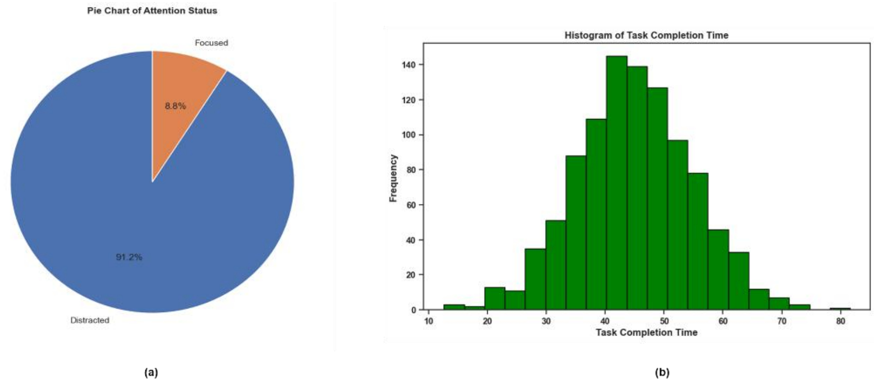


Fig.3: Analysis of employee attention and task completion timein (a) Pie chart of attention status, (b) Histogram of task completion time

Figure 4 depicts the correlations between micro-movements, task completion time, and attention states. (a) Depicts a hexbin plot in which higher concentrations of micro-movements correlate with longer task completion durations, demonstrating that distraction affects efficiency. (b) Depicts Andrews's curves, which clearly differentiate engaged and distracted personnel via discrete pattern clusters, underlining behavioural isolation.

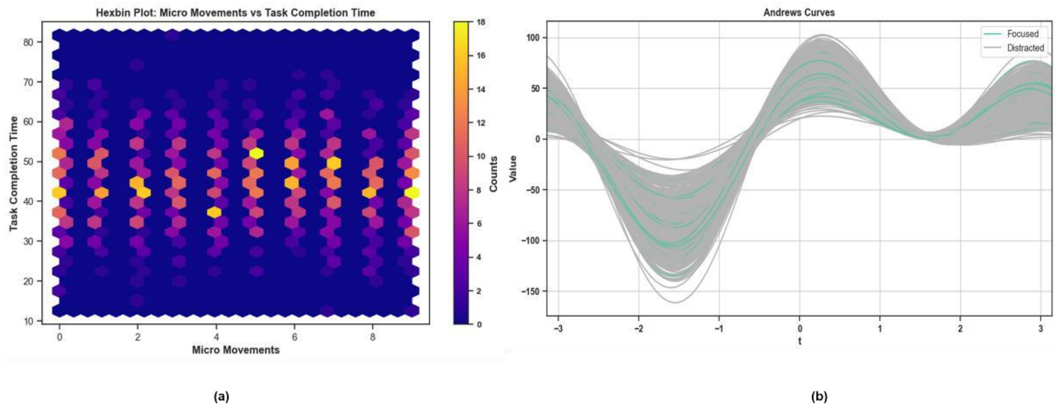


Fig.4: Visualization of (a) Hexbin plot: micro movements vs task completion time (b) Andrews curves

4.2 Evaluation metrics

Accuracy assesses a model's overall correctness by computing the ratio of correctly predicted instances (both attentive and distracted) to the total number of cases, indicating the overall performance of distraction detection across all observations. **Precision** is defined as the fraction of accurately detected distracted states among all instances projected to be distracting. It demonstrates how dependable the system is in avoiding false alarms when identifying attention loss, ensuring interventions are initiated only when essential. **Recall** measures the percentage of actual distracted states successfully detected by the system. It reflects the model's sensitivity in detecting actual distraction events, resulting in fewer occasions where employees lose attention while performing digital activities. The **F1-score** calculates the harmonic mean of precision and recall. It gives a single performance metric that takes into account both missed distractions and false alarms, which is especially valuable when the distraction and attention classes are imbalanced.

4.3 Comparison analysis

The proposed DT-S2Net is compared with the existing Original Residual Networks (ResNet-18) [10], Fine-Tuning Residual Networks (FT ResNet-18) [10], and User-Extended Training Residual Network (UET ResNet-18) [10] methods with four performance metrics: accuracy, precision, recall, and F1-score for employee attention & task efficiency. Table 1 and Figure 5 represent the outcome of the proposed and the existing methods, indicating that the proposed method has an accuracy of 99.01%, a precision of 99.57%, a recall of 98.77%, and an F1-score of 99.0% for employee attention & task efficiency.

Table 1: Performance metrics of the proposed and the existing method

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Original ResNet-18 [10]	78.08%	74.38%	82.08%	77.97%
FT ResNet-18 [10]	98.52%	99.19%	97.78%	98.43%
UET ResNet-18 [10]	96.04%	94.78%	95.86%	95.32%
DT-S2Net [Proposed]	99.01%	99.57%	98.77%	99.0%

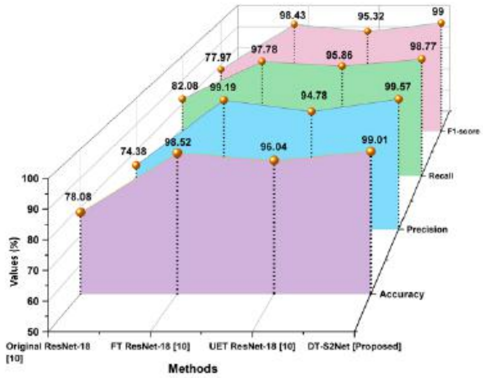


Fig.5: Demonstration of the evaluation metrics

4.4 Discussion

The original ResNet-18 [10] seeks to provide a foundational deep Convolutional framework for obtaining spatial and temporal features in distraction detection, as well as stable training with residual connections, but it suffers from adaptability, computational demands, and poor generalization across drivers. The FT ResNet-18 [10] increases performance by fine-tuning pretrained weights on task-specific data, increasing accuracy and reducing training time; nevertheless, it relies heavily on labeled data, risks overfitting, and requires costly retraining for new settings. The UET ResNet-18 [10] combines user-specific adaptation with transfer learning to reduce inter-individual variability and improve real-time detection robustness; nevertheless, its reliance on customization adds complexity and computing overhead. The suggested DT-S2Net approach efficiently combines interpretable decision tree classification with robust temporal pattern recognition in spiking networks, allowing for accurate detection of distracting activities. Its quick digital interventions reduce interruptions, promote sustained employee focus, and greatly increase task efficiency in digital workplaces.

V. CONCLUSION

Research seeks to improve employee attention and task efficiency by developing an effective digital distraction blocking system that recognizes distracted behaviours and reacts to keep concentration. The proposed DT-S2Net method combines DT for interpretable categorization with S2NN for time-based identification of attention and distraction behaviours. Research employs an employee attention & task efficiency dataset, and the pre-processing was done by the frame blocking method. After that, VGG16 is utilized for feature extraction. Its architecture allows the capture of facial emotions, gaze movement, head positioning, and minute movements that indicate attentiveness or distraction. Performance analysis was employed by four evaluation metrics, and the proposed method achieved higher accuracy (99.01%), precision (99.57%), recall (98.77%), and F1-score (99.0%) than the existing methods. The proposed DT-S2Net technique is limited by its reliance on visual signals, which can overlook slight cognitive distractions and necessitate high-quality recordings. The future scope involves integrating multimodal data, improving real-time adaptability, and tailoring interventions for varied employment situations to maximize attention management.

REFERENCES

[1] Farivar, F., Esmaeelinezhad, O. and Richardson, J., 2022. Digital intrusions or distractions at work and work-life conflict. *New Technology, Work and Employment*, *37*(3), pp.363-380http://dx.doi.org/10.1111/ntwe.12235

[2] Yemoh, M. and Amitai, F.I., 2022. The cost of workplace distractions to labor productivity, personal use of social media, and hand-held devices during working hours. *International Journal of Advanced Scientific Innovation*, *4*(5)

[3] Narwal, M., 2023. Digital distractions in the workplace: exploring cyberloafing impact on employee behaviour and innovation. *Virtual Economics*, *6*(4), pp.7-24.http://dx.doi.org/10.34021/ve.2023.06.04(1)

[4] Susanto, P.C., Syailendra, S. and Suryawan, R.F., 2023. Determination of motivation and performance: Analysis of job satisfaction, employee engagement, and leadership. *International Journal of Business and Applied Economics*, *2*(2), pp.59-68.https://10.0.218.119/ijbae.v2i2.2135

[5] Kuang, Y., Tian, S., Li, H., Yuan, C. and Chen, L., 2024. EEG-Based Measurement for Detecting Distraction in Coal Mine Workers. *Applied Sciences*, *15*(1), p.273.https://doi.org/10.3390/app15010273

[6] Sulistiani, L.S. and Faozanudin, M., 2022. Effectiveness Analysis of the Employee Work Performance Assessment System–A Critical Three-Component Approach. *Expert Journal of Business and Management*, *10*(2).

[7] Rashid, M.O., Hossen, M.M., Nashbat, M., Hasan-Zia, M., Kunju, A.K.A., Khandakar, A., Ashraf, A., Majid, M.E., Kashem, S.B.A. and Chowdhury, M.E., 2025. Self-DSNet: A Novel Self-ONNs Based Deep Learning Framework for Multimodal Driving Distraction Detection. *IEEE Access*.http://dx.doi.org/10.1109/ACCESS.2025.3545359

[8] Misra, A., Samuel, S., Cao, S. and Shariatmadari, K., 2023. Detection of driver cognitive distraction using machine learning methods. *IEEE Access*, *11*, pp.18000-18012.http://dx.doi.org/10.1109/ACCESS.2023.3245122

[9] Li, T., Zhang, Y., Li, Q. and Zhang, T., 2022. AB-DLM: An improved deep learning model based on attention mechanism and BiFPN for driver distraction behavior detection. *Ieee Access*, *10*, pp.83138-83151.http://dx.doi.org/10.1109/ACCESS.2022.3197146

[10] Fresta, M., Bellotti, F., Bochenko, I., Lazzaroni, L., Merlhiot, G., Tango, F. and Berta, R., 2025. Deep learning-based real-time driver cognitive distraction detection. *IEEE Access*.https://doi.org/10.1109/ACCESS.2025.3539392