Journal of Future Artificial Intelligence and Technologies

E-ISSN: 3048-3719

Research Article

Detecting Driver Fatigue Using Artificial Intelligence on a Realistic Driving Images

Elham T. Yasin [1], **Talha Alperen Cengel** [1], **Serkan Gerz** [2], **Selime Sinem Bahar** [2], **Bunyamin Gencturk** [1], **Ahmet Goktas** [2], and **Murat Koklu** [3,*]

[1] Graduate School of Natural and Applied Sciences, Department of Computer Engineering, Selcuk University, Konya 42130, Türkiye;
e-mail: ilham.tahsen@gmail.com; talhacengel@gmail.com; bunyamin.gencturk.58@gmail.com
[2] Alisan Logistics Inc., Istanbul 34752, Türkiye;
e-mail: serkan.gerz@alisangroup.com; sinem.bahar@alisangroup.com; ahmet.goktas@alisangroup.com
[3] Technology Faculty, Department of Computer Engineering, Selcuk University, Konya 42130, Türkiye;
e-mail: mkoklu@selcuk.edu.tr
* Corresponding Author: Murat Koklu

Abstract: Fatigue-related impairment is a major contributing factor to road accidents; however, detecting early visual indicators of driver tiredness remains challenging under realistic driving conditions. This study introduces an artificial intelligence–based system for distinguishing between alert and fatigued drivers using facial images captured in natural driving environments. A total of 41,793 annotated facial images from the Driver Drowsiness Dataset (DDD) were used in the experiments. Although the dataset reflects realistic driving scenarios captured by dashboard-mounted cameras, the proposed system was evaluated offline and not tested in live traffic environments. Deep visual features were extracted using the SqueezeNet architecture and subsequently classified using three supervised learning models: Artificial Neural Networks (ANN), Random Forests (RF), and Support Vector Machines (SVM). Among the evaluated classifiers, ANN achieved the highest performance with an accuracy of 99.97%, followed by RF with 99.78% and SVM with 96.33%. The results indicate that combining lightweight deep feature extraction with classical machine learning classifiers can yield highly accurate fatigue detection while maintaining computational efficiency. The proposed framework provides valuable insights into the development of efficient, real-time driver fatigue monitoring systems with potential applications in accident prevention and road safety enhancement.

Keywords: Artificial neural networks; Deep feature extraction; Driver drowsiness dataset; Driver fatigue detection; Embedded vision systems; Machine learning classification; Road safety; SqueezeNet.

Received: November, 28th 2025
Revised: December, 28th 2025
Accepted: December, 30th 2025
Published: January, 10th 2026
Curr. Ver.: January, 10th 2026

1. Introduction

Driver tiredness has long been recognized as a key risk factor impacting road safety, as sleepy drivers exhibit delayed reflexes, diminished situational awareness, and impaired decision-making [1]. These behavioral changes considerably increase the likelihood of traffic accidents, particularly during long-distance travel and nighttime driving [2], [3]. With the rapid advancement of artificial intelligence (AI) and computer vision technologies, monitoring driver attention through visual signals has become an increasingly practical approach for enhancing real-time safety systems. Recent studies have demonstrated that facial dynamics—such as eye closure duration, blinking frequency, yawning patterns, and head position—serve as key indicators of driver fatigue [4], [5]. Consequently, machine learning and deep learning techniques have attracted significant attention for their ability to extract subtle visual cues. Compared with traditional rule-based systems, AI-driven models can learn complex patterns directly from data and operate effectively across diverse drivers, lighting conditions, and environmental settings.

DOI : 10.62411/faith.3048-3719-299

Various studies have attempted to detect driver drowsiness using convolutional neural networks (CNNs) [3], handcrafted features [6], video-based temporal modeling [7], and hybrid fusion approaches [8]. While these techniques have reported encouraging results, many of them rely on computationally intensive architectures or controlled laboratory conditions. As a result, their applicability in real-world driving environments remains limited. To bridge the gap between research prototypes and deployable fatigue-monitoring systems, there is a clear need for lightweight feature extractors combined with efficient classification algorithms.

This paper advances the field by employing SqueezeNet, a compact neural architecture designed to minimize model size while preserving representational capacity, to extract deep facial features from a large-scale, realistic driving dataset [9], [10]. These features are subsequently classified using three widely adopted supervised learning methods: Artificial Neural Networks (ANN), Random Forests (RF), and Support Vector Machines (SVM). The proposed approach aims to provide a practical solution for real-time driver fatigue detection by integrating a lightweight deep feature extractor with computationally efficient classifiers.

This study evaluates the robustness of deep features extracted using the lightweight SqueezeNet architecture across multiple machine learning classifiers. Rather than conducting a competitive comparison among classifiers, the objective is to assess whether compact deep representations can maintain effective performance under different classification paradigms.

2. Literature Review

Driver fatigue detection has evolved from early handcrafted-feature approaches toward data-driven machine learning and deep learning methods. Traditional techniques based on eye or head movement cues are computationally efficient but sensitive to variations in illumination, facial appearance, and camera viewpoints [11], [12]. Recent studies predominantly employ convolutional neural networks to extract discriminative facial representations from images or video sequences, enabling improved robustness under diverse driving conditions. These methods demonstrate strong performance but often rely on deep architectures and complex processing pipelines.

To address deployment constraints, several studies have explored lightweight models, transfer learning, and embedded-oriented designs, frequently combined with explicit face detection or region-of-interest extraction. While such pipelines improve accuracy, they increase system complexity and computational overhead [13]. Moreover, limited attention has been given to evaluating whether compact deep feature representations remain robust across different classification paradigms when face detection is omitted. Table 1 summarizes representative state-of-the-art studies and highlights this methodological gap, which motivates the present work.

Dua et al. [14] proposed a driver drowsiness detection system using four deep learning models—AlexNet, VGG-FaceNet, FlowImageNet, and ResNet—to analyze RGB video data. The models extracted features related to facial expressions, hand movements, head movements, and behavioral cues, which were categorized into four classes: no drowsiness, blinking, yawning, and head nodding. A SoftMax classifier achieved an accuracy of 85%.

Nasri et al. [15] introduced DistractNet, a CNN-based model for detecting driver distraction, trained on the State Farm Driver Distraction Dataset, which contains 22,424 images across 10 activity classes. The model was compared with several pre-trained networks, including ResNet-50, GoogLeNet, InceptionV3, and AlexNet, and achieved the highest accuracy of 99.32% [10]. In a related study, the same authors proposed a CNN-based approach for detecting and predicting driver sleepiness using facial features extracted from the Real-Life Drowsiness Dataset (RLDD), achieving a classification accuracy of 96%.

Magán et al. [16] developed an Advanced Driver Assistance System (ADAS) to detect driver drowsiness using 60-second video sequences in which the driver's face is visible. Two alternative solutions were proposed: one combining recurrent and convolutional neural networks, and another using deep feature extraction followed by a fuzzy logic–based classifier. Both approaches achieved similar accuracy levels, approximately 65% on training data and 60% on test data.

Hossain et al. [17] evaluated multiple CNN architectures, including Simple CNN, VGG-16, ResNet50, and MobileNetV2, for detecting driver distraction. Their dataset comprised 102,150 images across ten classes. Experimental results showed that MobileNetV2 achieved the highest accuracy of 98.12%, followed by Simple CNN at 97.45% and ResNet50 at 94.50%.

Ahmed et al. [3] presented a real-time driver drowsiness detection system based on machine learning techniques using eye and mouth features as primary indicators. Their dataset consisted of 2,900 images with four extracted features, and the proposed CNN model achieved an accuracy of 97.23%.

Andrean et al. [2] investigated driver drowsiness detection using facial features extracted from a dataset containing 22,348 images of drowsy drivers and 19,445 images of alert drivers. The study compared the Haar Cascade and YOLO-face segmentation methods, followed by decision-tree–based eye-focused classification. YOLO-face segmentation produced higher-quality regions of interest, achieving an accuracy of 98.54%, compared with 98.03% for Haar Cascade.

Table 1. Summary of previous related studies in the literature.

References	Dataset	Records	Class	Methods	Accuracy
Dua et al., [14]	National Tsing Hua University(NTHU)-driver drowsiness standard video datasets	200 training video clips from 25 subjects and 50 testing clips from 6 subjects	2	SoftMax	85.00%
Nasri et al., [1]	State Farm Distracted Driver Detection | Kaggle	22,424	10	DistractNet ResNet-50 GoogLeNet InceptionV3 AlexNet	99.32% 98.16% 97.26% 98.02% 98.44%
Nasri et al., [15]	Real-Life Drowsiness Dataset RLDD	41,793	2	CNN	96.00%
Magán et al., [16]	Yawning Detection Video Dataset	30 individuals 18 subjects' videos used for training, 6 subjects for validation, 6 subjects for testing	2	CNN	65.00%
Hossain et al., [17]	StateFarm Distracted Driver Detection Dataset	22,424	10	CNN ResNet50 MobilNetV2	97.45% 94.50% 98.12%
Ahmed et al., [3]	Annotated Drowsiness Detection Dataset	2900	4	CNN	97.23%
Andrean et al., [2]	Real-Life Drowsiness Dataset RLDD	41,793	1	YOLO-Face Haar Cascade	98.54% 98.03%

As summarized in Table 1, prior studies demonstrate that deep learning–based approaches can achieve high accuracy, particularly when using large models, explicit face detection, or extensive preprocessing. However, these requirements often hinder real-time deployment in practical driving scenarios. Moreover, limited attention has been given to evaluating whether lightweight deep feature representations remain robust across different classification paradigms when explicit face detection or region extraction is omitted. This gap motivates the present study, which assesses the effectiveness of compact deep features extracted by SqueezeNet when combined with efficient machine learning classifiers under realistic driving conditions.

3. Proposed Method

In this study, classification was performed using cross-validation on the "Driver Drowsiness Dataset." The classification process was conducted using machine learning algorithms, including ANN, RF, and SVM, with cross-validation. SqueezeNet image embedding was utilized for the process. The methodology followed in the study is provided in Figure 1. The proposed framework follows a two-stage design consisting of deep feature extraction from original facial images and subsequent classification using machine learning models, without relying on data augmentation or end-to-end network retraining.

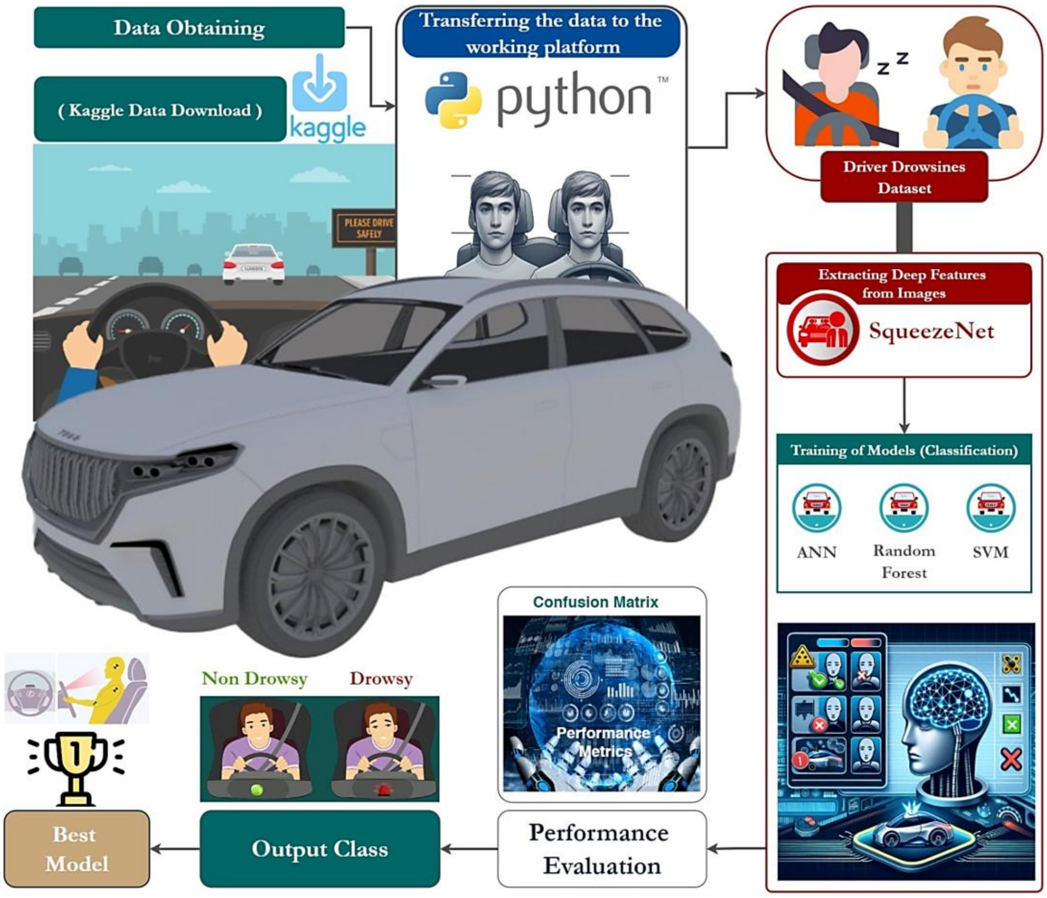


Figure 1. Illustrative conceptual diagram summarizing the main processing stages (Overview of the proposed driver fatigue detection framework)

3.1. Dataset

The Driver Drowsiness Dataset (DDD) was obtained directly from a publicly available Kaggle repository. As described by the dataset creators, the DDD was constructed from video recordings captured using dashboard-mounted cameras in real driving scenarios. Individual image samples were extracted from these videos using VLC software, after which facial regions were automatically localized and cropped using the Viola–Jones algorithm [1], [18]. All images were subsequently resized to a uniform resolution of 227 × 227 pixels during dataset preparation.

The DDD is publicly released only in the form of extracted image frames, while the original driving videos are not included in the dataset distribution. Consequently, the dataset authors do not explicitly provide detailed information on the original video resolution, frame rate, and frame sampling strategy. This limitation is inherited from the dataset and therefore applies to the present study. The dataset consists of 41,793 RGB facial images categorized into two classes: drowsy and non-drowsy. The class distribution, along with representative raw image samples from both categories obtained directly from the Kaggle repository, is illustrated in Figure 2. These examples reflect the typical visual characteristics of the dataset, including predominantly frontal facial views and relatively controlled acquisition conditions.

In the present study, no additional preprocessing steps were applied beyond the dataset's original preparation. The images were used exactly as provided by the dataset authors, without further frame extraction, face detection, cropping, alignment, enhancement, or data augmentation. Deep visual features were directly extracted from the provided images using the SqueezeNet architecture and subsequently classified using machine learning models (ANN, RF, and SVM). Due to its large scale and origin from real driving recordings, the DDD provides a realistic benchmark for evaluating driver fatigue classification methods at the frame level under practical data constraints.

3.2. Deep Feature Extraction Using SqueezeNet

SqueezeNet was selected due to its compact architecture and favorable trade-off between model size and representational capacity. Its Fire module design enables efficient feature learning while maintaining strong discriminative power, making the network well suited for embedded and resource-constrained systems. SqueezeNet is a convolutional neural network that employs Fire modules to significantly reduce the number of parameters while achieving performance comparable to larger architectures. Each Fire module first applies 1×1 convolutional filters to reduce channel dimensionality, followed by a combination of 1×1 and 3×3 convolutional filters to expand the feature representation.

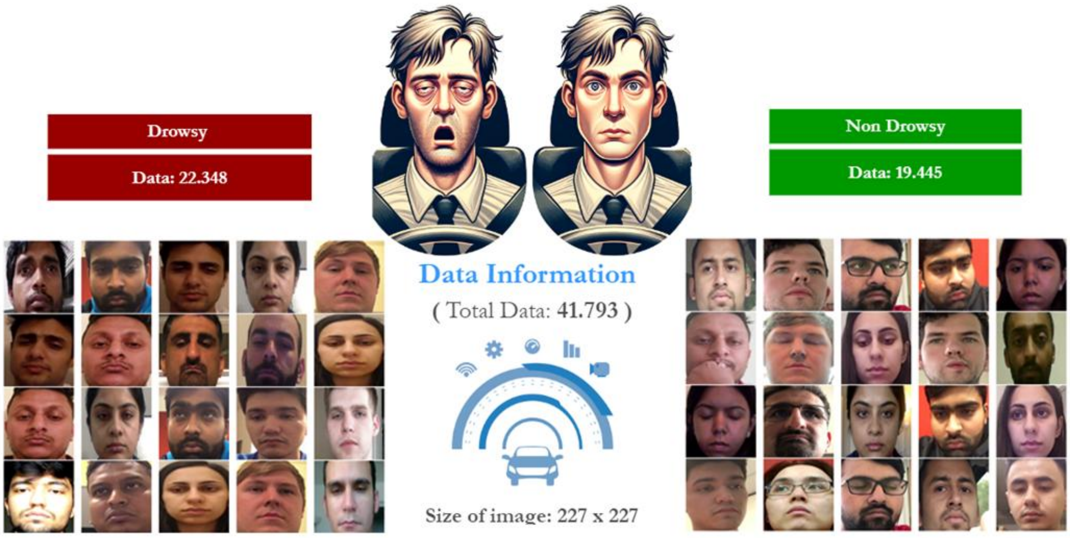


Figure 2. Class distribution and representative raw image samples from the Driver Drowsiness Dataset (DDD)

3.2.1. Lightweight Architecture

SqueezeNet contains approximately 1.2 million parameters, substantially fewer than those of ResNet-50 (approximately 25 million) or MobileNetV2 (approximately 3.4 million). Due to its compact size, SqueezeNet requires less memory and enables faster inference, making it particularly suitable for embedded and in-vehicle applications with limited computational resources [19].

3.2.2. Preservation of Discriminative Facial Features

Despite its relatively small size, the squeeze–expand architecture of SqueezeNet enables it to retain strong representational power. This property is especially beneficial for facial analysis tasks that require capturing subtle visual cues, such as eye narrowing, eyelid drooping, and micro-expressions associated with driver fatigue. As a result, SqueezeNet can effectively encode fatigue-related facial characteristics while maintaining computational efficiency [20].

3.2.3. Compatibility with Classical Machine Learning Classifiers

The final convolutional outputs of SqueezeNet provide compact, low-dimensional feature embeddings that integrate effectively with conventional machine learning classifiers such as ANNs, RFs, and SVMs. Unlike larger architectures, whose high-dimensional feature vectors often require additional dimensionality reduction, SqueezeNet embeddings can be used for classification without further processing.

In this study, SqueezeNet was employed as a pretrained lightweight deep feature extractor. Deep features were extracted from the final global average pooling layer (avgpool), which captures high-level semantic representations of facial characteristics relevant to driver fatigue. The resulting feature vector had a dimensionality of 1000 features per image. Prior to classification, the extracted features were normalized using standard feature normalization to ensure consistent scaling and to improve classifier stability. These normalized feature vectors were subsequently used as input to the ANN, Random Forest, and Support Vector Machine classifiers.

3.3. Training Models

After extracting deep feature embeddings with SqueezeNet, three supervised learning algorithms were used to classify facial images into drowsy and non-drowsy categories: ANN, RF, and SVM. ANN was employed in this study as a nonlinear classifier to model the complex relationships present in the deep feature embeddings extracted by SqueezeNet. The extracted features form high-dimensional representations in which class boundaries are not strictly linear, making ANN suitable for capturing subtle variations between drowsy and non-drowsy facial patterns. Figure 3 illustrates a conceptual structure of the ANN used for binary fatigue classification, consisting of an input layer, one or more hidden layers, and an output layer. An ANN is composed of interconnected processing units, known as artificial neurons, where each neuron processes incoming signals through weighted connections and activation functions before propagating the output to subsequent layers, enabling effective information flow through the network [21], [22]. This layered structure allows ANN to learn nonlinear decision boundaries and adapt to complex feature interactions, which are common in facial fatigue analysis tasks [23], [24].

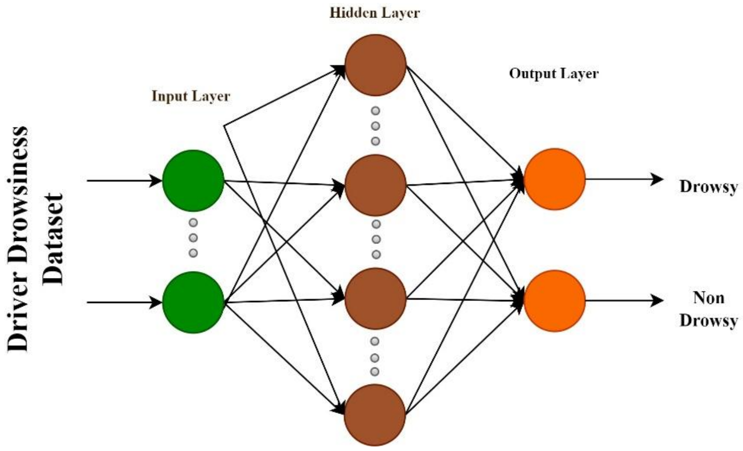


Figure 3. Conceptual illustration of the ANN structure used for binary classification (drowsy vs. non-drowsy).

RF was employed in this study as an ensemble-based classifier to evaluate the robustness of the deep feature embeddings extracted by SqueezeNet. The extracted features exhibit variability across samples, and RF effectively handles this variability by aggregating the predictions of multiple decision trees trained independently on randomly selected subsets of the data and features. Figure 4 illustrates a conceptual representation of the Random Forest decision mechanism based on majority voting. By combining multiple weak learners, RF reduces the risk of overfitting and improves generalization performance, particularly in high-dimensional feature spaces. This ensemble strategy increases model diversity and robustness across different data distributions [22], [25], [26], making RF well suited for classifying drowsy and non-drowsy facial representations in this study.

SVM was employed in this study using a radial basis function (RBF) kernel due to its effectiveness in modeling nonlinear class boundaries. The deep feature embeddings extracted by SqueezeNet exhibit complex and non-linearly separable distributions, which linear decision boundaries cannot adequately handle. Figure 5 illustrates a nonlinear decision boundary formed by an RBF-based SVM, where samples from one class can be enclosed within a flexible boundary to achieve effective separation. By applying the RBF kernel, SVM implicitly maps the input feature vectors into a higher-dimensional space, allowing the classifier to form flexible, nonlinear decision boundaries that better separate drowsy and non-drowsy facial representations [25], [27]. This property makes RBF-based SVM particularly suitable for facial fatigue detection, where subtle visual cues may vary across subjects and illumination conditions.

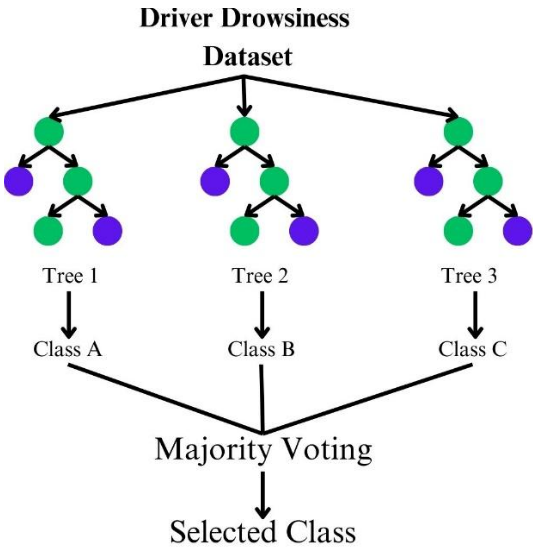


Figure 4. Conceptual illustration of the RF classification mechanism based on majority voting.

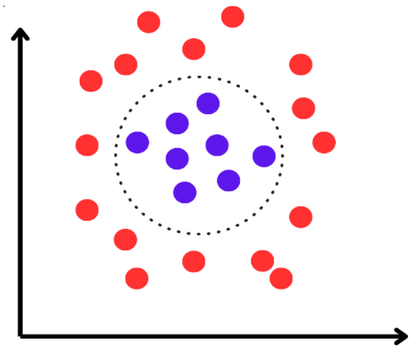
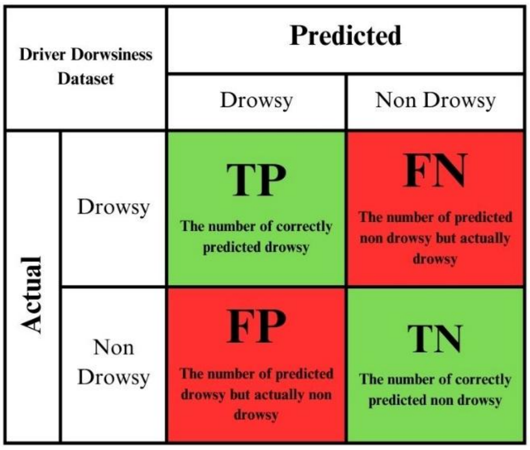


Figure 5. Illustrative example of the SVM with RBF kernel for decision boundary for binary classification.

3.4. Performance Metrics and Confusion Matrix

To evaluate the classification performance of the proposed framework, a confusion matrix was first used to analyze the predictions for each class. The confusion matrix provides a comprehensive overview of the relationship between predicted labels and ground-truth labels, enabling detailed interpretation of correct and incorrect classifications for drowsy and non-drowsy samples [28], [29]. A schematic illustration of the confusion matrix used in this study is shown in Figure 6.



Driver Dorwsiness Dataset		Predicted	
		Drowsy	Non Drowsy
Actual	Drowsy	**TP** The number of correctly predicted drowsy	**FN** The number of predicted non drowsy but actually drowsy
	Non Drowsy	**FP** The number of predicted drowsy but actually non drowsy	**TN** The number of correctly predicted non drowsy

Figure 6. Confusion matrix structure for binary driver fatigue classification.

Based on the confusion matrix, four fundamental quantities are defined: true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). These quantities form the basis for computing standard performance metrics that quantitatively assess classification effectiveness. Using TP, TN, FP, and FN, several widely adopted evaluation metrics were calculated, including accuracy, precision, recall, and F1-score. Accuracy measures the overall correctness of the classifier, while precision and recall focus on the reliability and completeness of positive predictions, respectively. The F1-score provides a balanced measure by combining precision and recall, which is particularly useful when class distributions are not perfectly balanced [30], [31]. The mathematical formulations of the evaluation metrics used in this study are defined as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{2}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3}$$

$$\text{F1} - \text{score} = \frac{2TP}{2TP + FP + FN} \tag{4}$$

These metrics were computed for each classification model to provide a consistent and fair comparison of their performance.

3.5. Data Splitting and Cross-Validation Strategy

Cross-validation is a commonly used technique for evaluating the performance of machine learning models. In conventional evaluation settings, a dataset is split into training and test subsets. In contrast, cross-validation partitions the dataset into multiple smaller subsets and iteratively uses some of these subsets for training while reserving the remaining subset for testing. This process allows the model performance to be evaluated across different data partitions, leading to a more reliable overall performance estimate [32], [33].

In this study, the dataset was divided into 10 mutually exclusive folds, with class distribution preserved in each fold via stratified sampling. During each iteration, nine folds were used for training, and the remaining fold was reserved for testing, ensuring no overlap between training and testing samples. This procedure was repeated ten times, and the performance metrics were averaged across all folds. Such a validation protocol helps reduce bias caused by random partitioning and provides a more stable estimate of model performance. Since subject-level or video-level identifiers are not available in the publicly released DDD dataset, stratified 10-fold cross-validation was conducted frame-wise using the available image samples. Although images extracted from the same video may exhibit visual similarity, this evaluation protocol is consistent with prior studies that use the DDD dataset and allows for fair comparison under the same data constraints.

4. Experimental Results

4.1. Setup and Model Configuration

A comprehensive set of experiments was conducted to evaluate the effectiveness of the proposed driver fatigue detection framework. The complete dataset, consisting of 41,793 facial images, was first processed using the SqueezeNet architecture for deep feature extraction. The resulting feature embeddings were then classified using three supervised ML models: ANN, RF, and SVM.

To ensure that the reported results were not influenced by sampling bias, all experiments were evaluated using stratified 10-fold cross-validation. In this validation scheme, the dataset was divided into 10 mutually exclusive folds, preserving the class distribution within each fold. During each iteration, nine folds were used for training, and one-fold was reserved for testing. Performance metrics were computed exclusively on unseen test samples, and the final results represent the average performance across all folds. The evaluation metrics included accuracy, precision, recall, and F1-score.

The internal parameter configurations of the three classifiers are summarized in Table 2. These parameter settings were intentionally selected based on commonly adopted configurations reported in the literature, aiming to balance computational efficiency and classification performance. Extensive hyperparameter tuning was deliberately avoided to reduce the risk of overfitting and to ensure a fair and reproducible comparison across models.

For the ANN classifier, a single hidden layer with 100 neurons was employed, using the ReLU activation function and the Adam optimization algorithm, with a maximum of 200 training iterations. The RF classifier was configured with 500 decision trees, considering five attributes at each split and enforcing a minimum node size of five samples. The SVM classifier utilized a RBF kernel with a regularization parameter of C = 1.0, epsilon (ε) = 0.10, automatic kernel scaling, a numerical tolerance of 0.001, and a maximum of 100 iterations.

Table 2. Parameter settings used for the classification models.

Model	Parameters
ANN	Hidden layer neurons: 100; Activation: ReLU; Optimizer: Adam; Maximum iterations: 200
RF	Number of trees: 500; Features per split: 5; Minimum samples per node: 5
SVM	Kernel: RBF; Cost (C): 1.0; Epsilon (ε): 0.10; Gamma: auto; Tolerance: 0.001; Maximum iterations: 100

4.2. Quantitative Performance Results

The quantitative performance results obtained from the stratified 10-fold cross-validation are presented in Table 3. Overall, all three classifiers achieved strong performance when trained on the deep feature embeddings extracted by SqueezeNet. Among the evaluated models, the ANN classifier achieved the highest accuracy of 99.97%, followed closely by the RF classifier at 99.78%. The SVM classifier achieved an accuracy of 96.33%, which, although lower than those of ANN and RF, remains competitive given the complexity of the fatigue classification task. The precision, recall, and F1-score values across all models were consistently high, indicating balanced classification performance.

Table 3. Performance metrics obtained for all classification models.

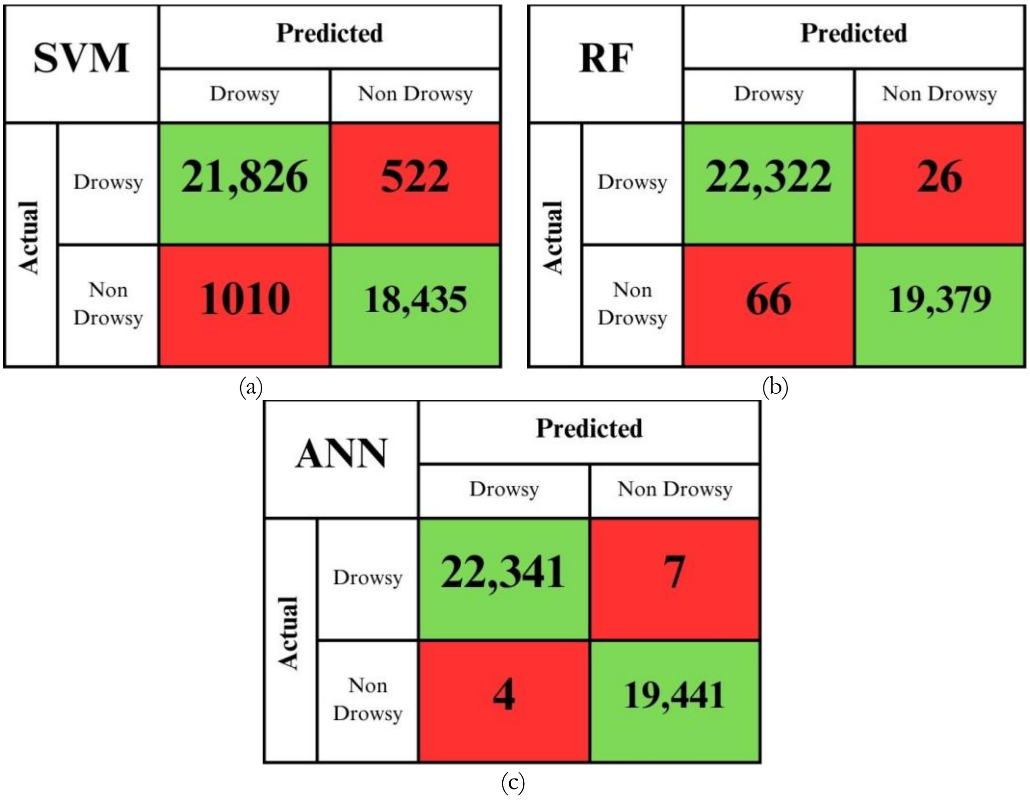
Model	Accuracy	Precision	Recall	F1-Score
ANN	99.97%	99.98%	99.97%	99.98%
RF	99.78%	99.71%	99.88%	99.79%
SVM	96.33%	95.58%	97.66%	96.61%

4.3. Confusion Matrix Analysis

To gain deeper insight into the classification behavior of each model, confusion matrices were generated for ANN, RF, and SVM, as illustrated in Figure 7.

The ANN classifier achieved the lowest number of misclassifications, correctly classifying 41,782 out of 41,793 samples, with only 11 incorrect predictions. The RF classifier followed closely, producing 41,701 correct classifications and 92 misclassifications. In contrast, the SVM classifier correctly classified 40,261 samples and misclassified 1,532. The comparatively higher error rate observed with the SVM classifier was primarily due to borderline cases in which fatigue-related facial cues were subtle, partially occluded, or inconsistently expressed. In contrast, ANN and RF demonstrated greater robustness in handling such variations, contributing to their superior overall performance.

The experimental results indicate that compact deep feature embeddings extracted by SqueezeNet can be highly effective when paired with classical machine learning classifiers. The consistently strong performance across ANN, RF, and SVM highlights the robustness and discriminative power of the extracted deep features across different learning paradigms. It is important to emphasize that the purpose of employing multiple classifiers was not to conduct a competitive comparison among machine learning algorithms, but rather to assess the robustness and effectiveness of the SqueezeNet-extracted features across diverse classification strategies. The superior performance of ANN suggests its advantage in modeling complex nonlinear relationships within the deep feature space, while the strong performance of RF further supports the generalizability of the proposed framework.



SVM		Predicted	
		Drowsy	Non Drowsy
Actual	Drowsy	21,826	522
	Non Drowsy	1010	18,435

(a)

RF		Predicted	
		Drowsy	Non Drowsy
Actual	Drowsy	22,322	26
	Non Drowsy	66	19,379

(b)

ANN		Predicted	
		Drowsy	Non Drowsy
Actual	Drowsy	22,341	7
	Non Drowsy	4	19,441

(c)

Figure 7. Confusion matrix results for each model (**a**) SVM; (**b**) RF; (**c**) ANN.

5. Discussion and Comparison

5.1. Discussion of Results

The findings of this study demonstrate that deep feature extraction using a lightweight architecture can effectively support accurate driver fatigue classification under realistic driving data conditions. Among the evaluated classifiers, ANN achieved the highest performance, followed closely by RF, while SVM showed competitive but comparatively lower accuracy. These results indicate that although SqueezeNet provides compact and discriminative feature embeddings, different classifiers exploit these representations in distinct ways.

The superior performance of ANN can be attributed to its ability to learn nonlinear decision boundaries in the high-dimensional deep feature space, enabling it to capture subtle fatigue-related facial cues, such as partial eye closure, reduced facial muscle tension, and micro-expressions, which may appear inconsistently across drivers. The RF classifier also performed well, suggesting that ensemble-based decision rules are effective for handling diverse, high-dimensional embedding features. In contrast, the SVM classifier, while effective overall, showed greater sensitivity to ambiguous or borderline samples, particularly when fatigue-related facial cues were weak or partially occluded.

5.2. Comparison with Previous Studies and Practical Implications

A quantitative comparison with prior studies using the same Driver Drowsiness Dataset (DDD) is provided in Table 4. Previous works reported accuracies ranging from approximately 96% using conventional CNN architectures to 98.54% with YOLO-Face–based pipelines and up to 100% using deeper transfer-learning models such as ResNet50V2. The proposed SqueezeNet-based feature extraction combined with ANN achieved 99.97% accuracy, matching or closely approaching the best reported results while using significantly fewer parameters.

The RF-based configuration also performed competitively with 99.78% accuracy, while the SVM-based approach achieved 96.33%, demonstrating that classical classifiers remain effective when paired with compact deep feature embeddings. Although near-perfect accuracy was achieved, potential overfitting cannot be entirely ruled out due to dataset homogeneity and controlled acquisition conditions. Nevertheless, the absence of data augmentation suggests that the achieved performance primarily stems from the discriminative capacity of the extracted features.

Overall, the lightweight nature of SqueezeNet, combined with low computational complexity, makes the proposed framework well-suited for future real-time deployment in embedded vehicle systems. Further validation on more diverse datasets encompassing different driving environments, vehicle types, and demographic variations would strengthen the applicability of the proposed approach in real-world scenarios.

Table 4. Quantitative comparison of the proposed method with previous studies using the Driver Drowsiness Dataset.

Study	Dataset	Method	Accuracy (%)	Remarks
Nasri et al. [1]	RLDD-derived images (basis of DDD)	CNN (from scratch + TL)	>96	Initial work by the dataset creators; accuracy reported on the Kaggle dataset page.
Andrean et al. [2]	DDD	YOLO-Face + Decision Tree	98.54	Face detection combined with a classical classifier.
Andrean et al. [2]	DDD	Haar Cascade + Decision Tree	98.03	Slightly lower performance compared to the YOLO-Face pipeline.
Proposed Study	DDD	SqueezeNet + ANN	99.97	Best performance among classical ML classifiers using deep features.
Proposed Study	DDD	SqueezeNet + RF	99.78	Ensemble-based classifier with excellent performance.
Proposed Study	DDD	SqueezeNet + SVM	96.33	Competitive performance compared to earlier CNN-based approaches.

5.3. Methodological Considerations and Limitations

An important implication of these findings is that high fatigue-detection accuracy does not strictly require computationally intensive end-to-end deep learning models. Instead, combining compact deep feature extractors with efficient classical machine learning classifiers can yield competitive performance while offering practical advantages for embedded vehicle systems.

The evaluation was conducted using frame-wise cross-validation due to the absence of subject-level or video-level identifiers in the publicly available DDD dataset. Therefore, the reported results should be interpreted as a frame-level proof-of-concept evaluation of feature robustness rather than subject-independent generalization. This evaluation protocol is consistent with prior studies using the same dataset and enables fair comparison under identical data constraints.

Despite the promising results, several limitations remain. The dataset primarily consists of frontal facial images, and variations such as extreme head poses, the use of sunglasses, and low-light driving conditions may reduce classification performance. Furthermore, the analysis was performed on individual image frames rather than temporal sequences, meaning that dynamic fatigue indicators such as blink duration or yawning frequency were not explicitly modeled.

6. Conclusions

This research introduced an efficient driver fatigue detection system that integrates SqueezeNet-based deep feature extraction with three supervised learning algorithms: ANN, RF, and SVM. When evaluated on a large, realistic driver dataset, the proposed system demonstrated a strong ability to distinguish fatigue-related facial patterns. The ANN classifier achieved 99.97% accuracy, with RF producing results that were closely comparable. These findings indicate that compact convolutional feature extractors can capture critical facial characteristics associated with driver fatigue without relying on large, computationally expensive network architectures. The proposed framework is particularly suitable for future real-time deployment in embedded vehicle systems, especially in environments with limited computational resources. By combining lightweight deep feature embeddings with classical machine learning classifiers, the system can significantly reduce inference time while maintaining high classification accuracy. Future work will focus on extending the proposed framework by incorporating temporal modeling, multimodal inputs (such as steering behavior and physiological signals), and domain adaptation techniques to improve robustness across diverse driving conditions and environmental contexts. These directions are expected to enhance further the adaptability and practical applicability of driver monitoring systems for modern transportation safety applications.

Author Contributions: Conceptualization: E.T.Y., T.A.C., and M.K.; Methodology: E.T.Y. and T.A.C.; Software: T.A.C. and E.T.Y.; Validation: E.T.Y., S.G., S.S.B., B.G., A.G., and M.K.; Formal analysis: E.T.Y.; Investigation: E.T.Y., S.G., S.S.B., B.G., A.G., and M.K.; Resources: E.T.Y.; Data curation: E.T.Y.; Writing—original draft preparation: E.T.Y. and T.A.C.; Writing—review and editing: E.T.Y., T.A.C., S.G., S.S.B., B.G., A.G., and M.K.; Visualization: B.G.; Supervision: M.K.; Project administration: M.K.. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The dataset used in this study is publicly available through the Kaggle Data repository at https://www.kaggle.com/datasets/ismailnasri20/driver-drowsiness-dataset-ddd.

Acknowledgments: The authors would like to acknowledge the administrative and technical support provided by Selçuk University and Alışan Logistics Inc. during the preparation of this study. Their assistance contributed to the completion and overall quality of the manuscript. In addition, during the preparation of this work, the authors used DeepL Translator and Grammarly to support language translation and grammar refinement. All content was subsequently reviewed and edited by the authors, who take full responsibility for the accuracy and integrity of the published article.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] I. Nasri, M. Karrouchi, H. Snoussi, K. Kassmi, and A. Messaoudi, "Detection and Prediction of Driver Drowsiness for the Prevention of Road Accidents Using Deep Neural Networks Techniques," in *Lecture Notes in Electrical Engineering*, 2022, pp. 57–64. doi: 10.1007/978-981-33-6893-4_6.

[2] M. N. Andrean *et al.*, "Comparing Haar Cascade and YOLOFACE for Region of Interest Classification in Drowsiness Detection," *J. Media Inform. Budidarma*, vol. 8, no. 1, p. 272, Jan. 2024, doi: 10.30865/mib.v8i1.7167.

[3] T. Ahmed, O. Jyoti, and T. H. Mou, "Drivers' Drowsiness Detection System Enhancement Using Deep Learning: CNN-based Approach," in *2023 26th International Conference on Computer and Information Technology (ICCIT)*, Dec. 2023, pp. 1–6. doi: 10.1109/ICCIT60459.2023.10441229.

[4] W. Deng and R. Wu, "Real-Time Driver-Drowsiness Detection System Using Facial Features," *IEEE Access*, vol. 7, pp. 118727–118738, 2019, doi: 10.1109/ACCESS.2019.2936663.

[5] H. A. Kassem, M. Chowdhury, and J. H. Abawajy, "Drivers Fatigue Level Prediction Using Facial, and Head Behavior Information," *IEEE Access*, vol. 9, pp. 121686–121697, 2021, doi: 10.1109/ACCESS.2021.3108561.

[6] V. Owen and N. Surantha, "Computer Vision-Based Drowsiness Detection Using Handcrafted Feature Extraction for Edge Computing Devices," *Appl. Sci.*, vol. 15, no. 2, p. 638, Jan. 2025, doi: 10.3390/app15020638.

[7] W. Qin, L. Liu, and L. Sun, "Video-based driver drowsiness detection using stacked landmark heatmaps and spatial–temporal attention," *Comput. Graph.*, vol. 131, p. 104336, Oct. 2025, doi: 10.1016/j.cag.2025.104336.

[8] M. H. Alkinani, W. Z. Khan, Q. Arshad, and M. Raza, "HSDDD: A Hybrid Scheme for the Detection of Distracted Driving through Fusion of Deep Learning and Handcrafted Features," *Sensors*, vol. 22, no. 5, p. 1864, Feb. 2022, doi: 10.3390/s22051864.

[9] J. Uddin, "A Novel Lightweight Deep Learning Approach for Drivers' Facial Expression Detection," *Designs*, vol. 9, no. 2, p. 45, Apr. 2025, doi: 10.3390/designs9020045.

[10] A. R. Shahid and H. Yan, "SqueezExpNet: Dual-stage convolutional neural network for accurate facial expression recognition with attention mechanism," *Knowledge-Based Syst.*, vol. 269, p. 110451, Jun. 2023, doi: 10.1016/j.knosys.2023.110451.

[11] I. Cinar, "Deep Learning and LSTM Integration for Analyzing Driver Behaviors," *Turkish J. Math. Comput. Sci.*, vol. 17, no. 1, pp. 191–211, Jun. 2025, doi: 10.47000/tjmcs.1644390.

[12] S. K. S. Al-doorı, Y. S. Taspınar, and M. Koklu, “Distracted Driving Detection with Machine Learning Methods by CNN Based Feature Extraction,” *Int. J. Appl. Math. Electron. Comput.*, vol. 9, no. 4, pp. 116–121, Dec. 2021, doi: 10.18100/ijamec.1035749.

[13] L. R. Zuama, D. R. I. M. Setiadi, A. Susanto, S. Santosa, H.-S. Gan, and A. A. Ojugo, “High-Performance Face Spoofing Detection using Feature Fusion of FaceNet and Tuned DenseNet201,” *J. Futur. Artif. Intell. Technol.*, vol. 1, no. 4, pp. 385–400, Feb. 2025, doi: 10.62411/faith.3048-3719-62.

[14] M. Dua, Shakshi, R. Singla, S. Raj, and A. Jangra, “Deep CNN models-based ensemble approach to driver drowsiness detection,” *Neural Comput. Appl.*, vol. 33, no. 8, pp. 3155–3168, Apr. 2021, doi: 10.1007/s00521-020-05209-7.

[15] I. Nasri, M. Karrouchi, H. Snoussi, K. Kassmi, and A. Messaoudi, “DistractNet: a deep convolutional neural network architecture for distracted driver classification,” *IAES Int. J. Artif. Intell.*, vol. 11, no. 2, p. 494, Jun. 2022, doi: 10.11591/ijai.v11.i2.pp494-503.

[16] E. Magán, M. P. Sesmero, J. M. Alonso-Weber, and A. Sanchis, “Driver Drowsiness Detection by Applying Deep Learning Techniques to Sequences of Images,” *Appl. Sci.*, vol. 12, no. 3, p. 1145, Jan. 2022, doi: 10.3390/app12031145.

[17] M. U. Hossain, M. A. Rahman, M. M. Islam, A. Akhter, M. A. Uddin, and B. K. Paul, “Automatic driver distraction detection using deep convolutional neural networks,” *Intell. Syst. with Appl.*, vol. 14, p. 200075, May 2022, doi: 10.1016/j.iswa.2022.200075.

[18] I. Nasri, “Driver Drowsiness Dataset (DDD),” *Kaggle.com*, 2022. https://www.kaggle.com/datasets/ismailnasri20/driver-drowsiness-dataset-ddd

[19] M. Koklu, R. Kursun, E. T. Yasin, and Y. S. Taspinar, “Detection of Defects in Soybean Seeds by Extracting Deep Features with SqueezeNet,” in *2023 IEEE 12th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, Sep. 2023, pp. 713–717. doi: 10.1109/IDAACS58523.2023.10348939.

[20] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, “SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size,” *arXiv*. Nov. 04, 2016. doi: 10.48550/arXiv.1602.07360.

[21] Y. S. Taspinar, “Machine Learning-Driven E-Nose-Based Diabetes Detection: Sensor Selection and Feature Reduction Study,” *Sensors*, vol. 25, no. 21, p. 6607, Oct. 2025, doi: 10.3390/s25216607.

[22] H. F. Sen and Y. S. Taspinar, “Obesity Prediction Using Machine Learning Methods: Performance Analysis of DT, RF, and ANN Models,” in *2025 7th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (ICHORA)*, May 2025, pp. 1–6. doi: 10.1109/ICHORA65333.2025.11016848.

[23] M. B. Yildiz, E. T. Yasin, and M. Koklu, “Fisheye freshness detection using common deep learning algorithms and machine learning methods with a developed mobile application,” *Eur. Food Res. Technol.*, vol. 250, no. 7, pp. 1919–1932, Jul. 2024, doi: 10.1007/s00217-024-04493-0.

[24] Y.-S. Park and S. Lek, “Artificial neural networks: Multilayer perceptron for ecological modeling,” in *Developments in environmental modelling*, vol. 28, Elsevier, 2016, pp. 123–140. doi: 10.1016/B978-0-444-63623-2.00007-4.

[25] K. Erdem, M. B. Yildiz, E. T. Yasin, and M. Koklu, “A Detailed Analysis of Detecting Heart Diseases Using Artificial Intelligence Methods,” *Intell. Methods Eng. Sci.*, vol. 2, no. 4, pp. 115–124, Dec. 2023, doi: 10.58190/imiens.2023.71.

[26] B. Isgor and M. Koklu, “Lightweight Hybrid Model for Bone Fracture Detection Using MobileNetV2 Feature Extraction and Ensemble Learning,” *J. Futur. Artif. Intell. Technol.*, vol. 2, no. 3, pp. 521–533, Dec. 2025, doi: 10.62411/faith.3048-3719-284.

[27] A. Khamis, A. Ishaq, M. E. Irhebhude, and D. T. Chinyio, “Dimensionality-Aware Dry Bean Classification Using Transfer Learning and SVM: Addressing Variety and Resolution Constraints,” *J. Futur. Artif. Intell. Technol.*, vol. 2, no. 3, pp. 417–431, Sep. 2025, doi: 10.62411/faith.3048-3719-140.

[28] N. D. Marom, L. Rokach, and A. Shmilovici, “Using the confusion matrix for improving ensemble classifiers,” in *2010 IEEE 26-th Convention of Electrical and Electronics Engineers in Israel*, Nov. 2010, pp. 000555–000559. doi: 10.1109/EEEI.2010.5662159.

[29] B. Gencturk *et al.*, “Detection of hazelnut varieties and development of mobile application with CNN data fusion feature reduction-based models,” *Eur. Food Res. Technol.*, vol. 250, no. 1, pp. 97–110, Jan. 2024, doi: 10.1007/s00217-023-04369-9.

[30] K. Tutuncu, I. Cinar, R. Kursun, and M. Koklu, “Edible and Poisonous Mushrooms Classification by Machine Learning Algorithms,” in *2022 11th Mediterranean Conference on Embedded Computing (MECO)*, Jun. 2022, pp. 1–4. doi: 10.1109/MECO55406.2022.9797212.

[31] D. Singh *et al.*, “Classification and Analysis of Pistachio Species with Pre-Trained Deep Learning Models,” *Electronics*, vol. 11, no. 7, p. 981, Mar. 2022, doi: 10.3390/electronics11070981.

[32] C. Schaffer, “Selecting a classification method by cross-validation,” *Mach. Learn.*, vol. 13, no. 1, pp. 135–143, Oct. 1993, doi: 10.1007/BF00993106.

[33] M. Koklu, I. Cinar, Y. S. Taspinar, and R. Kursun, “Identification of Sheep Breeds by CNN- Based Pre-Trained Inceptionv3 Model,” in *2022 11th Mediterranean Conference on Embedded Computing (MECO)*, Jun. 2022, pp. 01–04. doi: 10.1109/MECO55406.2022.9797214.