# Artificial Intelligence with Deep Learning Driven Attention-Guided Temporal Convolutional Network for Skin Cancer Detection in Biomedical Images

**Manar Ahmed Hamza**
Department of Computer and Self Development, Preparatory Year Deanship, Prince Sattam bin Abdulaziz University, AlKharj, Saudi Arabia
ma.hamza@psau.edu.sa (corresponding author)

**Ameerah Almutiry**
Department of Information Technology, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia
amalmutiry@kau.edu.sa

*Received: 25 February 2026 | Revised: 22 March 2026 | Accepted: 1 April 2026*

 *DOI: https://doi.org/10.48084/etasr.18371*

**ABSTRACT**

**Skin diseases are a widespread global health concern, as pathogenic agents lead to both physical illness and psychological distress and, in severe cases, may result in skin cancer. Sustainable Development Goal (SDG) 3, which refers to Good Health and Well-Being, aims to ensure healthy lives and promote well-being for people of all ages, with a focus on reducing mortality from diseases. However, accurate identification of skin conditions from clinical imagery remains a significant challenge in medical image analysis, which has led to the development of various early detection technologies. Recently, image-based detection has witnessed considerable progress due to advances in Deep Learning (DL). DL-enabled approaches have shown outstanding performance in classifying and segmenting skin lesions due to their potential to capture intricate features from skin lesion imagery with high precision. This article presents a Feature Fusion and Attention-Guided Temporal Convolutional Neural Network for Skin Cancer Detection (FFATCN-SCD) framework, which aims to develop an efficient and intelligent system for accurate skin cancer identification and classification. The FFATCN-SCD technique initially performs image preprocessing using a Bilateral Filter (BF) to enhance image quality by reducing noise, followed by feature extraction using the fusion of EfficientNetV2 and InceptionV3 to capture detailed features that contribute to efficient detection. Subsequently, a Temporal Convolutional Network (TCN) with an Attention Mechanism (AM) is utilized to effectively classify skin cancer, and Antlion Optimization (ALO) is applied for optimal parameter tuning of the classification network to ensure improved model performance. To demonstrate the superior performance of the FFATCN-SCD approach, comprehensive simulations are conducted, and the results are evaluated using several metrics. Comparative analysis further demonstrates the superiority of the FFATCN-SCD model across various evaluation metrics.**

***Keywords-skin cancer; EfficientNetV2; Antlion Optimization (ALO); Temporal Convolutional Network (TCN); Deep Learning (DL); InceptionV3***

## I. INTRODUCTION

Cancer is described as the uncontrolled development of abnormal tissue in a particular body region. Globally, skin cancer is one of the most rapidly spreading diseases, in which abnormal skin cells arise in an uncontrolled manner [1]. Accurate early identification and precise diagnosis are essential for effective cancer treatment. Melanoma**,** the deadliest type of skin cancer, is responsible for most skin cancer-related deaths in advanced countries [2]. Early skin cancer detection is critical for reducing mortality and improving prognosis. However, skin cancer identification typically relies on clinical imaging with limited sensitivity and is subsequently validated using medical samples [3]. This method is generally not suitable for assessing cancer treatment and screening responses. Computer-Aided Diagnosis (CAD) may rapidly, consistently, and reliably diagnose several diseases [4]. CAD also offers opportunities for sophisticated tumor detection and prevention that is both accurate and cost-efficient. Human organ diseases are normally evaluated using image technology. Dermatoscopic imaging

analysis, clinical screening, Computed Tomography (CT), and other methods are primarily utilized for the visual diagnosis of skin lesions [5]. However, dermatologists with less experience may demonstrate decreased precision in skin lesion diagnosis, highlighting the need for automated diagnostic systems. Numerous CAD methods have been developed using various border extraction, classification, detection, and selection techniques [6]. Several studies have investigated image preprocessing methods for diagnosing skin cancer; additionally, these studies compare CAD and Artificial Intelligence (AI) model performance with the diagnostic precision of experienced dermatologists [7]. Hence, Deep Learning (DL) techniques have displayed promising outcomes and high efficiency for data analysis and image processing [8, 9].

This paper presents a Feature Fusion and Attention-Guided Temporal Convolutional Neural Network for Skin Cancer Detection (FFATCN-SCD) framework. The key contributions are given as follows:

- Introduces a fusion-based feature representation model involving EfficientNetV2 and InceptionV3 to capture detailed features contributing to effective detection.
- Applies a Temporal Convolutional Network (TCN) with an Attention Mechanism (AM) to efficiently classify skin cancer, and Antlion Optimization (ALO) is implemented for optimal parameter tuning of the classification process.
- The proposed model follows a multifaceted approach integrating Bilateral Filter (BF)-based preprocessing, a feature fusion process, an attention-aware TCN, and ALO-based hyperparameter tuning for skin cancer classification, demonstrating the novelty of the work.
- To demonstrate the improved performance of the FFATCN-SCD system, extensive simulation studies are conducted, and the results are evaluated using diverse metrics.

## II. PREVIOUS STUDIES ON SKIN CANCER DETECTION

A summary of prior research related to skin cancer detection is presented in this section. Authors in [10] proposed an approach that may be integrated into a digital health environment, effectively saving lives by enabling early intervention and diagnosis, particularly in underserved areas. This research employs DL and few-shot learning within a Machine Learning (ML) framework for skin cancer diagnosis. The input images are preprocessed using a DL-based method for effective identification and prediction. Authors in [11] suggested a novel DL technique that depends on the MetaFormer structure, enhanced particularly for skin cancer detection. This method features a hybrid design that substitutes conventional self-attention models with a new focal self-attention module, improving its capability to identify crucial areas, reduce noise, and extract features more efficiently, finally boosting diagnostic precision. In [12], a robotic CAD system was presented for the early identification of skin cancer. The study aims to improve the efficacy of skin cancer detection. Primarily, the input imagery undergoes preprocessing to enhance image quality and extract relevant features, followed by classification of skin lesions.

Authors in [13] presented Skin-DeepNet, an innovative DL-driven framework intended for the automatic early detection of skin cancer lesions from medical imaging. Skin-DeepNet integrates a dual-branch preprocessing phase for enhancing image contrast, accomplished by robust skin lesion segmentation employing the GrabCut and Mask Region-based Convolutional Neural Network (Mask R-CNN) methods to achieve excellent segmentation precision. In [14], a fuzzy logic-driven image segmentation approach along with a modified DL technique is presented for skin cancer identification. The preprocessing stage is designed to enhance lesion visibility by eliminating artifacts such as dermoscopic scales and hair follicles. Authors in [15] proposed the utilization of Deep Block Convolutional Neural Networks (DB-CNNs) as a novel approach to detect skin cancer. The presented method was trained and tested using a large dataset containing images of various skin lesions. DB-CNNs are designed to capture the advantages of hierarchical features of skin lesions, including intricate patterns and textures that are vital for precise diagnosis.

## III. MATERIALS AND METHODS

In this study, an FFATCN-SCD framework is proposed. The FFATCN-SCD methodology begins with an image preprocessing phase, where a BF is implemented to improve image quality by reducing noise while preserving critical skin lesion details. Following preprocessing, a feature extraction stage is conducted using a fusion of EfficientNetV2 and InceptionV3 models to extract significant attributes from the input imagery. The extracted features are further passed to a classification stage, where a TCN integrated with an AM efficiently detects and classifies skin cancer types. To further improve model performance, an optimization step is utilized using the ALO algorithm for tuning the classification network parameters. Figure 1 illustrates the overall process of the FFATCN-SCD model.

### *A. Image Preprocessing*

Initially, image preprocessing is performed using a BF to enhance image quality while reducing noise. The BF is applied to the raw skin lesion images. It levels the data by combining intensity and spatial information. The BF is an effective tool for image preprocessing as it smooths the input data while preserving vital features and edges [16]. In this approach, BF is modified by incorporating magnitude awareness, ensuring that the preprocessing adapts dynamically to the varying contrast and color intensity of skin cancer regions. Let $H(a)$ denote the input data at position $a$. The filtered output of the BF, $H_i(a)$, is given as follows:

$$H_i(a) = \frac{1}{W(a)} \sum_{b \in C(a)} H(b) \cdot g_y(\|a-b\|) \cdot g_x(|H(a)-H(b)|) \quad (1)$$

Here, $C(a)$ denotes the neighborhood around pixel $a$, and $H_i(a)$ represents the output of the BF at position $a$. The spatial Gaussian function $g_y(\|a-b\|)$ is defined as:

$$g_y(\|a-b\|) = \exp\left(-\frac{\|a-b\|^2}{2\sigma_y^2}\right) \quad (2)$$

Here, $\sigma_y$ represents the spatial variance controlling the extent of spatial smoothing. The range Gaussian function $g_x(|H(a)-H(b)|)$ is specified as:

$$g_x(|H(a)-H(b)|) = \exp\left(-\frac{\|H(a)-H(b)\|^2}{2\sigma_x^2}\right) \quad (3)$$

where $\sigma_x$ denotes intensity-based smoothing. Finally, $W(a)$ represents the normalization factor:

$$W(a) = \sum_{b\in C(a)} g_y(\|a-b\|) \cdot g_x(|H(a)-H(b)|) \quad (4)$$

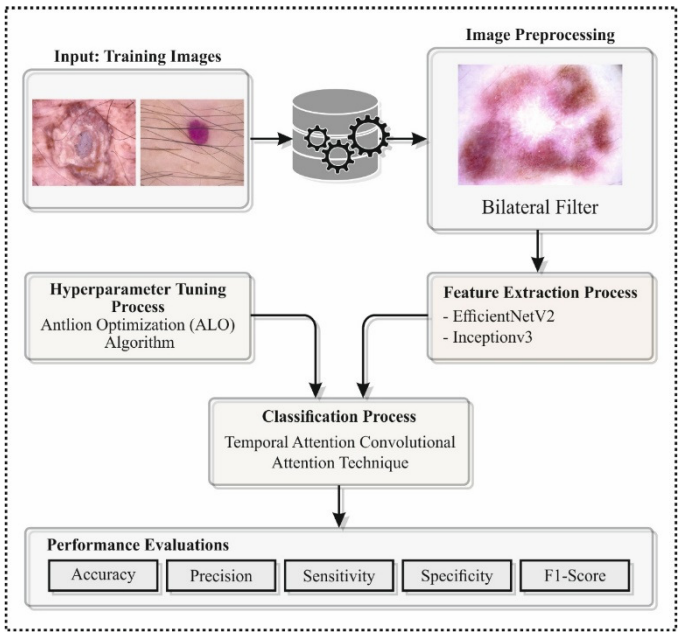


Fig. 1. Overall process of the FFATCN-SCD framework.

## B. Fusion Techniques

The feature extraction is carried out through the fusion of EfficientNetV2 and InceptionV3 to extract detailed features for accurate identification.

### 1) EfficientNetV2

EfficientNet-V2 is a ConvNet architecture recognized for its improved speed–precision trade-off. It incorporates MBConv and Fused-MBConv blocks, allowing faster training in initial layers and effective feature extraction in the deeper layers [17]. This method utilizes compound scaling, which equally scales the depth, width, and resolution to maximize performance under constrained computing resources. For this task, EfficientNet-V2-S, with nearly 24 million parameters and an input image dimension of 384 × 384, provides a good balance between model accuracy and complexity. Training uses transfer learning, where the model is initialized with ImageNet-pretrained weights and subsequently fine-tuned. Methods such as data augmentation (including zooms, flips, and rotations) and the Swish activation function improve the model's generalization and robustness. The model is typically trained using the Adam optimizer. The output layer uses a Softmax function to map the extracted features into four classes. The structural efficacy and scalable design of EfficientNet-V2 make it well suited for real-time, higher-precision classification methods.

### 2) InceptionV3

The InceptionV3 network is continually developed and refined based on GoogLeNet (InceptionV1). Through the continuous optimization of several Inception modules within the network, the architecture has been progressively improved, most notably in InceptionV3 [18]. The key innovations are as follows:

- Inception module A: A large convolutional kernel is divided into numerous smaller kernels, such as splitting a 5 × 5 kernel into two 3 × 3 kernels and a 7 × 7 kernel into three 3 × 3 kernels. This significantly reduces the number of parameters while preserving the receptive field.
- Inception module B: A convolutional kernel is divided into asymmetric convolution kernels, also known as spatially separable convolutions. For instance, a 3 × 3 convolution is replaced with a combination of 1 × 3 and 3 × 1 convolutions. This approach reduces both the number of parameters and computational complexity.
- Inception module C: Asymmetric convolutional kernels are further factorized to increase the network's representational capacity and capture more diverse image features.
- Grid size reduction module: This module performs parallel convolution and pooling operations to achieve downsampling while increasing the number of channels, improving the model's computational efficiency.

Due to the significant advantages of InceptionV3 in extracting digital image features, this study focuses on removing the network's fully connected layer outputs (predictions) and utilizing the extracted features for medical image analysis.

## C. Skin Cancer Detection Model

A TCN-AM is utilized to efficiently classify skin cancer. In comparison with classical Recurrent Neural Networks (RNNs), 1D convolutional networks enable stable gradient propagation, reducing issues related to vanishing or exploding gradients and offering strong parallel computation capabilities [19]. However, convolution layers, particularly Causal Convolution (C-Conv), are limited by a local receptive field, making them less effective for long sequences and long-term dependency modeling. By eliminating more detailed temporal information, increasing the receptive field or deepening the network can enhance feature extraction. Nevertheless, these changes substantially increase the training time of the system and are often followed by overfitting. This necessitates the exploration of enhanced frameworks to tackle these challenges. To lessen these concerns, C-Conv is effectively enhanced to Dilated Causal Convolution (DC-Conv), which enlarges the receptive field by increasing the dilation of the convolution kernel. This approach expands the convolution kernel without increasing the network depth, thereby improving the model's ability to process longer sequence data more effectively.

The TCN attention block is mainly intended for processing sequential information and capturing long-term dependencies in time series. The receptive field is influenced by factors such as kernel size, network depth, and dilation rate. This module consists of three sub-blocks, including weight normalization, dropout, convolution, and the ReLU activation function. These components are encapsulated into a single TCN block. Then, the dilation coefficient and causality are integrated to expand the convolutional receptive field, thus improving the model's ability to process long sequences. In this context, $w[m]$ denotes the convolution kernel weights, $y[x]$ represents the output of the convolution operation, $d$ denotes the dilation rate, and $x[t - d \cdot m]$ refers to the input sequence component.

$$y[x] = \sum_{m=0}^{k-1} w\,[m] \cdot x[t - d \cdot m] \quad (5)$$

Meanwhile, to prevent network degradation and mitigate the vanishing gradient problem, these layers are stacked and connected via residual connections, enabling the TCN to capture features at multiple time scales and integrate them effectively. The ReLU activation function is defined as:

$$ReLU(x) = \max(0, x) \quad (6)$$

$$f(x) = ReLU\big(Weight_{Norm}(y[x])\big) \quad (7)$$

$$output_{TCN} = TCN(x) = dropout\big(f(x)\big),\ n \in \{1,2,3\} \quad (8)$$

*D. Antlion Optimization Algorithm*

Lastly, ALO is leveraged for optimal parameter tuning of the classification model, ensuring enhanced model performance. ALO is a population-based metaheuristic algorithm inspired by the hunting behavior of antlions [20]. The stepwise process for ALO execution is defined as follows. The algorithm simulates the interaction between ants and antlions through a trapping mechanism. The motion of ants is modeled using a random walk as formulated below:

$$X(t) = [0, \text{cumsum}(2r(t_1) - 1), \text{cumsum}(2r(t_2) - 1), \ldots, \text{cumsum}(2r(t_n) - 1)] \quad (9)$$

where cumsum denotes the cumulative sum. The function $r(t)$ is defined as:

$$r(t) = \begin{cases} 1 & \text{if } rand > 0.5 \\ 0 & \text{if } rand \le 0.5 \end{cases} \quad (10)$$

where $rand$ is a randomly generated number in the range [0,1], and $t$ represents the time step of the random walk. During initialization, the positions of ants $M_a$ are represented using the following matrix:

$$M_a = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,d} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,d} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,d} \end{pmatrix} \quad (11)$$

For evaluating each ant, a fitness value $P_t$ is assigned and is represented as:

$$M_{oa} = \begin{pmatrix} F_t([a_{1,1}, a_{1,2}, \ldots a_{1,d}]) \\ F_t([a_{2,1}, a_{2,2}, \ldots a_{2,d}]) \\ \vdots \\ F_t([a_{n,1}, a_{n,2}, \ldots a_{n,d}]) \end{pmatrix} \quad (12)$$

In real-world terms, ants explore the search space, and maintaining their positions along with fitness values is essential for the optimizer. The following matrices represent the positions $M_{al}$ and fitness values $M_{oal}$ of antlions:

$$M_{al} = \begin{bmatrix} al_{1,1} & al_{1,2} & \cdots & al_{1,d} \\ al_{2,1} & al_{2,2} & \cdots & al_{2,d} \\ \vdots & \vdots & \vdots & \vdots \\ al_{n,1} & al_{n,2} & \cdots & al_{n,d} \end{bmatrix} \quad (13)$$

$$M_{oal} = \begin{bmatrix} F_t([al_{1,1}, al_{1,2}, \ldots, al_{1,d}]) \\ F_t([al_{2,1}, al_{2,2}, \ldots, al_{2,d}]) \\ \vdots \\ F_t([al_{n,1}, al_{n,2}, \ldots, al_{n,d}]) \end{bmatrix} \quad (14)$$

In the search space, ants update their positions by performing a random walk at each iteration, followed by normalization:

$$x_i^t = \frac{(x_i^t - a_i) \times (d_i - c_i^t)}{(d_i^t - a_i)} + c_i^t \quad (15)$$

Here, $a_i$ and $d_i$ denote the minimum and maximum values of the random walk for variable $i$. In the search space, ants perform random walks and may be trapped by antlions. This behavior is modeled as:

$$c_i^t = al_j^t + c^t \quad (16)$$

$$d_i^t = al_j^t + d^t \quad (17)$$

To simulate the hunting capability of antlions, the model incorporates roulette wheel selection. In the ALO method, this operator is used to select an antlion based on its fitness value during the optimization process. It increases the probability of stronger antlions capturing ants.

Using this mechanism, antlions build traps whose dimensions depend on their fitness, while ants move randomly. When an ant is trapped, the antlion narrows the search region by reducing the boundaries of the random walk:

$$c^t = \frac{c^t}{I} \quad (18)$$

$$d^t = \frac{d^t}{I} \quad (19)$$

where $I$ denotes a ratio, and $c^t$ and $d^t$ represent the upper and lower bounds at the $t$-th iteration. When an ant reaches the bottom of the pit, the hunting process is completed, and the ant is captured by the antlion. The antlion then updates its position to match the captured ant if the ant has better fitness:

$$al_j^t = a_i^t, \quad \text{if } f(a_i^t) > f\big(al_j^t\big) \quad (20)$$

Elitism is a key feature of evolutionary algorithms that ensures the preservation of the best solution during the optimization process. The ALO method defines a fitness

function to achieve optimal classification performance. It assigns a positive value representing the effectiveness of the candidate solution. In this work, the fitness function is defined as the classification error rate:

$$fitness(x_i) = ClassifierErrorRate(x_i) = \frac{number\ of\ misclassified\ instances}{Total\ number\ of\ instances} * 100 \quad (21)$$

## IV. EXPERIMENTAL RESULTS AND ANALYSIS

The proposed model is simulated using the Python 3.8.5 tool on a PC with Intel Core i5-8600K CPU, GeForce GTX 1050 Ti (4 GB), 16 GB RAM, 250 GB SSD, and 1 TB HDD. The parameter settings are given as follows: learning rate: 0.01, dropout: 0.5, batch size: 5, epoch count: 50, and activation: ReLU. In this section, the experimental analysis of the FFATCN-SCD framework is investigated using the Skin Cancer ISIC dataset [21]. The dataset comprises a total of 2,239 dermatological image samples, classified into 9 distinct skin lesion classes. The experimental results are examined in terms of accuracy ($accur_y$), precision ($preci_n$), sensitivity ($sens_y$), specificity ($spec_y$), and F1-score ($F1_{score}$).

The distribution across classes is as follows: Actinic Keratosis (C1) 114 images, Basal Cell Carcinoma (C2) 376 images, Dermatofibroma (C3) 95 images, Melanoma (C4) 438 images, Nevus (C5) 357 images, Pigmented Benign Keratosis (C6) 462 images, Seborrheic Keratosis (C7) 77 images, Squamous Cell Carcinoma (C8) 181 images, and Vascular Lesion (C9) 139 images. This diverse and moderately imbalanced dataset is suitable for training and assessing machine learning algorithms for skin lesion detection and classification tasks. Figure 2 illustrates the sample images.

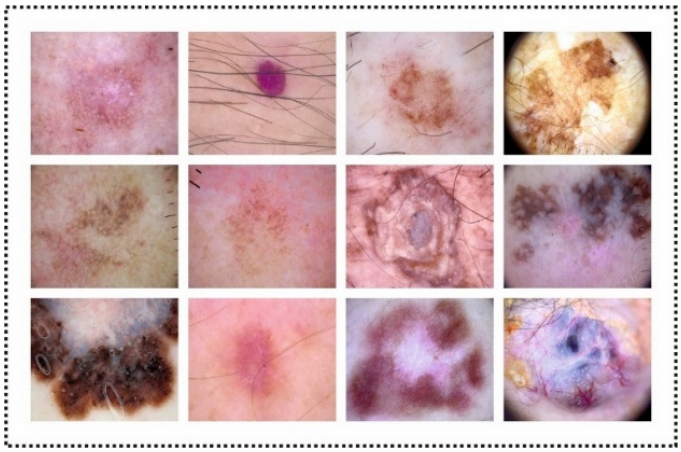

Fig. 2. Sample images.

Figure 3 presents the skin cancer detection performance of the FFATCN-SCD approach using a 70% training and 30% testing split under several measures. On the 70% training phase, the proposed FFATCN-SCD model achieves average $accur_y$, $preci_n$, $sens_y$, $spec_y$, and $F1_{score}$ of 98.23%, 91.78%, 85.80%, 98.96%, and 88.23%, respectively. Furthermore, on 30% testing, the proposed FFATCN-SCD system gains average $accur_y$, $preci_n$, $sens_y$, $spec_y$, and $F1_{score}$ of 98.45%, 91.84%, 87.59%, 99.09%, and 89.31%, respectively.

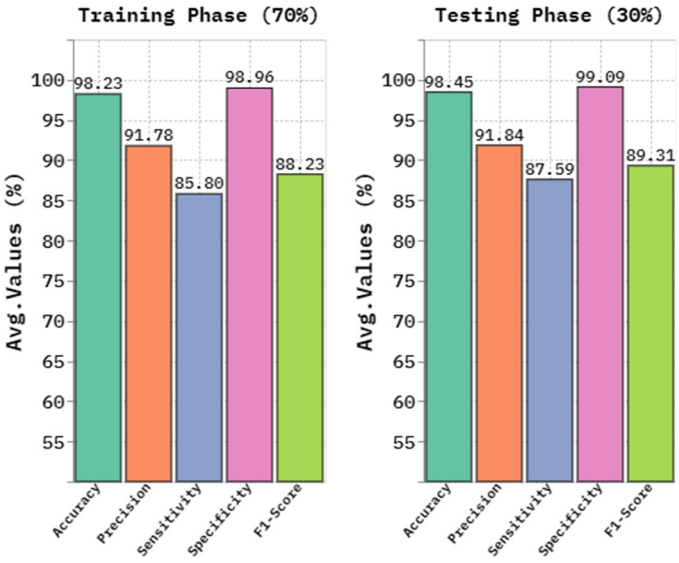


Fig. 3. Average values of the FFATCN-SCD model on 70% training and 30% testing splits.

Table I presents the k-fold Cross Validation (CV) analysis of the FFATCN-SCD approach using different measures. With CV-1, the FFATCN-SCD approach attains $accur_y$ of 97.56%, $preci_n$ of 73.27%, $sensi_y$ of 81.61%, and $speci_y$ of 89.92%. Also, for CV-5, the FFATCN-SCD approach attains $accur_y$ of 97.24%, $preci_n$ of 86.59%, $sensi_y$ of 81.58%, and $speci_y$ of 94.48%. For CV-8, the FFATCN-SCD approach attains $accur_y$ of 87.35%, $preci_n$ of 65.28%, $sensi_y$ of 76.82%, and $speci_y$ of 85.17%. Finally, with CV-10, the FFATCN-SCD approach attains $accur_y$ of 94.56%, $preci_n$ of 68.89%, $sensi_y$ of 73.50%, and $speci_y$ of 92.91%.

TABLE I. K-FOLD CV ANALYSIS OF THE FFATCN-SCD APPROACH

| K-fold CV | $Accur_y$ (%) | $Preci_n$ (%) | $Sensi_y$ (%) | $Speci_y$ (%) |
|---|---|---|---|---|
| CV-1 | 97.56 | 73.27 | 81.61 | 89.92 |
| CV-2 | 87.12 | 68.04 | 73.18 | 88.63 |
| CV-3 | 91.35 | 89.16 | 74.84 | 91.35 |
| CV-4 | 87.57 | 69.50 | 75.92 | 88.51 |
| CV-5 | 97.24 | 86.59 | 81.58 | 94.48 |
| CV-6 | 96.79 | 67.06 | 84.08 | 86.96 |
| CV-7 | 97.73 | 81.80 | 86.78 | 90.65 |
| CV-8 | 87.35 | 65.28 | 76.82 | 85.17 |
| CV-9 | 93.95 | 80.66 | 84.43 | 85.28 |
| CV-10 | 94.56 | 68.89 | 73.50 | 92.91 |

Table II presents the comparative analysis of the FFATCN-SCD model with existing approaches [22-24]. The values indicate that the presented model has achieved superior performance. The existing approaches such as ResNetv250, DenseNet121, Swinv-base, MaxViT-base, Efficient-B4, MobileNetv3-large-075, ShuffleNet, 3D-CNN, SVM-CNN, IoT-DL, and ConvNeXtV2 have achieved lower performance,

with $accur_y$ of 84.93%, 86.35%, 91.79%, 90.84%, 88.27%, 88.77%, 89.31%, 96.56%, 97.11%, 90.67%, and 96.34%, respectively. In contrast, the proposed FFATCN-SCD methodology achieves the highest performance, with $accur_y$ of 98.45%, $preci_n$ of 91.84%, $sensi_y$ of 87.59%, and $speci_y$ of 99.09%.

TABLE II. COMPARATIVE PERFORMANCE ANALYSIS OF THE FFATCN-SCD MODEL WITH EXISTING APPROACHES

| Approach | $Accur_y$ (%) | $Preci_n$ (%) | $Sensi_y$ (%) | $Speci_y$ (%) |
|---|---|---|---|---|
| ResNetv250 [24] | 84.93 | 78.09 | 74.06 | 75.71 |
| DenseNet121 [24] | 86.35 | 81.26 | 77.98 | 79.32 |
| Swinv-base [24] | 91.79 | 90.09 | 85.57 | 88.93 |
| MaxViT-base [24] | 90.84 | 87.83 | 86.23 | 87.37 |
| Efficient-B4 [24] | 88.27 | 82.10 | 80.35 | 81.10 |
| MobileNetv3-large-075 [24] | 88.77 | 84.42 | 80.77 | 82.49 |
| ShuffleNet [22] | 89.31 | 85.03 | 80.57 | 91.60 |
| 3D-CNN [22] | 96.56 | 91.36 | 81.86 | 96.77 |
| SVM-CNN [23] | 97.11 | 91.61 | 84.74 | 96.97 |
| IoT-DL [23] | 90.67 | 89.34 | 81.30 | 85.64 |
| ConvNeXtV2 [23] | 96.34 | 89.64 | 84.58 | 95.48 |
| FFATCN-SCD (proposed) | 98.45 | 91.84 | 87.59 | 99.09 |

## V. CONCLUSION

In this manuscript, a Feature Fusion and Attention-Guided Temporal Convolutional Neural Network for Skin Cancer Detection (FFATCN-SCD) framework is presented for accurate skin cancer detection and classification. The FFATCN-SCD method primarily performs image preprocessing utilizing a Bilateral Filter (BF) to enhance image quality while reducing noise. Following that, the feature representation is carried out through the fusion of EfficientNetV2 and InceptionV3 to extract detailed features for effective identification. Then, a Temporal Convolutional Network with Attention Mechanism (TCN-AM) is leveraged for effective classification of skin cancer, and Antlion Optimization (ALO) is utilized for optimal parameter tuning of the classification model, ensuring enhanced model performance.

To validate the superior performance of the FFATCN-SCD model, a comprehensive simulation was conducted, and the results were evaluated using several performance metrics. The comparative analysis indicates that the FFATCN-SCD technique consistently outperforms existing methods across all evaluation metrics.

## DECLARATION OF COMPETING INTERESTS

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this manuscript.

## ACKNOWLEDGMENT

The authors extend their appreciation to Prince Sattam bin Abdulaziz University for funding this research work through the project number (PSAU/2025/01/37859).

## DATA AVAILABILITY

The data that support the findings of this study are openly available [21].

## REFERENCES

[1] M. Kumar, M. Alshehri, R. AlGhamdi, P. Sharma, and V. Deep, "A DE-ANN Inspired Skin Cancer Detection Approach Using Fuzzy C-Means Clustering," *Mobile Networks and Applications*, vol. 25, no. 4, pp. 1319–1329, Aug. 2020, https://doi.org/10.1007/s11036-020-01550-2.

[2] M. Wang, X. Gao, and L. Zhang, "Recent global patterns in skin cancer incidence, mortality, and prevalence," *Chinese Medical Journal*, vol. 138, no. 2, pp. 185–192, Jan. 2025, https://doi.org/10.1097/CM9.0000000000003416.

[3] O. T. Jones, C. K. I. Ranmuthu, P. N. Hall, G. Funston, and F. M. Walter, "Recognising Skin Cancer in Primary Care," *Advances in Therapy*, vol. 37, no. 1, pp. 603–616, Jan. 2020, https://doi.org/10.1007/s12325-019-01130-1.

[4] V. A. Rajendran and S. Shanmugam, "Automated Skin Cancer Detection and Classification using Cat Swarm Optimization with a Deep Learning Model," *Engineering, Technology & Applied Science Research*, vol. 14, no. 1, pp. 12734–12739, Feb. 2024, https://doi.org/10.48084/etasr.6681.

[5] M. S. A. Huda, T. E. Shrestha, A. Hossain, N. B. Sharif, M. A. Ali, and T. I. Erdei, "DeepMelaNet: Advancing Melanoma Stage Classification in Skin Cancer Diagnosis," *Engineering, Technology & Applied Science Research*, vol. 15, no. 1, pp. 19627–19635, Feb. 2025, https://doi.org/10.48084/etasr.8336.

[6] P. N. Srinivasu, J. G. SivaSai, M. F. Ijaz, A. K. Bhoi, W. Kim, and J. J. Kang, "Classification of Skin Disease Using Deep Learning Neural Networks with MobileNet V2 and LSTM," *Sensors*, vol. 21, no. 8, Apr. 2021, Art. no. 2852, https://doi.org/10.3390/s21082852.

[7] A. Mahbod, G. Schaefer, C. Wang, G. Dorffner, R. Ecker, and I. Ellinger, "Transfer learning using a multi-scale and multi-network ensemble for skin lesion classification," *Computer Methods and Programs in Biomedicine*, vol. 193, Sept. 2020, Art. no. 105475, https://doi.org/10.1016/j.cmpb.2020.105475.

[8] N. Razmjooy *et al.*, "Computer-aided diagnosis of skin cancer: A review," *Current Medical Imaging*, vol. 16, no. 7, pp. 781–793, 2020, https://doi.org/10.2174/1573405616666200129095242.

[9] S. K. Bandyopadhyay, P. Bose, A. Bhaumik, and S. Poddar, "Machine Learning and Deep Learning Integration for Skin Diseases Prediction," *International Journal of Engineering Trends and Technology - IJETT*, vol. 70, pp. 11–18, 2022, https://doi.org/10.14445/22315381/IJETT-V70I2P202.

[10] O. Akinrinade and C. Du, "Skin cancer detection using deep machine learning techniques," *Intelligence-Based Medicine*, vol. 11, Jan. 2025, Art. no. 100191, https://doi.org/10.1016/j.ibmed.2024.100191.

[11] I. Pacal, B. Ozdemir, J. Zeynalov, H. Gasimov, and N. Pacal, "A novel CNN-ViT-based deep learning model for early skin cancer diagnosis," *Biomedical Signal Processing and Control*, vol. 104, June 2025, Art. no. 107627, https://doi.org/10.1016/j.bspc.2025.107627.

[12] L. Zhang, J. Zhang, W. Gao, F. Bai, N. Li, and N. Ghadimi, "A deep learning outline aimed at prompt skin cancer detection utilizing gated recurrent unit networks and improved orca predation algorithm," *Biomedical Signal Processing and Control*, vol. 90, Apr. 2024, Art. no. 105858, https://doi.org/10.1016/j.bspc.2023.105858.

[13] A. S. Al-Waisy, S. Al-Fahdawi, M. I. Khalaf, M. A. Mohammed, B. Al-Attar, and M. N. Al-Andoli, "A deep learning framework for automated early diagnosis and classification of skin cancer lesions in dermoscopy images," *Scientific Reports*, vol. 15, no. 1, Aug. 2025, Art. no. 31234, https://doi.org/10.1038/s41598-025-15655-9.

[14] S. K. Singh, V. Abolghasemi, and M. H. Anisi, "Fuzzy Logic with Deep Learning for Detection of Skin Cancer," *Applied Sciences*, vol. 13, no. 15, Aug. 2023, Art. no. 8927, https://doi.org/10.3390/app13158927.

[15] R. Mittal, V. Malik, J. Singh, S. Gupta, A. P. Srivastava, and A. Sankhyan, "Skin Cancer Detection Using Deep Block Convolutional Neural Networks," in *2023 10th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer*

*Engineering (UPCON)*, Gautam Buddha Nagar, India, 2023, pp. 1187–1191, https://doi.org/10.1109/UPCON59197.2023.10434682.

[16] A. A. Mohamed, S. Alasmari, S. K. Sharma, G. G. Tejani, and S. J. Mousavirad, "Bionic hand movement recognition using BioMotionNet with U3AR-Net segmentation and hybrid optimization," *Systems and Soft Computing*, vol. 7, Dec. 2025, Art. no. 200383, https://doi.org/10.1016/j.sasc.2025.200383.

[17] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 2818–2826, https://doi.org/10.1109/CVPR.2016.308.

[18] M. Tan and Q. Le, "EfficientNetV2: Smaller Models and Faster Training," in *Proceedings of the 38th International Conference on Machine Learning*, Virtual, 2021, pp. 10096–10106.

[19] S. Bai, J. Z. Kolter, and V. Koltun, "An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling." arXiv, Apr. 19, 2018, https://doi.org/10.48550/arXiv.1803.01271.

[20] S. Mirjalili, "The Ant Lion Optimizer," *Advances in Engineering Software*, vol. 83, pp. 80–98, May 2015, https://doi.org/10.1016/j.advengsoft.2015.01.010.

[21] "Skin Cancer ISIC." Kaggle. [Online]. Available: https://www.kaggle.com/datasets/nodoubttome/skin-cancer9-classesisic.

[22] A. D. Khalaf, H. Hamdan, A. Abdul Halin, and N. Manshor, "Segmentation and Classification of Skin Cancer Diseases Based on Deep Learning: Challenges and Future Directions," *IEEE Access*, vol. 13, pp. 90163–90184, 2025, https://doi.org/10.1109/ACCESS.2025.3569170.

[23] J. Yang *et al.*, "IoT-Driven Skin Cancer Detection: Active Learning and Hyperparameter Optimization for Enhanced Accuracy," *IEEE Journal of Biomedical and Health Informatics*, 2025, https://doi.org/10.1109/JBHI.2025.3578419.

[24] B. Ozdemir and I. Pacal, "A robust deep learning framework for multiclass skin cancer classification," *Scientific Reports*, vol. 15, no. 1, Feb. 2025, Art. no. 4938, https://doi.org/10.1038/s41598-025-89230-7.