



Article

# Enhancing Fingerprint Liveness Detection Accuracy Using Deep Learning: A Comprehensive Study and Novel Approach

Deep Kothadiya <sup>1</sup>, Chintan Bhatt <sup>2,\*</sup>, Dhruvil Soni <sup>2</sup>, Kalpita Gadhe <sup>2</sup>, Samir Patel <sup>2</sup>, Alessandro Bruno <sup>3,\*</sup> and Pier Luigi Mazzeo <sup>4</sup>

<sup>1</sup> U & P U Patel Department of Computer Engineering, CHA-RUSAT Campus, Charotar University of Science and Technology, Petlad 388421, India; deepkothadiya.ce@charusat.ac.in

<sup>2</sup> Department of Computer Science and Engineering, School of Technology, Pandit Deendayal Energy University, Gandhinagar 382007, India

<sup>3</sup> Department of Business, Economics, Law, Consumer Behaviour, IULM University, 20143 Milan, Italy

<sup>4</sup> ISASI Institute of Applied Sciences & Intelligent Systems-CNR, 73100 Lecce, Italy

\* Correspondence: chintan.bhatt@sot.pdpu.ac.in (C.B.); alessandrobruno10@gmail.com (A.B.)

**Abstract:** Liveness detection for fingerprint impressions plays a role in the meaningful prevention of any unauthorized activity or phishing attempt. The accessibility of unique individual identification has increased the popularity of biometrics. Deep learning with computer vision has proven remarkable results in image classification, detection, and many others. The proposed methodology relies on an attention model and ResNet convolutions. Spatial attention (SA) and channel attention (CA) models were used sequentially to enhance feature learning. A three-fold sequential attention model is used along with five convolution learning layers. The method's performances have been tested across different pooling strategies, such as Max, Average, and Stochastic, over the LivDet-2021 dataset. Comparisons against different state-of-the-art variants of Convolutional Neural Networks, such as DenseNet121, VGG19, InceptionV3, and conventional ResNet50, have been carried out. In particular, tests have been aimed at assessing ResNet34 and ResNet50 models on feature extraction by further enhancing the sequential attention model. A Multilayer Perceptron (MLP) classifier used alongside a fully connected layer returns the ultimate prediction of the entire stack. Finally, the proposed method is also evaluated on feature extraction with and without attention models for ResNet and considering different pooling strategies.

**Keywords:** liveness detection; deep learning; computer vision; attention model; ResNet50



**Citation:** Kothadiya, D.; Bhatt, C.; Soni, D.; Gadhe, K.; Patel, S.; Bruno, A.; Mazzeo, P.L. Enhancing Fingerprint Liveness Detection Accuracy Using Deep Learning: A Comprehensive Study and Novel Approach. *J. Imaging* **2023**, *9*, 158. <https://doi.org/10.3390/jimaging9080158>

Academic Editor: Alessandro Piva

Received: 8 June 2023

Revised: 31 July 2023

Accepted: 2 August 2023

Published: 7 August 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Over the last few decades, much scholarly attention has been paid to the weaknesses of biometric systems, which are victims of sensor attacks (conducted using synthetic biometric characteristics such as sticky fingers or high-quality printed photographs of the iris). In the digital age, automatic access to services is becoming increasingly crucial. As a result, the field of technology known as “biometric recognition” has emerged.

In this regard, various standardization initiatives at the international level have been established to address security evaluation in biometric systems, such as the Common Standards through various Supporting Documents or the Biometric Evaluation Methodology.

Subsequently, efforts in analyzing the direct attack weaknesses of automatic recognition systems have improved the security level provided by biometric systems.

Various liveness detection techniques have been presented in the past few years. Anti-spoofing algorithms can differentiate between genuine and fake traits using multiple physiological properties.

Due to their well-known uniqueness and persistence, fingerprints are one of the most significant biometric characteristics. In addition to being employed by forensic and law

enforcement organizations worldwide, fingerprint recognition systems are also used in mobile devices for widespread use.

Most fingerprint capture techniques need the finger to contact the capturing apparatus's surface. Some sensor-based systems suffer from technical issues, such as low contrast from dirt or moisture on the plate of the capturing device or latent fingerprints from past users (ghost fingerprints). Hygiene issues reduce the attractiveness of contact-based fingerprint systems, which restricts their adoption, especially in multiuser applications.

Priesnitz et al. [1] surveyed touchless fingerprint systems and compared them to traditional security measures based on something you know (PIN, password, etc.) or something you have; biometric technology offers several advantages (key, card, etc.) [2]. Conventional authentication methods cannot distinguish between imposters who have unlawfully obtained access rights and legitimate users and thus cannot substantiate false identity claims.

Moreover, biometric systems do not require users to carry keys that may be misplaced or stolen or to memorize complicated PINs that are quickly forgotten. However, biometric facilities also have some drawbacks. Because external attacks on biometric systems could lower their level of security, it is essential to comprehend the risks they face. The vulnerability analysis plays the same importance in identifying potential attack points and suggests additional defences to increase their benefits to authentic users. Particular focus has been placed on direct attacks against the fingerprint recognition mechanism among the examined vulnerabilities. These attack techniques involve supplying a sensor with a fake fingerprint to identify the user as legitimate and give access [3].

The International Fingerprint Liveness Detection Competition (LivDet) is one of many projects highlighting the limitations and prospects of presentation attack detectors (PADs). The tournament developed seven editions between 2009 [4] and 2021 [5], presenting brand-new, complex falsification techniques, scanner types, and spoofs materials.

In this article, various software-based liveness detection systems are compared. To this end, several experiments have been carried out with several methods to estimate the top static and dynamic features for vitality identification and assess their effectiveness on the task. The same finger must be placed on the sensor and pulled off it more than once to obtain two or more static characteristics. Moreover, the dynamic features are taken from numerous image frames (i.e., several photos are captured once the finger is positioned on the sensor for a while) [6].

This work focuses on fingerprint liveness detection by investigating the effectiveness of ResNet50 architecture, attention principles, pooling strategies, and several classifiers on different datasets.

Therefore, the contribution is threefold: (1) a proposed novel attention-based ResNet architecture for fingerprint liveness detection, (2) a thorough testing study to check the effectiveness of different pooling strategies, and (3) a comparative analysis with standard computer vision models and classifiers.

The remainder of the article is organized as follows: Section 2 provides a literature survey and comparative analysis for liveness detection; Section 3 describes the proposed methodology to recognize the realness of finger impression; Section 4 deepens the description of the experiment; and Section 5 draws a line on discussions and conclusions.

### Literature Review

Using quality metrics to detect liveness has a new fingerprint parameterization that authors [7] have suggested. They have built the unique system and tested it on a development set of the LivDET competitions dataset, which comprises 4500 live and unreal images procured from three different kinds of sensors. The suggested method demonstrates robustness and accuracy in 93% of properly classified samples. Frassetto et al. [8] suggest using Local Binary Patterns and CNN with random weights. Both of these techniques are used in combination with Support Vector Machine (SVM). They have conducted several experiments on the dataset of the LivDET competition of the years 2009, 2011, and

2013 which comprised about 50,000 live and forgery fingerprint impressions taken from different sensors. When compared to previously reported findings, they reduced test error by 35%. Agarwal and A. Bansal [9] have proposed the fusion of pores perspiration and texture features in a static software-based approach. They carried out experiments on the LivDET 2013 and LivDET 2015 dataset, and their methods have also shown improvement in comparison to the state-of-the-art methods.

Tan [10] used 58 live, 80 spoof, and 25 cadaver participants for three separate scanners. They achieved an accuracy of around 90.9% to 100% using this particular dataset of spoof and live fingerprints. Sequeira et al. [11] introduced an automation-based image segmentation phase to separate the fingerprint impression from the backdrop. They have compared supervised learning approaches with image feature analysis methods. Dubey et al. [12] proposed to combine low-level gradient features with Speeded-Up Robust Features (SURF). They essentially extracted these features from a single fingerprint image to overcome the issues with dynamic software methodologies. Their results outperform the existing best average ERR by 9.625%.

Koshy et al. [13] designed a robust and multi-scenario dataset-based solution that achieves an accuracy of around 90% on classification. They only require one image from a finger to determine whether the input is live or phony. Ali et al. [14] analyzed SOTA (state-of-the-art) fingerprint recognition systems. They broke down the phases of fingerprint identification step by step and summarized fingerprint databases with relevant properties. The fingerprint liveness detection-based approach, based on DCNN and voting strategy, has been presented by Wang et al. [15], which outperforms handcrafted features methods and simultaneously optimizes feature extraction and classifier training.

Chowdhury et al. [16] worked with deep learning-based methods and models to benchmark them on the task. Eight additional scientific articles were compared, and they conducted their investigation using three distinct types of methodologies.

Ahmad et al. [17] surveyed the scientific literature on the topic by looking at 146 crucial studies. Nur-A-Alam et al. [18] trained a multiclass classifier on specific properties has been proposed by the authors. They have performed various experiments that showed that the proposed methodology showed better performance with about 99.87% accuracy. Comparisons have also been conducted against recent machine learning classification techniques, such as SVM (97.86%) and Random Forest (95.47%).

Win et al. [19] targeted the field of criminal investigation by reviewing recent literature on fingerprint classification techniques and applications. They have also analyzed and compared the different computer vision and deep learning algorithms of finger impression images for classification. Priesnitz et al. [20] introduced a touchless fingerprint recognition-based methodology. They published approaches that range from self-identification finger print recognition to practice. At each stage of the recognition process, they additionally provide an overview of the state-of-the-art in the area of touchless 2D fingerprint recognition. Boero et al. [21] designed an Intrusion detection system (IDS) to analyze and detect security problems. The method focuses on anomaly detection and statistical analysis. They obtained the results which allow the best classifiers and show the performance that exploits the decisions of a bank of classifiers acting in parallel. Table 1 illustrates additional comparisons for reference.

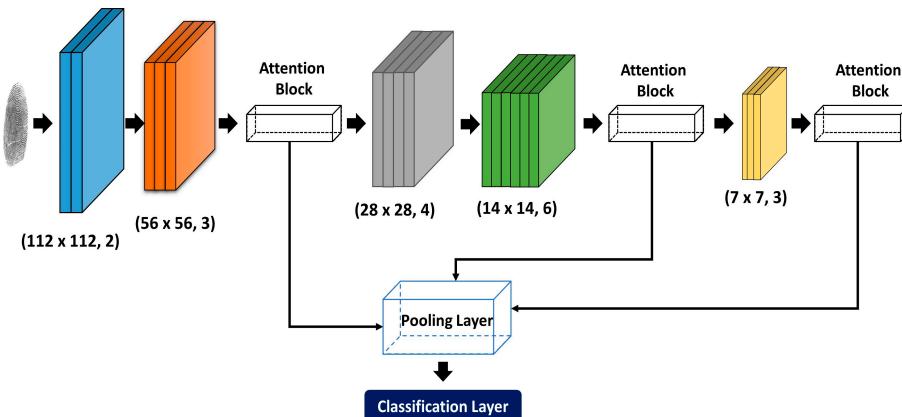
**Table 1.** Comparative analysis of fingerprint presentation attack detection using different deep learning models.

Author	Year	Approach	Result
Y Zhang et al. [22]	2020	FLDNet	91% Acc
Y Zhang et al. [23]	2019	JLW_B	95.25 Acc
C Yuan et al. [24]	2020	DCNN	95.35 Acc
Z Xia et al. [25]	2018	FLD + SVM	5.69 CE

## 2. Materials and Methods

### 2.1. Methodology

The proposed methodology used ResNet50 architecture with an Attention mechanism to extract deep features from fingerprint images. Therefore, the proposed architecture is divided into three main modules (i) ResNet module, (ii) Pooling strategy, and (iii) Visual attention. Figure 1 shows the proposed methodology architecture consisting of five convolution ResNet blocks and three attention blocks featuring different pooling strategies.

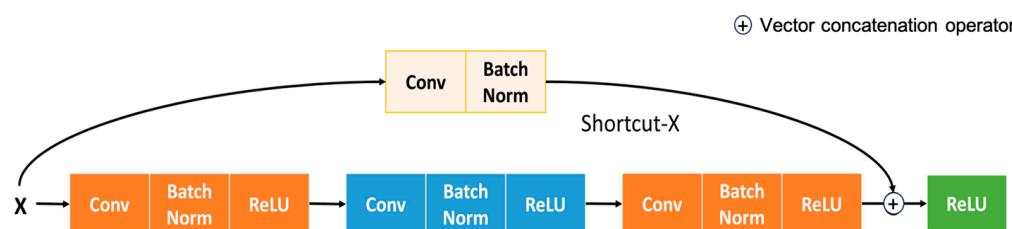


**Figure 1.** The architecture of the proposed attention-based convolution learning methodology.

### 2.2. Convolution Learning with ResNet50

In this subsection, insights into Convolution Learning with ResNet50 are given. Vanilla ResNet architecture has been used for convolution learning, with skip connection to counter overfitting and gradient explosion. The residual module aims to add the features extracted from the neural networks. On top of that, short connections are used to reduce the vanishing gradient problem. Figure 2 represents the internal structure of convolution learning with ResNet as a backbone. The introduction of batch normalization allows for increasing convolution speed and stability. Equation (1) shows the normalization process adopted in ReLU's activation to normalize features.

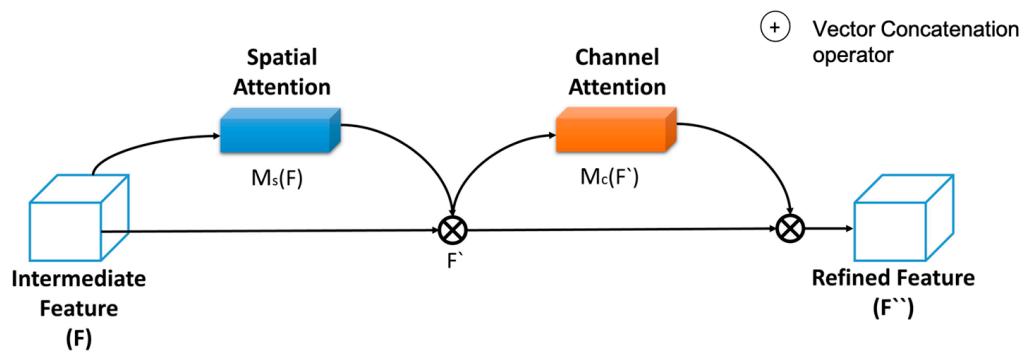
$$f(x) = \begin{cases} x, & \text{if } x > 0 \\ 0, & \text{otherwise} \end{cases} \quad (1)$$



**Figure 2.** Functional architecture residual module with shortcut connection.

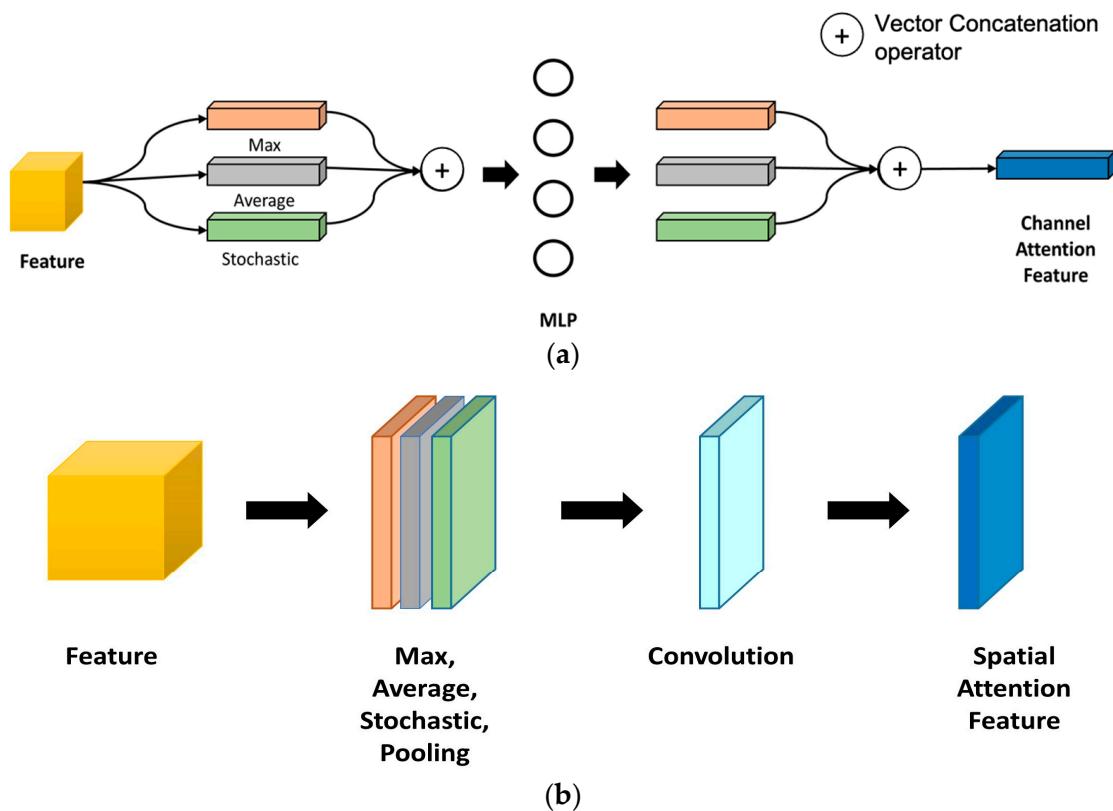
### 2.3. Visual Attention

The proposed methodology relies on attention models to enhance feature learning capabilities and prevent the loss of features caused by single pooling. The proposed method adopts spatial- and channel-based sequential attention modules. In particular, three sequential attention modules represent the backbone network of ResNet. Figure 3 provides a graphical depiction of the functional structure of the spatial-channel attention block.



**Figure 3.** Proposed channel attention and spatial attention sequential model.

The concatenation of basic channel attention modules (CA) and spatial attention modules (SA) relies on mixed attention [26]. The internal structure of the CA and SA blocks are displayed in Figures 4a and 4b, respectively.



**Figure 4.** (a) Functional structure of proposed channel attention block; (b) Functional structure of proposed spatial attention block.

CA module allows to compute the relevant links between distinct channels. In addition, the SA block accounts for executing multiple types of visual attention to obtain information on content, texture, and background.

Feature Map  $F(C \times H \times W)$  is passed through the global average pooling to generate  $F_c(1 \times 1 \times C)$ , for  $C$  channels. Then, it is forwarded to MLP (Multilayer Perceptron) with four (4) hidden layers. The activation of the hidden layer was set to scale the output as  $R^{(1 \times 1 \times \frac{C}{r})}$  to adjust the ResNet map according to the compression rate.

The final map  $M_c(F)$  is computed as in Equation (2), where  $W$  and  $b$  are weights and biases for channel attention  $R^c$ .

$$M_c(F) = \text{Norm}(W_1(W_0 \text{AvgPool}(F_c) + b_0) + b_1) \quad (2)$$

SA generates intermediate spatial features  $F_s(H \times W \times 1)$  to empower spatial locations. Input feature  $F_c$  adopts full  $3 \times 3$  convolution to extract the information from the input matrix. A higher spatial attention dimension adopts  $R^{(H \times W \times 1)}$  through  $3 \times 3$  convolution.

Finally, extracted SA output  $M_s(F)$  is generated with Batch normalization as in Equation (3).

$$M_s(F) = \text{Norm}\left(\text{AvgPool}\left(f_0^{3 \times 3}(F_c)\right)\right) \quad (3)$$

Here, Equation (3) is generalized with global average pooling.

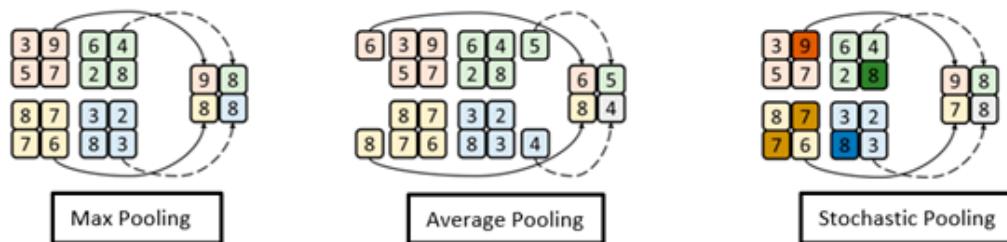
In addition, the simulation of the model also explores different pooling strategies to find the most appropriate based on input image property. Element-wise multiplication is used to produce refined intermediate features from the attention blocks. They are represented as spatial attention features  $F(s)$  and channel attention features  $F_c$  and used in Equations (4) and (5) to calculate the final output of the attention block  $F''$  as an intermediate feature, while  $F$  is considered an input feature for the attention block.

$$F' = M_c(F) \otimes F \quad (4)$$

$$F'' = M_s \otimes F' \quad (5)$$

#### 2.4. Pooling Operation

The proposed methodology also tests three pooling strategies to optimize performance over fingerprint liveness detection. That is motivated by pooling method performances being dependent on the nature of targeted images. Pooling plays a critical role in convolution learning in extracting information from learned features. It also brings in a dimension reduction. CNN-based architecture has common pooling strategy as max pooling, average pooling, and stochastic pooling, as illustrated in Figure 5. The pooling strategy can be selected based on image background, channel type, texture, and many more features. For instance, non-maximum values in the pooling kernel are discarded by max pooling, while the maximal feature values are not retained by average pooling. Conversely, retention of features in a certain direction is not the primary objective of stochastic pooling. Therefore, sticking to one strategy of pooling may lead to limited classification performance.



**Figure 5.** Various pooling strategies employed in the proposed architecture.

#### 2.5. Algorithm

The computational features that have been previously described are processed by a three-layer MLP [27] network to detect the liveness of fingerprints. For the sake of clarity, a description of the training procedure for the Dual Attention model is provided in the form of Algorithm 1 down below. The scripting structure of the proposed dual attention-based methodology demonstrates the major functional component of the architecture.

**Algorithm 1:** Training procedure for Dual Attention model

---

Output: Attention features  $F$

1.  $\text{Pol}(x) \leftarrow P\{\text{Max}, \text{Avg}, \text{Stoch}\}$
2.  $\text{Mc} \leftarrow \text{Channel Attention Model}$
3.  $\text{Ms} \leftarrow \text{Spatial Attention Model}$
4.  $\text{model} \leftarrow \text{ResNet}()$
5.      $\text{model.FcLayer} \leftarrow \text{ReLU}()$
6.      $\text{mode.Attention} \leftarrow f(\text{Mc} \otimes \text{Ms})$
7.      $\text{model.Pooling} \leftarrow f(\text{Pol}_x)$
8.     for epoch =0 to epoch:
9.         for  $(x_{\text{train}}, y_{\text{train}})$ :
10.              $x \leftarrow \text{model.normalization}(x_{\text{train}}, \text{kernel})$
11.              $y \leftarrow \text{model.Conv}(\sum f(\text{Weight}_x \otimes \text{Input}_x))$
12.              $a \leftarrow \text{model.attention}(\text{pol}(y)) =$
13.              $z \leftarrow \text{model.Pooling}(\sum_0^2 f(\text{Pol}_a_x))$
14.              $\alpha. \leftarrow \text{model.FcLayer}(\text{MLP}(z))$
15.              $f(x) \leftarrow \text{optimization.adam}(\alpha)$
16.              $loss \leftarrow \text{loss} \cdot \sum_i^n y_i \log(\alpha_i)$
17.         end for
18.     end for
19.     Return:  $f(x)$ , Feature of training dataset

---

### 3. Results

#### 3.1. Dataset

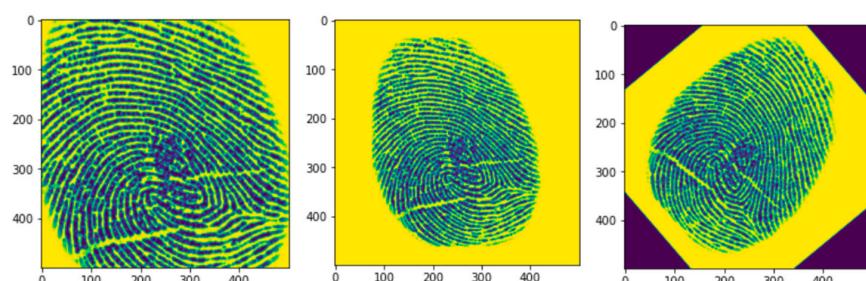
The proposed methodology has been tested over two different datasets: LivDet DB [28] and ATVS DB [29]. The former consists of three different types of images: (i) Biometrika, (ii) CrossMatch, and (iii) Identix. Fack fingerprints are designed with silicone, gelatine, and playdoh. ATVS DB consists of different datasets having three different techniques: (i) flat optical Biometrika, (ii) flat capacitive Precise, and (iii) thermal sweeping Yubee (demonstrated in Table 2).

**Table 2.** Describes the characteristics of the targeted dataset.

Dataset	Class	Real/Live	Fake	Avg. Resolution
LivDet [28]	2	5000	3000	580 dpi
ATVS [29]	2	4800	4000	520 dpi

#### 3.2. Data Augmentation

The proposed methodology relies on data augmentation for targeted datasets to increase the knowledge inference capabilities of the model. Generally, data augmentation methods include rotation, scaling, flip, crop, and many more. Among them, angular rotation ( $\pm 20^\circ$ ) and scaling were used to increase the training dataset. Figure 6 represents some augmented samples from the training dataset.



**Figure 6.** Manual Augmented dataset used in the proposed methodology.

### 3.3. Experiments and Result

This section provides the reader with the experimental settings and results. Several trials have been carried out with ResNet architecture with different depths: ResNet50 and ResNet34. Dual attention blocks were embedded at a different level of convolution. As shown In Table 3, the model stack includes convolution layers, SA, CA, pooling, and activation functions. It features an initial kernel size set to  $7 \times 7$ , and the final fully connected layer is characterized by ReLU activation.

**Table 3.** ResNet34 and ResNet50 network stacks and the corresponding Channel attention (CA) and Spatial attention (SA) modules are described below.

Conv-Layer	Output Size	ResNet34	ResNet50
01	$112 \times 112$	$7 \times 7, 64$	$7 \times 7, 64$
02	$56 \times 56$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
Attention (CA + SA)			
03	$28 \times 28$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$
04	$14 \times 14$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$
Attention (CA + SA)			
05	$7 \times 7$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
Attention (CA + SA)			
Fully Connected	$1 \times 1$	Pooling, Activation	

The training phase has been carried out with the following parameters and settings: 20 epochs; Adam optimizer; learning rate set to 10 to the power (-3); binary cross-entropy as loss function and accuracy metrics to evaluate performances; batch size set to 16; dropout equals to 0.3. The training and testing tasks have been run on an NVIDIA GeForce RTX 3080 (AIDA Lab KSA, Riyadh Saudi Arabia) and are coded with Python 3.10.

The proposed dual attention-based methodology has achieved a 97.78% accuracy rate over the LivDet dataset. The proposed model was evaluated using Sensitivity [30], Precision [31], and F1-score [32], mathematically described in Equations (6)–(8).

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (6)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (7)$$

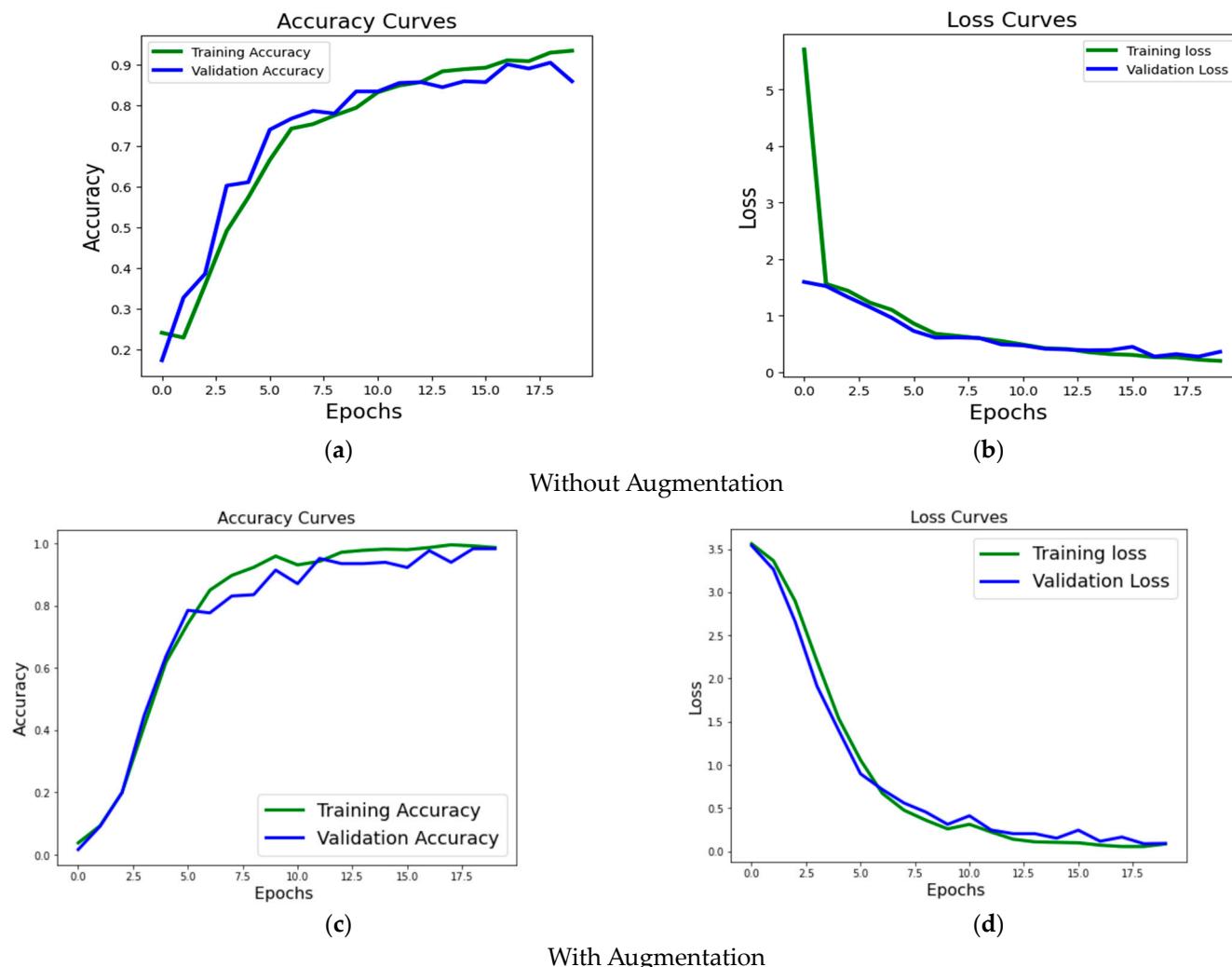
$$\text{F1-score} = \frac{2 * \text{TP}}{2 * \text{TP} + \text{FP} + \text{FN}} \quad (8)$$

Table 4 represents a statistical evaluation of the proposed methodology with ResNet34 and ResNet50 on two different datasets (LivDet and ATVS).

**Table 4.** Comparative analysis of ReNet34 and ResNet50 with attention modules over both datasets are given below.

Model	Dataset	Sensitivity	Precision	F1-Score	Accuracy
ResNet34 + Attention	LivDet	0.94	0.95	0.95	95.81%
	ATVS	0.95	0.95	0.95	95.52%
ResNet50 + Attention	LivDet	0.97	0.97	0.97	97.78%
	ATVS	0.96	0.97	0.96	97.05%

The accuracy metrics and loss function (Figure 7) allow benchmarking training and validation steps for the proposed method. In particular, Figure 7a,b illustrate the results of the proposed methodology for the ResNet50 model without augmentation. At the same time, subplots (c) and (d) draw the corresponding curves for the training with augmented data.



**Figure 7.** Accuracy and loss evaluation of the proposed methodology over the LivDet dataset without augmentation are given in subfigures (a,b). The same metrics are graphically depicted in subfigures (c,d) for the experiments with augmentation.

Table 5 shows ResNet50 and ResNet34 architectures are compared with and without attention modules. Experiments have been run using three different pooling strategies

(Table 6). Stochastic pooling offers remarkable accuracy over average and max pooling for the proposed Res-Net50-based methodology on the LivDet dataset.

**Table 5.** Comparative analyses of ResNet variants with and without attention models for the LivDet dataset are given below.

Model	Attention	Avg. Precision	Avg. Recall	Avg. F1-Score
ResNet34	Yes	0.95	0.96	0.95
	No	0.83	0.84	0.84
ResNet50	Yes	0.97	0.97	0.97
	No	0.87	0.86	0.86

**Table 6.** Comparative analysis is given for different pooling strategies for the Attention-based ResNet50 model on the LivDet dataset.

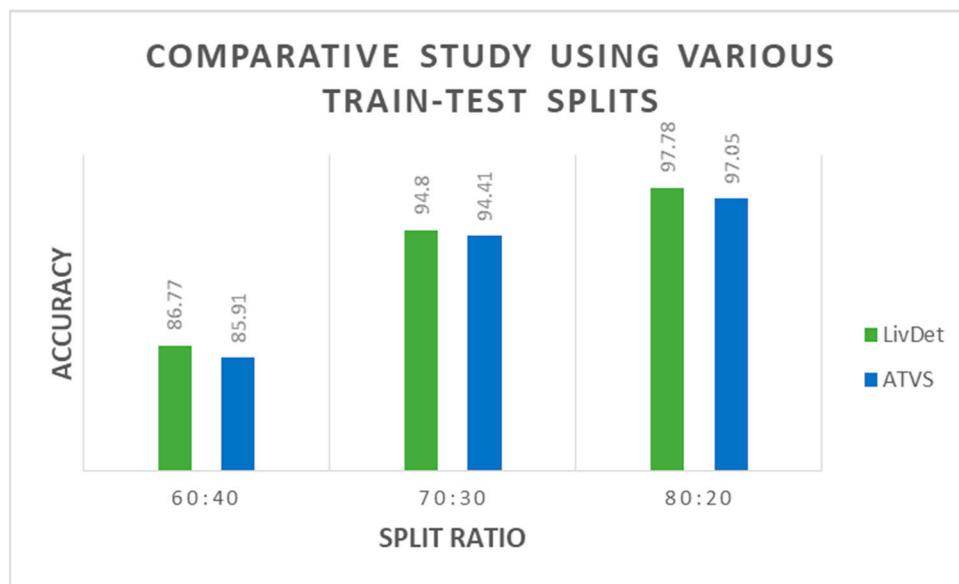
Database	Max Pooling	Average pooling	Stochastic Pooling
LivDet	97.23%	97.14%	97.78%
ATVS	96.00%	96.33	97.05%

However, the selection of the pooling strategy highly depends on the input type.

Table 7 shows a comparative analysis of ResNet50 with different learning rates and Dropout values on both datasets. The experiments reveal a 0.0001 Learning Rate and a dropout of 0.3, being the combination performing higher. Some other tests have been run with various train test split ratios, as shown in Figure 8.

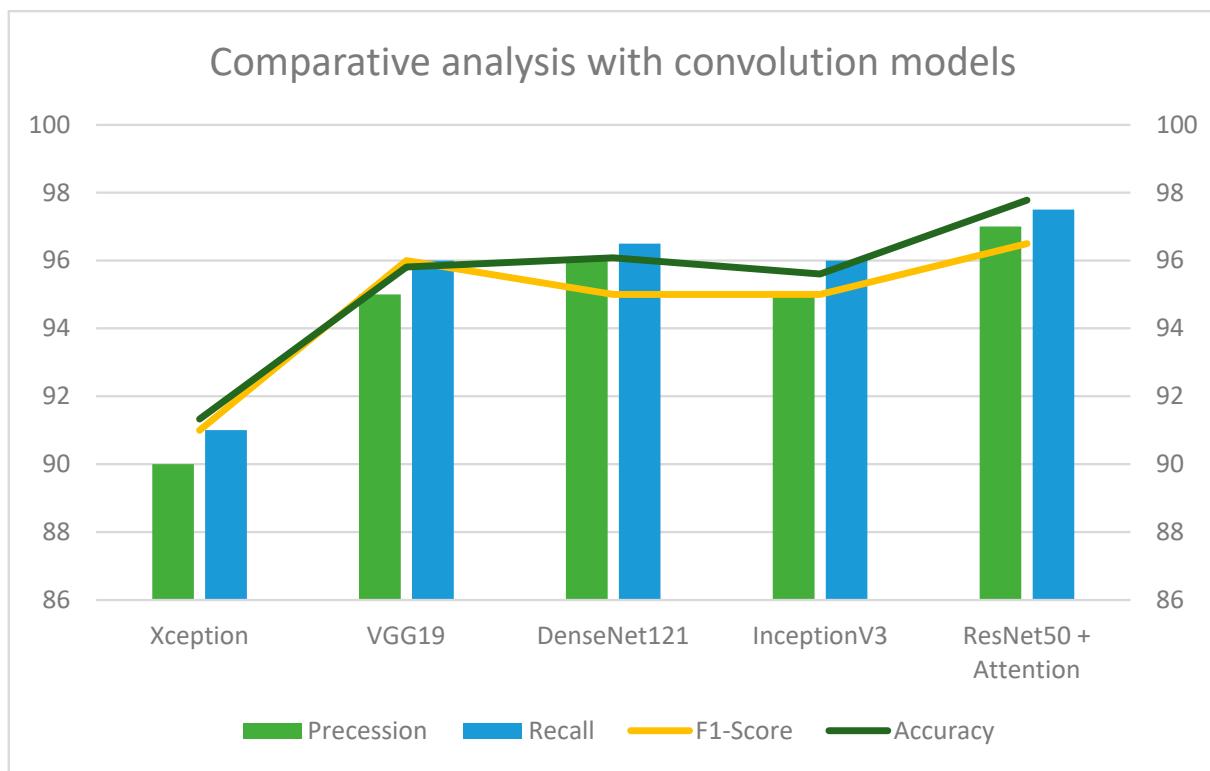
**Table 7.** Accuracy analysis with a combination of different Learning rates and dropout rates for proposed architecture with ResNet50.

Learning Rate	0.01			0.001			0.0001		
	Dropout	0.1	0.2	0.3	0.1	0.2	0.3	0.1	0.2
LivDet	90.03	94.5	95.22	94.23	96.61	97.33	96.86	96.96	97.78
ATVS	88.64	92.74	94.12	93.88	96.35	96.98	96.41	96.52	97.05



**Figure 8.** Comparison study considering various train-test ratios for the proposed architecture.

Figure 9 represents a comparative analysis of the proposed sequential attention model against other state-of-the-art convolution models such as VGG19, Densenet121, and InceptionV3, and different optimizers such as Gradient Descent, Stochastic Gradient Decent (SGD), and Adam [33]. Table 8 shows a comparative analysis with different optimizers and different deep learning models.



**Figure 9.** The proposed methodology has been compared against other deep learning models on the LivDet dataset.

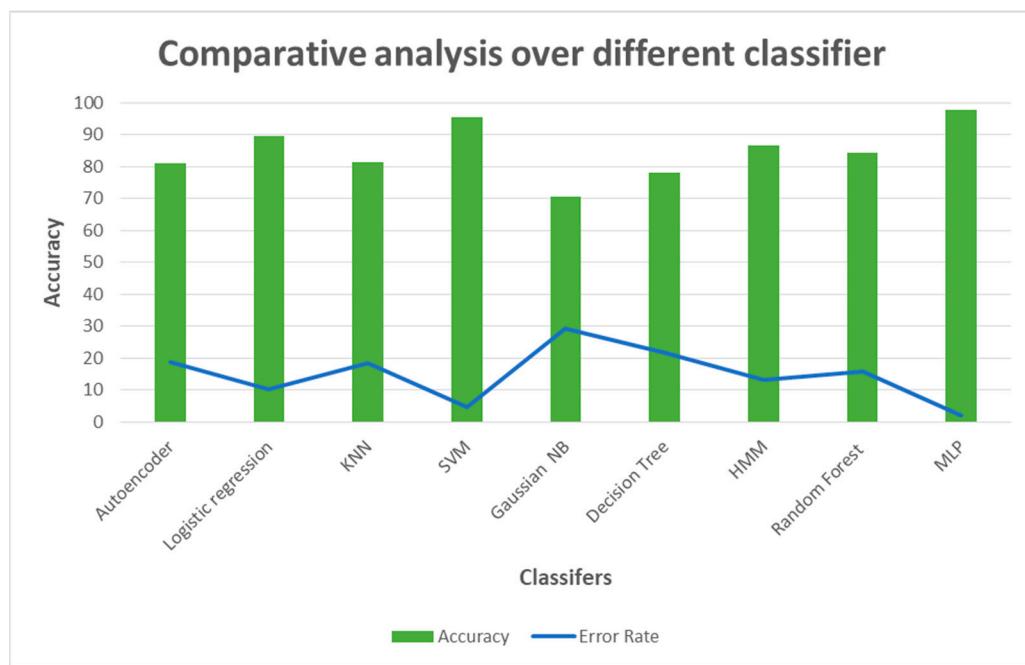
**Table 8.** Comparative analysis with different optimizers and models trained over both the scenarios, with and without association of attention model for ResNet34 and ResNet50.

Model	Optimizer	Accuracy	Precision	Recall
ResNet34	GD	77.46%	0.79	0.75
	SGD	82.60%	0.83	0.83
	Adam	83.56%	0.83	0.84
ResNet34 + attention	GD	85.30%	0.88	0.85
	SGD	94.21%	0.94	0.94
	Adam	95.81%	0.95	0.96
Resnet50	GD	71.05%	0.71	0.72
	SGD	84.71%	0.85	0.85
	Adam	87.21%	0.87	0.86
Resnet50 + Attention	GD	88.49%	0.88	0.87
	SGD	95.90%	0.96	0.97
	Adam	97.78%	0.97	0.97

#### 4. Discussion

The proposed study adopts an attention-based technique to detect liveness from fingerprint images and recognize them as real or fake. The proposed methodology uses a unique approach to see the realness of figure print from pictures as dual attention-based Resnet50 architecture has achieved 97.78% remarkable accuracy on the LivDet fingerprint dataset. Experiments have been carried out on ATVS for a more extensive analysis of the proposed technique.

Furthermore, the proposed methodology evaluates the robustness of different pooling strategies to extend accuracy or avail results. In particular, the experimental campaign in this work reveals the stochastic pooling method outperforming Max and Average pooling on the fingerprint recognition task. The architecture proposed heavily relies on attention blocks: spatial and channel attention modules are employed to enhance feature extraction. The proposed methodology is superior to other deep learning convolution models, such as Xception, InceptionV3, VGG19, DenseNet121, and InceptionV3, with around 0.7% accuracy. Further investigations have been conducted concerning the employment of classifiers (Figure 10), such as Random Forest (RF) [34], Linear Regression (LR) [35], KNN [36], SVM [37], Gaussian NB [37], Decision Tree [38], HMM [39], Autoencoder [40] and Support Vector Machine (SVM) [41]. Convolution layer five output from ResNet50 is converted into a 2048-dimensional array, which feeds all comparison classifiers. In that case, the ResNet50 stack is used until convolutional layer five only for feature extraction, represented by the 2048-dimensional array.

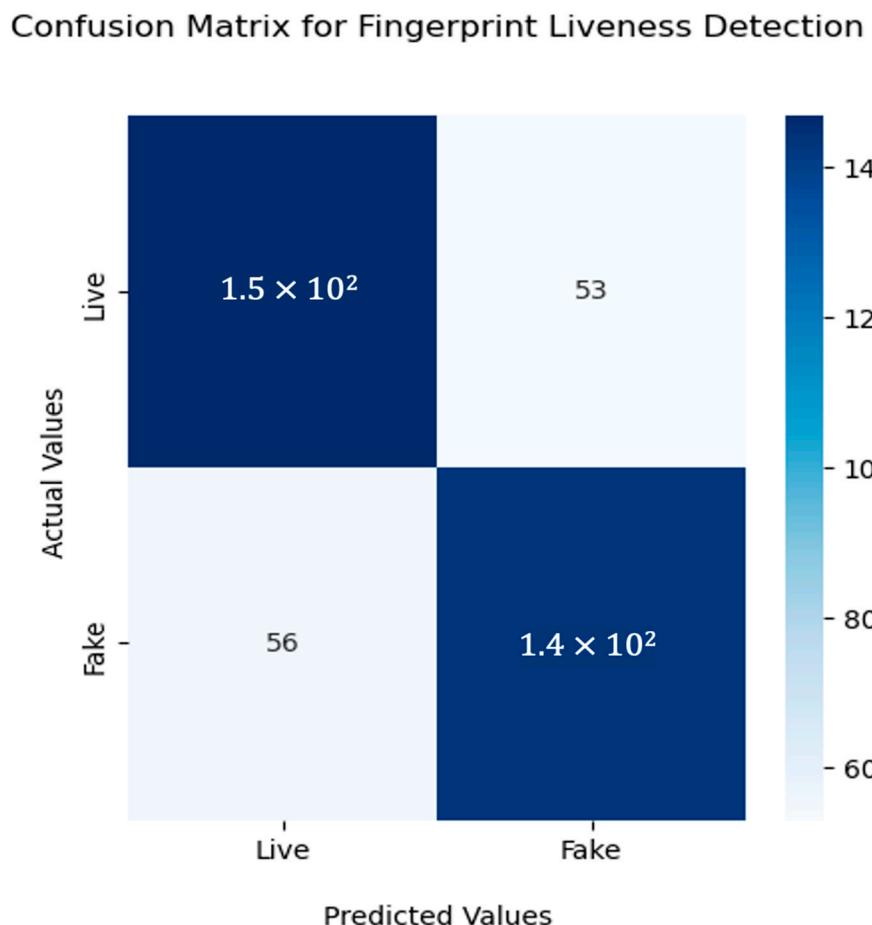


**Figure 10.** Comparative analysis of the proposed dual attention-based model with different classifiers.

As noticed in Figure 10, Multilayer Perceptron (MLP) has proven higher accuracy than other machine learning classifiers.

Figure 11 illustrates the confusion matrix of the proposed methodology on the LiveDet dataset, with 53 images misclassified as fake out of 1500 images.

The proposed dual attention-based ResNet50 accounts for 50 layers in the neural network, bringing in not negligible computational costs.



**Figure 11.** Confusion matrix for fingerprint liveness detection using Resnet50 with sequential attention model.

## 5. Conclusions

The digital fingerprint is a powerful applicable tool for a plethora of scenarios, such as person authentication in commercial business, civil and forensic usage, etc. Most fingerprint-based scanners proved robust and accurate in fingerprint recognition. They stand almost 100% accuracy during authentication.

In this study, a vision-based methodology has been introduced to focus on fingerprint liveness detection. As previously mentioned, the contribution of this work is threefold: (1) the introduction of a novel attention-based ResNet architecture for fingerprint liveness detection, (2) a thorough testing study to check the effectiveness of different pooling strategies, (3) a comparative analysis with standard computer vision models and classifiers.

An attention-based learning approach to recognize the liveness of fingerprint images is proposed to tackle fingerprint liveness detection. To this end, the methodology relies on ResNet as the architecture backbone for convolution learning. In particular, two attention module channels followed by spatial attention are featured in the architecture. The architecture has been tested on three different pooling strategies with, interestingly, some positive outcomes achieved with Stochastic pooling.

Furthermore, the extensive experimental campaign conducted over two different datasets showed the positive impact of dual attention-based learning on the correctness of the results. Comparisons to other machine learning and deep learning models, such as Random Forest and CNN variants, such as VGG19 and DenseNet121, also confirm that.

Feature work for this study to enhance the model toward more accuracy and robustness for real-time scenarios. The proposed methodology is meant to be extended to detect presentation attacks using other biometrics, such as retina and face. The proposed study is

also planned to be improved by integrating Explainable AI to interpret the prediction of the proposed deep learning model.

**Author Contributions:** D.K.: Data Curation, Investigation, Writing—original draft, C.B.: Methodology, Project administration, Writing—review and editing, S.P.: Data curation, Investigation, Software, D.S.: Data curation, Investigation, Software, K.G.: Methodology, Writing—review and editing, A.B.: Methodology, P.L.M.: Writing—review and editing. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data are available from the LivDet 2021 challenge upon reasonable request.

**Conflicts of Interest:** The authors declare that there are no conflict of interest regarding the publication of this article.

## References

- Priesnitz, J.; Huesmann, R.; Rathgeb, C.; Buchmann, N.; Busch, C. Mobile Contactless Fingerprint Recognition: Implementation, Performance and Usability Aspects. *Sensors* **2022**, *22*, 792. [[CrossRef](#)]
- Adler, A. Vulnerabilities in biometric encryption systems. In Proceedings of the IAPR Audio- and Video-Based Biometric Person Authentication (AVBPA), Hilton Rye Town, NY, USA, 20–22 July 2005; Springer LNCS-3546; pp. 1100–1109.
- Galbally, J.; Alonso-Fernandez, F.; Fierrez, J.; Ortega-Garcia, J. A high performance fingerprint liveness detection method based on quality related features. *Futur. Gener. Comput. Syst.* **2012**, *28*, 311–321. [[CrossRef](#)]
- Marcialis, G.M.; Lewicke, A.; Tan, B.; Coli, P.; Grimberg, D.; Congiu, A.; Tidu, A.; Roli, F.; Schuckers, S. First International Fingerprint Liveness Detection Competition—LivDet 2009. In Proceedings of the International Conference on Image Analysis and Processing, Vietri sul Mare, Italy, 8–11 September 2009; Springer: Berlin/Heidelberg, Germany, 2009; pp. 12–23.
- Lee, Y.K.; Jeong, J.; Kang, D. An Effective Orchestration for Fingerprint Presentation Attack Detection. *Electronics* **2022**, *11*, 2515. [[CrossRef](#)]
- Micheletto, M.; Orrù, G.; Casula, R.; Yambay, D.; Marcialis, G.L.; Schuckers, S. Review of the Fingerprint Liveness Detection (LivDet) competition series: From 2009 to 2021. In *Handbook of Biometric Anti-Spoofing: Presentation Attack Detection and Vulnerability Assessment*; Springer: Berlin/Heidelberg, Germany, 2023; pp. 57–76.
- Galbally, J.; Alonso-Fernandez, F.; Fierrez, J.; Ortega-Garcia, J. Fingerprint liveness detection based on quality measures. In Proceedings of the 2009 First IEEE International Conference on Biometrics, Identity and Security (BIdS), Tampa, FL, USA, 22–23 September 2009. [[CrossRef](#)]
- Nogueira, R.F.; Lotufo, R.d.A.; Machado, R.C. Evaluating software-based fingerprint liveness detection using Convolutional Networks and Local Binary Patterns. In Proceedings of the 2014 IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications (BIOMS) Proceedings, Rome, Italy, 17 October 2014. [[CrossRef](#)]
- Agarwal, D.; Bansal, A. Fingerprint liveness detection through fusion of pores perspiration and texture features. *J. King Saud Univ. Comput. Inf. Sci.* **2022**, *34*, 4089–4098. [[CrossRef](#)]
- Tan, B. New approach for liveness detection in fingerprint scanners based on valley noise analysis. *J. Electron. Imaging* **2008**, *17*, 011009. [[CrossRef](#)]
- Sequeira, A.; Cardoso, J. Fingerprint Liveness Detection in the Presence of Capable Intruders. *Sensors* **2015**, *15*, 14615–14638. [[CrossRef](#)]
- Dubey, R.K.; Goh, J.; Thing, V.L.L. Fingerprint Liveness Detection From Single Image Using Low-Level Features and Shape Analysis. *IEEE Trans. Inf. Forensics Secur.* **2016**, *11*, 1461–1475. [[CrossRef](#)]
- Koshy, R.; Mahmood, A. Enhanced Deep Learning Architectures for Face Liveness Detection for Static and Video Sequences. *Entropy* **2020**, *22*, 1186. [[CrossRef](#)] [[PubMed](#)]
- Ali, M.M.; Mahale, V.H.; Yannawar, P.; Gaikwad, A.T. Overview of fingerprint recognition system. In Proceedings of the 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), Chennai, India, 3–5 March 2016. [[CrossRef](#)]
- Wang, C.; Li, K.; Wu, Z.; Zhao, Q. A DCNN Based Fingerprint Liveness Detection Algorithm with Voting Strategy. In *Biometric Recognition*; Springer International Publishing: Cham, Switzerland, 2015; pp. 241–249. [[CrossRef](#)]
- Chowdhury, A.M.M.; Imtiaz, M.H. Contactless Fingerprint Recognition Using Deep Learning—A Systematic Review. *J. Cybersecur. Priv.* **2022**, *2*, 714–730. [[CrossRef](#)]
- Ahmad, A.S.; Hassan, R.; Ahmad, M.N. Fake Fingerprint Detection Approaches: Systematic Literature Review. *Int. J. Innov. Technol. Explor. Eng.* **2019**, *8*, 1–8.

18. Alam, N.A.; Ahsan, M.; Based, M.; Haider, J.; Kowalski, M. An intelligent system for automatic fingerprint identification using feature fusion by Gabor filter and deep learning. *Comput. Electr. Eng.* **2021**, *95*, 107387. [[CrossRef](#)]
19. Win, K.N.; Li, K.; Chen, J.; Viger, P.F.; Li, K. Fingerprint classification and identification algorithms for criminal investigation: A survey. *Futur. Gener. Comput. Syst.* **2020**, *110*, 758–771. [[CrossRef](#)]
20. Priesnitz, J.; Rathgeb, C.; Buchmann, N.; Busch, C.; Margraf, M. An overview of touchless 2D fingerprint recognition. *EURASIP J. Image Video Process.* **2021**, *2021*, 8. [[CrossRef](#)]
21. Boero, L.; Cello, M.; Marchese, M.; Mariconti, E.; Naqash, T.; Zappatore, S. Statistical fingerprint-based intrusion detection system (SF-IDS). *Int. J. Commun. Syst.* **2016**, *30*, e3225. [[CrossRef](#)]
22. Zhang, Y.; Pan, S.; Zhan, X.; Li, Z.; Gao, M.; Gao, C. FLDNet: Light Dense CNN for Fingerprint Liveness Detection. *IEEE Access* **2020**, *8*, 84141–84152. [[CrossRef](#)]
23. Zhang, Y.; Shi, D.; Zhan, X.; Cao, D.; Zhu, K.; Li, Z. Slim-ResCNN: A deep residual convolutional neural network for fingerprint liveness detection. *IEEE Access* **2019**, *7*, 91476–91487. [[CrossRef](#)]
24. Yuan, C.; Xia, Z.; Sun, X.; Wu, Q.M.J. Deep Residual Network With Adaptive Learning Framework for Fingerprint Liveness Detection. *IEEE Trans. Cogn. Dev. Syst.* **2020**, *12*, 461–473. [[CrossRef](#)]
25. Xia, Z.; Yuan, C.; Lv, R.; Sun, X.; Xiong, N.N.; Shi, Y.-Q. A Novel Weber Local Binary Descriptor for Fingerprint Liveness Detection. *IEEE Trans. Syst. Man Cybern. Syst.* **2020**, *50*, 1526–1536. [[CrossRef](#)]
26. Woo, S.; Park, J.; Lee, J.Y.; Kweon, I.S. CBAM: Convolutional block attention module. In Proceedings of the European Conference on Computer Vision, Munich, Germany, 8–14 September 2018; Springer International Publishing: Cham, Switzerland, 2018; pp. 3–19.
27. Kothadiya, D.R.; Bhatt, C.M.; Saba, T.; Rehman, A.; Bahaj, S.A. Signformer: DeepVision Transformer for Sign Language Recognition. *IEEE Access* **2023**, *11*, 4730–4739. [[CrossRef](#)]
28. Casula, R.; Micheletto, M.; Orrù, G.; Delussu, R.; Concas, S.; Panzino, A.; Marcialis, G.L. LivDet 2021 Fingerprint Liveness Detection Competition—Into the unknown. In Proceedings of the International IEEE Joint Conference on Biometrics (IJCB) 2021, Shenzhen, China, 4–7 August 2021; pp. 1–6.
29. Galbally, J.; Fierrez, J.; Alonso-Fernandez, F.; Martinez-Diaz, M. Evaluation of direct attacks to fingerprint verification systems. *Telecommun. Syst.* **2010**, *47*, 243–254. [[CrossRef](#)]
30. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.C. MobileNetV2: Inverted Residuals and Linear Bottlenecks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018.
31. Chollet, F. Xception: Deep Learning with Depthwise Separable Convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017.
32. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015.
33. Kothadiya, D.R.; Bhatt, C.M.; Rehman, A.; Alamri, F.S.; Saba, T. SignExplainer: An Explainable AI-Enabled Framework for Sign Language Recognition with Ensemble Learning. *IEEE Access* **2023**, *11*, 47410–47419. [[CrossRef](#)]
34. Song, Q.; Liu, X.; Yang, L. The random forest classifier applied in droplet fingerprint recognition. In Proceedings of the 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), Zhangjiajie, China, 15–17 August 2015; pp. 722–726.
35. Gowthami, A.; Mamatha, H. Fingerprint Recognition Using Zone Based Linear Binary Patterns. *Procedia Comput. Sci.* **2015**, *58*, 552–557. [[CrossRef](#)]
36. Gnanasivam, P.; Muttan, D.S. Estimation of age through fingerprints using wavelet transform and singular value decomposition. *Int. J. Biom. Bioinform.* **2012**, *6*, 58–67.
37. Hong, J.-H.; Min, J.-K.; Cho, U.-K.; Cho, S.-B. Fingerprint classification using one-vs-all support vector machines dynamically ordered with naive Bayes classifiers. *Pattern Recognit.* **2008**, *41*, 662–671. [[CrossRef](#)]
38. Zhao, L.; Lee, S.; Jeong, S.-P. Decision Tree Application to Classification Problems with Boosting Algorithm. *Electronics* **2021**, *10*, 1903. [[CrossRef](#)]
39. Guo, J.M.; Liu, Y.F.; Chang, J.Y.; Lee, J.D. Fingerprint classification based on decision tree from singular points and orientation field. *Expert Syst. Appl.* **2014**, *41*, 752–764. [[CrossRef](#)]
40. Razzak, M.I.; Hussain, S.A.; Belaïd, A.; Sher, M. Multi-font numerals recognition for Urdu script based languages. *Int. J. Recent Trends Eng.* **2009**, *2*, 70.
41. Qi, Y.; Qiu, M.; Jiang, H.; Wang, F. Extracting Fingerprint Features Using Autoencoder Networks for Gender Classification. *Appl. Sci.* **2022**, *12*, 10152. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.