ISSN: 2424-8460(O)
2251-2608(P)

Original Research Article

Optimizing different palmprint recognition models trained via transfer learning using attention mechanisms

Hao Jiang [1], Nghia Thi Mai [2], Md Abdus Samad Kamal [1], Iwanori Murakami [1], Kou Yamada [1]

[1] *Gunma University, 1-5-1 Tenjincho, Kiryu, Gunma 376-8515, Japan*

[2] *Posts and Telecommunications Institute of Technology, Km10, Nguyen Trai, Ha Dong District, Hanoi*

Abstract: Palmprint recognition is an authentication technology based on human biometrics, which has gained wide attention in the biometrics field due to its uniqueness, stability and non-invasiveness. And now it is widely used in the fields of financial payment, security monitoring, and intelligent medical treatment. Furthermore, this technology has already been widely adopted in China's daily life, particularly in mobile payment systems. This study proposes an innovative palmprint recognition method that combines data preprocessing, attention mechanism, and migration learning to optimize feature extraction and classification performance. The results show that the method in this paper outperforms traditional methods on multiple datasets, exhibiting high robustness and adaptability. It provides some ideas for future palmprint research.

Keywords: palmprint; recognition; attention; mechanism; transfer; learning; image processing

1. Introduction

In recent years, biometrics have received widespread at-tention for their unique ability to provide secure and reliable authentication. Biometric identification has become deeply embedded in various facets of modern life, including health-care and financial security. According to market statistics from GrandView Research (2023), the global biometric pay-ment market is projected to expand from USD 42.9 billion in 2023 to USD 127 billion by 2030, achieving a compound an-nual growth rate (CAGR) of 16.7%.[1]Thus, biometric iden-tification holds promising prospects for future development.

Among conventional biometric technologies—Including fingerprint, iris, and facial recognition—Fingerprint recog-nition stands out as one of the most mature. Its core ad-vantages lie in exceptional uniqueness and operational prac-ticality. Research indicates that fingerprint ridge patterns remain invariant throughout an individual's lifetime, with a false acceptance rate (FAR) as low as 0.001%. Match-ing typically completes in under 1 second, making it ideal for high-throughput applications. However, its small capture area renders it susceptible to interference from oil, moisture, or surface contaminants.[3]

Iris recognition, widely regarded as the most accurate bio-metric modality, is predominantly deployed in high-security environments such as nuclear facilities and border control. Its non-contact operation (with a detection range of up to 1 meter) mitigates hygiene concerns. Nevertheless, the tech-nology faces limitations due to high infrastructure costs and performance dependency on optimal conditions.

Facial recognition has gained widespread adoption owing to its passive data collection capability. When integrated with specialized cameras (e.g., thermal imaging), it enables multimodal data acquisition. However, this technology ex-hibits significant sensitivity to illumination conditions, with recognition performance potentially degrading by 25-40% under strong backlighting or extreme lighting conditions. Also despite advancements in AI, concerns persist regard-ingits reliability and security vulnerabilities, particularly in

ISSN: 2424-8460(O)
2251-2608(P)

dynamic real-world scenarios.

Among various biometric modalities, palmprint recogni-tion stands out as a promising technology due to its rich tex-tural features, high degree of uniqueness, and ease of acqui-sition. Unlike other biometric traits such as fingerprints or facial features, palmprints are less prone to wear and tear, do not cause interference when captured, and contain multiple types of recognition features, including principal lines, wrin-kles, and ridges.This ensures high accuracy while maintain-ing robustness against external contaminants (e.g., oil, mois-ture) due to its larger capture area. Requiring only standard RGB cameras for data acquisition, the technology demon-strates significantly lower implementation costs compared to other biometric modalities. These inherent advantagesposi-tion it as a particularly competitive solution in the field of biometric identification.[4]

Conventional palmprint recognition methods typically rely on hand-crafted features, such as Gabor filters or local binary patterns, combined with statistical classifiers. How-ever, these methods face challenges in adapting to com-plex real-world scenarios, such as changing lighting con-ditions, partial occlusion, and image noise. The emer-gence of deep learning, especially convolutional neural net-works (CNNs), has revolutionized biometric identification by enabling automatic feature extraction and learning robust representations.[5]

This research aims to utilize advanced deep learning tech-niques to address the limitations of existing palmprint recog-nition systems.highlight the challenges of dynamic environ-ments in palmprint recognition, which our attention-based approach aims to address. We propose a novel framework that combines migration learning with pre-trained models (e.g., ResNet101 and VGG16) to efficiently extract local and global features. In addition, an innovative attention mecha-nism is introduced to dynamically focus on key palmprint regions to enhance the model's ability to recognize fine tex-ture details. By combining powerful data preprocessing, en-hancement techniques and comprehensive evaluation strate-gies, this study aims to improve recognition accuracy and robustness across different datasets

He main innovations of this paper are twofold:

1) Conducting comparative analysis of multiple models through transfer learning and deep learning to determine the most suitable architecture for palmprint recognition;

2) Performing model comparisons after introducing atten-tion mechanisms to verify their optimization effects and in-vestigate potential impacts on model suitability for palmprint recognition tasks.[6, 7, 8]

This paper is structured as follows: Section 1 introduces fundamental biometric recognition technologies and the unique characteristics of palmprint identification. Section 2 elaborates on the theoretical foundations of the methodolo-gies employed in this study. Section 3 details experimen-tal data definitions, computational methods, and prepara-tory procedures. Section 4 presents comprehensive analysis of the experimental results. The concluding section evalu-ates both findings and limitations, while discussing future prospects for this technology.

1.1. Research procedure

In this research, we will divide the steps into the following to carry out the research.

1) Image processing: Since different AI models have dif-ferent requirements for images, we need to process therel-evant images before training before use, i.e., checking and eliminating corrupted images, unifying the imagesize, nor-malizing the pixel values, and carrying out data enhancement such as random rotation, translation, scaling, etc.

2) Data set division: Divide training, validation, and test sets proportionally (e.g. 70%/20%/10%) to ensure consistent category proportions.

3) Model design: classical pre-training models such as ResNet101 and VGG16 are used to efficiently extract

ISSN: 2424-8460(O)
2251-2608(P)

the lo-cal texture and global structural characteristics of palmprint images through migration learning to improve the feature ex-traction efficiency, and at the same time, the network weight distribution is adjusted according to the data characteristics to optimize the adaptation capability. The attention mecha-nismis introduced into the model to enhance the accuracy and robustness of palmprint feature expression by dynami-cally optimizing the model's ability to pay attention to key regions, significantly improving the model's ability to recog-nize fine textures, thus improving classification and recogni-tion performance.

4) Model training: using cross-entropy loss and optimizer (e.g., Adam), combined with learning rate scheduler and dropout to prevent overfitting.

5) Model Evaluation: Evaluate the performance of the model using metrics such as accuracy, F1 score, etc., and draw a confusion matrix to analyze the classification effect.

2. Technical introduction

In this section, we will briefly explain the relevant tech-niques which I used in this study.

2.1. Use of modules

In this study, we have experimented with the follow-ing four given models, InceptionV3 , Xception, ResNet101, VGG19. Each of them will be described below.

1) InceptionV3

InceptionV3 is a deep convolutional neural network proposed by Google in 2015, which is an improved version of the Inception series of networks, mainly used for image classification and feature extraction tasks. It is based on the original Inception architecture, and fur-ther improves the performance and computational effi-ciency of the network through a series of optimization designs, while reducing the demand for computational resources. The design concept of InceptionV3 is to re-duce the computational complexity while maintaining the depth of the network through a refined modular de-sign.

InceptionV3 achieves extremely high accuracy in the ImageNet classification task while significantly reducing the number of parameters and computational re-source requirements compared to similar deep network architectures such as VGG and ResNet. This makes In-ceptionV3 a representative network that balances high efficiency with high performance and is widely used in various computing platforms and real-world applications. Its related schematic is shown in **Figure 1** [9]

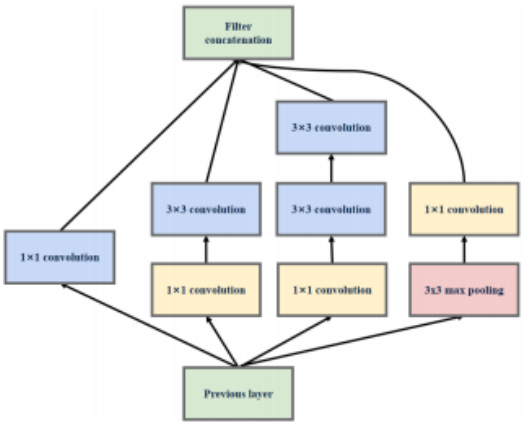

Figure 1. Inception V3.

2) Xception

TheXception model is a deep convolutional neural net-work architecture proposed by Google in 2017, whose

ISSN: 2424-8460(O)
2251-2608(P)

name is derived from "Extreme Inception", which rep-resents the extreme expansion of the Inception architecture.The core idea of Xception is to redesign the tra-ditional convolution operation as Depthwise Separable Convolution, which significantly improves the compu-tational efficiency and feature extraction capability of the model. The core idea of Xception is to redesign the traditional convolution operation as Depthwise Separa-ble Convolution, which significantly improves the com-putational efficiency and feature extraction capability of the model.

In addition, the Xception model retains key features of modern deep networks, such as the use of the ReLU activation function, Batch Normalization, and a glob-ally averaged pooling layer. It also employs Resid-ual Connection to mitigate the gradient vanishing prob-lem and accelerate network convergence. These de-signs allow Xception to be both computationally effi-cient and powerful in feature learning, while also ex-hibiting good robustness in real-world applications. Its related schematic is shown in **Figure 2** [10]

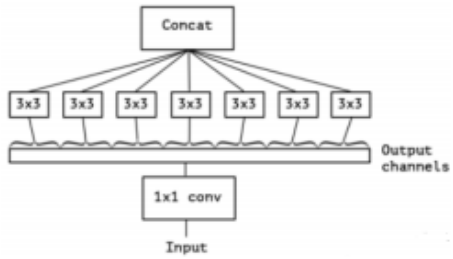

Figure 2. Xception.

3) ResNet101

ResNet101 is a member of the deep residual network (Residual Network) family, which was proposed by Mi-crosoft Research in 2015 to solve the problem of gra-dient disappearance and degradation that deep neural networks are prone to during the training process. Its core idea is to optimize the training efficiency and performance of deep networks by introducing a Residual Block, which allows information to propagate through the network indirect jumps.

Compared to other members of the ResNet family, such as ResNet50 or ResNet34, ResNet101 exhibits higher classification accuracy due to its deeper network structure capable of capturing more complex and abstract features. Meanwhile, by optimizing the design of the residual module, ResNet101 maintains a relatively manageable level of computational cost. Its excellent performance on large-scale datasets such as Image Net has made it an important benchmark model in the field of deep learning, and it is widely used in tasks such as image classification, target detection, and semantic segmentation. [11]

4) Visual Geometry Group (VGG) model

The VGG model is a convolutional neural network (CNN) architecture proposed by the Computer Vision Group at the University of Oxford in the ImageNet im-age classification task in 2014. Its main feature is to build networks with greater depth by stacking smaller 3×3 convolutional kernels and 2×2 maximum pooling layers. This design idea is simple but efficient, giving the model a strong representation in capturing image features.

5) The core innovation of the VGG model is its increased

network depth (typically 16 or 19 layers), which uses smaller convolutional kernels instead of larger ones (e.g., 7×7 or 5×5) compared to earlier networks such as AlexNet. This design reduces the number of parame-ters while maintaining the size of the receptive field.The VGG model allows the network to better extract high-level features through the superposition of multiple lay-ers of convolutions, while simplifying the design logic of the model.

ISSN: 2424-8460(O)
2251-2608(P)

The VGG network structure consists of a series of con-volutional and pooling layers, and finally a fully con-nected layer is attached for classification tasks. The input is usually a fixed-size image (e.g., 224 × 224 × 3), which is spread and passed to the fully connected layer after multiple layers of convolution and pooling operations. The output of each layer is passed through a ReLU activation function to increase nonlinearity, while Dropout techniques are used to prevent overfit-ting.

Although the VGG model performed well in the Im-ageNet competition, its drawbacks are more obvious: the huge number of parameters (especially in the fully connected layer) leads to high computational cost and long training time. Nevertheless, the idea of the VGG model has been widely applied and optimized in subse-quent model designs, and modern architectures such as ResNet and DenseNet, for example, have borrowed its modular design.[12, 13]

2.2. Attention mechanisms

Attention Mechanism (AM) is a technique widely used in deep learning, originally originated in the field of natural lan-guage processing to improve the performance of sequence-to-sequence models (Seq2Seq). Its core idea is to improve the expressiveness and processing efficiency of the model by dynamically assigning attention weights, allowing the model to focus on key parts of the input while ignoring relatively unimportant information when processing data.

The attention mechanism was first applied to machine translation tasks, where a weighted representation is gener-ated by aligning the output of each step of the decoder with the hidden state of the encoder, and calculating the size of each input word's contribution to the current output. The computation of the weights is usually done through a score function, such as dot product, addition, or a learnable feed-forward network, and the score is then normalized to a prob-ability distribution by a Softmax function.

In the field of image processing, attention mechanisms also perform well. For example, in convolutional neural networks, the attention mechanism dynamically adjusts the weights of each feature channel or spatial location to en-hance the model's ability to perceive key features. Such mechanisms are usually realized in the form of channel at-tention, spatial attention, or a combination of both, such as SE module, CBAM, etc.

In this study we use the SE. The Squeeze-and-Excitation (SE) block is an architectural unit designed to improve the representational power of a network by explicitly model-ing channel-wise relationships. The mechanism operates through three key operations:

The first is the Squeeze Operation. Its primary function is to compress spatial features into channel descriptors by pooling global averages. The squeeze operation compresses spatial information into a channel descriptor using global av-erage pooling. For a feature map $U \in R^{H\times W\times C}$ with C channels, we calculate:

$$z_c = F_{sq}(U_c) = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} u_c(i,j) \tag{1}$$

Where uc (i,j) represents the activation at spatial location (i,j) in channel c.

Then is the Excitation Operation.Its primary function is to learn nonlinear inter-channel relationships and output channel-wise weights ranging from 0 to 1 The excitation operation learns a nonlinear interaction between channels through two fully-connected layers:

$$s = F_{ex}(z, W) = \sigma(W_2 \delta(W_1 z)) \tag{2}$$

Where:

• $W_1 \in R^{\frac{C}{r}\times C}$ and $W_2 \in R^{\frac{C}{r}}$ are learnable weight matrices

ISSN: 2424-8460(O)
2251-2608(P)

• r is the reduction ratio (typically set to 16)

• δ denotes the ReLU activation function

• σ represents the sigmoid function

Last is the Scale Operation. Its primary function is fea-ture recalibration.The final output is obtained by rescaling the original feature map U with the learned channel weights:

$$\tilde{u}c = F_{scale}(U_c, s_c) = s_c \cdot u_c \quad (3)$$

Where s_C is the excitation weight for channel c.

Since the Transformer architecture was proposed, Multi-Head Self-Attention (MHSA) has become one of the core applications of the attention mechanism. It captures global dependencies by computing the similarity with other posi-tions at each position of the input sequence, thus addressing the limitations of traditional RNNs in long sequence process-ing. The multi-head mechanism further enhances the expres-sive power of the model by computing attention in parallel in different subspaces.[13, 14]

The advantages of the attention mechanism lie in its flex-ibility and versatility, which not only improves the perfor-mance of the model for a specific task, but also intuitively ex-plains the decision-making process of the model. It is widely used in image classification, target detection, text generation, speech recognition, etc., which promotes the development of deep learning in multimodal tasks. As we can see in the **Figure 3**, we can know some project that already using the attention mechanisms.[15]

Figure 3. Attention Mechanisms use in other project.

2.3. Transfer learning

Migration learning is a machine learning approach that aims to extract knowledge acquired in one task to another related task to improve the performance of the model on the new task. While traditional machine learning typically requires large amounts of labeled data, migration learning can significantly reduce the dependence of training on data from a new task by leveraging existing models or data, thus speeding up training and improving performance. The core idea of migration learning is the commonalities that may ex-ist between different tasks, and by sharing these common-alities, pre-existing models can be used as a starting point for new tasks. For example, in computer vision, a model trained on ImageNet is able to extract common image lan-guage features that can be migrated to new tasks such as medical image classification. Similarly, in nature process-ing, predefined language models such as BERT can be mi-grated, to meet specific tasks such as sentiment analysis and question and answer systems. Migration learning performs urgently in scenarios where data is scarce or computational resources are limited, and is one of the important directions in the development of modern artificial intelligence.[16]

In this study, transfer learning was employed to adapt pre-trained models for palmprint recognition tasks. Given that the four benchmark models were originally designed for distinct purposes (e.g., Xception for face recognition)By comparing post-transfer performance metrics across archi-tectures, this approach enables empirical determination of optimal model characteristics for palmprint biometrics, ful-filling our comparative

ISSN: 2424-8460(O)
2251-2608(P)

evaluation objectives.

3. Research process

3.1. Innovation point realization

In this paper, we introduced the attention mechanism to optimize the model for image processing and recognition. The model with the introduction of the attention mechanism has significant advantages over the model without the intro-duction, which are mainly reflected in the extraction ability, accuracy and robustness. Attention mechanism models are able to dynamically assign weights and pay more attention to task-relevant key regions in the image, thus avoiding wasting computational resources on relevant details. This ability en-hances the discriminative power of the model, and as a result, it exhibits higher accuracy in classification, detection, and segmentation tasks, especially in complex scenarios, such as when the image contains multiple targets or backgrounds with interfering variables, where the model is able to more accurately perform region delineation and identify important information. Attention still enhances the robustness of the model to bias and noise, and even if the target is partially biased, the model can still focus on the significant features of the bias mechanism and maintain good performance. In addition, the attention mechanism helps the model to bet-ter capture local details and global attention information by dynamically adjusting the weights on multi-pixel features, which is a critical key in solving highly variable tasks. [17]

3.2. Model training

The experiment utilized a combined dataset consisting of the PolyU Palmprint Dataset and a custom PTIT-Hand dataset, totaling 15,200 palmprint images captured by RGB cameras at 640×480 resolution. The data cov-ers varying il-lumination conditions (200-1000 Lux) and hand rotation an-gles (±30°). [18]

The preprocessing pipeline included: (1) ROI extrac-tion using OTSU's algorithm, (2) image normalization to 256×256 pixels, and (3) channel-wise standardization (mean=[0.485,0.456,0.406], std=[0.229,0.224,0.225]). The dataset was partitioned into training/validation/test sets at a 7:2:1 ratio while maintaining consistent class distributions. In this time, before the model training needs to be carried out on the various images for the relevant processing, in the actual situation, we will not be facing the camera recogni-tion, Palmprint will be flipped, missing and so on, and be-fore the training needs to be carried out on the images for the relevant processing, in this study first of all the image for the relevant books of measurements such as grayscale, color channel, etc., in order to draw its edge contour, in order to determine the attention mechanism attention to the The scope of the attention mechanism is determined. The results are shown in **Figure 4** below. This module transforms palm images captured by RGB cameras through standardized pre-processing, including grayscale conversion, color normaliza-tion, and resizing, to generate training-optimized input data.

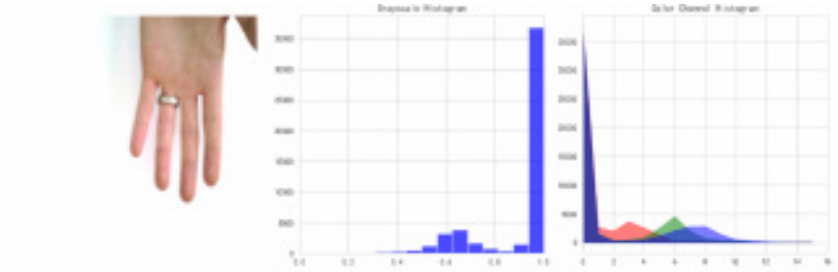

Figure 4. Draw its edge contour.

After processing the picture, we here first used the above four models more common four models for

ISSN: 2424-8460(O)
2251-2608(P)

training, but be-cause of its four models commonly used in the face and so on, so in the study utilized migration learning, one is to save time, reduce time costs, two is the use of the ex-isting model for migration learning its and faster conver-gence speed with, more accurate weight parameters, more conducive to our research The first is to save time .For fine-grained features in palmprint images, we adopted a hier-archical fine-tuning strategy: freezing the bottom convolu-tional layers of ResNet101 (preserving general edge detec-tion capabilities) while fine-tuning only the last three resid-ual blocks and classification layer. A tiered learning rate de-cay was applied (1e-5 for bottom layers, 1e-3 for top lay-ers) to prevent disruption of pretrained features. Experi-mental results demonstrated a 2.1.% accuracy improvement over global fine-tuning approaches. Here we decided to uti-lize four common models, namely InceptionV3, Xception, ResNet-101, VGG, and test their accuracy by comparing them with relevant data after training, as shown in **Figure 5** below.

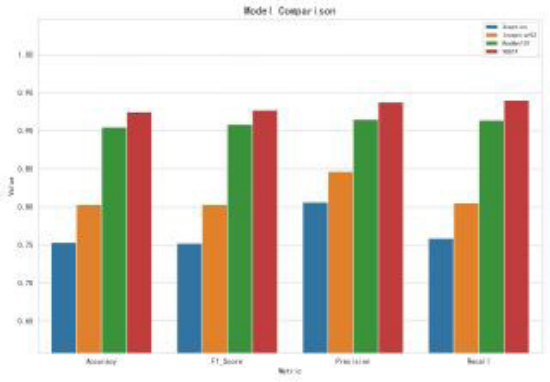

Figure 5. Histogram of model accuracy.

On the way it can be seen that for training after migra-tion learning, for the model ResNet-101, VGG is more ac-curate and basically achieves more than 90 accuracy and precision. But but for the model InceptionV3 and Xcep-tion effect is worse, These models demonstrated relatively low recognition accuracy, typically achieving less than 80% correct identification rates,For the Xception model, the ac-curacy was below 80%. During experimentation, we ob-served poor transfer learning efficacy for palmprint recogni-tion tasks. Notably, they required substantially longer train-ing times compared to the two better-performing architec-tures. Even after complete training, the models frequently exhibited matching failures, including both false rejections and incorrect identifications.

In this table ACCURACY is one of the most commonly used metrics in the performance evaluation of classification models,indicating the proportion of samples correctly pre-dicted by the model to the total samples. It is suitable for scenarios where the distribution of categories is more bal-anced, and can reflect the overall prediction performance of the model, which is calculated by (4).

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}, \tag{4}$$

Where We decide followings,

True Positive(TP): the number of samples correctly pre-dicted by the model as positive categories,

TN (True Negative): the number of samples correctly pre-dicted by the model as negative,

FP (False Positive): the number of samples that the model incorrectly predicts as positive,

FN (False Negative): the number of samples that the model incorrectly predicts as negative. [19] Precision is one of the important metrics to evaluate the performance of a classifica-tion model. It focuses on how many of the samples predicted by the model as positive classes are correct, and is often used to measure the accuracy of the model in predicting positive classes. It is calculated as (5).

ISSN: 2424-8460(O)
2251-2608(P)

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{5}$$

Recall, also known as sensitivity or check all, measures the model's ability to cover real positive class samples. It de-scribes the proportion of all real positive class samples that are successfully identified by the model. In other words, Re-call focuses on the portion of the positive class that has not been missed. It is calculated as $\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$. (6)

The F1-Score is a composite metric for the performance evaluation of classification models, which is a reconciled av-erage of Precision and Recall, used to find a balance between these two metrics and given by $\text{F1-Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$. (7)

Afterwards, we add the attention mechanism on this basis and then train the four models, and then compare the training results, we can know how to optimize the effect, in which for the implementation of the attention mechanism, the main use of Squeeze-and-Excitation Attention simple attention mech-anism, in short,the main way to improve the feature compu-tation capability of the network is by compressing the feature vector of each channel through the Squeeze component, and weighting the feature representation of each channel through the Excitation component, so as to improve the feature com-putation capability of the network. In simple terms, it is mainly by compressing the feature vector of each channel through the Squeeze component, and weighting the feature representation of each channel through the Excitation com-ponent, so as to improve the feature computation ability of the network. In order to achieve what we call the attention mechanism, so that it focuses on a certain part and features, and combined with related information can be seen that it does not produce too much burden on the operation of the model. And after the actual use of its confusion series of each model with the final accuracy of the histogram. Then the data can be visualized in a more intuitive way. The data are in **Figure 6** to **Figure 9**.

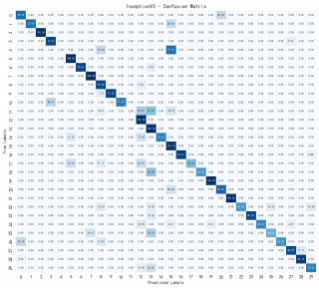

Figure 6. InceptionV3 using Attention mechanisms.

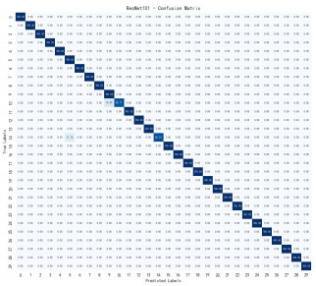

Figure 7. ResNet-101 using Attention mechanisms.

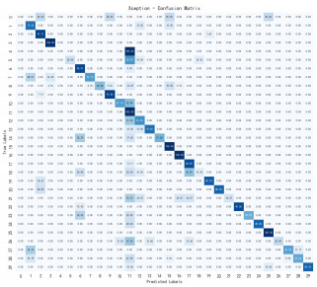

Figure 8. Xception using Attention mechanisms.

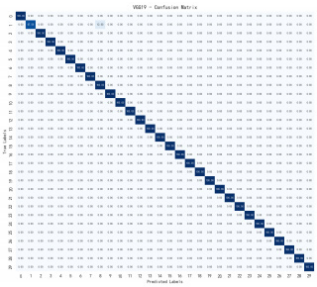

Figure 9. VGG using Attention mechanisms.

4. Analysis of results

First, the experimental results, as visualized in the confu-sion matrix, demonstrate the recognition performance across 29 individual palmprint classes. The principal diagonal elements represent correct classification rates, while off-diagonal entries indicate discrepancies between predicted and actual matches. from the picture of the confusion ma-trix can be seen, but InceptionV3 and Xception still exist in the accuracy of the situation is not high,The evaluation demonstrated InceptionV3 achieved up to 28% error rate, with Xception failing completely (100% error) on certain samples its error is large, from the confusion matrix in the heterochromatic differences can also be seen in therecogni-tion of its existence in the error is more, The remaining two models achieved superior recognition accuracy, consistently maintaining less than 95% correct classification rates with no more than 10% error margins, demonstrating particularly effective performance for

ISSN: 2424-8460(O)
2251-2608(P)

palmprint identification tasks.

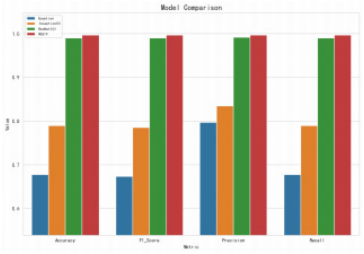

Figure 10. Accuracy Improvements with Attention Mecha-nisms.

As evidenced by the histogram distributions in **Figure 10** compared to previous Figure. 5, it can be clearly found that after the introduction of the attention mechanism, the accuracy of the four models has been improved by nearly 5%, and for the two models that are used in a better state, their accuracy has reached 98%, which is a high degree of improvement.These results provide empirical evidence that our introduced atten-tion mechanism effectively optimizes recognition accuracy across diverse palmprint identification models.

Based on the histogram and **Figure 5** above, the follow-ing conclusions can be drawn: First, after introducing the attention mechanism, all four models showed significant im-provements in accuracy. For the two poorly performing models,the attention mechanism improved their accuracy by 20%, while the two better-performing models even achieved 98% accuracy. This indicates that the introduction of the attention mechanism has significantly optimized palmprint recognition.

Second, regarding the original design intentions of In-ceptionV3 and Xception, even after applying transfer learn-ing and deep learning, their performance in capturing fine-grained features of palmprints remains unsatisfactory. These models require substantial time and cost to train effectively. Therefore, for practical deployment, it is recommended to use the more mature VGG model, which is commonly used in fingerprint recognition. Since the VGG model was de-signed as a large-scale solution for fingerprint recognition, its feature extraction for fingerprint-related patterns is simi-lar to that for palmprints, resulting in better performance af-ter transfer learning. Additionally, the VGG model demands less training data and shorter training times, making it a more practical choice for real-world applications.

This concludes the study. [20]

5. Conclusion and future directions

In this experiment, we have used four models to com-pare the effects and added an attention mechanism to opti-mize their recognition and compare the effects during the experiment. In this experiment used a large dataset, We thought that due to inceptionV3 and Xception need larger data. With InceptionV3 (approximately 23M parameters) and Xception (about 23M parameters) are more lightweight compared to VGG16 (about 138M parameters), but they re-quire large-scale training data to achieve full performance potential. Hereford, in this experiment, the limited train-ing data resulted in suboptimal performance for InceptionV3 and Xception, which typically require large-scale datasets for effective training., but after the actual training of its still appear poor results, indicating that the training of these two models, one is the need for large amounts of data, equipment, environment, data quality requirements are higher, and the second is that it may not be adapted to the palm of our hand recognition, although the use of the migration learning is used to train them, but InceptionV3 and Xception were orig-inally designed for processing complex natural images (such as object classification in ImageNet). Their multi-scale con-volution and depthwise separable convolution modules may be insufficient for effectively

ISSN: 2424-8460(O)
2251-2608(P)

capturing fine-grained texture features like palmprint wrinkles andridge patterns. So,the effect is much less than that the other two models. There-fore, we can use ResNet-101 and VGG model in the case of small arithmetic power and little data in the real application.

This study has several methodological limitations in data sampling. First, we only controlled for angle and illumina-tion variations while neglecting other potential influencing factors such as shadows and surface stains. Second, due to computational resource and time constraints, the experiment was limited to 29 palmprint classes without more extensive or complex training data. Finally, the data collection focused exclusively on Asian skin tones without incorporating other demographic variations, representing an area for potential optimization in future studies.

The proposed palmprint recognition framework, integrat-ing attention mechanisms with transfer learning, demon-strates significant potential for future advancements in bio-metrics and related fields. One promising direction is the ex-tension to multi-modal biometric systems, combining palm-print data with other biometric traits (e.g., fingerprints, iris, or facial features) to enhance security and robustness in real-world applications such as mobile payment authentication and border control.

Additionally, the deployment of lightweight models op-timized for edge devices (e.g., smartphones and IoT sys-tems) could further improve accessibility and real-time per-formance. Future research may explore efficient attention variants (e.g., MobileViT or EfficientNet-based attention) to reduce computational overhead while maintaining high ac-curacy.

Beyond biometrics, the techniques developed in this study could be adapted for medical imaging analysis (e.g., palm vein recognition for patient identification) and industrial automation (e.g., robotic hand-object interaction based on palmprint texture). The attention mechanism's ability to fo-cus on discriminative features also makes it applicable to general fine-grained recognition tasks, such as material clas-sification and defect detection in manufacturing.

Finally, federated learning could be explored to enhance privacy-preserving palmprint recognition, enabling decen-tralized model training without exposing sensitive biomet-ric data. These advancements would align with growing de-mands for secure, efficient, and scalable AI-driven authenti-cation systems.[21]

References

[1] Diego Maureira, Oscar Romero, Andrés Illanes, Lorena Wil-son, Carminna Ottone, Industrial bioelectrochemistry for waste valorization: State of the art and challenges,Biotechnology Advances,Volume64,2023,108123,ISSN 0734-9750, https://doi.org/10.1016/j.biotechadv.2023.1.08123.

[2] Zhang, D., Kong, W. K., You, J., Wong, M. (2003). Online palmprint identification. IEEE Transactions on Pattern Analy-sis and Machine Intelligence, 25(9), 1041-1050.

[3] Jia, W., Zhang, B., Lu, J., Zhu, Y., Zhao, Y., Zuo, W. Ling, H. (2018). Palmprint recognition based on complete direction representation. IEEE Transactions on Image Processing, 26(9), 4483-4498.

[4] Jain, A. K., Nandakumar, K., Ross, A. (2016). 50 years of bio-metric research: Accomplishments, challenges, and opportuni-ties. Pattern Recognition Letters, 79, 80-105.

[5] Zhang, David Zuo, Wangmeng Yue, Songfeng. (2012). A Com-parative Study of Palmprint Recognition Algorithms. ACM Comput. Surv.. 44. 2. 10.1145 2071389.2071391.

[6] J.Lee, S.Jang, Y. Lee and S. Lee, ”One-Stage Mobile Palmprint Recognition via Keypoint Detection Network,"2023 Interna-tional Technical Conference on Circuits/Systems, Computers, and Communications (ITC CSCC), Jeju, Korea,

Republic of, 2023, pp. 16, doi: 10.1109/ITCCSCC58803.2.023.1.0212697.

[7] P.Hennings,M.Savvides and B. V. K. V. Kumar,”Palm-print Recognition with Multiple Correlation Filters Using Edge Detection for Class-Specific Segmentation,"2007 IEEE Workshop on Automatic Identification Advanced Technologies, Alghero, Italy, 2007, pp. 214219, doi:10.1109/AU-TOID.2007.380622.

[8] Amrouni, Nadia, Amir Benzaoui, and Abdelhafid Zeroual. 2024. ”Palmprint Recognition:ExtensiveExploratio n of Databases,Methodologies,Comparative Assessment,and FutureDirections"Applied Sciences14,no.1: 153. https://doi.org/10.3390/app14010153.

[9] N.A.Nainan, J.H,R.Jalan, R.S.N,K.V.C and S. P. Shankar, ”Real Time Face Mask Detection Using MobileNetV2 and In-ceptionV3 Models,"2021 IEEE Mysore Sub Section Interna-tional Conference (MysuruCon), Hassan, India, 2021, pp. 341-345, doi: 10.1109/MysuruCon52639.2021.9641675.

[10] Rismiyati, S.N.Endah, Khadijah and I.N.Shiddiq, ”Xception Architecture Transfer Learning for Garbage Classification,"2020 4th International Conference on Informatics and Com-putational Sciences (ICICoS), Semarang, Indonesia, 2020, pp. 1-4, doi: 10.1109/ICICoS51170.2020.9299017.

[11] Q.Zhou andY.Li, ”Method for Tool Power Recognition Based on improved ResNet50 and Transfer learning,"2024 4th Inter-national Conference on Electronics, Circuits and Information Engineering (ECIE), Hangzhou, China, 2024, pp. 572-576, doi: 10.1109/ECIE61885.2.024.1.0627101.

[12] Y.Tao, ”Image Style Transfer Based on VGG Neural Net-work Model,"2022 IEEE International Conference on Ad-vances in Electrical Engineering andComputer Applica-tions (AEECA), Dalian, China, 2022, pp.1475-1482, doi: 10.1109/AEECA55500.2022.9918891.

[13] Z.Zheng and N.Slam,”A VGG Fire Image Classifica-tion Model with Attention Mechanism,"2024 5th International Conference on Computer Vision, Image and Deep Learning (CVIDL), Zhuhai, China, 2024, pp. 873-877, doi: 10.1109/CVIDL62147.2024.1.0603830.

[14] Gao,Y.;Zhang,J.;Wang,M.;Tan,Z.;Liang,M. A Lightweight Load Identification Model Update Method Based on Channel Attention. Energies 2025,18,2885. https://doi.org/10.3390/en18112885.

[15] H. Mo, Z. Yang, P. Li and Q. Wang, ”The Method of Portrait Segmentation Based on Efficient Channel Attention Mecha-nism,"2024 4th International Conference on Computer Sci-ence and Blockchain (CCSB), Shenzhen, China, 2024,pp. 380-385, doi: 10.1109/CCSB63463.2.024.1.0735544.

[16] L. Shao, F. Zhu and X. Li, ”Transfer Learning for Visual Categorization: A Survey,"in IEEE Transactions on Neural Networks and Learning Systems, vol. 26, no. 5, pp. 1019-1034, May 2015, doi:10.1109/TNNLS.2014.2.330900. key-words:{Knowledge transfer;Visualization;Training;Training data;Adaptation models;Learnin.

[17] F. Yang, ”AttenVGG: Integrating Attention Mechanisms with VGGNet for Facial Expression Recognition,"2024 5th Inter-national Conference on Electronic Communication and Artifi-cial Intelligence (ICECAI), Shenzhen, China, 2024, pp. 332-337, doi: 10.1109/ICECAI62591.2.024.1.0674809.

[18] Genovese, A., Piuri, V., Plataniotis, K. N., Scotti, F. (2016). PalmNet:Gabor-PCAconvolutional networks for touch-less palmprint recognition. IEEE Transactions on Information Forensics and Security, 11(3), 553-567.

[19] M.Z.Hossain, F.Uz Zaman and M. R. Islam, ”Advancing AI-Generated Image Detection: Enhanced Accuracy through CNN and Vision Transformer Models with Explainable AI Insights,"2023 26th International Conference on Computer and Infor-mation Technology (ICCIT), Cox's Bazar, Bangladesh, 2023, pp. 1-6, doi: 10.1109/ICCIT60459.2023.1.0440990.

[20] L.Fei, G. Lu, W.Jia, S.Teng and D.Zhang, ”Feature Extrac-tion Methods for Palmprint Recognition: A Survey and Eval-uation,"in IEEE Transactions on Systems, Man, and Cyber-netics: Systems, vol. 49, no. 2, pp. 346-363, Feb. 2019, doi: 10.1109/TSMC.2018.2795609.

[21] Lee, J., et al. (2023). ”One-Stage Mobile Palmprint Recogni-tion via Keypoint Detection Network."ITC-CSCC.