University of Mohamed Boudiaf - M'sila

FACULTY OF TECHNOLOGY

DEPARTEMENT OF ELECTRONICS

Serial number……..…

Inscription number……..……

Thesis

Presented for the Degree of

DOCTORAT LMD

Field: Electronics

Specialty: Embedded systems electronics

THEME

Deep Learning Approach for Multimodal Biometric Recognition System

Presented By

NADIR Cheyma

Graduated on: 2 1 /04/2025

Before the jury composed of:

Name & Surname	Grade	Establishment	Quality
BOURAS Mounir	**Professor**	Univ. Of **M'sila**	President
ATTALLAH Bilal	**Professor**	Univ. Of **M'sila**	Supervisor
BRIK Youcef	**MCA**	Univ. Of **M'sila**	Co-Supervisor
BENZAOUI Amir	**Professor**	Univ. Of **Skikda**	Examiner
ASBAI Nassim	**MCA**	Univ. Of **Bordj Bou Arréridj**	Examiner
DJERIOUI Mohamed	**MCA**	Univ. Of **M'sila**	Examiner
KHENNOUF Salah	**MCA**	Univ. Of **M'sila**	Guest

Academic Year: 2024/2025

DEDICATION

To my family, who made this journey possible.

To my parents thank you for always supporting me, guiding me, and believing in me. Your sacrifices and love kept me going.

To my twin brother, we have shared everything, grown together, and supported each other through every challenge.

To my sisters, your love and encouragement mean everything to me.

To my nieces, Aya and Meriem, may this work remind you that learning never stops.

This achievement is for all of you a reflection of family, love, and shared dreams.

NADIR CHEYMA

Acknowledgments

First of all, I thank Allah for giving me the strength and patience to complete this thesis.

I would like to express my deepest gratitude to my thesis supervisors, **Professor ATTALLAH Bilal** and **Dr. BRIK Youcef**, for proposing this research topic and for their continuous support and guidance over the past four years.

Professor ATTALLAH Bilal, thank you for your help, kindness, understanding, and availability. I'm especially thankful for the chance to co-supervise four Master's theses with you. This helped me a lot in my academic and research skills. Your advice on writing and using LaTeX improved my work a lot. Your suggestions and follow-up on both the theory and practical parts of this thesis were very important.

Dr. BRIK Youcef deserves special thanks for his exceptional help, guidance, kindness, patience, understanding, and availability. I am particularly grateful for his significant role in my professional development. With his support, I progressed from being a beginner in programming to achieving competence in this field. I deeply appreciate his patience in explaining difficult concepts and his confidence in my abilities. Your influence on my academic journey has been significant.

I also want to thank the jury members for their time and useful comments on this work:

Professor BOURAS Mounir (University of M'sila), Thank you for accepting to be the President of the jury and for your presence and guidance.

Professor BENZAOUI Amir (University of Skikda), Thank you for your detailed review, valuable feedback, and interesting questions.

Dr. ASBAI Nassim (University of Bordj Bou Arreridj), Thank you for your valuable feedback.

Dr. DJERIOUI Mohamed (University of M'sila), thank you for your questions and review.

Dr. KHENNOUF Salah (University of M'sila),We sincerely thank you for honoring us with your presence and for your valuable support.

I also thank all the teachers in the Department of Eletronic at the University of M'sila who helped me during my studies. Finally, I thank my family and friends for their strong support, encouragement, love, and patience during this journey.

Abstract

Multimodal biometric recognition has emerged as a powerful approach to enhance recognition performance by leveraging multiple data sources. This thesis investigates unimodal and multimodal systems, focusing on finger vein (FV) and palmprint recognition, which are highly secure and reliable modalities due to their unique characteristics and resistance to forgery. Finger vein patterns, located beneath the skin, are invisible to the eye and difficult to replicate, making them ideal for high-security applications such as access control and financial transactions. Similarly, palmprint recognition relies on the distinct patterns of lines, wrinkles, and textures on the palm, which remain consistent throughout an individual's life, even among identical twins. However, both modalities face challenges: finger vein recognition is affected by variations in lighting and image quality, while palmprint systems struggle with illumination changes, rotation, and scale variations, particularly in contactless environments.

To address these challenges, we propose a series of experiments. First, we develop unimodal systems for each modality, evaluating their performance using transfer learning with three CNN models (VGG16, VGG19, and MobileNetV2). The best-performing model, MobileNetV2, is selected to extract more relevant features, and we fine-tune the top layers of the model. Additionally, we enhance the images using preprocessing techniques such as histogram equalization and contrast-limited adaptive histogram equalization (CLAHE). To account for potential issues like injuries or dirt, we propose leveraging information from both the left and right instances of each modality. This system is further evaluated by replacing the classifier with machine learning classifiers. All experiments are conducted under three protocols: Protocol 1 (P1), which uses 70% of the dataset for training and the remaining 30% for testing; Protocol 2 (P2), which uses 50% for training and the remaining 50% for testing; and Protocol 3 (P3), which uses 90% for training and the remaining 10% for testing. For finger vein recognition, we utilize the SDUMLA and FV-USM datasets, achieving first-rank accuracies of 99.57%, 97.58%, and 99.76% for SDUMLA and 99.90%, 99.60%, and 99.80% for FV-USM. For palmprint recognition, we employ the CASIA and IIT Delhi datasets, achieving accuracies of 99.13%, 99.69%, and 100% for CASIA and 99.46%, 98.50%, and 99.35% for IIT Delhi.

Next, we preprocess finger vein and palmprint images to enhance system performance. For multimodal recognition, we propose a method to isolate the region of interest (ROI) from raw images, improve image quality using techniques such as CLAHE and data augmentation, and extract features from both modalities. These features are concatenated and fed into machine learning classifiers. To further improve performance, we apply score-level fusion using the outputs of MobileNetV2 and the classifiers. Experiments are conducted using the SDUMLA-HMT dataset for finger vein recognition and the IIT Delhi dataset for palmprint recognition. Three protocols (P1, P2, and P3) are evaluated. For finger vein recognition, first-rank accuracy rates reach 97.88%, 98.27%, and 99.06% for the left finger vein and 97.41%, 96.23%, and 97.64% for the right finger vein across the three protocols. Palmprint recognition achieves first-rank accuracies of 98.35%, 96.70%, and 99.06% for the left palmprint and 99.06%, 95.30%, and 99.50% for the right palmprint. The P3 protocol consistently demonstrates the highest performance.

The proposed multimodal fusion approach achieves perfect recognition accuracy (100%) at both feature and score levels, significantly outperforming unimodal methods across all protocols. This superior performance is measured in precision, recall, F1 score, AUC, R1, and R5 metrics, with the fusion system consistently achieving optimal results. These findings underscore the effectiveness of our multimodal approach in enhancing the accuracy and reliability of biometric recognition systems.

Keywords: Multimodal, Finger Vein, Palmprint, Unimodal, Feature fusion, Score fusion.

Résumé

La reconnaissance biométrique multimodale est devenue une approche puissante pour améliorer les performances de reconnaissance en exploitant plusieurs sources de données. Cette thèse étudie les systèmes unimodaux et multimodaux, en se concentrant sur la reconnaissance des veines des doigts (VD) et des empreintes palmaires, qui sont des modalités hautement sécurisées et fiables en raison de leurs caractéristiques uniques et de leur résistance à la falsification. Les motifs des veines des doigts, situés sous la peau, sont invisibles à l'œil et difficiles à reproduire, ce qui les rend idéaux pour des applications de haute sécurité telles que le contrôle d'accès et les transactions financières. De même, la reconnaissance des empreintes palmaires repose sur les motifs distincts de lignes, de rides et de textures sur la paume, qui restent cohérents tout au long de la vie d'un individu, même chez des jumeaux identiques. Cependant, les deux modalités sont confrontées à des défis : la reconnaissance des veines des doigts est affectée par les variations d'éclairage et de qualité d'image, tandis que les systèmes d'empreintes palmaires ont du mal à gérer les changements d'éclairage, la rotation et les variations d'échelle, en particulier dans les environnements sans contact.

Pour relever ces défis, nous proposons une série d'expériences. Tout d'abord, nous développons des systèmes unimodaux pour chaque modalité, en évaluant leurs performances à l'aide de l'apprentissage par transfert avec trois modèles CNN (VGG16, VGG19 et MobileNetV2). Le modèle le plus performant, MobileNetV2, est sélectionné pour extraire des caractéristiques plus pertinentes, et nous affinons les couches supérieures du modèle. De plus, nous améliorons les images à l'aide de techniques de prétraitement telles que l'égalisation d'histogramme et l'égalisation d'histogramme adaptative à contraste limité (CLAHE). Pour tenir compte des problèmes potentiels tels que les blessures ou la saleté, nous proposons d'exploiter les informations des instances gauche et droite de chaque modalité. Ce système est ensuite évalué en remplaçant le classificateur par des classificateurs d'apprentissage automatique.

Toutes les expériences sont menées selon trois protocoles : Protocole 1 (P1), qui utilise 70 % de l'ensemble de données pour l'entraînement et les 30 % restants pour les tests ; Protocole 2 (P2), qui utilise 50 % pour l'entraînement et les 50 % restants pour les tests ; et Protocole 3 (P3), qui utilise 90 % pour l'entraînement et les 10 % restants pour les tests. Pour la reconnaissance des veines des doigts, nous utilisons les ensembles de données SDUMLA et FV-USM, atteignant des précisions de premier ordre de 99,57 %, 97,58 % et 99,76 % pour SDUMLA et de 99,90 %, 99,60 % et 99,80 % pour FV-USM. Pour la reconnaissance des empreintes palmaires, nous utilisons les ensembles de données CASIA et IIT Delhi, atteignant des précisions de 99,13 %, 99,69 % et 100 % pour CASIA et de 99,46 %, 98,50 % et 99,35 % pour IIT Delhi.

Ensuite, nous prétraitons les images des veines des doigts et des empreintes palmaires pour améliorer les performances du système. Pour la reconnaissance multimodale, nous proposons une méthode pour isoler la région d'intérêt (ROI) des images brutes, améliorer la qualité de l'image à l'aide de techniques telles que CLAHE et l'augmentation des données, et extraire des caractéristiques des deux modalités. Ces fonctionnalités sont concaténées et introduites dans des classificateurs d'apprentissage automatique. Pour améliorer encore les performances, nous appliquons une fusion au niveau du score en utilisant les sorties de MobileNetV2 et des classificateurs. Des expériences sont menées à l'aide de l'ensemble de données SDUMLA-HMT pour la reconnaissance des veines des doigts et de l'ensemble de données IIT Delhi pour la reconnaissance des empreintes palmaires. Trois protocoles (P1, P2 et P3) sont évalués. Pour la reconnaissance des veines des doigts, les taux de précision de premier rang atteignent 97,88 %, 98,27 % et 99,06 % pour la veine du doigt gauche et 97,41 %, 96,23 % et 97,64 % pour la veine du doigt droit dans les trois protocoles. La reconnaissance des empreintes palmaires atteint des précisions de premier rang de 98,35 %, 96,70 % et 99,06 % pour l'empreinte palmaire gauche et de 99,06 %, 95,30 % et 99,50 % pour l'empreinte palmaire droite. Le protocole P3 démontre systématiquement les meilleures performances.

L'approche de fusion multimodale proposée atteint une précision de reconnaissance parfaite (100 %) aux niveaux des caractéristiques et du score, surpassant considérablement les méthodes unimodales dans tous les protocoles. Cette performance supérieure est mesurée en termes de précision, de rappel, de score F1, d'AUC, de R1 et de R5, le système de fusion obtenant systématiquement des résultats optimaux. Ces résultats soulignent l'efficacité de notre approche multimodale pour améliorer la précision et la fiabilité des systèmes de reconnaissance biométrique.

Mots-clés : Multimodal, Veine du doigt, Empreinte palmaire, Unimodal, Fusion de caractéristiques, Fusion de scores.

الملخص

لقد برز التعرف البيومتري المتعدد الوسائط كنهج قوي لتعزيز أداء التعرف من خلال الاستفادة من مصادر بيانات متعددة. تدرس هذه الأطروحة الأنظمة أحادية الوسائط ومتعددة الوسائط، مع التركيز على التعرف على أوردة الإصبع وبصمات اليد، وهي طرق آمنة وموثوقة للغاية بسبب خصائصها الفريدة ومقاومتها للتزوير. إن أنماط أوردة الإصبع، الموجودة تحت الجلد، غير مرئية للعين ويصعب تكرارها، مما يجعلها مثالية للتطبيقات عالية الأمان مثل التحكم في الوصول والمعاملات المالية. وبالمثل، يعتمد التعرف على بصمات اليد على الأنماط المميزة للخطوط والتجاعيد والملمس على راحة اليد، والتي تظل ثابتة طوال حياة الفرد، حتى بين التوائم المتطابقة. ومع ذلك، يواجه كلا الطريقتين تحديات: يتأثر التعرف على أوردة الإصبع باختلافات في الإضاءة وجودة الصورة، بينما تكافح أنظمة بصمات اليد مع تغييرات الإضاءة والدوران واختلافات الحجم، خاصة في البيئات غير التلامسية.

ولمعالجة هذه التحديات، نقترح سلسلة من التجارب. أولاً، نقوم بتطوير أنظمة أحادية النمط لكل نمط، وتقييم أدائها باستخدام التعلم الانتقالي مع ثلاثة نماذج VGG16 وVGG19 وMobileNetV2 يتم اختيار النموذج الأفضل أداءً، MobileNetV2، لاستخراج المزيد من الميزات ذات الصلة، ونقوم بضبط الطبقات العليا من النموذج. بالإضافة إلى ذلك، نقوم بتعزيز الصور باستخدام تقنيات المعالجة المسبقة مثل histogram equalization وcontrast-limited adaptive histogram equalization CLAHE. لمراعاة المشكلات المحتملة مثل الإصابات أو الأوساخ، نقترح الاستفادة من المعلومات من كل من الحالات اليمنى واليسرى لكل نمط. يتم تقييم هذا النظام بشكل أكبر عن طريق استبدال المصنف بمصنفات التعلم الآلي. يتم إجراء جميع التجارب بموجب ثلاثة بروتوكولات: البروتوكول 1 (P1)، والذي يستخدم 70٪ من مجموعة البيانات للتدريب و30٪ المتبقية للاختبار؛ البروتوكول 2 (P2)، والذي يستخدم 50٪ للتدريب و50٪ المتبقية للاختبار؛ والبروتوكول 3 (P3)، والذي يستخدم 90% للتدريب والـ 10% المتبقية للاختبار. للتعرف على أوردة الإصبع، نستخدم مجموعات بيانات SDUMLA وFV-USM، ونحقق دقة من الدرجة الأولى بنسبة 99.57% و97.58% و99.76% لـ SDUMLA و99.90% و99.60% و99.80% لـ FV-USM. للتعرف على بصمات اليد، نستخدم مجموعات بيانات CASIA وIIT Delhi، ونحقق دقة 99.13% و99.69% و100% لـ CASIA و99.46% و98.50% و99.35% لـ IIT Delhi.

بعد ذلك، نقوم بمعالجة مسبقة لصور أوردة الإصبع وبصمات اليد لتحسين أداء النظام. للتعرف المتعدد الوسائط، نقترح طريقة لعزل منطقة الاهتمام (ROI) من الصور الخام، وتحسين جودة الصورة باستخدام تقنيات مثل CLAHE وزيادة البيانات، واستخراج الميزات من كلتا الطريقتين. يتم ربط هذه الميزات وإدخالها في مصنفات التعلم الآلي. لتحسين الأداء بشكل أكبر، نطبق دمج مستوى النتيجة باستخدام مخرجات MobileNetV2 والمصنفات. يتم إجراء التجارب باستخدام مجموعة بيانات SDUMLA-HMT للتعرف على أوردة الإصبع ومجموعة بيانات IIT Delhi للتعرف على بصمة اليد. يتم تقييم ثلاثة بروتوكولات (P1 وP2 وP3). بالنسبة للتعرف على أوردة الإصبع، تصل معدلات الدقة من الدرجة الأولى إلى 97.88٪ و98.27٪ و99.06٪ لوريد الإصبع الأيسر و97.41٪ و96.23٪ و97.64٪ لوريد الإصبع الأيمن عبر البروتوكولات الثلاثة. يحقق التعرف على بصمة اليد دقة من الدرجة الأولى بنسبة 98.35% و96.70% و99.06% لبصمة اليد اليسرى و99.06% و95.30% و99.50% لبصمة اليد اليمنى. يُظهِر بروتوكول P3 باستمرار أعلى أداء.
يحقق نهج الاندماج المتعدد الوسائط المقترح دقة التعرف المثالية (100%) على مستوى السمات والدرجات، متفوقًا بشكل كبير على الطرق أحادية النمط عبر جميع البروتوكولات. يتم قياس هذا الأداء المتفوق من حيث الدقة والتذكر ودرجة F1 وAUC وR1 وR5، مع تحقيق نظام الاندماج باستمرار لنتائج مثالية. تؤكد هذه النتائج على فعالية نهجنا المتعدد الوسائط في تعزيز دقة وموثوقية أنظمة التعرف البيومترية.

الكلمات الرئيسية: متعدد الوسائط، وريد الإصبع، وبصمة اليد، أحادي النمط، اندماج السمات، اندماج الدرجات.

Contents

List of Figures

List of Tables

List of Abbreviations

CNN	Convolutional Neural Network
DL	Deep Learning
KNN	K-Nearest Neighbors
SVM	Support Vector Machine
ML	Machine Learning
RF	Random Forest
CLAHE	Contrast-Limited Adaptive Histogram Equalization
ROI	Region of Interest
TL	Transfer Learning
FT	Fine-Tuning
FV	Finger Vein
MI	Multiple-Instance Learning
AI	Artificial Intelligence
AUC	Area Under Curve
CMC	Cumulative Match Characteristic
PIN	Personal Identification Number
FRR	False Rejection Rate
FAR	False Acceptance Rate
EER	Equal Error Rate
ROC	Receiver Operating Characteristic

Chapter 1

Introduction

The need for security increases rapidly as automation becomes more integrated into our daily lives. Many questions arise daily, such as textbf Is this person authorized to execute a specific transaction? , Do you think this individual should have access to a secure system? These questions center on a single security issue: accurately identifying individuals. There are three common ways to handle these security problems [1]:

The first method relies on **something you know**, such as passwords and Personal Identification Numbers (PINs), while the second method links to **something you have**, such as credit cards or keys.

Both methods give authorization to media such as passwords or keys, representing entities other than the end users. A user gets authority if he obtains the password or other media if not, he loses access. People must keep many cards and memorize passwords under such security schemes. Those who forget their cards or their passwords could get into serious trouble.

Researchers are discovering several approaches to solving this issue, with the biometrics approach showing the most promise. Biometrics should be a system that identifies or verifies individuals based on their unique physical characteristics or behaviors. It can distinguish between an authentic individual and a fraudulent impostor because it uses the third way, **something that you are**, to determine personal identification [2].

Biometric systems have gained significant attention due to their ability to authenticate individuals based on unique physical traits, such as fingerprints, iris patterns, palmprints, or vein structures, or behavioral traits, such as voice or gait. These systems offer unparalleled convenience and security by eliminating the need for external tokens or memorization. By leveraging something that you are, biometrics addresses critical vulnerabilities in traditional security systems and enhances the overall user

experience.

1.1 Problem Statement

Despite their potential, unimodal biometric systems, which depend on a single biometric trait for identification, face several limitations [3]:

- **Noisy Data**: Biometric samples may be affected by poor quality, environmental conditions, or sensor issues.
- **Non-Universality**: Some individuals may lack a required biometric trait due to physical or medical conditions.
- **Vulnerability to Spoofing**: Unimodal systems can be deceived by fake biometric samples, such as artificial fingerprints or photos.
- **Intra-Class Variations**: Differences in biometric samples from the same individual due to factors like lighting, posture, aging, or sensor noise can make consistent recognition difficult.
- **Inter-Class Variations**: Similar biometric traits across different individuals (e.g., identical twins) can lead to false matches.

To overcome these challenges, researchers have developed multimodal biometric systems that combine multiple biometric traits [4]. These systems integrate data from different sources to enhance recognition accuracy, robustness, and security. However, multimodal systems introduce additional complexities, such as selecting suitable fusion techniques at sensor, feature, score, or decision levels. Traditional machine learning algorithms used in biometric systems also face limitations. They often require extensive preprocessing and handcrafted feature extraction, which may not generalize well across different biometric traits or datasets [5]. Addressing these issues necessitates more advanced approaches that leverage recent advancements in artificial intelligence and deep learning.

1.2 Motivation and Research Objectives

1.2.1 Motivation

There are numerous reasons that motivate our interest in enhancing the performance of unimodal biometric identification approaches. Biometric features are not always stable when captured. For example, palmprint images may vary due to factors such as hand placement or skin conditions like dryness or moisture. Similarly, finger vein

patterns can be affected by environmental factors such as temperature or lighting, as well as physiological changes like blood flow variations, aging, or health conditions. These gaps degrade the performance of unimodal systems, leading to higher error rates and reduced reliability.

Moreover, unimodal systems are vulnerable to spoofing attacks, where a person may use fake biometric samples (e.g., a printed palmprint or a synthetic finger vein image) to deceive the system. They also struggle with intra-class variations, where the same individual's biometric data may differ across captures due to physiological changes, scars, aging, or temporary conditions like swollen fingers or behavioral inconsistencies like variations in hand placement.

To address these limitations, multimodal biometric systems have emerged as a promising solution. By integrating multiple biometric traits, multimodal systems provide complementary information, enhancing recognition accuracy and robustness. Palmprint patterns are rich in texture and unique to individuals, while finger vein patterns are under the skin and highly secure, making them difficult to forge. Combining these modalities mitigates the weaknesses of each individual trait, resulting in a more reliable and secure identification system.

In our research, we focus on these two modalities due to their complementary characteristics. To the best of our knowledge, there is limited research on combining these using deep learning-based feature extraction and fusion techniques. To enhance feature extraction, we propose using **MobileNetV2**, a lightweight and efficient deep learning model, fine-tuned for palmprint and finger vein recognition. After feature extraction, we explore both feature-level and score-level fusion to combine the information from both modalities. Finally, machine learning classifiers, such as Support Vector Machines (SVM), k-Nearest Neighbors (KNN), and Random Forests, are employed to classify the fused features.

1.2.2 Research Objectives

The primary objectives of this research are as follows: **Develop a multimodal biometric system using palmprint and finger vein modalities:**

- Investigate the unique characteristics of these biometrics and their complementary strengths.
- Design a system architecture that integrates both modalities to enhance recognition accuracy and security.

Enhance MobileNetV2 for biometric feature extraction:

- Adapt the pre-trained model to extract discriminative features through fine-tuning.

- Utilize Region of Interest (ROI) techniques to focus on the most relevant parts of the images.

Explore fusion strategies:

- Implement and compare feature-level and score-level fusion techniques.
- Analyze the impact of different fusion strategies on overall system performance.

Evaluate system performance:

- Conduct experiments using publicly available palmprint and finger vein datasets.
- Assess the system using comprehensive metrics: precision, recall, F1 score, Area Under the Curve (AUC), Rank 1, and Rank 5.
- Compare performance against existing unimodal and multimodal biometric systems.

Address technical challenges:

- Propose an enhanced Region of Interest (ROI) extraction technique for palmprint images.
- Develop methods to handle scars and injuries by leveraging features extracted from MobileNetV2.
- Implement advanced preprocessing techniques such as histogram equalization and Contrast Limited Adaptive Histogram Equalization (CLAHE) to enhance image quality.

1.3 Methodology and Contributions

This thesis develops a robust multimodal biometric system through three progressive chapters, starting with unimodal analysis and advancing to multimodal integration.

1.3.1 Unimodal Performance

The initial phase evaluates finger vein and palmprint modalities to establish baseline recognition rates and identify modality-specific challenges.

- **Finger Vein Preprocessing:** Histogram equalization enhances contrast in finger vein images, improving visibility of vein patterns under varying lighting conditions.

- **Palmprint Preprocessing:** Contrast Limited Adaptive Histogram Equalization (CLAHE) locally adjusts contrast in palmprint images, enhancing line features despite variations in hand position or rotation in contactless scenarios.
- **Deep Feature Extraction:** VGG16, VGG19, and MobileNetV2, pre-trained on large datasets, are fine-tuned to extract features from enhanced images, identifying the best-performing model for each modality.
- **Instance Learning and Classification:** Two strategies are evaluated:
  - One Instance Learning: Treats each biometric instance independently.
  - Multiple Instance Learning: Combines multiple instances from the same subject for improved robustness.

  Machine learning classifiers process extracted features, with Softmax scores used to compare performance across strategies.

1.3.2 Multimodal Biometric System

The second phase integrates finger vein and palmprint modalities into a multimodal system.

- **ROI Extraction:** For palmprints, the proposed ROI extraction method is compared with the original dataset's ROIs to ensure optimal feature capture. ROIs are also extracted for finger vein images.
- **Data Preprocessing:** Palmprint images undergo data augmentation and CLAHE to enhance dataset variability and contrast. Finger vein images are processed with CLAHE for consistent enhancement across modalities.
- **Feature Concatenation and Classification:** Features from both modalities are combined into a unified vector, processed by classifiers (SVM, KNN, Random Forest) for multimodal recognition.
- **Score-Level Fusion:** Matching scores from classifiers are combined using methods like average, weighted sum, product, Borda count, max, harmonic, and Sugeno to optimize recognition accuracy.

1.3.3 Contributions

- **One Vs Multiple-Instance Learning:** Demonstrates that multiple-instance learning enhances system robustness compared to one-instance learning.

- **Region of Interest (ROI) Extraction:** Advanced algorithms are implemented to extract the most informative regions from finger vein and palmprint images, ensuring that subsequent processing focuses on the most discriminative biometric features.
- **Image Enhancement using CLAHE:** The application of Contrast Limited Adaptive Histogram Equalization enhances the quality of extracted ROIs, improving the contrast and visibility of critical features in both modalities.
- **Fine-tuning of the top layers:** is performed to adapt the model to the specific characteristics of the biometric data, ensuring optimal feature representation.
- **Fusion Techniques:** Advances multimodal biometrics through feature concatenation and score-level fusion, leveraging complementary modality information.
- **Feature-Level Fusion:** Features extracted from both modalities are fused through concatenation at the feature level, creating a comprehensive representation that captures the complementary information inherent in finger vein and palmprint data.
- **Classification:** The fused feature vectors are classified using SVM, RF, and KNN algorithms, selected for their robustness in handling high-dimensional feature spaces and their proven efficacy in biometric recognition tasks.

1.4 Thesis Organization

This thesis is structured as follows:

- **Chapter 1: Introduction** Provides an overview of biometric systems, highlighting the limitations of unimodal systems and the motivation for developing a multimodal system using palmprint and finger vein modalities. It outlines the research objectives, contributions, and thesis organization.
- **Chapter 2: Background and Related Work** Introduces the fundamental components of biometric systems, including data acquisition, preprocessing, feature extraction, and classification.
- **Chapter 3: Implementation of Finger Vein and Palmprint Recognition** Describes the implementation of the finger vein and palmprint recognition system. It details image preprocessing techniques, feature extraction using cnn models, and classification with machine learning algorithms. The system's performance is evaluated, comparing one-instance and multiple-instance recognition approaches.

- **Chapter 4: Feature-Level Fusion for Multimodal Biometrics** Presents the integration of palmprint and finger vein modalities into a unified multimodal system. Feature-level fusion techniques are explored, combining extracted features from both modalities for a more robust representation. The impact on system performance is evaluated, and results are discussed in detail.

- **Chapter 5: Score-Level Fusion for Enhanced Recognition** Extends the multimodal system by investigating score-level fusion techniques, such as weighted sum, averaging, and product of scores. The effectiveness of score-level fusion in improving recognition accuracy and robustness is analyzed and compared with feature-level fusion.

- **Chapter 6: Conclusion and Future Work** Summarizes the research findings and key contributions. It also discusses the system's limitations and suggests future research directions, including exploring additional modalities and advanced fusion techniques.

Chapter 2

Fundamentals of Biometric Systems

2.1 Introduction

Biometrics is an evolving technology that has become increasingly integrated into our daily lives. Its primary goal is to reliably establish a person's identity using their unique biological traits, thereby enhancing security in public spaces. A biometric system operates by capturing, processing, and matching these biological features to verify or identify individuals.

The structure of this chapter is as follows: Section 2 provides a definition and comparison of biometric traits and characteristics, offering a foundational understanding of the key concepts. Section 3 presents an overview of a biometric system's architecture and operating modes, detailing the components and functionalities that define such systems. Section 4 focuses on evaluating the effectiveness of a biometric system, including the criteria and metrics used to assess performance. Section 5 delves into feature extraction methods, exploring the techniques used to extract features from biometric traits, and discusses classification methods, highlighting the algorithms and approaches employed to process and match biometric data. Section 6 examines the types of biometric systems, and the difference between unimodal and multimodal systems, and introduces the concept of fusion levels (e.g., sensor-level, feature-level, score-level, and decision-level fusion) in multimodal systems. Finally, Section 7 concludes the chapter by summarizing the key points.

2.2 Biometric Traits and Characteristics

The term biometrics originates from the Greek words bios (life) and metron (measure). Biometrics utilizes an individual's physical or behavioral characteristics to digitally authenticate and grant access to devices and systems. It involves the study of an individual's unique physical and behavioral traits through measurement and statistical

analysis. Every individual possesses distinctive characteristics that cannot be easily stolen, forgotten, replicated, or shared [6].

Examples of biometric traits include face [7], palmprint [8], iris [9], fingerprint [10], voice [11], gait [12], hand geometry [13], and ECG [14]. These traits encompass biological, physiological, and behavioral aspects.

According to [15], the selection of biometric characteristics for various applications is based on several key criteria:

- **Universality**: The trait must be present in all individuals.
- **Uniqueness**: The characteristic should vary sufficiently across individuals.
- **Permanence**: The trait should remain stable over time.
- **Measurability**: It must be easily captured and processed.
- **Performance**: The trait should enable accurate and efficient identification.
- **Acceptability**: Users must be comfortable providing the biometric data.
- **Circumvention**: The trait should be difficult to forge or mimic.

Biometric traits are of two types, called physiological and behavioral traits as shwo in Table 2.1. [16]

-Physiological characteristics refer to an individual's physical traits, including voice and speech, iris, fingerprint, hand and palm geometry, ear, face, and finger vein. These characteristics are typically recognized for their collectability, uniqueness, permanence, and cost-effectiveness in identification and verification.

- Behavioral traits such as gait, signature, handwriting, and speech, describe aspects of an individual's personality and behavior. [16]

Biometric features play a crucial role in the development of various recognition systems, offering significant advantages across multiple applications. The selection of biometric traits depends on the specific application, the availability of dataset samples, the complexity of implementation, and the acceptable tolerance levels.

For an optimal biometric authentication system, both the false rejection rate (FRR) and false acceptance rate (FAR) should ideally be below 0.1 percent. Several factors influence the selection of a biometric trait, including distinctiveness, permanence, collectability, performance, susceptibility to spoofing, processing speed, and cost efficiency [16].

Table 2.2 provides a comparative overview of various biometric traits across key attributes such as acceptability, performance, universality, distinctiveness, permanence, and collectability. The values are categorized as high (H), medium (M), or low (L).

Biometric Trait	Important Features and Characteristics	Applications
Fingerprint	• Secure, reliable, highly accurate, and inexpensive • Fast matching process • Low memory usage • Vulnerable to cuts, scars, dust, and dirt • Requires physical contact with the system	• Driver Authentication • Forensics • License/visa authentication • Access control
Ear	• Stable, low computational complexity • Fast identification • Errors due to hairstyle, jewelry, hats, etc.	• Law Enforcement • Forensics Surveillance
Hand Geometry	• Durable and compact • Low processing needs • Prone to errors from injuries or jewelry	• Military access control
Palm Print	• Wide range of features • High reliability and permanence • Good recognition even with low-resolution scans	• Personal Identification • Medical diagnosis • Blood relation ID • Athlete selection
Hand Vein	• Noninvasive and highly secure • Difficult to forge • Unique, affected by temperature	• Door security • Banking • Login authentication • Transportation systems • Hospitals, schools
Finger Vein	• Highly secure and difficult to replicate • Non-contact and hygienic • Unaffected by external skin conditions • Requires special infrared sensors	• Access control • ATM and banking security • Identity verification in high-security facilities
Voice	• Easy to implement • Cost-effective • Sensitive to noise and throat conditions	• Web-based transactions • Health services
Signature	• Low false acceptance rate • High accuracy • Easy retrieval	• Banking system
Iris	• High scalability and accuracy • Fast processing but costly • No physical contact, requires user cooperation • Sensitive to diseases and distance	• Access control • National security (air, land)

Table 2.1: Comparison of Biometric Modalities

For palmprint and finger-vein biometrics, survey results indicate that palmprint has medium universality, high permanence, medium collectability, high performance, and medium acceptability, while finger vein demonstrates medium universality, high permanence, medium collectability, high performance, and high acceptability

2.3 Overview of a Biometric System's Architecture and Operating Modes

A general biometric device consists of an integrated functionality of five primary components or modules. as shown in Figure 2.1 below. The components of biometric systems may vary between different implementations. The initial five steps include data

Biometric	Universal	Permanence	Collectability	Performance	Acceptability
Fingerprint	M	H	M	H	M
Face	H	M	H	L	H
Ear	H	H	M	M	H
Hand Geometry	M	M	H	M	M
Palm Print	M	H	M	H	M
Finger Vein	M	H	M	H	H
Voice	M	L	H	L	H
Signature	L	L	H	L	H
Iris	H	H	M	M	L

Table 2.2: Comparative Evaluation of Biometric traits

acquisition through sensor collection, signal processing, data storage (also referred to as template storage), the implementation of comparison algorithms for matching, and decision-making.

1-Data Acquisition: This subsystem is designed to capture a sample of an individual's biometric characteristic, such as an image or signal. The sample, referred to as raw biometric data, is uncompressed and collected through a sensor [17]. This stage, where the user interacts with the biometric system, is commonly known as biometric presentation [18].

2-Preprocessing and Feature Extraction: To create a digital representation known as a biometric template or reference, which captures the individual's unique characteristics and remains relatively stable across multiple samples, this system component extracts features from the biometric sample [[19], [20]]. This stage involves sample enhancement, feature extraction, and quality assessment (segmentation). Quality control tests (segmentation and feature extraction) yield a quality score, which indicates the sample's suitability for feature extraction [17].

3-Data Storage: In the biometrics field, this part of the system is often referred to as the reference because it contains the biometric template for usage in upcoming procedures [19]. During the enrollment process, certain templates are created and saved in the enrollment database.

4-Comparison (Matching) Algorithm: This algorithm is a comparison (matching) algorithm. The third step involves comparing each freshly created sample template on this section of the system with one or more reference templates, via a comparison algorithm. The number of reference templates used is determined by the application. The act of comparison results in the generation of a comparison score, which is also referred to as a similarity score. This score is used to determine the degree of similarity that exists between the templates [17]. Following this, the comparison score is incorporated into a module that is used for decision-making.

5-Decision : The score from the comparison component serves as input for this system segment to be compared against the threshold for verification or identification tasks. The biometric system administrator generally determines the threshold, a predefined value. The compared templates are considered matched if the comparator's

score above the threshold; if it falls below, the templates are not matched [18]. According to reference [21], thresholds are essential for system security. Systems can be highly secure or entirely vulnerable, contingent upon the configuration of their threshold settings.

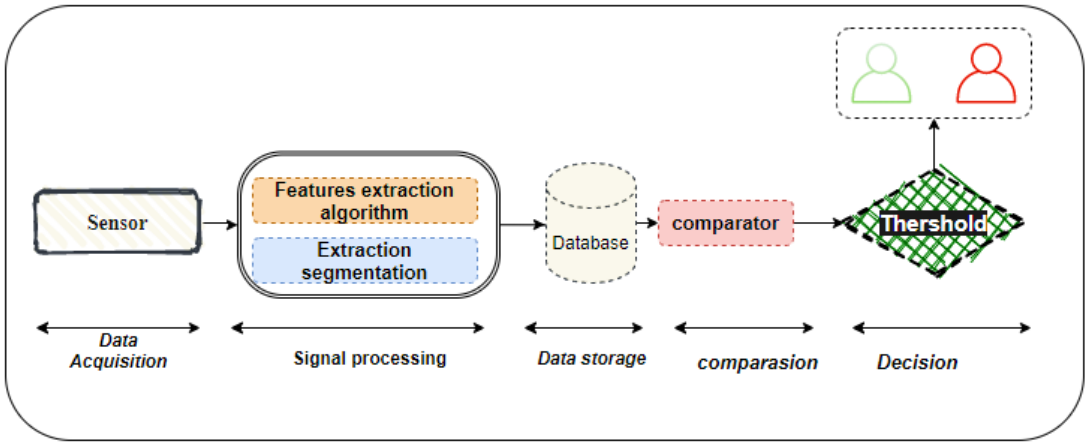


Figure 2.1: Components of Biometric Systems

Biometric systems can operate in three modes: **Enrollment, Verification, and Identification**. Subsequently, the figures will illustrate an example of a biometric system using fingerprints as the modality. [22]

1-Enrollment:

Enrollment see figure 2.2 might be considered the initial stage of any biometric system. When a user is registered in the system for the first time, this is the stage in which one or more biometric modalities are recorded and saved in a database. Additionally, this is the stage in which the user is verified. There is a possibility that this registration will also involve the incorporation of biography material into the database.

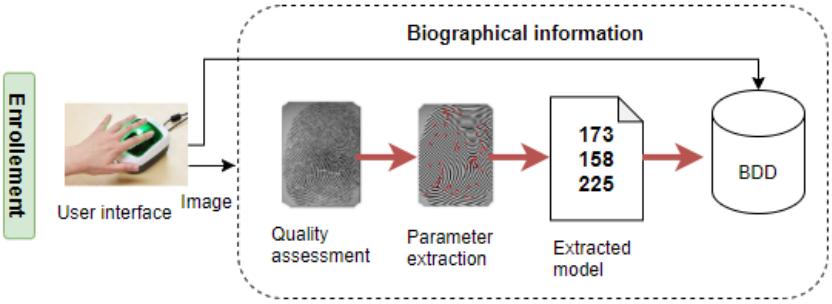


Figure 2.2: Enrollment of a person in a biometric system.

2- Verification:

When a biometric system operates in authentication mode (Fig. 2.3), the user asserts their identity, and the system verifies whether this assertion is valid. For example,

if a user (X) wishes to withdraw money from an ATM, they enter their personal identification number (PIN) and provide a biometric sample. The system then acquires the biometric data and performs a 1:1 match, comparing the input only with the stored template corresponding to user X.

If the similarity score between the user's biometric sample and the registered template exceeds a predefined threshold, the assertion is validated, and the user is authenticated. Otherwise, the system rejects the assertion, classifying the user as an impostor. In summary, a biometric system operating in verification mode answers the question: "Am I really X?"

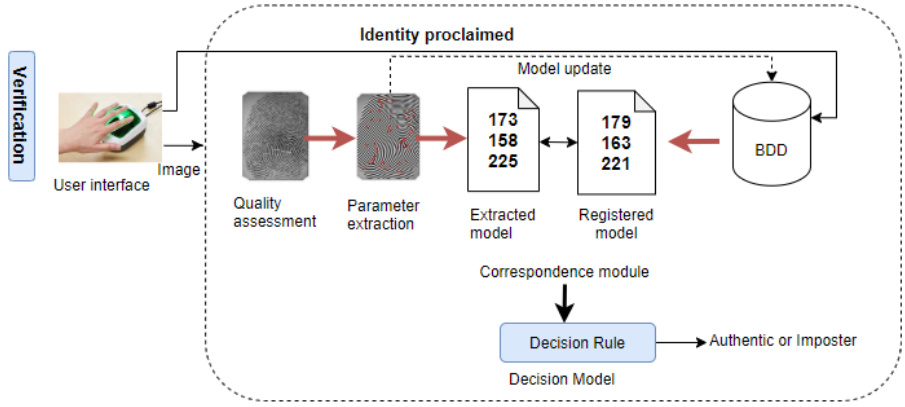


Figure 2.3: Authentication of an individual in a biometric system.

3- Identification:

In a biometric system functioning in identification mode (Fig. 2.4), the user does not explicitly claim an identity. Instead, the system performs a 1:N match, comparing the biometric sample against all stored templates in the database. The goal is to determine whether the user is already enrolled in the system and, if so, to identify them.

The system assigns a similarity score to each comparison and identifies the individual whose stored template exhibits the highest similarity to the input biometric sample. If the maximum similarity score falls below a predefined security threshold, the system rejects the individual, indicating that they are not among the enrolled users. Otherwise, the system accepts the individual and assigns them the identity corresponding to the highest similarity score.

2.4 Evaluating the Effectiveness of a Biometric System

Evaluating the performance of a biometric system is crucial for determining its effectiveness and suitability for various applications. Here are the key metrics and methods used:

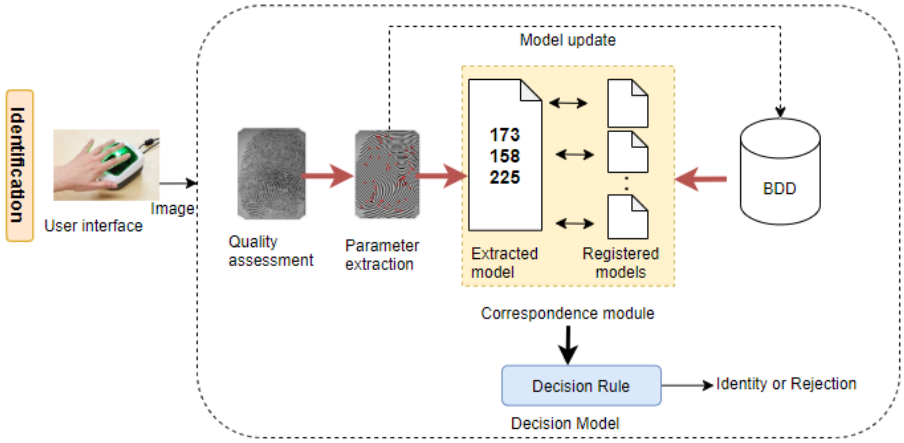


Figure 2.4: Identification of an individual in a biometric system.

The initial criterion is referred to as the False Reject Rate (FRR). This rate indicates the proportion of individuals who ought to be acknowledged but are excluded by the system [22].

The second criterion is the False Acceptance Rate (FAR). This rate indicates the proportion of individuals who ought not to be identified yet are nonetheless acknowledged by the system [22].

The third criterion is referred to as the Equal Error Rate (EER). This rate is derived from the initial two criteria and serves as a standard performance measurement metric. This point represents the equilibrium where the False Rejection Rate (FRR) equals the False Acceptance Rate (FAR), indicating the optimal balance between false rejections and false acceptances. A lower EER signifies superior overall system performance [22].

Figure 2.5 as mentioned in [22] illustrates the relationship between FRR and FAR, where the threshold determines the balance between false acceptances and false rejections.

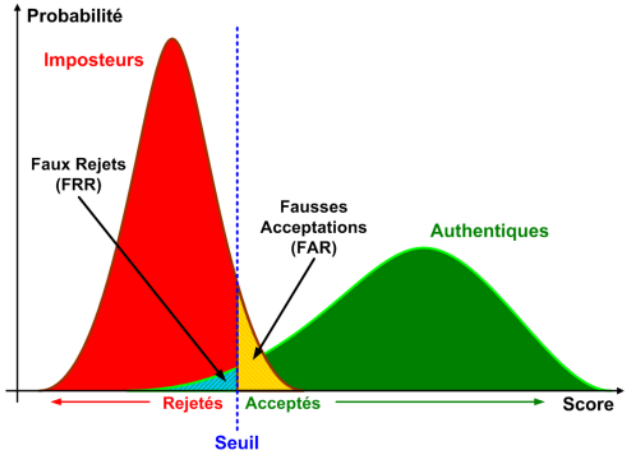


Figure 2.5: Illustration of FRR and FAR.

Depending on the nature of the biometric system (authentication or identification), there are two ways to measure its performance:

When the system operates in authentication mode, a Receiver Operating Characteristic (ROC) curve is used. The ROC curve plots the false reject rate (FRR) against the false accept rate (FAR) [23]. The closer the curve approaches the axes, the more efficient the system is, meaning it has a high overall recognition rate.

The Area Under the Curve (AUC) is a crucial metric derived from the ROC curve, representing the overall ability of the biometric system to discriminate between genuine and impostor attempts. A higher AUC value (closer to 1) indicates a more reliable system, whereas a value near 0.5 suggests poor discriminatory ability. The AUC provides a comprehensive measure of classifier performance across different threshold settings.

In contrast, for a system used in identification mode, a Cumulative Match Characteristic (CMC) curve is used. The CMC curve shows the percentage of people recognized as a function of a variable called rank [24,25]. Besides, the Rank-one (R1) and the Rank-five (R5) recognition rates are used to evaluate the identification rate returned by a set of inputs as the best first and the best five matches, respectively.

Figure 2.6 presents an example of a CMC curve, demonstrating the recognition rates at different ranks.

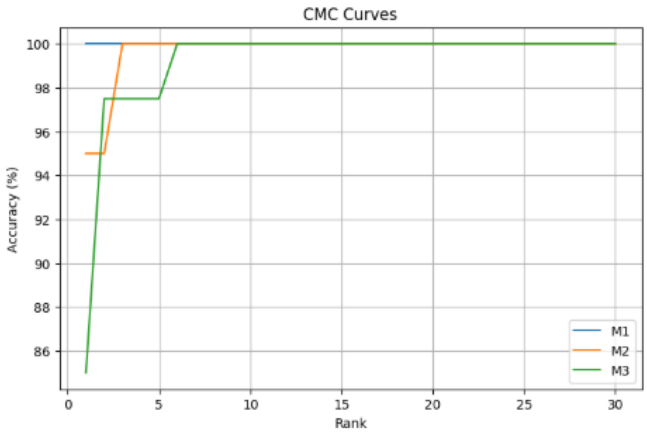


Figure 2.6: Example of a CMC curve .

Additionally, classification metrics such as precision (PR), recall (RC), accuracy, and F1-score (F1) are commonly used, as described in Equations 2.1, 2.2, 2.3, and 2.4. These metrics are derived from the confusion matrix [26], which provides a summary of the predictive results in a classification problem. Correct and incorrect predictions are categorized by class and summarized in a Fig 2.7, as shown below:

- **True Negative (TN):** The number of correct predictions for the negative class.
- **False Positive (FP):** The number of incorrect predictions where the actual class is negative but predicted as positive.

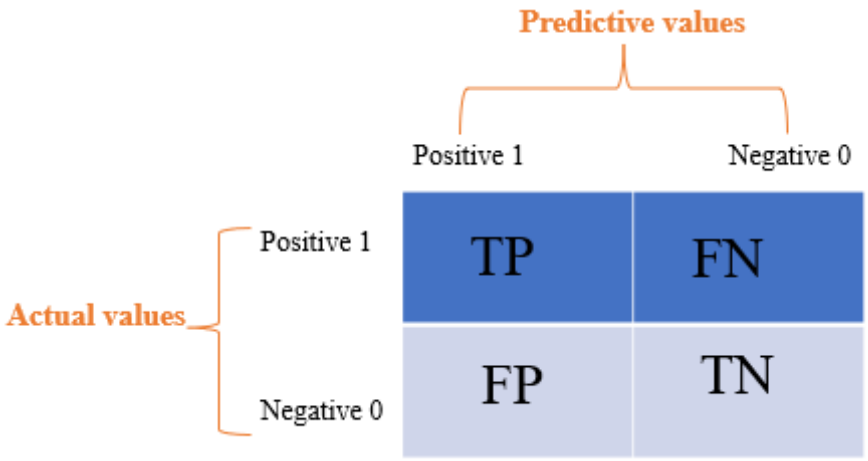


Figure 2.7: The confusion matrix.

- **False Negative (FN):** The number of incorrect predictions where the actual class is positive but predicted as negative.
- **True Positive (TP):** The number of correct predictions for the positive class.

Using these values, the classification metrics are calculated as follows:

- **Precision (Pr):**

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2.1}$$

- **Recall (Rec):**

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{2.2}$$

- **Accuracy:**

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{2.3}$$

- **F1-Score (F1):**

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{2.4}$$

The confusion matrix, along with these metrics, provides a comprehensive evaluation of the model's performance, highlighting both the correct and incorrect predictions across different classes.

2.5 Features Extraction Methods

Feature extraction is a critical step in biometrics systems. It involves transforming raw data into a set of features that can be used for tasks like classification. Among the various feature extraction methods handcrafted based methods and deep learning-based methods as illustrated in figure 2.8:

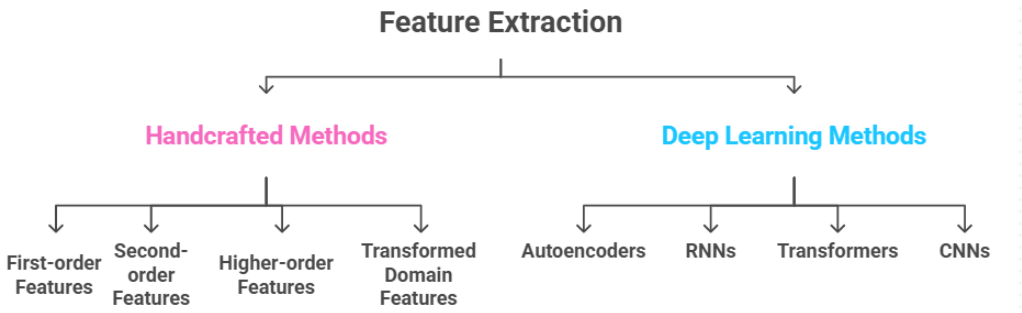


Figure 2.8: Features extraction methods .

2.5.1 Handcrafted Based Features Extraction

Handcrafted features play a critical role in image processing, as they summarize the intensity, texture, and spatial relationships. These features are grouped into different classes based on their mathematical and statistical properties. This work [27] provides a detailed classification of these features, along with their corresponding types and escriptions. This categorization helps illustrate how these features contribute to various applications, such as texture analysis, segmentation, and classification.

• **First-order features:** These features summarize the intensity distribution of pixel values using statistical measures like mean, variance [28], skewness, and kurtosis [29].

• **Second-order features:** These features, such as the Gray-level Co-occurrence Matrix (GLCM) [30], capture spatial relationships between pixel pairs, providing insights into texture properties like contrast, correlation, and homogeneity.

• **Higher-order features:** These features analyze complex spatial relationships and dependencies among pixels. For example, the Gray-level Size Zone Matrix (GLSZM) and Gray-level Distance Zone Matrix (GLDZM) measure homogeneity and clustering of pixels with similar intensities [31], while the Neighboring Gray-level Dependence Matrix (NGLDM) captures rotationally invariant properties [32] .The Gray-level Run-Length Matrix (GLRLM) [33–35] quantifies runs of consecutive pixels with the same intensity, highlighting texture uniformity and directionality.

• **Transformed domain features:** These features utilize transformations to analyze textures and patterns in images. Histograms of oriented gradients (HOG): considering the occurrences of gradient orientation in localized portions of an image [36]. Local Binary Patterns (LBP) encode texture by assigning binary values based on intensity comparisons with neighbors [37]. Gabor Transform [38] extracts texture information in specific orientations and scales, while Wavelet Transform decomposes images into multi-resolution subbands to capture both frequency and spatial details [39].

2.5.2 Deep Learning-Based Feature Extraction

Feature extraction in deep learning refers to the automated process by which neural networks learn to identify and represent the most relevant characteristics or patterns from raw input data. Unlike traditional approaches that relied on manually engineered features (e.g., edges, corners, textures), deep learning models discover these features in an end-to-end fashion during training. This capability is demonstrated across a variety of architectures tailored to different data types and tasks.

Several deep learning methods can be employed for feature extraction:

- **Autoencoders**

Autoencoders are unsupervised models designed to compress data into a lower-dimensional latent space and then reconstruct the original input. The latent representations capture the essential features of the data and can be used for tasks such as dimensionality reduction and anomaly detection [40].

- **Recurrent Neural Networks (RNNs) and Their Variants**

RNNs, including Long Short-Term Memory (LSTM) networks and Gated Recurrent Units (GRUs), excel at extracting temporal features from sequential data such as time-series or text. In applications like video analysis, RNNs can complement spatial feature extractors by capturing the temporal evolution of features [41].

- **Transformers**

Originally developed for natural language processing, transformers use self-attention mechanisms to capture dependencies across input sequences. Recently, Vision Transformers (ViTs) have adapted this methodology for image processing, treating an image as a sequence of patches to learn global relationships without relying solely on localized operations [42].

- **Convolutional Neural Networks (CNNs)**

CNNs have become the cornerstone for feature extraction in image classification. Their layered architecture automatically learns hierarchical representations from raw pixel data. Early layers capture low-level features (e.g., edges and textures), intermediate layers build up mid-level features (e.g., shapes and motifs), and deeper layers form high-level, abstract concepts (e.g., objects or scenes) [43, 44].

A CNN architecture typically comprises several key operations, including convolution, batch normalization, activation functions (e.g., ReLU), pooling, fully connected layers, and dropout, . Each component plays a critical role in the network's ability to learn and classify images effectively. Figure 2.9 illustrates a simplified CNN architecture for classification tasks.

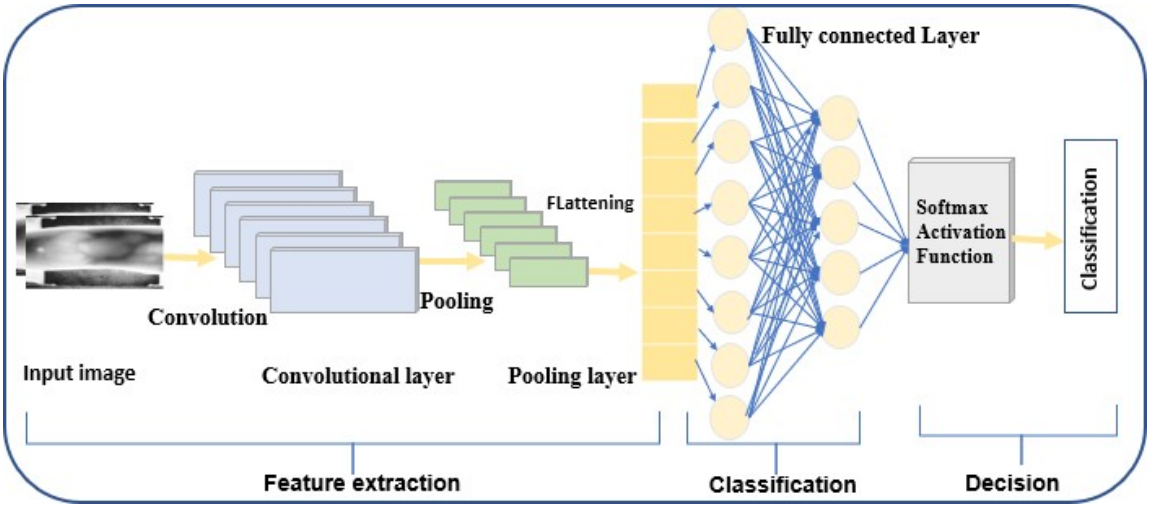


Figure 2.9: Structure of a Convolutional Neural Network.

- Convolution Layers

The convolution layer is the foundational block of a CNN and is responsible for detecting features in an image. It operates by sliding a filter (or kernel) across the input image at a specified stride. Filters can be initialized randomly, based on prior knowledge, or through transfer learning. During training, these filters are updated to optimize feature detection. The output volume of the convolution layer is influenced by several hyperparameters:

- **Filter size:** Commonly, 3×3 filters are used, though larger filters like 5×5 or 7×7 may be employed depending on the application. Filters are 3D, with a depth matching the input volume.
- **Padding:** Zero-valued pixels are added around the image borders to control the spatial dimensions of the output.
- **Stride:** The stride determines the step size of the filter as it moves across the image. A stride of 2 means the filter shifts two pixels at a time.
- Batch Normalization

Batch normalization is a technique used to normalize the inputs or activations from the previous layer, ensuring zero mean and unit variance. This process accelerates training, reduces generalization errors, and improves the stability of the network [45].

- Activation Functions: ReLU

The Rectified Linear Unit (ReLU) is a non-linear activation function that outputs the input directly if it is positive; otherwise, it outputs zero. ReLU has become the standard activation function in neural networks due to its ability to speed up training and improve model performance [46].

- Pooling Layers

Pooling layers reduce the dimensionality of feature maps, introducing translation invariance and decreasing the number of learnable parameters. Common types of pooling include:

- **Max pooling:** Outputs the maximum value within each patch of the feature map.
- **Average pooling:** Computes the average value within each patch. Global average pooling reduces the feature map to a 1×1 array while preserving depth, enabling CNNs to handle variable-sized inputs.

- Fully Connected Layers

Fully connected (FC) layers flatten the features extracted by the final convolution layer into a 1D array and connect them to the network's outputs. Each input is linked to each output through learnable weights, enabling the network to make predictions based on the learned features.

- Dropout

Dropout is a regularization technique that randomly deactivates neurons during training. This prevents overfitting by ensuring that neurons do not become overly reliant on specific features [47].

Table 2.3 provides the components of a CNN and their respective roles.

Component	Description
Convolution Layer	Detects image features using sliding filters.
Batch Normalization	Normalizes activations to accelerate training and minimize errors.
ReLU	Introduces non-linearity; outputs positive inputs or zero.
Pooling Layer	Reduces dimensionality; max or average pooling variants.
Fully Connected Layer	Flattens features and connects to outputs through learnable weights.
Dropout	Randomly deactivates neurons to prevent overfitting.
SoftMax Activation Function	Converts outputs into a probability distribution.

Table 2.3: Components of a CNN

2.6 Classification

After the feature extraction stage, the next step in the image classification pipeline is classification, where a classifier assigns the extracted feature representations to predefined categories. Various classification methods can be used, ranging from deep learning-based approaches to traditional machine learning classifiers.

One of the most commonly used classifiers in deep learning architectures is the Softmax classifier. It is typically placed in the final layer of a neural network and converts the raw output logits into a probability distribution over the target classes. The class with the highest probability is selected as the predicted label, making Softmax ideal for multi-class classification tasks [43].

Alternatively, traditional machine learning classifiers can be applied to the deep features extracted from models like CNNs, autoencoders, or transformers. k-Nearest Neighbors (k-NN) is a simple, non-parametric method that classifies a new instance based on the majority label of its nearest neighbors in the feature space. While effective for small datasets, k-NN suffers from high computational costs as dataset size increases [48].

Another powerful classifier is the Support Vector Machine (SVM), which finds an optimal hyperplane to separate different classes in high-dimensional feature spaces. When combined with deep learning-based feature extraction, SVM can effectively handle complex classification tasks, especially with non-linear kernels such as the Radial Basis Function (RBF) [49, 50].

Ensemble learning methods like Random Forest (RF) and Decision Trees (DT) are also widely used for classification tasks. Random Forest is an ensemble of decision trees that aggregates multiple predictions to enhance accuracy and reduce overfitting. It is particularly useful for handling deep features extracted from CNNs or other deep networks [51]. On the other hand, Decision Trees split feature spaces based on information gain or entropy, making them easy to interpret but prone to overfitting unless regularized [52].

By selecting an appropriate classifier based on dataset size, feature complexity, and computational constraints, classification performance can be significantly optimized for various image recognition tasks.

2.6.1 Unimodal VS Multimodal Biometric System

Biometric recognition systems can be classified into two types: unimodal and multimodal. Unimodal systems use a single biometric trait for user recognition. While they are reliable and have advantages over traditional methods, they face limitations such as noise in the sensed data, non-universality issues, vulnerability to spoofing attacks, and intra-class and inter-class similarity problems [3].

Multimodal biometric systems, in contrast, utilize multiple biometric traits for user recognition [4]. These systems are widely applied in real-world scenarios due to their ability to overcome the limitations of unimodal systems [3]. Different traits in multimodal biometric systems can be fused using information from one of the biometric system's modules. The four types of fusion include sensor-level fusion, feature-level

fusion, score-level fusion, and decision-level fusion.

Multimodal biometric systems integrate multiple sources of biometric data to enhance recognition accuracy, security, and robustness. They combine different modalities or multiple samples of the same modality to address challenges such as sensor noise, spoofing, and variations in data acquisition [53]. These systems can be classified into five main types [54], as summarized in Figure 2.10.

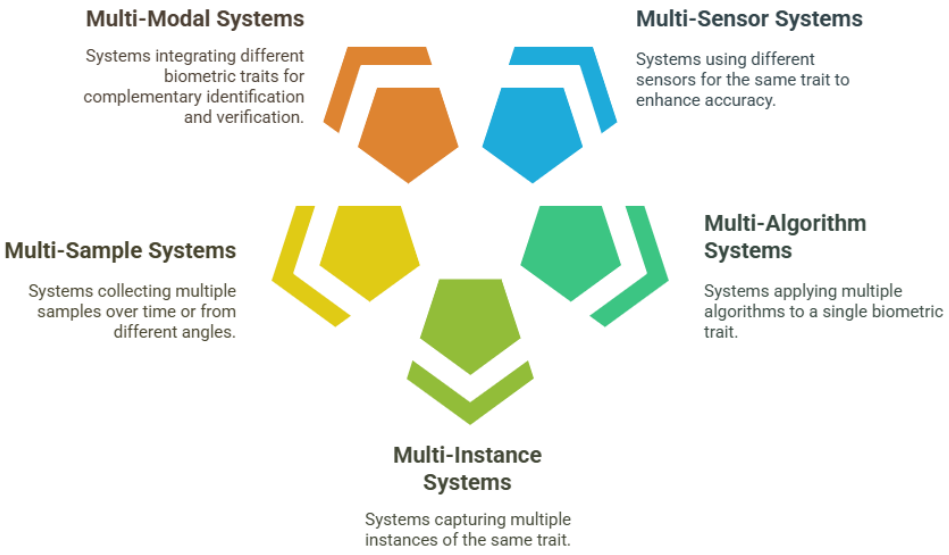


Figure 2.10: Multimodal Biometric Systems Types.

- **Multi-Sensor Systems:** These use different sensors to capture the same biometric trait, such as an infrared and visible-light camera for face recognition. This approach enhances recognition accuracy by capturing additional information.
- **Multi-Algorithm Systems:** In these systems, a single biometric trait is processed using multiple algorithms to improve performance. For example, different fingerprint matching algorithms can be applied to enhance template comparison.
- **Multi-Instance Systems:** This category captures multiple instances of the same biometric trait, such as using both irises for iris recognition or multiple fingerprints from different fingers.
- **Multi-Sample Systems:** Multiple samples of the same biometric trait are collected at different times or from different perspectives, such as various facial angles for face recognition. This helps address variations due to environmental conditions or user positioning.

- **Multi-Modal Systems:** These systems integrate completely different biometric traits, such as a combination of face and fingerprint recognition, to provide complementary identification and verification.

2.6.1.1 Types of Fusion Levels in Multimodal Biometric Systems

In multimodal biometric systems, different biometric traits can be fused at various stages of processing. The fusion process enhances accuracy and resilience to variations in biometric data. The four primary levels of fusion [55] are:

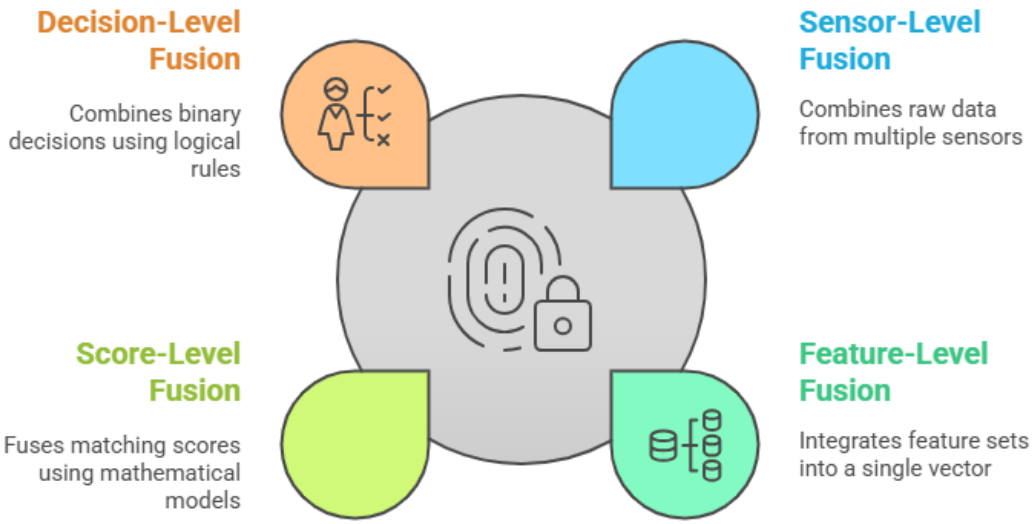


Figure 2.11: Levels of fusion.

- **Sensor-Level Fusion:** This is the earliest stage of fusion, where raw biometric data from multiple sensors are combined before feature extraction. This method retains the most information but requires compatibility in data formats.
- **Feature-Level Fusion:** In this approach, feature sets extracted from different biometric modalities are combined into a single feature vector. Since this method integrates rich information from multiple sources, it can significantly improve recognition accuracy, but it requires proper feature normalization and dimensionality reduction techniques.
- **Score-Level Fusion:** Each biometric trait generates a matching score based on similarity with stored templates. These scores are then fused using mathematical models such as weighted sums, machine learning, or statistical approaches

to produce a final score. Score-level fusion is widely used due to its balance between computational efficiency and accuracy.

- **Decision-Level Fusion:** At this highest level, each biometric trait independently provides a classification result (e.g., accept or reject), and the final decision is made based on majority voting or logical rules. While easy to implement, decision-level fusion is less flexible because it operates on binary outputs rather than rich feature data .

By integrating multiple modalities and fusion techniques, multimodal biometric systems achieve superior performance compared to unimodal systems, making them suitable for high-security applications such as border control, financial transactions, and forensic identification.

2.7 Conclusion

This chapter outlined biometric systems, covering concepts, architecture, performance metrics, feature extraction, classification methods, and a comparison between unimodal and multimodal biometric systems through fusion levels: sensor, feature, score, and decision . In the following chapter, we will implement finger vein and palmprint biometrics, analyzing their unique traits and extraction,classification techniques to enhance accuracy, security, and robustness.

Chapter 3

One VS Multi-instance Biometrics System using Finger Vein and Palmprint

3.1 Introduction

This chapter focuses on finger vein and palmprint recognition, two secure and unique biometric methods. While effective, these systems face challenges such as lighting changes, image quality issues, and variations in hand position. To improve image clarity, we apply tailored enhancement techniques: histogram equalization for finger veins and CLAHE for palmprints.

We examine two learning strategies: one-instance (S1) and multiple-instance (S2). Single-instance uses one image per person, which keeps processing simple but may be affected by noise such as dirt or injuries, in that image. Multiple-instance addresses this limitation by combining information from several images, making the system more robust to such issues and capturing greater feature variability. Feature extraction is done using transfer learning with pre-trained deep models (VGG16, VGG19, and MobileNetV2), with fine-tuning applied to adapt these models to biometric data. The extracted features are classified using SVM, KNN, and Random Forest. This chapter includes a review of related work, the proposed methods, experiments, and key findings. The chapter is organized as follows: Section 2 reviews related work, Section 3 details the proposed methodologies, Section 4 presents the experimental setup and results, and Section 5 concludes with key findings.

3.2 Literature Survey

In the field of finger vein recognition, researchers have made significant strides in addressing image quality issues and enhancing pattern-generation techniques. Miura et al. pioneered two influential methods: a line-tracking approach for extracting vein patterns from unclear images [56], and a technique for calculating local maximum curvatures in cross-sectional profiles to consistently extract vein centerlines [57]. Qin et al. further improved pattern extraction accuracy by implementing a region growth method [58]. Recent advancements have largely focused on deep learning techniques, which have substantially improved verification performance [59, 60]. Tang et al. proposed a CNN-based matching approach with supervised discrete hashing, while He et al. developed a multi-layer classifier using backpropagation neural networks enhanced by principal component analysis [61]. Zhao et al. introduced the Intensity Orientation Vector (IOV) and a novel learning scheme called Semantic Similarity Preserved Discrete Binary Feature Learning (SSP-DBFL) [62]. Fang et al. designed a lightweight Siamese network with self-attention mechanisms [63], and Boucherit et al. proposed a Merge Convolutional Neural Network combining multiple CNNs [64]. Shaheed et al. explored Curvelet transform and multi-scale augmentation for image enhancement, followed by feature extraction using depth-wise separable convolution layers [65]. These diverse approaches have collectively advanced the accuracy, efficiency, and robustness of finger vein recognition systems, establishing them as a reliable biometric authentication method.

In the other hand Palmprint is a biometric modality that has recently received heightened interest. Palmprints include details features, geometry-based features, delta points, principal lines, and wrinkles [66]. Each element of a palmprint displays unique features, including texture, ridges, lines, and Wrinkles. The temporal stability of palmprint Wrinkles enhances their extraction [67]. The collection of palmprints requires specialized devices, making the process more complex than that of other biometric modalities, including fingerprints, iris scans, and facial recognition.Investigations in palmprint recognition have examined various handcrafted features, including basic component analysis (PCA), independent component analysis (ICA), The Fourier transform, wavelet transform, line combination, and deep scattering features [68–70].Samai et al.introduced a deep learning model for recognizing both 2D and 3D palmprints [71].Zhong et al. presented a palmprint recognition algorithm utilizing a Siamese network [72]. Two VGG-16 networks with shared parameters were employed to extract features from two input palmprint images. A supplementary network was established to directly assess the similarity between the two palmprints utilizing their convolutional features. Izadpanahkakhk et al. [73] introduced a transfer learning method for palmprint verification that concurrently extracts regions of inter-

est and features from the images. A pre-trained convolutional network was employed alongside SVM for predictive analysis.

3.3 Overview of The Proposed

This section provides a comprehensive overview of the proposed method's workflow. It begins with a description of the datasets used, followed by a detailed explanation of the preprocessing steps applied to the finger vein and palmprint images. Next, the feature extraction process is carried out using three CNN models—MobileNetV2, VGG19, and VGG16 in two distinct approaches: transfer learning without preprocessing and transfer learning with preprocessing. Subsequently, the features extracted from the best-performing model for the left and right finger vein and palmprint images are used to create a highly relevant feature set. Finally, the classification step is performed using multiple machine learning algorithms, including Support Vector Machine (SVM), Random Forest (RF), and K-Nearest Neighbors (KNN). The proposed recognition framework is illustrated in Fig. ref Fig1.

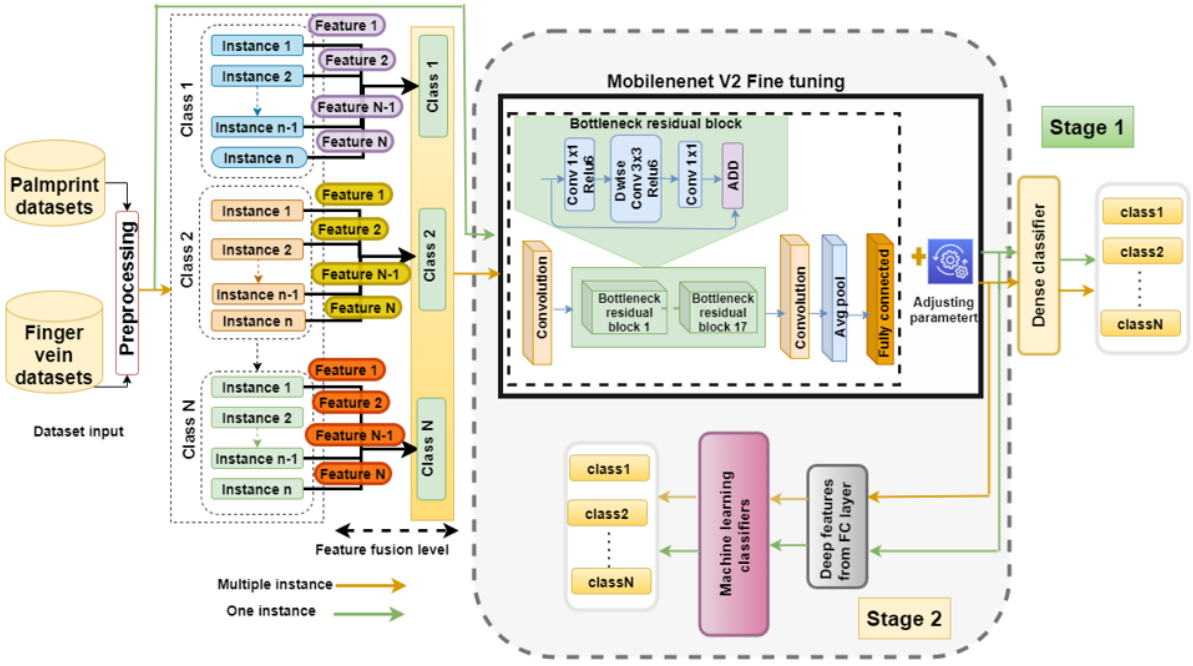


Figure 3.1: Proposed System.

3.3.1 Data Description

The effectiveness of the proposed method is evaluated using four publicly available datasets, comprising two finger vein datasets and two palmprint datasets. These

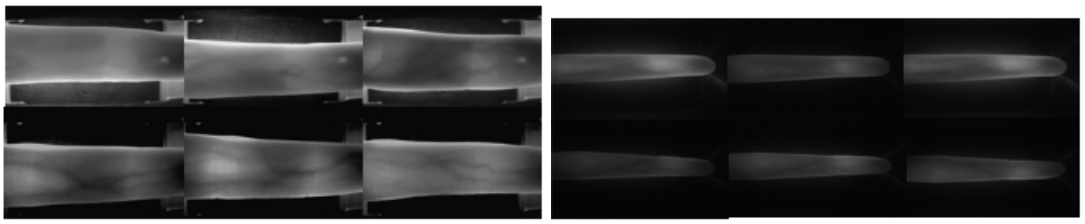
datasets were selected to ensure a comprehensive assessment of the method's performance across different biometric modalities. The finger vein datasets include the FV-USM dataset from Universiti Sains Malaysia [74] and the SDUMLA dataset [75] from Shandong University. For palmprint recognition, the CASIA dataset [76] from the Chinese Academy of Sciences and the IITD dataset [77] from the Indian Institute of Technology Delhi are utilized. Below, each dataset is described in detail.

3.3.1.1 FV-USM Dataset

The Universiti Sains Malaysia Finger Vein (FV-USM) dataset [74], consists of 5,904 finger vein images collected from 123 subjects, including 83 males and 40 females. Each subject contributed images from four fingers: the left index, left middle, right index, and right middle fingers, resulting in a total of 492 classes. Each finger was imaged six times across two sessions, with each session containing 2,952 images. This dataset provides a robust foundation for evaluating finger vein recognition methods.

3.3.1.2 SDUMLA Dataset

The Shandong University Machine Learning and Applications Homologous Multimodal Traits (SDUMLA-HMT) dataset [75], comprises 636 finger vein images collected from 106 individuals. The images were captured six times per finger in grayscale BMP format at a resolution of 320x240 pixels. Data was collected from both hands, covering the left index (LI), left middle (LM), left ring (LR), right index (RI), right middle (RM), and right ring (RR) fingers. This dataset offers a diverse set of samples for testing the proposed method.



(a) samples from the SDUMLA FV. **(b)** samples from the USM FV.

Figure 3.2: Sample images from the SDUMLA and USM finger vein datasets.

3.3.1.3 CASIA Dataset

The Chinese Academy of Sciences Institute of Automation (CASIA) Palmprint database [76] contains 5,502 images collected from 312 individuals. Both the Left Palmprint (LP)and Right Palmprint (RP) of each subject were systematically captured.

3.3.1.4 IITD Dataset

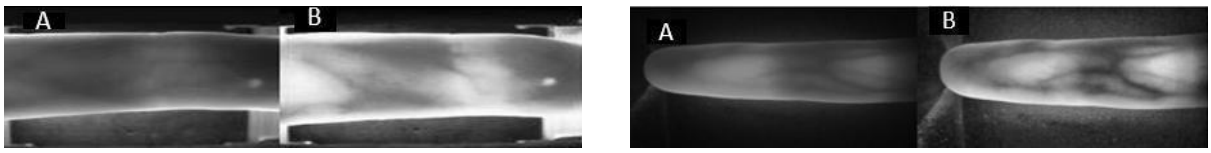
The Indian Institute of Technology Delhi (IITD) palmprint database [77] provides a valuable resource for palmprint recognition research, featuring 2,601 images from 460 unique palms. The dataset includes contributions from 230 individuals aged 14 to 15 years. Each participant provided five to six samples from both the left and right palms. The region of interest (ROI) images in the dataset are standardized to a size of 150 x 150 pixels.

(a) samples from the IITD . **(b)** samples from the CASIA.

Figure 3.3: Sample images from the IITD and CASIA palmprint datasets.

3.3.2 Histogram Equalization

The primary advantage of histogram equalization is the enhancement of visual quality, facilitating recognition. This technique enhances the contrast in photographs. It achieves this by uniformly dispersing the predominant intensity values, hence enhancing contrast in regions with poor local contrast. Figure 3.4 (a and b) illustrates instances from the used datasets before to and subsequent to histogram equalization [78].



(a) Sample from SDUMLA dataset. **(b)** Sample from FV-USM dataset.

Figure 3.4: Histogram equalization influence: A) before and B) after.

3.3.3 Contrast Limited Adaptive Histogram Equalization

CLAHE is a line contrast enhancement pre-processing tool. CLAHE involves conducting histogram equalization on non-overlapping picture sub-areas and applying interpolation to repair irregularities across boundaries [79] Figure 3.5 (a and b) presents examples from the used datasets before and after CLAH.

(a) Sample from IITD dataset **(b)** Sample from CASIA dataset

Figure 3.5: CLAHE influence using both datasets: (a) before, (b) after.

3.3.4 Features Extraction and Classification

Transfer learning is an effective technique in deep learning that leverages the capabilities of pre-trained models to enhance performance on new related tasks [80]. When applying transfer learning in convolutional neural networks, multiple strategies can be used, such as fine-tuning the entire network or fine-tuning only its upper layers [81].

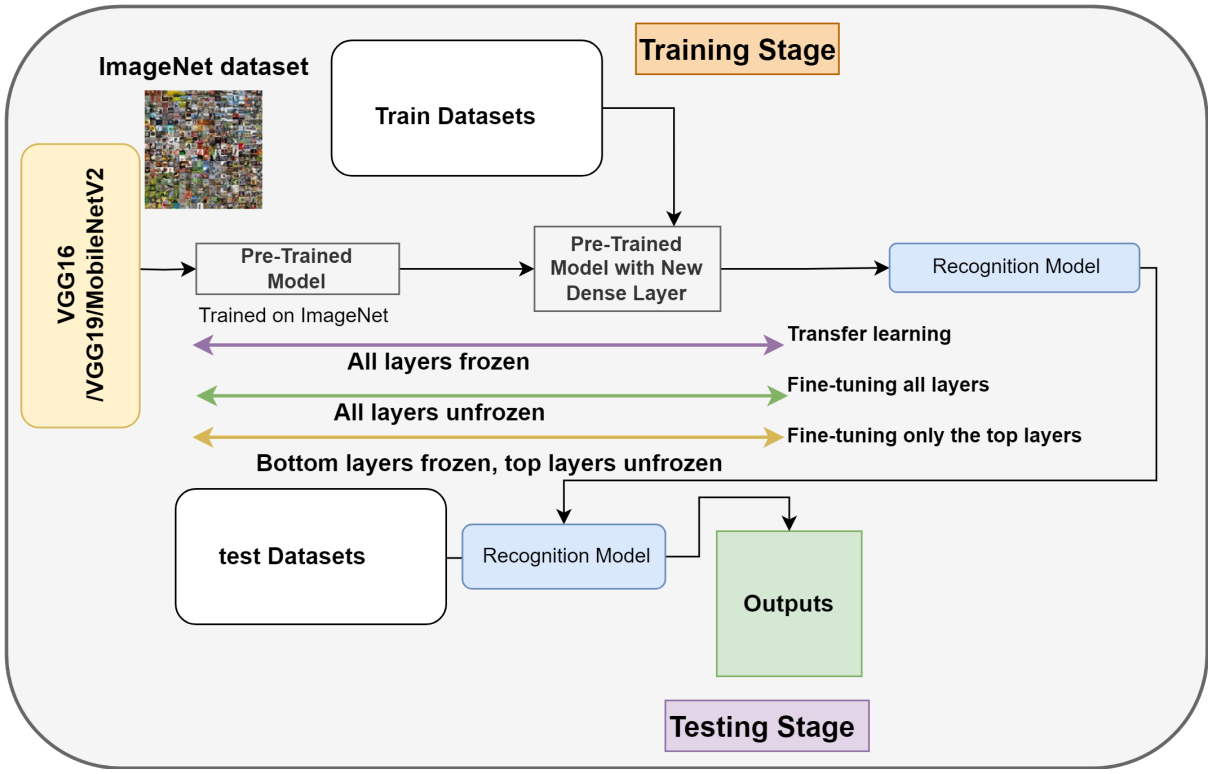


Figure 3.6: The process of transfer learning and the different fine-tuning strategies.

- **Transfer Learning**: This technique uses a CNN pre-trained on a large dataset as an initial base for a new task. The pre-trained weights provide an initial set of

parameters that capture general features, which can enhance performance and reduce training time for the new task [82].

- **Fine-tuning all layers**: Known as full fine-tuning, this approach allows the update of all layers in the pre-trained model while training on the new task. This technique enables the entire network to adapt to the new dataset. However, it requires more computational resources and may result in overfitting when applied to small datasets [83].

- **Fine-tuning top layers**: This strategy involves freezing the network's bottom layers, which capture more general features, while only adjusting the weights of the upper layers. It is frequently used with limited data, as it decreases the risk of overfitting [81].

3.3.4.1 CNN architectures

To implement transfer learning in our experiments, we used three state-of-the-art CNN architectures:

- **MobileNetV2**: An efficient model that uses depth-wise separable convolutions, resulting in fewer parameters and lower computing costs than typical convolutional layers. It includes the concept of inverted residuals and linear bottlenecks, which increase the network's flexibility while preserving a lightweight topology, making it suitable for mobile and embedded applications [84].

- **VGG16 and VGG19**: These models are defined by their simplicity and depth. Both use a series of convolutional blocks that include 3x3 convolutional layers, followed by ReLU activation and max-pooling layers. VGG16 includes 16 weight layers, comprising 13 convolutional layers and three fully connected layers. VGG19 consists of 19 weight layers, including 16 convolutional layers and 3 fully connected layers [85].

In this work, we use three pre-trained models: VGG16, VGG19, and MobileNetV2 for feature extraction and classification in two modalities: finger vein and palmprint images. Initially, we use transfer learning with these models to extract features, using a SoftMax classifier to classify the extracted features into their respective categories. MobileNetV2 demonstrates the best performance among the three models and is selected for further refinement. In the second experiment, we fine-tune the top layers of MobileNetV2, unfreezing some upper layers and training them on our datasets to adapt the model more effectively to our specific task. Additionally, we apply preprocessing techniques to enhance the quality of the input images, further improving the model's performance.

Finally, we explore an approach in the third experiment by replacing the SoftMax classifier with traditional machine learning classifiers, including Support Vector Machine (SVM), k-nearest Neighbors (KNN), and Random Forest (RF). Features are extracted from the global average pooling layer of MobileNetV2, which captures high-level representations of the input images and are then fed into these classifiers. SVM is included due to its effectiveness in handling both linear and non-linear classification tasks, its strong generalization capabilities, and its compatibility with dropout regularization to further mitigate overfitting [10, 86]. KNN is evaluated for its simplicity and ability to capture local patterns in feature space [87, 88], while Random Forest [89, 90] is tested for its robustness to high-dimensional data and its ensemble-based reduction of overfitting.

3.3.5 Strategies for Preparing Data and Protocols

Our study evaluates the performance of our proposed system using two datasets for each of the two biometric modalities: finger vein and palmprint. For finger vein recognition, we utilize the SDUMLA and FV-USM datasets, while for palmprint recognition, we use the CASIA and IITD datasets.

We use two strategies, S1 and S2, which are applied in both modalities. the first strategy (S1) assesses the biometric data independently for each instance, whereas the second strategy (S2) aggregates features from multiple instances to create a unified subject representation.

For the finger vein modality, in the first strategy (S1), we selected 106 subjects from the SDUMLA dataset, with each subject providing six images per finger, and 123 subjects from the FV-USM dataset, each contributing six samples per finger. In the second strategy (S2), we used 106 subjects from the SDUMLA dataset, with 36 samples per class, and 123 subjects from the FV-USM dataset, with each subject providing 24 images.

In the palmprint modality, in the first strategy (S1), we selected 230 subjects from the IITD dataset, with each subject providing between five to six images, and 312 subjects from the CASIA dataset, each contributing six samples. while in the second strategy (S2) we employed 230 subjects from the IITD dataset, with each subject providing 11 or 12 images, and 312 subjects from the CASIA dataset, with each subject providing 16 or 17 images.

To create the training and testing sets, we used three different protocols: Protocol 1 (P1) with 70% of the data for training and 30% for testing, Protocol 2 (P2) with an equal split of 50% for both training and testing and Protocol 3 (P3) with 90% for training and 10% for testing.

Table 3.1 provides a detailed overview of the dataset preparation for each protocol

and modality using the training sets. By applying the same strategies (S1 and S2) to both finger vein and palmprint modalities, our study ensures a consistent and comprehensive evaluation of our proposed technique across different biometric data types.

Table 3.1: Dataset Distribution for Finger Vein and Palmprint Modalities.

	Finger Vein				Palmprint			
	SDUMLA		FV-USM		IITD		CASIA	
	S1	S2	S1	S2	S1	S2	S1	S2
P1	424	2650	492	841	1682	1380	1872	3744
P2	318	1908	369	609	1218	1034	1248	2496
P3	530	3180	615	1069	2138	2070	2184	4368

3.3.6 Evaluation Criteria

The proposed system is evaluated using multiple metrics, as described in detail in Subsubsection 2.4. These metrics include precision (PR), recall (Rc), F1-score (F1), and AUC (Area Under the Curve). Additionally, we analyze the Cumulative Match Characteristics (CMC) curves and report the Rank-one (R1) and Rank-five (R5) recognition rates to comprehensively assess the system's performance.

3.3.7 Results and Discussion

This section outlines the experiments conducted to evaluate the proposed system. In the first experiment, the baseline performance is assessed using VGG16, VGG19, and MobileNetV2 on finger vein (FV) and palmprint datasets, without preprocessing or fine-tuning. The MobileNetV2 model, identified as the top performer, is then fine-tuned in the second experiment for feature extraction and classification, using S1 and S2 strategies. In the third experiment, the softmax classifier is removed from MobileNetV2, using it as a feature extractor, and the extracted features are classified using SVM, KNN, and RF under the same protocols and learning strategies.

The performance is evaluated across three protocols. Throughout all training phases, a trial-and-error method was adopted to optimize hyperparameters, using the RMSprop optimizer (batch sizes: [16, 32, 64]; epochs: +100).

3.3.7.1 First Experiment

The first experiment focused on evaluating the baseline performance of the system by training three deep learning models: VGG16, VGG19, and MobileNetV2 using transfer learning (TL). We used Protocol P1, a widely used method in research. This split ensures the model has sufficient data to learn effectively while also having sufficient

data to test its performance on unseen samples. The goal of this experiment is to determine the best-performing model within our study by assessing its ability to learn patterns.

The models were initialized using pre-trained weights from the ImageNet dataset, enabling the transfer of low-level and mid-level feature representations to our specific task. For classification, dense layers with the SoftMax activation function were utilized. The experiment was conducted on finger vein (FV) and palmprint datasets without any preprocessing or fine-tuning to create a baseline performance. This step ensures that the initial results reflect the raw capability of the models before any optimization or preprocessing techniques are applied. The results of this baseline evaluation, including performance metrics for both datasets, are presented in Tables 3.2 and 3.3.Also the CMC curves are plotted in Figures 3.7 to 3.14

Table 3.2: Performance metrics for the SDU and USM datasets resulting from transfer learning technique.

Data	Model	Inst	Pr (%)	Rc (%)	F1 (%)	R1 (%)	R5 (%)	AUC (%)
SDU	VGG16	LI	95	93	93	93.40	97	97
		LM	86	84	83	84.43	93	97
		LR	85	85	84	85.38	90	98
		RI	91	91	89	91.00	97	98
		RM	93	90	90	90.57	94	94
		RR	88	88	86	88.21	97	96
	VGG19	LI	89	90	88	90.09	94	99
		LM	83	85	83	85.85	93	99
		LR	87	86	85	86.79	95	99
		RI	89	86	86	86.73	93	97
		RM	87	85	84	85.38	95	99
		RR	82	83	88	83.49	89	98
	MobileNetV2	LI	97	95	95	94.80	97	98
		LM	85	86	85	86.32	92	98
		LR	93	92	92	87.26	92	98
		RI	93	92	92	91.46	96	98
		RM	93	92	92	92.45	94	94
		RR	87	90	87	89.62	92	95
USM	VGG16	LI	95	96	95	95.93	98	98
		LM	93	92	91	92.68	97	97
		RI	93	92	92	91.46	96	97
		RM	95	94	94	94.72	97	97
	VGG19	LI	94	94	93	94.31	97	97
		LM	90	90	88	90.24	97	97
		RI	89	88	87	88.35	96	97
		RM	94	93	93	93.90	97	97
	MobileNetV2	LI	97	96	96	96.34	98	99
		LM	96	97	96	96.75	98	98
		RI	94	95	94	95.12	98	99
		RM	97	96	97	96.34	98	99

As shown in Tables 3.2 and 3.3, MobileNetV2 consistently outperformed VGG16 and VGG19 across both datasets in precision, recall, F1 score, Rank 1, Rank 5, and AUC. This superiority, further visualized in the CMC curves (Figures 3.7–3.14), highlights its robustness in classifying biometric patterns. The baseline performance of MobileNetV2, particularly in handling raw, unprocessed data positions it as the optimal model for subsequent experiments involving preprocessing and fine-tuning.

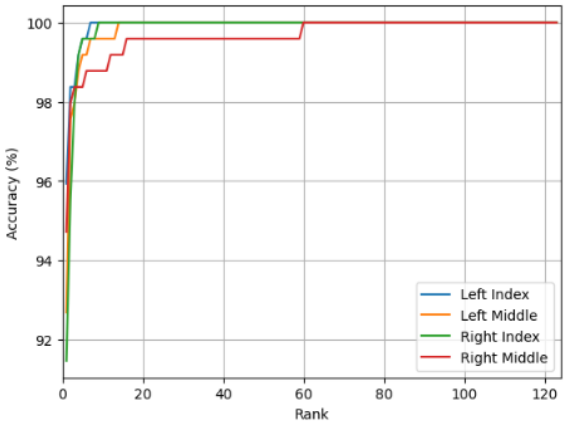


Figure 3.7: CMC for VGG16 using TL on FV-USM.

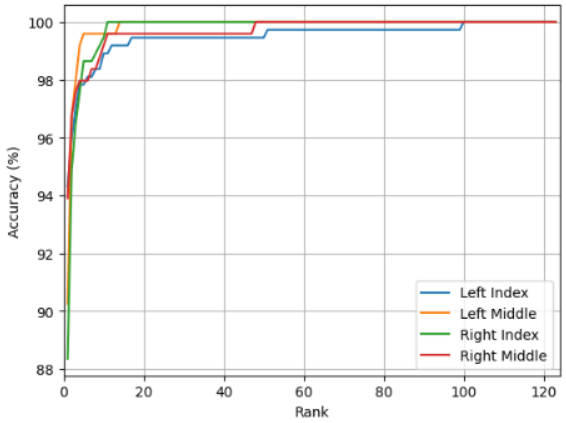


Figure 3.8: CMC for VGG19 using TL on FV-USM.

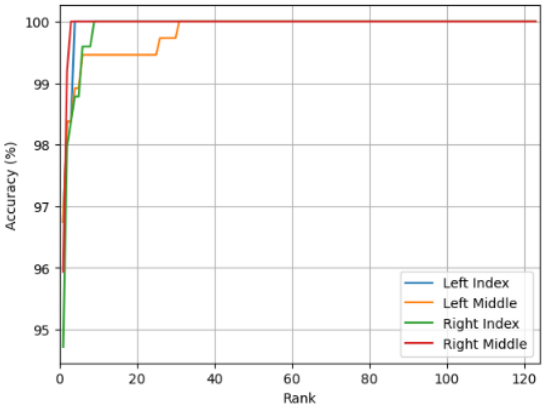


Figure 3.9: CMC for MobileNetV2 using TL on FV-USM.

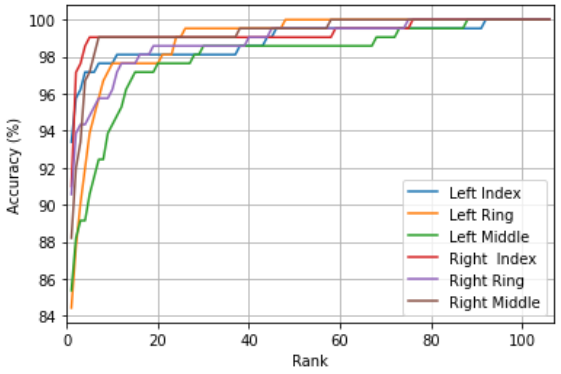


Figure 3.10: CMC for VGG16 using TL on SDUMLA.

3.3.7.2 Second Experiment

In the second experiment, we conducted a series of experiments to evaluate the effectiveness of fine-tuning MobileNetV2 and compare it with transfer learning results in the first experiment 3.3.7.1, both without preprocessing, for finger vein and palmprint recognition tasks. After that, we aimed to enhance the quality of finger vein and palmprint images by applying histogram equalization techniques as a preprocessing step to the dataset. For finger vein images, we used standard histogram equalization, while for palmprint images, we employed Contrast-Limited Adaptive Histogram Equalization (CLAHE). The goal of these experiments is to determine the optimal combination of techniques to address challenges such as low image quality, uneven illumination, and pattern variability.

For both finger vein and palmprint recognition, fine-tuning MobileNetV2 consistently outperformed transfer learning on without preprocessing datasets: Tables 3.2 using transfer learning and 3.5 and 3.4 using fine-tuning part (A) show that fine-tuning improved accuracy this superiority can be justified by the fact that fine-tuned models

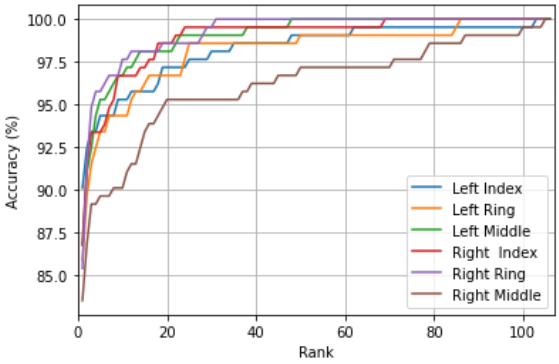


Figure 3.11: CMC for VGG19 using TL on SDUMLA.

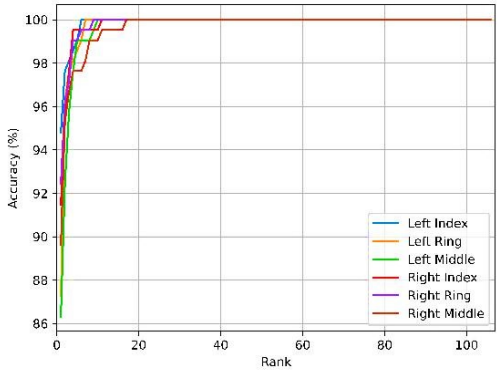


Figure 3.12: CMC for MobileNetV2 using TL on SDUMLA.

Table 3.3: Performance metrics for the IITD and CASIA datasets resulting from transfer learning technique.

Dataset	Models	instances	Pr (%)	Rc (%)	F1 (%)	R1 (%)	R5 (%)	AUC (%)
IITD	VGG16	LP	83.39	80.87	79.28	80.87	96.09	92.18
		RP	96.04	93.91	93.72	93.91	99.35	97.28
	VGG19	LP	88.27	85.22	84.26	85.22	97.61	94.55
		RP	93.42	91.09	90.53	91.09	99.35	93.57
	MobileNetV2	LP	97.10	95.65	95.51	95.65	99.57	99.39
		RP	96.99	95.43	95.18	95.43	99.13	99.57
CASIA	VGG16	LP	94.61	91.53	91.37	91.53	96.30	92.71
		RP	93.64	89.29	89.01	89.29	95.59	90.41
	VGG19	LP	95.26	92.70	92.40	92.70	97.25	91.70
		RP	93.68	89.81	89.60	89.81	96.22	89.61
	MobileNetV2	LP	97.33	96.51	96.25	96.51	98.94	99.45
		RP	97.90	97.16	96.98	97.16	99.47	99.63

can capture higher-level features specific to the patterns found in finger vein images compared to pre-trained models. We also plot the CMC curves in Figs.3.15–3.18. Similarly, Tables 3.3 using transfer learning and 3.6 and 3.7 using fine-tuning, part (A) demonstrate that fine-tuning better captured palmprint-specific features like principal lines and texture. Also, we plot CMC curves Figs.3.23–3.24.

These results highlight fine-tuning's advantage in learning domain-specific features, unlike transfer learning, which relies on generic pre-trained weights. Based on these results, we henceforth use the fine-tuned MobileNetV2 to test the rest of our experiments.

Based on the results presented in Tables 3.4(B) and 3.5(B), it is evident that our finger vein recognition system with histogram equalization preprocessing achieved superior performance compared to the same system without histogram equalization. Similarly, palmprint images processed with CLAHE, as shown in Tables 3.6(B) and 3.7(B), outperform those without preprocessing.

The Cumulative Match Characteristic (CMC) curves for both methods are further illustrated in Figs. 3.19–3.22 and 3.25–3.26.

From these findings, we conclude that enhancing finger vein and palmprint im-

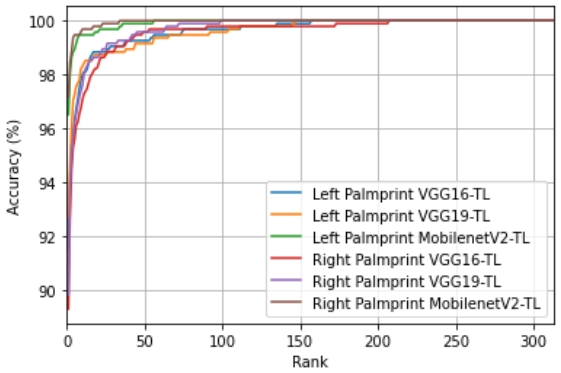


Figure 3.13: CMC for 3 models using TL on the CASIA.

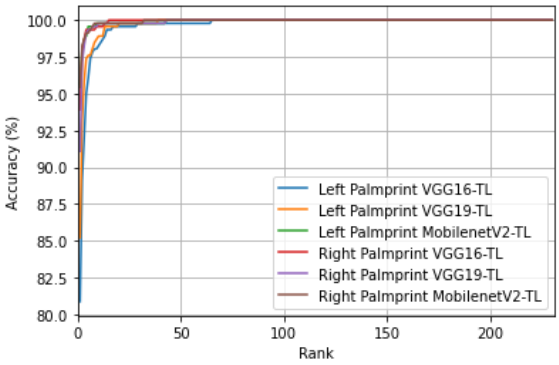


Figure 3.14: CMC for 3 models using TL on the IITD .

Table 3.4: The performance metrics for the FV-USM dataset resulting from: A) fine-tuning without preprocessing, and B) fine-tuning with preprocessing.

		A						B					
		Metrics (%)						Metrics (%)					
		PR	RC	F1	R1	R5	AUC	PR	RC	F1	R1	R5	AUC
P1	LI	98	97	97	97.15	99	98	99	99	99	98.78	100	99
	LM	98	98	98	98.38	99	100	100	100	100	100	100	100
	RI	98	97	97	97.56	99	100	99	99	99	98.78	100	100
	RM	99	99	99	98.78	99	100	100	100	100	99.59	99.6	100
P2	LI	97	97	97	97.02	99	99	99	98	99	98.37	100	100
	LM	99	99	99	98.92	99	100	100	99	99	99.45	100	99
	RI	97	95	95	95.39	98	98	97	96	97	96.21	98.4	98
	RM	94	95	94	95.12	98	98	98	97	97	96.75	98.7	98
P3	LI	99	99	99	99.19	99	100	99	98	99	99	100	100
	LM	99	99	99	99.19	100	100	100	100	100	100	100	100
	RI	98	98	98	98.37	100	100	99	98	99	99	99.6	100
	RM	99	99	99	99.19	100	100	100	100	100	100	100	100

ages is essential to address challenges such as poor quality, shading variations, and low contrast. Consequently, only preprocessed images will be utilized in the third experiment.

3.3.7.3 Third Experiment

In the third and final experiment of our study, we integrate deep features extracted from a fine-tuned MobileNetV2 with three machine learning classifiers—SVM, KNN, and RF using two learning strategies: single instance (S1) and multiple instance (MI) (S2), applied to both finger vein and palmprint recognition. We also plotted the Cumulative Match Characteristic (CMC) curves for multiple-instance learning scenarios. For the finger vein dataset, the CMC curves are presented in Fig. [3.39] to Fig. [3.42], while for the palmprint dataset, the curves are shown in Fig. [3.49] and Fig. [3.54].

Table 3.5: The performance metrics for the SDUMLA dataset resulting from: A) fine-tuning without preprocessing, and B) fine-tuning with preprocessing.

		A						B					
		Metrics (%)						Metrics (%)					
		PR	**RC**	**F1**	**R1**	**R5**	**AUC**	**PR**	**RC**	**F1**	**R1**	**R5**	**AUC**
P1	LI	97	96	96	95.75	98	98	97	97	96	96.70	99.50	99
	LR	96	96	94	93.87	97	97	94	95	94	94.81	97	98
	LM	93	92	91	91.98	95	97	93	93	92	93.40	98	99
	RI	94	92	92	92.42	97	98	93	94	93	94.34	98	99
	RR	95	95	94	94.81	98	99	94	95	94	95.28	98	100
	RM	96	94	94	93.87	96	99	93	94	93	93.87	98	100
P2	LI	91	88	87	88.36	94	97	91	91	89	90.88	95	98
	LR	90	86	85	86.16	92	96	87	88	86	88.05	94	97
	LM	91	89	89	89.31	92	95	92	91	90	90.88	95	97
	RI	93	89	88	88.99	92	94	93	91	90	91.19	96	96
	RR	91	89	88	88.99	92	96	90	91	89	91	95	96
	RM	88	85	84	85.22	90	94	89	89	88	88.99	94	96
P3	LI	90	93	91	93.40	96	96	92	94	92	94.34	99	97
	LR	93	95	94	95.28	97	98	95	97	96	97.17	97	98
	LM	96	97	96	97.17	99	98	98	99	98	99.06	100	99
	RI	93	95	94	95.28	98	99	94	96	94	96.23	100	100
	RR	94	96	96	96.23	98	99	97	98	98	98.11	99	100
	RM	90	92	91	92.45	96	99	91	94	92	94.34	98	99

Using the FV-USM dataset, we conducted multiple tests to determine the best accuracy through an iterative process of trial and error. Across three protocols, we achieved Rank 1 accuracy values of 99.79%, 99.24%, and 99.80% for one-instance learning. For multiple-instance learning we got 99.90%, 99.60%, and 99.80% for the three protocols, respectively.

Similarly, with the SDUMLA dataset, our trials resulted in Rank 1 scores of 95,90%, 90.52%, and 99,37% for one-instance learning, and 99,57%, 97,58%, and 99.76% for multiple-instance learning for the three protocols, respectively. Figures [3.27 to 3.30] depict the Cumulative Match Characteristic (CMC) curves of each trial one instance scenario.

The performance metrics for the FV-USM and SDUMLA datasets using MobileNetV2-based KNN are presented in the table 3.9. For the FV-USM dataset, the results demonstrate high accuracy across various protocols. In one-instance learning, Rank 1 accuracy values reached 98.15%, 97.22%, and 99.17% for the three protocols, respectively. For multiple-instance learning, the Rank 1 accuracy was even higher, achieving 99.85%, 97.80%, and 99.23% for the same protocols. These results highlight the robustness of the model, particularly in multiple-instance scenarios.

On the SDUMLA dataset, the performance was also strong but slightly lower com-

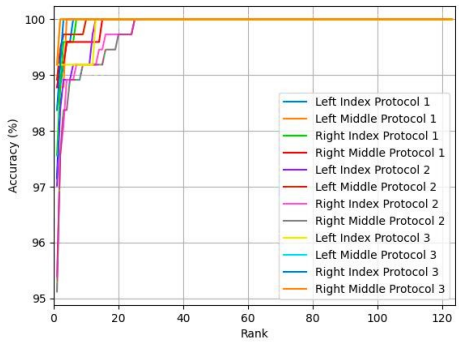


Figure 3.15: CMC for MobileNetV2 using FT on FV-USM without preprocessing.

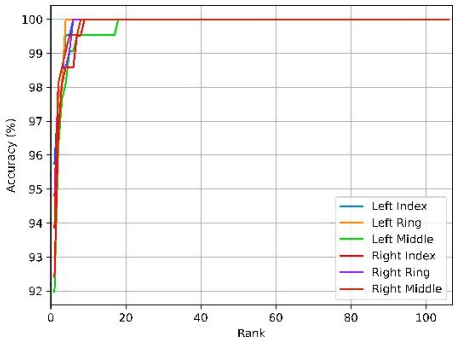


Figure 3.16: CMC for MobileNetV2 using FT and P1 on SDUMLA without preprocessing.

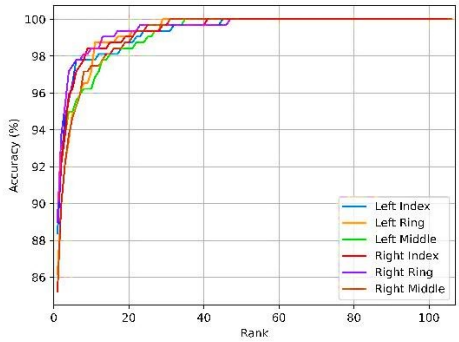


Figure 3.17: CMC for MobileNetV2 using FT and P2 on SDUMLA without preprocessing.

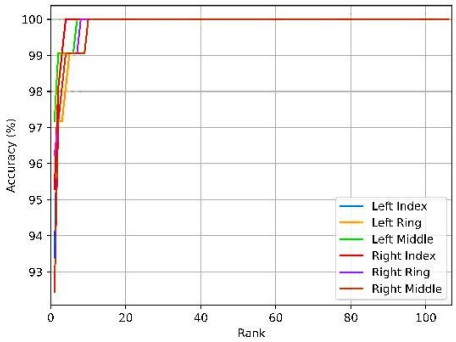


Figure 3.18: CMC for MobileNetV2 using FT and P3 on SDUMLA without preprocessing.

pared to FV-USM. For one-instance learning, Rank 1 accuracy values were 94.31%, 88.63%, and 98.12% across the three protocols. In multiple-instance learning, the accuracy improved significantly, with values of 99.25%, 96%, and 99.60% for the respective protocols. This indicates that the model performs exceptionally well in multiple-instance learning scenarios, even on more challenging datasets like SDUMLA.

The CMC curves one instance, depicted in Figures [3.43 to 3.34], further illustrate the model's performance across different trials.

The performance metrics for the FV-USM and SDUMLA datasets using MobileNetV2-based Random Forest (RF) are summarized in the table 3.10. For the FV-USM dataset, the model demonstrates strong performance across all protocols. In one-instance learning, Rank 1 accuracy values reached 98.15%, 97.15%, and 98.70% for the three protocols, respectively. For multiple-instance learning, the accuracy was even higher, achieving 99%, 98.40%, and 98.70% for the same protocols. These results indicate that the RF-based model is highly effective, particularly in multiple-instance scenarios, with near-perfect accuracy in some cases.

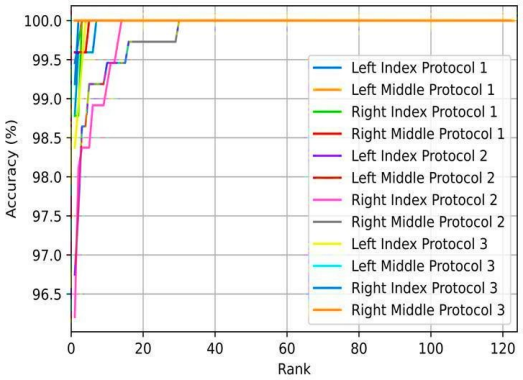


Figure 3.19: CMC for MobileNetV2 using FT on FV-USM with preprocessing.

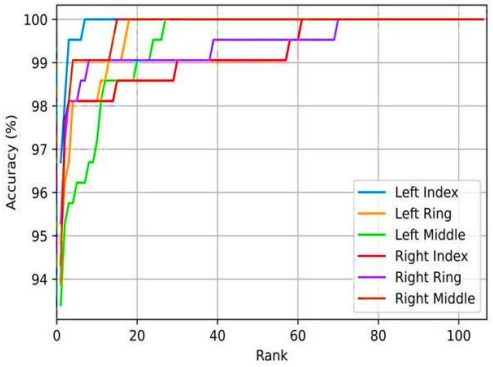


Figure 3.20: CMC for MobileNetV2 using FT and P1 on SDUMLA with preprocessing.

Table 3.6: The performance metrics for the CASIA dataset resulting from: A) fine-tuning without preprocessing, and B) fine-tuning with preprocessing.

		A						B					
		Metrics (%)						Metrics (%)					
		PR	RC	F1	R1	R5	AUC	PR	RC	F1	R1	R5	AUC
P1	LP	98.98	98.62	98.53	98.62	99.13	98.84	98.71	99.00	98.84	98.84	99.27	99.87
	RP	98.89	98.48	98.34	98.41	99.00	99.00	98.84	98.55	98.47	98.55	99.27	99.00
P2	LP	95.49	97.85	97.38	97.85	99.00	99.76	98.44	97.99	97.84	98.00	99.00	99.90
	RP	98.60	98.26	98.16	98.26	99.00	99.70	98.64	99.06	99.05	98.70	99.00	99.91
P3	LP	99.00	98.70	98.67	98.69	100.00	100.00	99.64	99.38	99.36	99.38	100.00	99.98
	RP	99.00	98.75	98.36	98.75	99.00	99.80	99.18	99.08	98.95	99.08	99.69	99.96

On the SDUMLA dataset, the performance was also robust but slightly lower compared to FV-USM. For one-instance learning, Rank 1 accuracy values were 94.04%, 87%, and 97.55% across the three protocols. In multiple-instance learning, the accuracy improved significantly, with values of 98.37%, 96%, and 99.20% for the respective protocols. This highlights the model's ability to perform well in multiple-instance learning, even on more challenging datasets like SDUMLA.

The CMC curves of one instance, depicted in Figures [3.35 to 3.38], further illustrate the model's performance across different trials.

Table 3.11presents comprehensive performance results for palmprint recognition systems, examining various machine learning approaches across different experimental conditions. Using two databases IITD and CASIA, the study evaluated three machine learning algorithms - Random Forest (RF), K-Nearest Neighbors (KNN), and Support Vector Machine (SVM) - analyzing both one instance performance and combined palmprint data through Multi-Instance (MI) analysis.

The experimental framework incorporated three protocols (P1, P2, and P3), with Protocol P3 consistently demonstrating superior performance across all testing scenar-

Table 3.7: Performance metrics for the IITD dataset resulting from: A) fine-tuning without preprocessing, and B) fine-tuning with preprocessing.

		A						B					
		Metrics (%)						Metrics (%)					
		PR	**RC**	**F1**	**R1**	**R5**	**AUC**	**PR**	**RC**	**F1**	**R1**	**R5**	**AUC**
P1	LP	93.48	93.99	93.48	95.28	97.5	99	97.39	96.04	97.39	96.52	99.35	99.84
	RP	94.33	95.36	94.35	93.72	98	99	95.34	9645	9543	9504	9804	9997
P2	LP	91.30	92.33	91.30	90.49	95.80	99.22	92.61	93.47	92.61	91.78	96.38	99.60
	RP	90.00	92.95	90.00	89.42	97.55	97.30	91.31	91.88	91.47	91.88	97.54	99.38
P3	LP	96.52	94.79	96.52	95.35	100.00	100.00	98.04	98.70	98.26	98.70	100.00	99.99
	RP	96.96	95.43	96.96	95.94	100.00	99.97	98.70	99.13	98.84	99.13	99.13	99.95

ios. Most notably, when implementing SVM classification with the CASIA database under Protocol P3, the system achieved perfect accuracy scores of 100% across all performance metrics.

Throughout the experiments, several significant patterns emerged in the data. Multi-Instance analysis consistently outperformed single palm analysis, indicating that incorporating data from both palms enhances recognition accuracy. The SVM classifier demonstrated superior performance compared to both RF and KNN across most test configurations, while the CASIA database generally yielded higher recognition rates than IITD, particularly when utilizing Protocol P3. These findings remained robust across different testing conditions, with most configurations maintaining performance metrics above 95%.

From Tables 3.8,3.9,3.10, and 3.11 we can observe that by using MobileNetV2 fine-tuning for feature extraction followed by various machine learning classifiers for both finger vein and palmprint modalities, we aim to leverage the advantages of this approach. While MobileNetV2 fine-tuning efficiently extracts relevant features from finger vein and palmprint images through adapting its weights to our specific biometric datasets, the machine learning classifiers are capable of learning decision boundaries and identifying pertinent features in the extracted feature representations from both modalities.

The fine-tuned MobileNetV2 demonstrated high performance in extracting relevant features from both finger vein and palmprint images due to its lightweight architecture and its capacity to capture discriminative characteristics specific to biometric patterns. Furthermore, integrating SVM, which is a binary classifier for differentiating between one class and the remaining ones, enhances the model's predictive capability by focusing on maximizing the margin between data points of different classes rather than minimizing cross-entropy. This is particularly effective when dealing with the distinct features extracted by MobileNetV2 from both finger veins and palmprints.

While the KNN classifier is simple to understand, it suffers when there are a lot

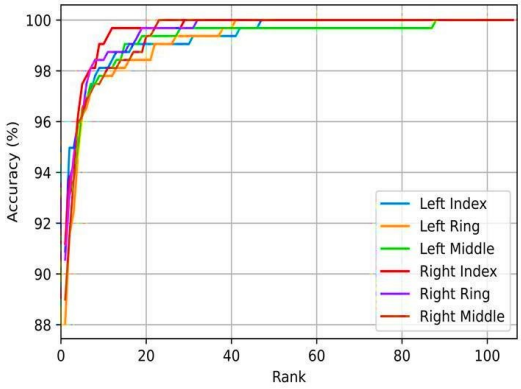


Figure 3.21: CMC for MobileNetV2 using FT and P2 on SDUMLA with preprocessing.

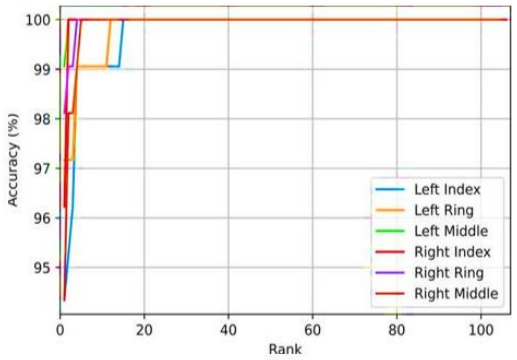


Figure 3.22: CMC for MobileNetV2 using FT and P3 on SDUMLA with preprocessing.

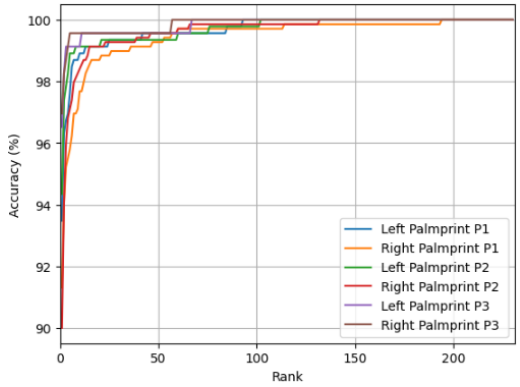


Figure 3.23: CMC for MobileNetV2 using FT on IITD without preprocessing.

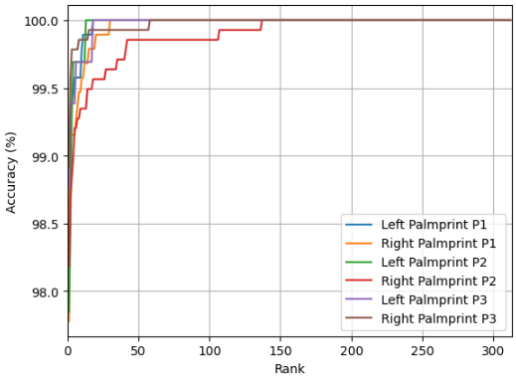


Figure 3.24: CMC for MobileNetV2 using FT on CASIA without preprocessing.

of features to consider. It relies on distance measurements, which are inefficient when considering the high-dimensional feature vectors extracted by MobileNetV2 from finger veins and palmprints. Random Forest, in contrast to KNN, is robust and adaptable, but it could miss differences that SVM can detect, especially when the boundaries between classes are complex across both biometric traits.

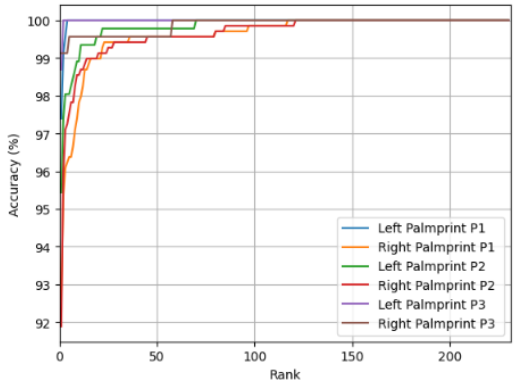


Figure 3.25: CMC for MobileNetV2 using FT on IITD with preprocessing.

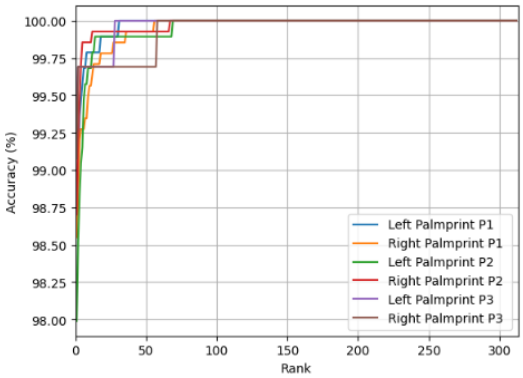


Figure 3.26: CMC for MobileNetV2 using FT on CASIA with preprocessing.

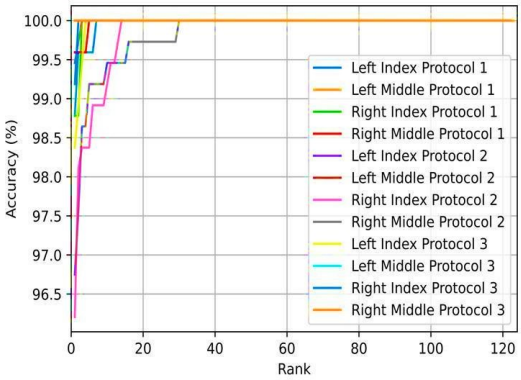


Figure 3.27: CMC for MobilenetV2-based SVM with FV-USM.

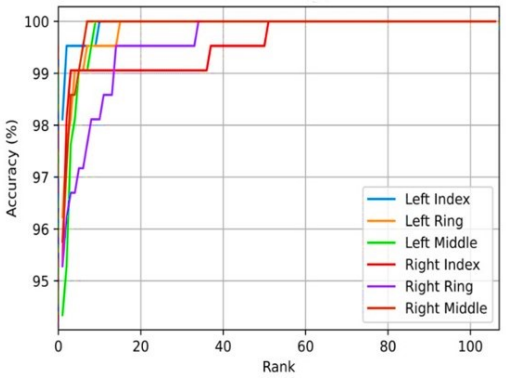


Figure 3.28: CMC for MobilenetV2-based SVM using P1 on SDUMLA.

Table 3.8: Performance Metrics for FV-USM and SDUMLA datasets using MobilenetV2 based SVM.

		FV-USM dataset							SDUMLA dataset						
		Metrics (%)							Metrics (%)						
		PR	**RC**	**F1**	**R1**	**R5**	**AUC**	**Acc avg**	**PR**	**RC**	**F1**	**R1**	**R5**	**AUC**	**Acc avg**
P1	LI	100	100	100	99.50	100	99		99	98	98	98.58	99.5	98	
	LR	-	-	-	-	-	-		95	96	95	96.22	99	97	
	LM	100	100	100	100	100	100		94	94	93	94.33	98	97	
	RI	100	100	100	100	100	100	99.79	94	95	94	95.28	99	97	95.90
	RR	-	-	-	-	-	-		94	95	94	95.28	96.5	98	
	RM	100	100	100	99.50	100	99		95	96	95	96	98	98	
	MI	100	100	100	99.90	100	100	99.90	100	100	100	99.57	99.80	100	99.57
P2	LI	100	99	99	99.45	99.7	99		94	94	94	92	96	96	
	LR	-	-	-	-	-	-		91	89	88	88.05	94	96	
	LM	99	99	99	99.18	99	99		95	93	93	89.62	96	0.95	
	RI	99	99	99	99.45	100	99	99.24	95	95	94	93.39	98	97	90.52
	RR	-	-	-	-	-	-		94	94	93	93	98	97	
	RM	99	99	99	98.91	100	99		94	91	91	87.42	94	97	
	MI	100	100	100	99.90	100	100	99.60	97	96	95	97.58	99	98	97.58
P3	LI	100	100	100	100	100	100		97	98	97	98.11	99.05	97	
	LR	-	-	-	-	-	-		99	99	99	99.05	100	99	
	LM	100	100	100	99.50	100	100		99	99	99	99.05	100	99	
	RI	100	100	100	99.50	100	100	99.80	99	99	99	99.05	100	99	99.37
	RR	-	-	-	-	-	-		100	100	100	100	100	100	
	RM	100	100	100	100	100	100		100	100	100	100	100	100	
	MI	100	100	100	99.59	100	100	99.59	99	99	99	99.76	100	99	99.76

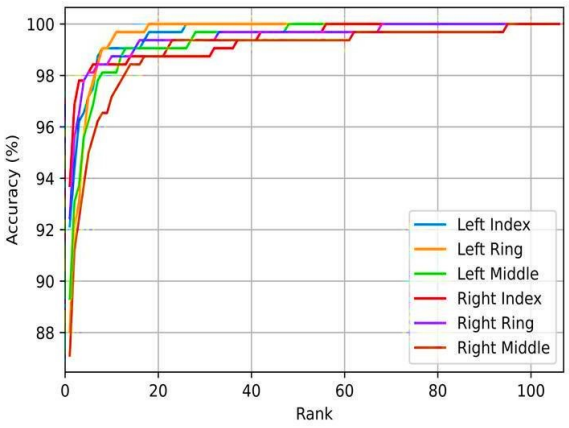


Figure 3.29: CMC for MobilenetV2-based SVM using P2 on SDUMLA.

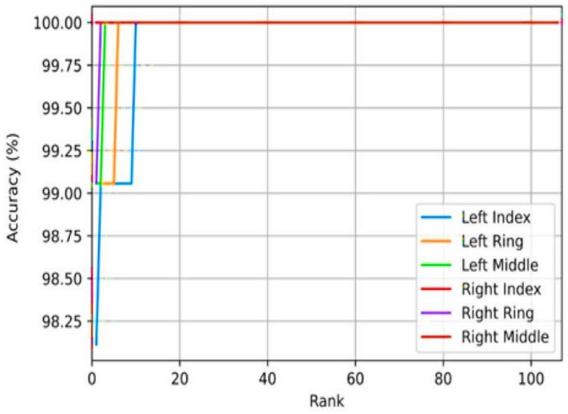


Figure 3.30: CMC for MobilenetV2-based SVM using P3 on SDUMLA.

Table 3.9: Performance Metrics for FV-USM and SDUMLA datasets using MobilenetV2 based KNN.

		FV-USM dataset							SDUMLA dataset						
		Metrics (%)							**Metrics (%)**						
		PR	**RC**	**F1**	**R1**	**R5**	**AUC**	**Acc avg**	**PR**	**RC**	**F1**	**R1**	**R5**	**AUC**	**Acc avg**
P1	LI	98	98	98	98.50	100	99		96	97	95	97.20	98	97	
	LR	-	-	-	-	-	-		95	95	94	95.20	98	96	
	LM	98	98	98	98.20	99	99		94	93	93	92.50	98	96	
	RI	96	97	96	98.30	99	99	98.15	93	94	93	94.30	98	96	94.31
	RR	-	-	-	-	-	-		91	92	93	91.50	97	95	
	RM	99	98	97	97.60	99	99		94	95	94	95.20	99	97	
	MI	100	100	100	100	99.85	100	99.85	99	99	99	99.25	100	100	99.25
P2	LI	97	98	97	97.50	99	98		93	92	92	91.80	95	94	
	LR	-	-	-	-	-	-		90	86	84	85.82	93	91	
	LM	98	98	98	97.80	99	99		86	85	86	86.16	90	94	
	RI	96	97	96	96.70	99	99	97.22	93	92	92	91.50	95	94	88.63
	RR	-	-	-	-	-	-		92	91	91	91.50	94	94	
	RM	96	97	96	96.60	99	99		87	85	83	85	92	93	
	MI	97	98	97	97.80	100	100	97.80	97	96	95	96	98	97	96
P3	LI	99	99	99	99.30	100	100		97	97	95	97.40	98	96	
	LR	-	-	-	-	-	-		97	98	97	97.80	99	98	
	LM	98	98	98	98.20	99	100		98	98	99	98.20	99	98	
	RI	99	99	99	99.20	100	100	99.17	98	97	97	97.20	99.50	98	98.12
	RR	-	-	-	-	-	-		99	98	99	98.60	99.50	100	
	RM	100	100	100	100	100	100		100	100	100	99.50	100	100	
	MI	99	99	99	99.23	100	100	99.23	99	99	99	99.60	99.75	99	99.60

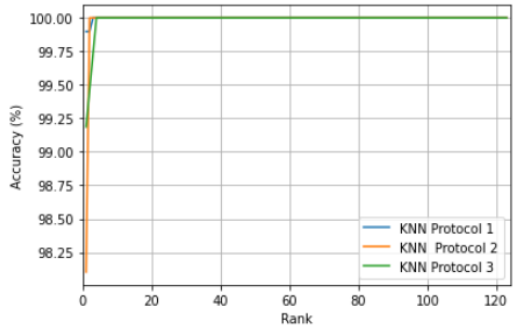


Figure 3.31: CMC for MobilenetV2-based KNN with FV-USM.

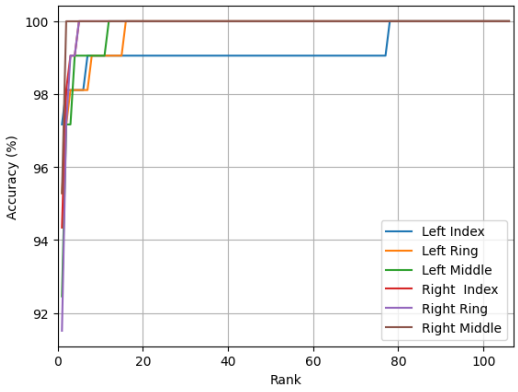


Figure 3.32: CMC for MobilenetV2-based KNN using P1 on SDUMLA.

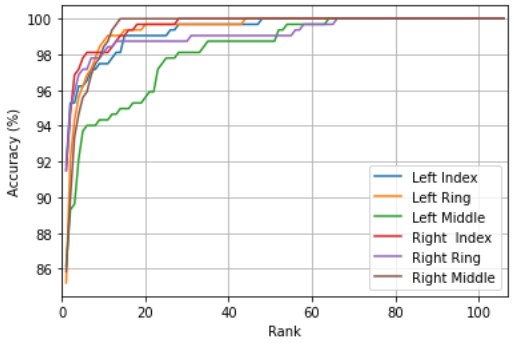


Figure 3.33: CMC for MobilenetV2-based KNN using P2 on SDUMLA.

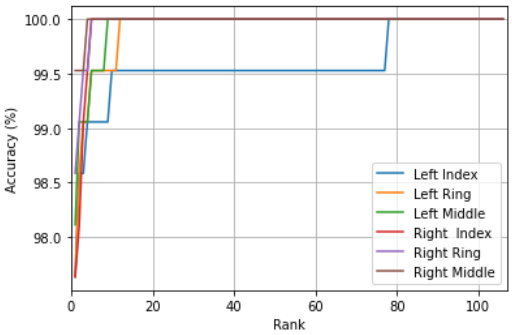


Figure 3.34: CMC for MobilenetV2-based KNN using P3 on SDUMLA.

Table 3.10: Performance Metrics for FV-USM and SDUMLA datasets using MobilenetV2 based RF.

		FV-USM dataset							SDUMLA dataset						
		Metrics (%)							**Metrics (%)**						
		PR	**RC**	**F1**	**R1**	**R5**	**AUC**	**Acc avg**	**PR**	**RC**	**F1**	**R1**	**R5**	**AUC**	**Acc avg**
P1	LI	98	99	98	98.50	99	99		97	98	96	97.60	99	97	
	LR	-	-	-	-	-	-		95	94	93	94.81	98	96	
	LM	97	98	97	98	99	99		92	93	93	92.60	96	96	
	RI	99	98	98	98.80	100	100	98.15	92	93	92	92.45	96	96	94.04
	RR	-	-	-	-	-	-		92	93	92	92.50	97	95	
	RM	96	97	96	97.30	99	99		94	95	94	94.34	98	97	
	MI	98	99	97	99	99	99	99	99	98	98	98.37	99	99	98.37
P2	LI	96	97	96	96.90	99	98		88	89	88	88.99	95	96	
	LR	-	-	-	-	-	-		82	83	84	83.33	90	94	
	LM	97	98	97	98.10	99	98		84	85	85	84.59	91	95	
	RI	94	95	94	95.20	98	98	97.15	88	89	89	89.31	95	96	87
	RR	-	-	-	-	-	-		90	91	91	90.88	95	96	
	RM	95	96	95	95.50	98	97		84	84	85	85	85.55	95	
	MI	97	98	97	98.40	99	99	98.40	96	96	96	96	96	98	96
P3	LI	98	99	98	99.10	100	100		98	97	96	97.60	98	96	
	LR	-	-	-	-	-	-		96	97	96	96.50	98	98	
	LM	97	98	97	98.20	99	100		98	99	97	98.20	99	98	
	RI	97	98	97	98.20	100	100	98.70	96	97	96	96.50	98	98	97.55
	RR	-	-	-	-	-	-		96	97	96	97.30	99.50	99	
	RM	99	99	99	99.30	100	100		99	99	99	99.20	100	100	
	MI	98	99	98	98.70	100	100	98.70	99	99	99	99.20	1005	100	99.20

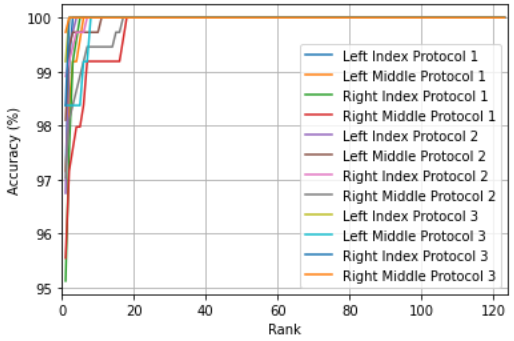


Figure 3.35: CMC for MobilenetV2-based RF with FV-USM.

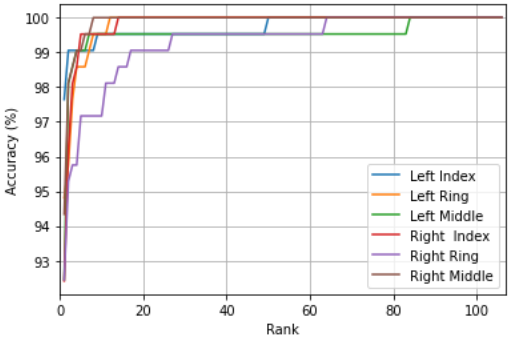


Figure 3.36: CMC for MobilenetV2-based RF using P1 on SDUMLA.

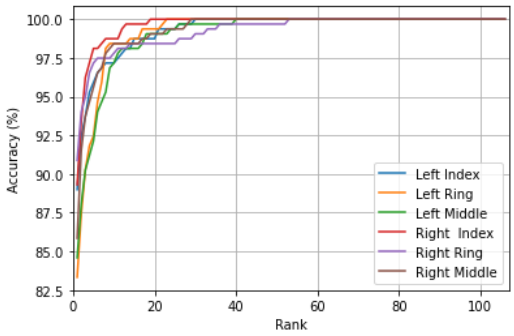


Figure 3.37: CMC for MobilenetV2-based KNN using P2 on SDUMLA.

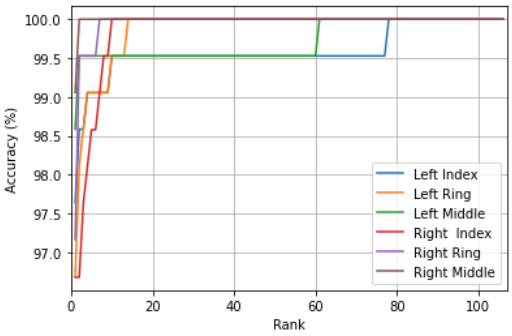


Figure 3.38: CMC for MobilenetV2-based RF using P3 on SDUMLA.

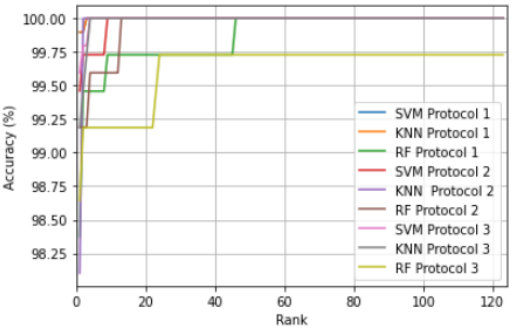


Figure 3.39: CMC for MobilenetV2-based ML using S2 on FV-USM.

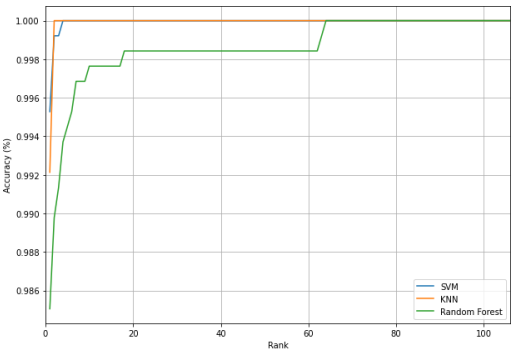


Figure 3.40: CMC for MobilenetV2-based ML using S2 on SDUMLA P1.

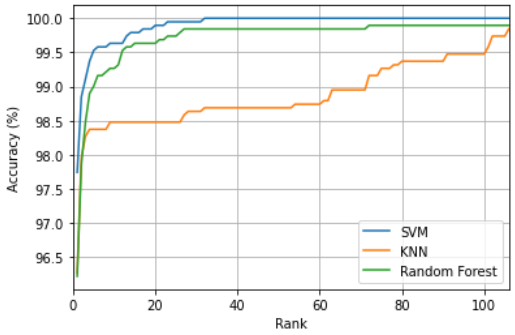


Figure 3.41: CMC for MobilenetV2-based ML using S2 on SDUMLA P2.

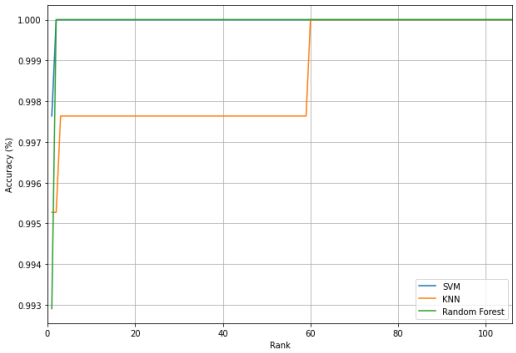


Figure 3.42: CMC for MobilenetV2-based ML using S2 on SDUMLA P1.

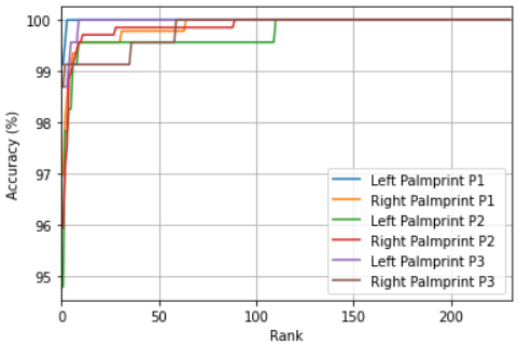


Figure 3.43: CMC for MobilenetV2-based SVM with IITD S1.

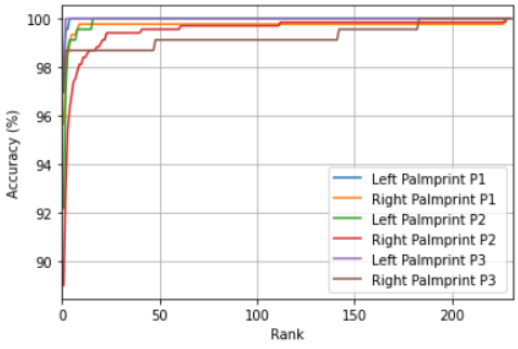


Figure 3.44: CMC for MobilenetV2-based RF with IITD S1.

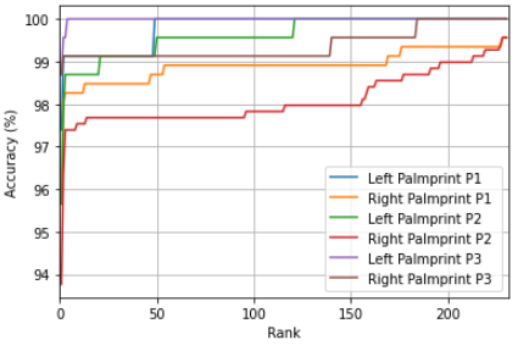


Figure 3.45: CMC for MobilenetV2-based KNN with IITD S1.

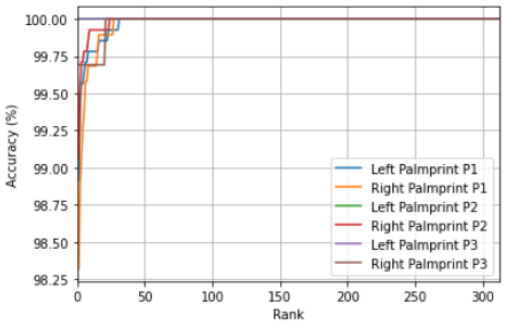


Figure 3.46: CMC for MobilenetV2-based SVM with CASIA S1.

Table 3.11: Performance comparison of different protocols and classifiers using Palmprint.

Protocol	ML	Data	Pr(%)	Rec(%)	F1(%)	R1(%)	R5(%)	AUC(%)
				P1				
		Left	96	97	96	97.39	99	99
	RF	Right	95	96	95	95.65	99	99
		MI	99	99	99	99.02	99.50	100
		Left	97	97	96	97.39	99	100
	KNN	Right	96	97	96	96.52	98	99
		MI	98	98	99	98.90	99.60	100
		Left	98	99	98	99.13	100	100
IITD	SVM	Right	97	97	97	97.73	99	99
		MI	100	100	100	99.46	99.70	100
		Left	97	98	97	98.30	99.50	99
	RF	Right	96	97	96	97.41	98.70	99
		MI	99	98	98	98.42	99.60	99
		Left	94	95	94	95.34	98	98
CASIA	KNN	Right	93	94	93	94.11	98	98
		MI	97	97	98	96.95	98.90	100
		Left	99	99	99	98.90	100	100
	SVM	Right	98	98	98	98.31	99	100
		MI	99	99	99	99.13	99.80	100
				P2				
		Left	91	92	91	92.17	98	99
	RF	Right	88	89	89	89	96	99
		MI	97	96	96	96.00	99.00	99
		Left	95	96	95	95.65	98	99
IITD	KNN	Right	93	94	93	93.60	97	99
		MI	97	97	97	97.03	99.30	99
		Left	94	95	95	94.65	99	99
	SVM	Right	95	96	95	96.08	99	100
		MI	99	99	98	98.55	99.90	100
		Left	97	98	97	98.50	100	100
	RF	Right	99	99	98	98.90	100	100
		MI	100	100	100	99.69	99.69	100
		Left	98	98	98	98.04	100	100
CASIA	KNN	Right	99	99	98	98.60	100	100
		MI	99	99	99	99.38	99.70	100
		Left	99	99	99	99.38	100	100
	SVM	Right	99	99	99	99.27	99	-
		MI	100	100	100	99.69	99.69	100
				P3				
		Left	98	98	98	97.89	99	100
	RF	Right	98	98	98	98.26	99	100
		MI	99	98	98	98.48	99.80	100
		Left	97	98	97	98	100	100
IITD	KNN	Right	98	99	98	98.70	99	100
		MI	98	98	98	98.04	99.60	100
		Left	98	99	98	98.70	99	100
	SVM	Right	98	99	98	98.70	99	100
		MI	99	99	99	99.35	100	100
		Left	100	99	99	99.38	100	100
	RF	Right	98	99	98	99.07	99	99
		MI	100	99	99	99.38	100	100
		Left	98	99	98	99.07	100	100
CASIA	KNN	Right	98	99	98	99.08	99	100
		MI	100	100	100	99.07	99.80	100
		Left	100	100	100	99.69	100	100
	SVM	Right	100	100	100	100	100	100
		MI	100	100	100	100	100	100

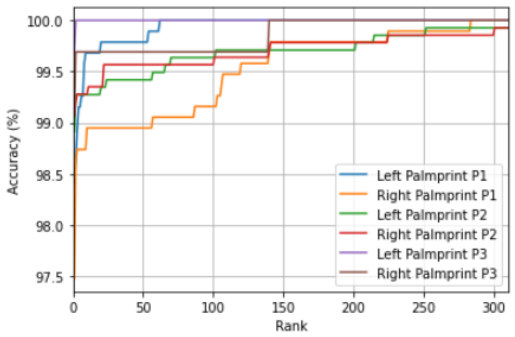


Figure 3.47: CMC for MobilenetV2-based RF with CASIA S1.

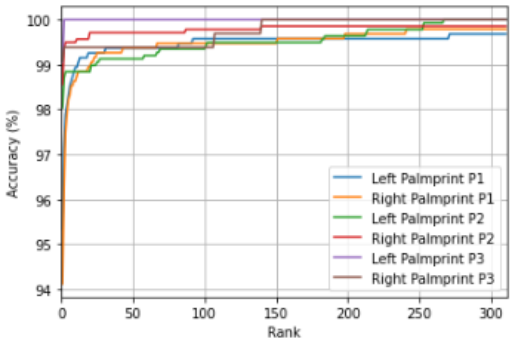


Figure 3.48: CMC for MobilenetV2-based KNN with CASIA S1.

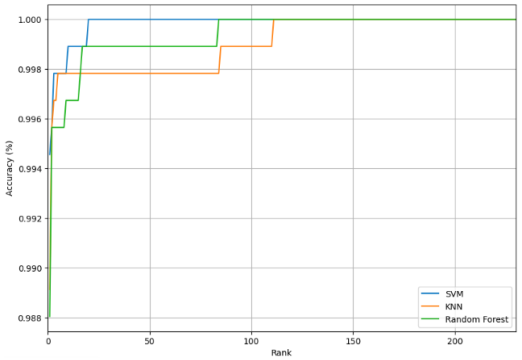


Figure 3.49: CMC for MobilenetV2-based ML P1 with IITD S2.

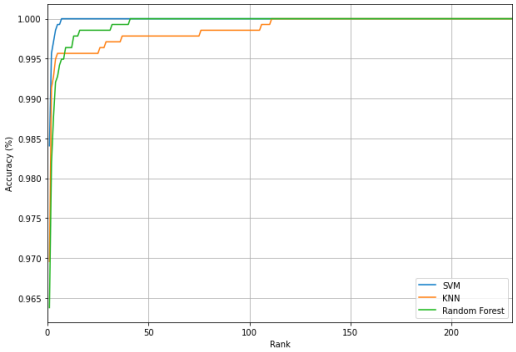


Figure 3.50: CMC for MobilenetV2-based ML P2 with IITD S2.

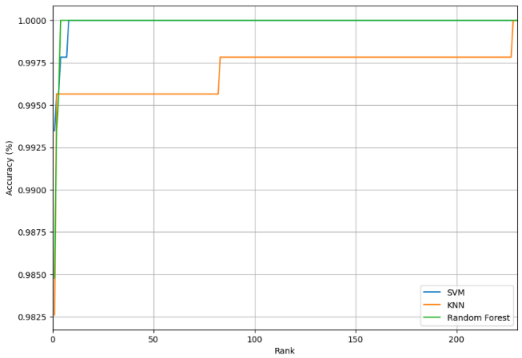


Figure 3.51: CMC for MobilenetV2-based ML P3 with IITD S2.

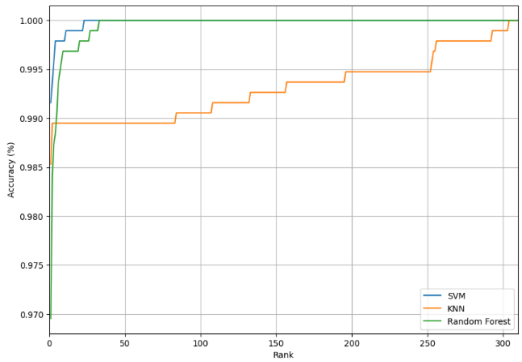


Figure 3.52: CMC for MobilenetV2-based ML P1 with CASIA S2.

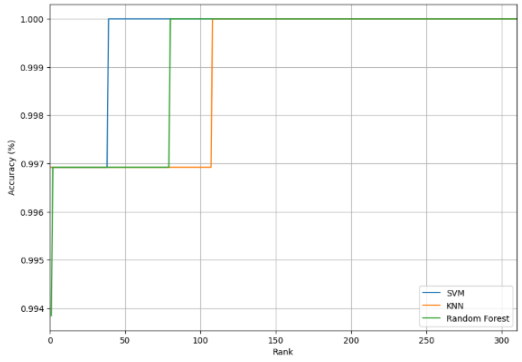


Figure 3.53: CMC for MobilenetV2-based ML P2 with CASIA S2.

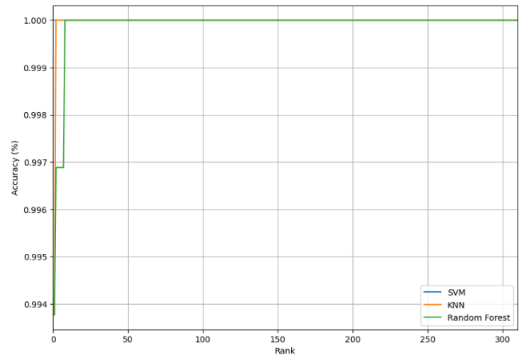


Figure 3.54: CMC for MobilenetV2-based ML P3 with CASIA S2.

Moreover, using multiple images from the same person (multiple-instance strategy) is better than using just one image (one-instance strategy) for both finger vein and palmprint recognition. Sometimes, health issues like anemia or heart problems can make vein patterns harder to see, while various factors such as illumination and hand positioning can affect palmprint image quality. By using multiple images of both modalities, we get more comprehensive information, making our system more accurate and robust through complementary biometric features extracted by the fine-tuned MobileNetV2.

3.3.8 Comparison with the State-of-the-Art Methods

We now compare our proposed method with existing approaches in Tables 3.12 and 3.13, which utilize the same dataset, protocols, and evaluation metrics, to demonstrate its effectiveness.

Table 3.12: Comparisons with works using the FV-USM and SDUMLA datasets.

Dataset	Ref	Year	P1	P2	P3
FV-USM	[91]	2019	98.58%	97.83%	-
	[60]	2020	97.95%	-	-
	[64]	2022	96.15%	94.21%	96.75%
	[92]	2022	97.84%	-	-
	[63]	2023	99.40%	-	-
	Our study	-	99.90%	99.60%	99.80%
SDUMLA	[91]	2019	98.58%	97.83%	-
	[57]	2007	86.01%	-	97.95%
	[93]	2023	-	-	98.11%
	[94]	2022	-	81.97%	-
	[95]	2022	98.70%	96%	98.1%
	[64]	2022	96.15%	94.21%	96.75%
	[96]	2022	-	96.61%	-
	[65]	2022	98.0%	-	98.8%
	[97]	2022	-	95.33%	-
	[98]	2022	-	-	92.33%
	[62]	2023	99.07%	-	-
	[63]	2023	99.06%	-	-
	Our study	-	99.57%	97.58%	99.76%

Table 3.12 compares the performance of our method with state-of-the-art approaches on the FV-USM and SDUMLA datasets under three protocols (P1, P2, P3). For FV-USM, our method achieves superior results with 99.90% (P1), 99.60% (P2), and 99.80% (P3), outperforming recent works like [63], 99.40% P1. On SDUMLA, we attain 99.57% (P1), 97.58% (P2), and 99.76% (P3), surpassing existing methods including [62] 99.07% P1 and [95], 98.70% P1. The comparison highlights consistent improvements across both datasets, particularly in P1 and P3 metrics.

Table 3.13 compares the performance of our method with state-of-the-art approaches on the CASIA and IITD datasets under three protocols (P1, P2, P3). For CASIA, our method achieves exceptional results with 99.13% (P1), 99.69% (P2), and 100% (P3), outperforming recent works like [99] (99.11% P1) and [100] (99.18% P2). On IITD, we

attain 99.46% (P1), 98.50% (P2), and 99.35% (P3), surpassing existing methods including [101] (99.61% P1) and [102] (98.10% P2). The comparison highlights consistent improvements across both datasets, particularly in P1 and P3 metrics, demonstrating the robustness and effectiveness of our approach.

Table 3.13: Comparisons with works using the CASIA and IITD datasets.

Year	Ref	Data	P1	P2	P3
2017	[103]	CASIA	98.00%	–	–
		IITD	97.98%	–	–
2018	[104]	CASIA	–	97.75%	–
		IITD	–	–	–
2019	[105]	CASIA	–	97.02%	–
		IITD	–	94.79%	–
2022	[99]	CASIA	99.11%	–	–
		IITD	98.43%	–	–
2020	[106]	CASIA	–	93.90%	–
		IITD	–	95.40%	–
2020	[107]	CASIA	–	–	–
		IITD	97.11%	–	–
2018	[108]	CASIA	–	98.06%	–
		IITD	–	–	–
2019	[109]	CASIA	–	–	–
		IITD	98.00%	–	–
2020	[110]	CASIA	–	98.66%	–
		IITD	–	97.85%	–
2022	[102]	CASIA	–	98.77%	–
		IITD	–	98.10%	–
2018	[111]	CASIA	–	–	–
		IITD	94.70%	–	–
2019	[101]	CASIA	97.65%	–	–
		IITD	99.61%	–	–
2019	[100]	CASIA	–	99.18%	–
		IITD	–	99.01%	–
2023	[112]	CASIA	–	–	–
		IITD	98.09%	–	–
–	**our study**	CASIA	99.13%	99.69%	100%
		IITD	99.46%	98.50%	99.35%

3.4 Conclusion

In this chapter we create a robust framework for finger vein and palmprint recognition, addressing challenges in illumination variability, feature extraction, and computational efficiency. By leveraging dual learning strategies: single and multiple-instance learning, and using preprocessing techniques such as histogram equalization for finger veins, CLAHE for palmprints, and fine-tuning the top layers with MobileNetV2 identified as the optimal architecture the system achieves enhanced accuracy and resilience in case of physical anomalies. Comparative evaluation of classifiers (SVM, KNN, RF) further refine recognition performance.

In the next chapter, we extend this framework by integrating feature-level fusion of finger vein and palmprint modalities. We exploit their complementary traits to strengthen recognition robustness and security in multi-modal biometric systems.

Chapter 4

Feature Level Fusion of Finger Vein and Palmprint ROIs for Multimodal Biometric Systems

4.1 Introduction

Following the detailed analysis of finger vein and palmprint recognition systems in the previous chapter, this chapter introduces a multimodal biometric framework that combines both modalities at the feature level. The aim is to improve system robustness, accuracy, and resistance to issues that affect single-modality systems, such as noise, variability, or partial degradation in individual samples.

The framework begins with preprocessing, where Region of Interest (ROI) extraction is performed using algorithms specifically developed in this research to isolate the most relevant biometric features. Image quality is further improved using CLAHE. MobileNetV2, identified in earlier chapters as an effective feature extractor, is used to obtain deep features from both modalities. These features are then fused into a single vector and classified using SVM, KNN, and Random Forest. This chapter is organized as follows: Section 2 reviews related works, with a focus on feature-level fusion in multimodal biometric systems. Section 3 provides a detailed description of the proposed framework, including the preprocessing, feature extraction, fusion, and classification stages. Section 4 describes the experimental methodology, encompassing the dataset and key findings. Finally, Section 5 concludes the chapter, summarizing the contributions.

4.2 Literature survey

Feature-level fusion has been extensively studied in multimodal biometric systems. Li et al. [113] proposed a Bimodal Palmprint Fusion Network (BPFNet) using palmprint and palm vein modalities due to their rich and complementary discriminative features that utilize attention mechanisms to focus on key discriminative regions, resulting in state-of-the-art results. Similarly, Zhang et al. [114] introduced a hybrid fusion approach combining feature-level and decision-level strategies for a multimodal biometric recognition system using iris, palm vein, and finger vein, achieving improved performance. These studies underscore the potential of feature-level fusion in advancing the performance of multimodal biometric systems. Kumar et al. [115] developed a hybrid multimodal biometric system that combines face and palmprint traits, leveraging deep learning models for feature extraction and fusion.In [116], the extracted elements of multi-stream CNNs with face, iris, and fingerprint modalities were fused by multi-level feature abstraction for identification.Ding et al. [117] proposed a CNN-based deep learning framework for multimodal face information-based recognition. A set of CNNs was implemented for extracting features of multiple face modalities.in [118], combined iris and periocular region modalities-based deep transfer learning methods were proposed for recognition. A VGG model was used for feature extraction, and the feature selection applied the binary particle swarm algorithm. The fusion of two modalities used matching score-level and feature-level fusion methods. Therar et al. [119] used a CNN with a transfer learning approach for multimodal right iris and left iris biometrics recognition. The extraction and classification of iris features were performed using the proposed deep learning and multi-class support vector machine (SVM) algorithm. In Ref. [120], finger vein and finger knuckle print features were extracted using three CNNs, alexnet, VGG16, and resnet50, interacting with transfer learning. The proposed fusion approaches combined these features for classification. The purpose of research [121] was to offer an innovative method for multimodal biometric systems called optimal feature level fusion. The suggested method, called optimal feature level fusion, uses OGWO, an optimization algorithm modeled after grey wolf behavior, to choose the most pertinent features. The outcomes show how the best feature level fusion technique can enhance the multimodal biometric system's performance in terms of recognition. The limitation was requirement for huge datasets for evaluation and training, or potential difficulties in choosing the best features for various biometric variables.The study objective [122] proposed a multimodal biometric identification technique that uses a person's face and iris to prove their identity. A system that uses a Rectangle Histogram of Oriented Gradient (R-HOG) to extract facial attributes before doing feature-level fusion using a special fusion method. The trial findings, the fusion time has been slashed by about 34.5 % in

comparison to other approaches. To recognize and address these issues in real-world applications, evaluation was required

4.3 Proposed System

This section provides a detailed description of the framework for the proposed approach. The process begins with preprocessing the datasets to extract regions of interest (ROI) from finger veins (FV) and palmprints. This step focuses on isolating the most relevant parts of the images. Once the ROIs are extracted, the images are enhanced using CLAHE. During the feature extraction phase, MobileNetV2 processes these enhanced images to obtain distinctive features. The system then creates representations from both left and right palmprints and finger veins, followed by feature concatenation. This feature-level fusion approach integrates complementary information from multiple biometric modalities, resulting in a comprehensive representation that substantially improves the model's performance. The concatenated feature vector is then fed into multiple classifiers.

The overall framework is illustrated in Fig. 4.1.

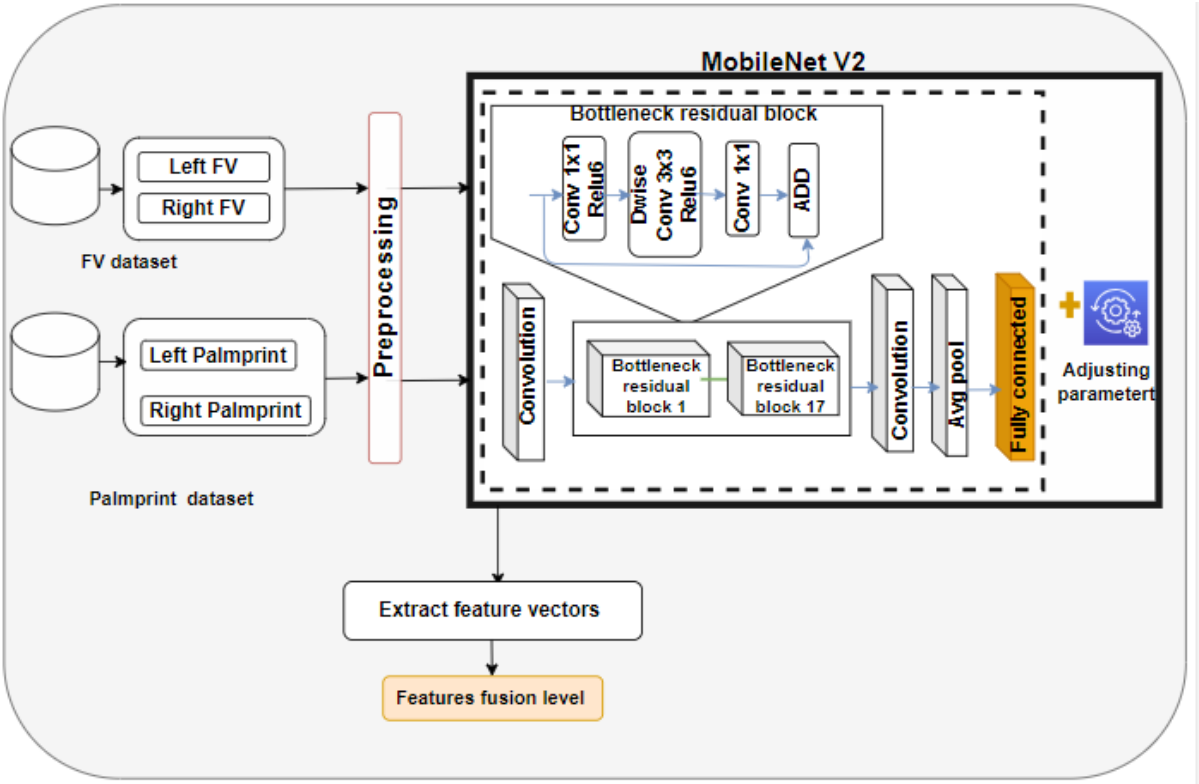


Figure 4.1: Proposed system.

The framework of feature-level fusion is illustrated in Fig. 4.2.

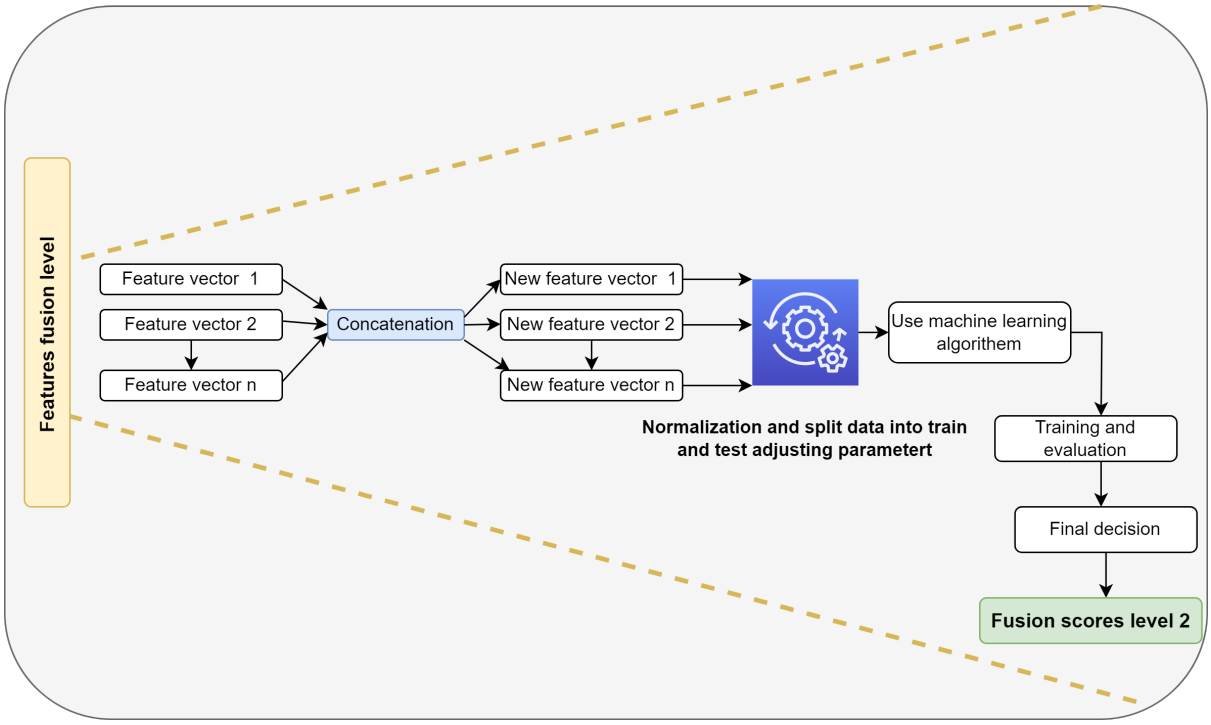


Figure 4.2: Proposed system.

4.3.1 Datasets Description

In this study, we developed a balanced multi-modal biometric framework using samples from two datasets. From the SDUMLA-HMT finger vein dataset, we selected 212 subjects, each providing 6 images per hand (left and right). To maintain modal symmetry, we selected an equivalent number of subjects (212) from the IIT Delhi palmprint dataset. This balanced selection ensures consistent subject distribution across both modalities, which is crucial for fusion-level methods.

To comprehensively evaluate system performance, we implemented three testing protocols with varying numbers of test images per subject. Protocol 1 utilized two images per subject for testing, yielding 424 test images per hand (848 total images across both hands). Protocol 2 employed three test images per subject, resulting in 636 test images per hand (1,272 total images). Protocol 3 implemented a minimal testing ratio with one image per subject, producing 212 test images per hand (424 total images).

Table 4.1 presents the image distribution.

Table 4.1: Testing Protocols Distribution for one modality

Protocol	Images per Subject	Number of Testing Images	Total Images
Protocol 1	2	424	848
Protocol 2	3	636	1272
Protocol 3	1	212	424

4.3.2 Preprocessing

The images of the finger veins and palmprint must first be preprocessed. The key benefit of the preprocessing stage is that it organizes the data, making recognition easier.This section discusses the stages of this operation in depth.

4.3.2.1 Region of Interest for Finger Veins

In this subsection, we describe the method for extracting the region of interest (ROI) of finger veins, as illustrated in Algorithm 1 and Fig. 4.3. The following paragraphs provide detailed explanations of these blocks:

-Grayscale image

Converting a color image to the grayscale domain reduces data complexity and improves processing performance by working with a single channel instead of three in the RGB (Red, Green, and Blue) domain. To convert the RGB-colored image to a grayscale image, a weighted average method was employed. This method was selected because it produces a more accurate grayscale image by assigning different weights to the RGB components as illustrated in Eq. (1) [123].

$$\text{Grayscale} = 0.2989 \cdot \text{Red} + 0.5870 \cdot \text{Green} + 0.1140 \cdot \text{Blue} \tag{1}$$

- Filter Gaussian blur

A Gaussian blur results from blurring an image with a Gaussian function in image processing. It is a common effect in graphics software for eliminating image noise and reducing detail. This blurring technique produces a smooth blur similar to that seen when viewing an image through a translucent screen, as opposed to the bokeh effect generated by an out-of-focus lens or the shadow cast by an object under standard lighting [124].

- Edge detection using the Laplacian operator

Edge detection is a process of capturing image characteristics in image classification, which includes recognizing image discontinuities. The second derivative operator utilized in edge detection is the Laplacian operator. The Laplacian operator can give better glob localization results compared to first derivative-based edge detectors like the Sobel operator [125].

-ROI research and cropping We have two boundaries in the region of interest (ROI) area. To determine these boundaries, we first scan the image from the centerline to the top of the image. A similar operation is performed to obtain the lower boundary of the rectangular area in the ROI, scanning the image from the centerline to the bottom. The average of these boundaries defines the ROI area.

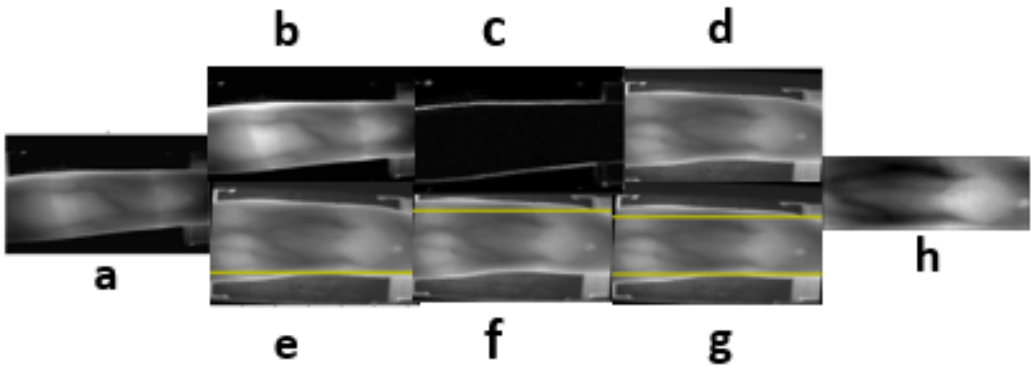


Figure 4.3: ROI extraction for finger vein images: (a) original image, (b) grayscale image, (c) edge detection, (d) rotation, (e, f, g) upper and lower boundaries, (h) ROI image.

Algorithm 4.1: Finger Vein ROI Extraction

Input: Image I

Output: Finger vein ROI

$G \leftarrow$ CONVERT TO GRAY(I)

$E \leftarrow$ EDGE DETECTION(G)

`// Get finger boundaries`

$B_{\text{upper}}, B_{\text{down}} \leftarrow$ GETBOUNDARIES(E);

`// Fit center line`

$k, b \leftarrow$ FIT CENTER LINE$(B_{\text{upper}}, B_{\text{down}})$;

$I_{\text{rot}} \leftarrow$ ALIGNFINGER(I, k, b);

$ROI \leftarrow$ EXTRACTROI$(I_{\text{rot}}, B_{\text{up}}, B_{\text{down}})$;

$ROI_{\text{final}} \leftarrow$ NORMALIZE ROI(ROI);

4.3.2.2 Region of Interest for Palmprints

This algorithm introduces a method for palmprint extraction, focusing on efficient Region of Interest (ROI) detection. Segmentation Method 1 demonstrates our proposed technique for extracting palmprint ROI with dimensions of 224x224 pixels. In comparison, Segmentation Method 2 represents the ROI extraction method documented in the literature, which produces a 150x150 pixel region. Figure 4.5 illustrates examples from both datasets.

-Binary Thresholding

To separate the hand region from the background, our implementation employs Otsu's thresholding method [126], which automatically determines the optimal threshold value by maximizing the inter-class variance between the foreground and background pixels. Given a grayscale image $I(x, y)$, the binary image $B(x, y)$ is obtained through:

$$B(x,y) = \begin{cases} 255, & \text{if } I(x,y) \geq T \\ 0, & \text{if } I(x,y) < T \end{cases}$$

where T is the optimal threshold value determined by Otsu's method. This threshold T is computed by minimizing the intra-class variance $\sigma_w^2(t)$, defined as:

$$\sigma_w^2(t) = w_1(t)\sigma_1^2(t) + w_2(t)\sigma_2^2(t)$$

where $w_1(t)$ and $w_2(t)$ are the probabilities of the two classes separated by threshold t, and $\sigma_1^2(t)$ and $\sigma_2^2(t)$ are the variances of these classes.

-Morphological Operations

Following binary thresholding, morphological operations are applied to enhance image quality [127]. We employ the closing operation with a rectangular structuring element K of size $k \times k$:

$$\text{Morphed}(x,y) = (B \bullet K)(x,y) = (B \oplus K \ominus K)(x,y)$$

-Contour Extraction and ROI Cropping

The convex hull effectively creates a polygon that encloses the entire hand contour, with special attention to the fingertips and valleys between fingers that serve as key reference points. After applying binary thresholding and morphological operations to clean up the image, the convex hull algorithm identifies the outermost points of the hand shape, creating a complete boundary around it. The rectangular boundary (bounding box) is then calculated based on the convex hull coordinates to define the general region of interest [128] [129]. However, we are interested in the palm area rather than the entire hand, so this bounding box needs refinement. The dimensions are adjusted using specific ratios, typically reducing the height to focus on the palm region below the fingers and adjusting the width to exclude edge areas. This refinement step is essential for isolating the palm region containing the distinctive vein patterns. The extracted palm region undergoes normalization for final processing and standard-

ization through resizing to fixed dimensions 224x224.

Algorithm 4.2: Palmprint Extraction Algorithm

Input: Image I, palm height ratio p_h, palm width ratio p_w, morphological kernel size k

Output: Palmprint image ROI

`ExtractPalmprint` (I, p_h, p_w, k) {

$G \leftarrow$ CONVERT TO GRAY(I)

$B \leftarrow$ BINARY THRESHOLDING(G)

$M \leftarrow$ MORPHOLOGICAL OPERATIONS (B, k)

$C_{\text{largest}} \leftarrow$ FIND LARGEST CONTOUR(M);

$H \leftarrow$ CONVEX HULL(C_{largest});

$(x_1, y_1, w_1, h_1) \leftarrow$ BOUNDING RECT(H);

`// Crop and resize palm region`

$P \leftarrow$ CROP(I, x_1, y_1, w_1, h_1);

$R \leftarrow$ RESIZE$(P, 224, 224)$;

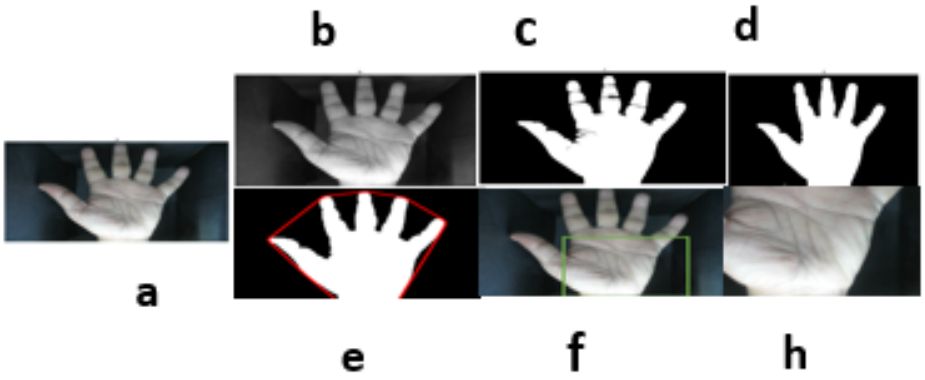
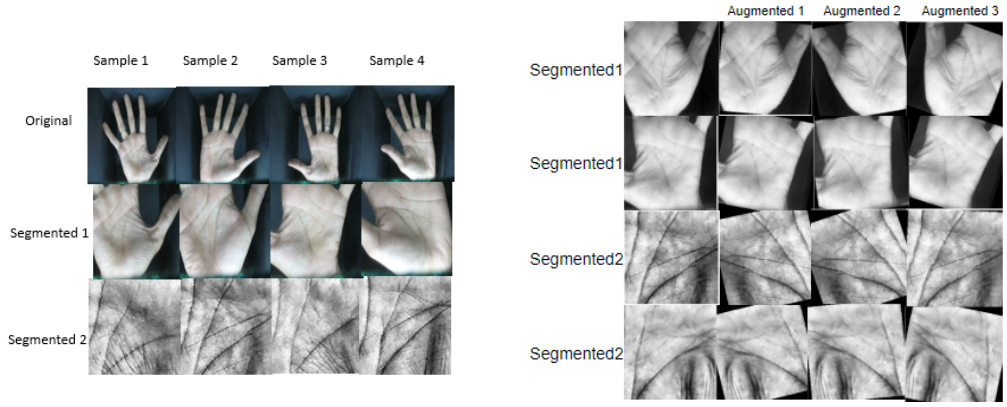


Figure 4.4: ROI extraction for Palmprint images: (a) original image, (b) grayscale image, (c) Binary image, (d) Morphological Operations, (e) Convex hull, (f) ROI image.

4.3.2.3 Data Augmenation

Data augmentation was implemented to enhance the palmprint dataset using two key transformations: random horizontal flipping with 50% probability and random rotation up to 30 degrees. Figure 4.5 illustrates the augmentation results, comparing original palmprint images with their transformed.



(a) Segmented Image **(b)** Augmented Image

Figure 4.5: Segmentation and Augmentation of Palmprint images.

4.4 Feature Level Fusion

We employ MobileNetV2 networks to extract discriminative features from both finger vein and palmprint modalities. Two separate MobileNetV2 networks serve as feature extractors. The first MobileNetV2 processes input from the finger vein modality, while the second processes input from the palmprint modality. After removing the classification layers from both networks, we extract features from the final layers, obtaining feature vectors of 1,280 dimensions for each modality. These high-dimensional feature representations are then concatenated along the channel dimension to create a unified feature vector of 2,560 dimensions, effectively combining complementary information from both modalities, as illustrated in Figure 4.6.

Feature vectors require normalization due to range variations. We use min-max normalization:

$$X' = \frac{X - \min(X)}{\max(X) - \min(X)} \tag{4.1}$$

where $X = \{x_1, x_2, x_3, \ldots, x_n\}$ represents the feature vector.

After normalization, the 1,280-dimensional feature vectors from each modality are concatenated into a fused vector, creating a comprehensive 2,560-dimensional feature space that preserves the raw information content from both biometric traits:

$$F_{\text{fused}} = [v_1, v_2, \ldots, v_{1280}, p_1, p_2, \ldots, p_{1280}] \tag{4.2}$$

where $V = \{v_1, v_2, \ldots, v_{1280}\}$ represents the feature vector of the finger vein modality, and $P = \{p_1, p_2, \ldots, p_{1280}\}$ represents the feature vector of the palmprint modality.

The concatenated feature vector F_{fused} combines the discriminative characteristics of both biometric modalities. This fused feature vector is subsequently used for classification using three machine learning algorithms: Support Vector Machine (SVM),

Random Forest (RF), and K-Nearest Neighbors (KNN).

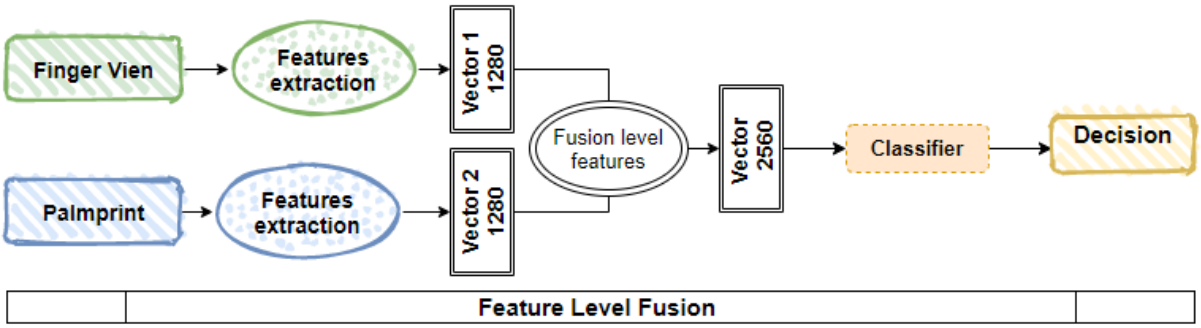


Figure 4.6: Multimodal Fusion at Feature Level.

4.5 Experimental Results and Discussion

In this study, we conducted three experiments to enhance multimodal biometric recognition. First, we explored left and right palmprint and finger vein patterns as baseline biometric systems. Second, we enhanced image quality using Contrast Limited Adaptive Histogram Equalization (CLAHE) preprocessing and augmented the dataset through data augmentation techniques. Finally, we implemented a feature-level fusion approach, combining palmprint and finger vein modalities.

The system optimization employed an RMSprop optimizer with batch sizes of 16 and 32, trained across 200 epochs. Performance evaluation was conducted using precision, recall, F1-score, Rank-1, Rank-5 recognition rates, and Area Under the Curve (AUC), as detailed in Section 2.4. System performance was visualized through Cumulative Match Characteristic (CMC) curves for each experimental trial.

4.5.1 First Experiment

In this experiment, we utilize biometric data from two modalities: palmprint and finger vein images, capturing samples from both the left and right hands. Based on findings from previous chapters, we employ MobileNet V2 fine-tuning, which demonstrated superior results in earlier experiments. Results obtained from both datasets are presented in Tables 4.2 and 4.3.

Next, we extract the region of interest (ROI) from both modalities and enhance the biometric images using contrast-limited Adaptive Histogram Equalization (CLAHE) to improve their quality and contrast. Additionally, we implement data augmentation techniques to expand our training dataset. Results obtained from both datasets after these enhancements are presented in Tables 4.2 and 4.4.

The results presented in Tables 4.2 (A) and 4.2 (B) show that the finger vein recognition system using CLHAE preprocessing achieves superior performance compared to the system without CLHAE. This finding underscores the importance of enhancing all finger vein images used in the recognition system to address poor quality and shading issues in various finger regions. Therefore, only the preprocessed finger vein images will be used in subsequent experiments.

Table 4.2: Performance Metrics for 3 Protocols for Finger Vein Using Two Scenarios: A) Without Preprocessing, B) With Preprocessing.

Protocol	Scenario	Instances	Pr(%)	Rec(%)	F1(%)	R1(%)	R5(%)	AUC(%)
P1	A	Left	98.00	97.00	97.00	96.70	99.00	98.00
		Right	97.00	96.00	96.00	95.90	98.00	97.00
		MI	96.82	95.75	95.54	96.78	98.98	99.98
	B	Left	97.56	97.88	97.44	97.88	99.29	100.00
		Right	98.27	97.41	97.23	97.41	98.58	99.97
		MI	98.43	98.35	98.08	98.35	99.29	99.99
P2	A	Left	96.00	95.00	95.00	95.30	98.00	97.00
		Right	96.00	95.00	95.00	95.70	97.50	95.00
		MI	96.31	96.23	95.75	98.58	99.65	100.00
	B	Left	96.46	97.64	96.86	98.27	99.21	99.99
		Right	97.23	96.23	96.08	96.23	98.27	99.94
		MI	97.26	96.78	96.72	95.75	98.82	100.00
P3	A	Left	97.00	98.00	97.00	97.80	98.00	96.00
		Right	97.00	98.00	97.00	98.20	98.50	97.00
		MI	97.96	97.64	97.60	96.23	99.06	99.96
	B	Left	98.58	99.06	98.74	99.06	99.53	100.00
		Right	98.71	98.27	98.24	97.64	98.58	100.00
		MI	98.87	98.58	98.52	98.58	99.65	100.00

Table 4.3: Performance Metrics for 3 Protocols using palmprint without preprocessing.

	Dataset	Instances	Pr(%)	Rec(%)	F1(%)	R1(%)	R5(%)	AUC(%)
P1	Segmented 1	Left	97	96	96	95.54	97.80	98
		Right	97	98	97	98.35	99	99
		MI	98.71	98.35	98.27	98.35	99	99
	Segmented 2	Left	94	93	93	93.20	96	96
		Right	94	95	95	94.50	98	96
		MI	95.61	92.92	93.01	93	98.47	97
P2	Segmented 1	Left	92	88	88	88.06	96	97.89
		Right	98	97	97	96.54	98.50	97
		MI	95.86	94.89	94.79	95	98.19	98
	Segmented 2	Left	89	86	87	86	96	93
		Right	95	94	94	94.34	98	95
		MI	94.78	93.40	93.33	94	97.64	97
P3	Segmented 1	Left	97	97	98	98.11	99.06	99.98
		Right	98	99	98	98.58	99.06	100
		MI	98.90	98.82	98.74	98.82	100	100
	Segmented 2	Left	96	97	97	97.17	99	99.50
		Right	94	95	95	94.50	98.58	99.94
		MI	97.72	96.70	96.49	96.70	98	98

Similarly, the palmprint modality shows good outcomes. Tables 4.3 and 4.4 present the results for two segmentations, each evaluated with and without preprocessing. Segmentation 1 consistently demonstrates superior performance across both techniques

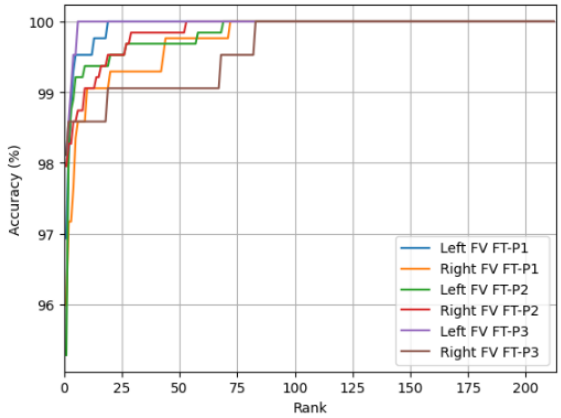


Figure 4.7: CMC curves for SDUMLA-FV without preprocessing.

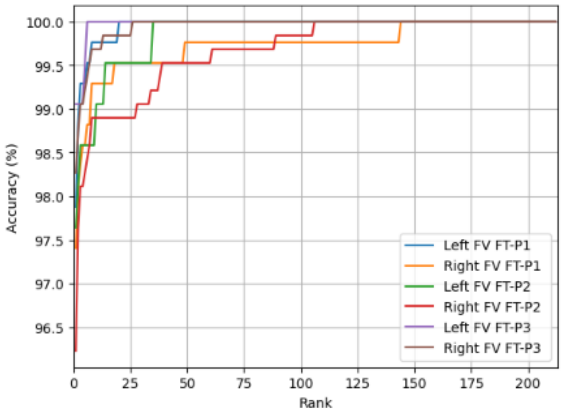


Figure 4.8: CMC curves for SDUMLA-FV with preprocessing.

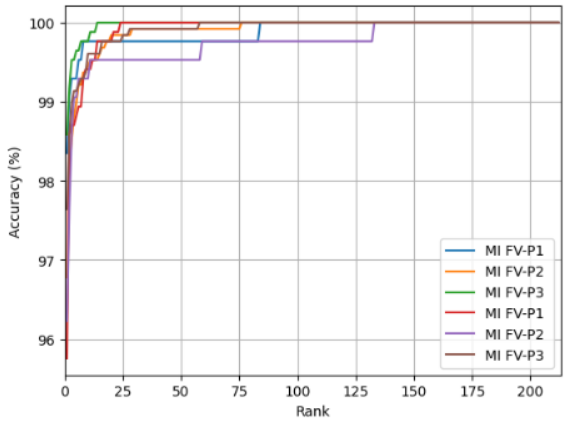


Figure 4.9: CMC curve for SDUMLA-FV-MI.

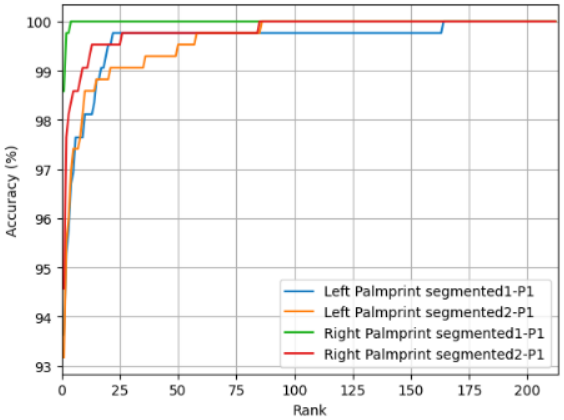


Figure 4.10: CMC curve for palmprint without preprocessing using P1.

compared to Segmentation 2. As a result, the next experiment will focus on using Segmentation 1 with preprocessing.

We also plot the CMC curves for the finger vein modality, both with and without preprocessing, in Figures [4.7 to 4.9]. Similarly, the CMC curves for the palmprint modality are presented in Figures [4.10 to 4.17].

Table 4.4: Performance Metrics for 3 Protocols using palmprint with preprocessing.

	Dataset	Instances	Pr(%)	Rec(%)	F1(%)	R1(%)	R5(%)	AUC(%)
P1	Segmented 1	Left	97.00	98.00	97.00	98.35	99.20	99.00
		Right	98.00	99.00	98.74	99.06	99.00	100.00
		MI	99.56	99.29	99.00	99.29	99.65	100.00
	Segmented 2	Left	96.46	96.70	96.04	96.71	98.59	99.96
		Right	97.72	97.17	96.86	97.17	98.00	98.00
		MI	98.58	98.11	98.12	98.11	99.76	99.97
P2	Segmented 1	Left	97.00	96.00	96.00	96.70	99.00	98.00
		Right	96.00	95.30	95.00	95.30	98.80	98.00
		MI	98.57	98.23	98.25	98.23	99.69	100.00
	Segmented 2	Left	94.89	93.87	93.49	93.88	97.49	98.81
		Right	95.11	94.60	94.89	94.60	97.00	98.00
		MI	98.12	97.29	97.27	97.29	99.65	99.94
P3	Segmented 1	Left	98.00	99.00	98.74	99.06	100.00	100.00
		Right	100.00	99.00	99.00	99.50	100.00	100.00
		MI	99.61	99.28	99.28	99.22	100.00	100.00
	Segmented 2	Left	98.11	98.58	98.27	98.58	100.00	100.00
		Right	98.58	99.06	98.74	99.06	99.00	99.00
		MI	99.37	99.06	98.99	99.06	100.00	100.00

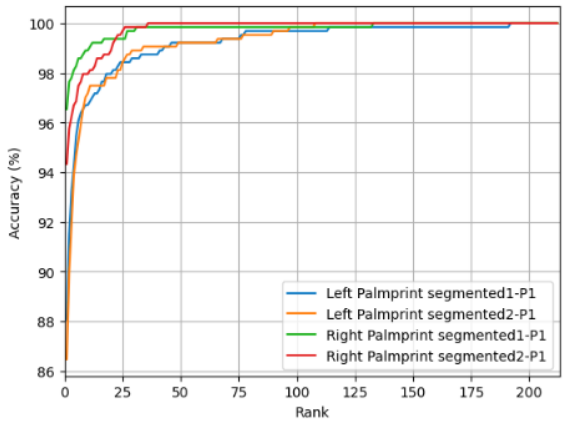


Figure 4.11: CMC curve for palmprint without preprocessing using P2.

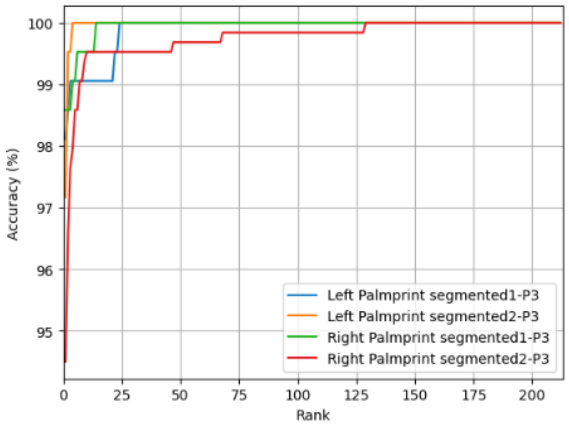


Figure 4.12: CMC curve for palmprint without preprocessing using P3.

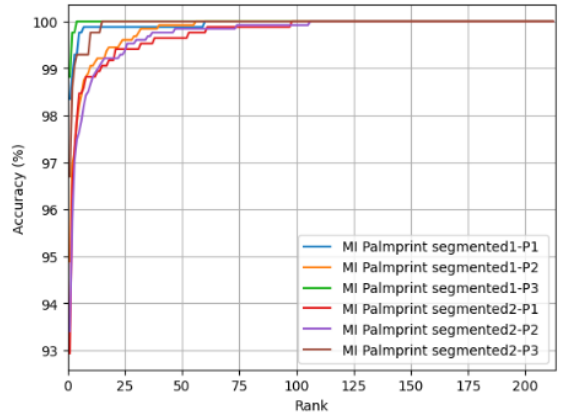


Figure 4.13: CMC curve for palmprint without preprocessing using MI.

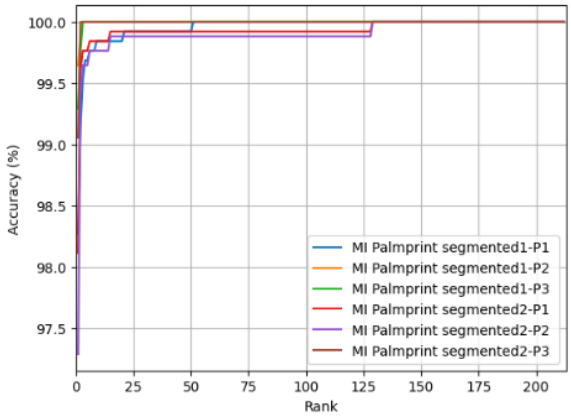


Figure 4.14: CMC curve for palmprint with preprocessing using MI.

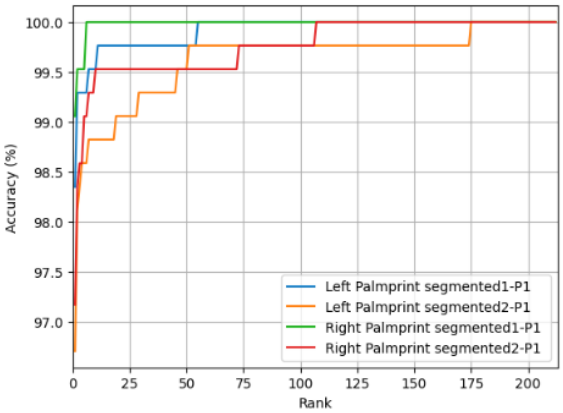


Figure 4.15: CMC curve for palmprint with preprocessing using P1.

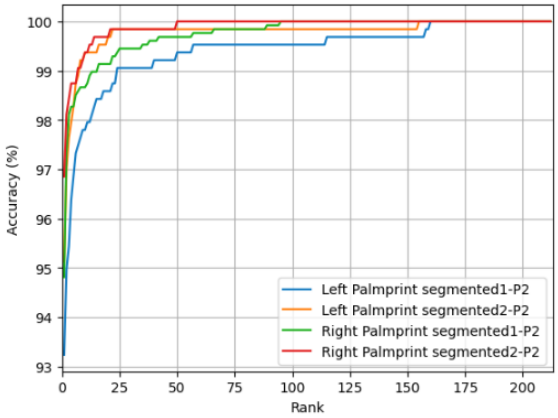


Figure 4.16: CMC curve for palmprint with preprocessing using P2.

4.5.2 Second Experiment

In this experiment, we focus on feature-level fusion for multimodal biometric recognition. We combine the extracted feature vectors from palmprint and finger vein modalities at the feature level before classification. This approach allows us to integrate the discriminative characteristics of both modalities into a unified feature vector, enhancing the overall recognition performance. Table 4.5 presents the types of fusion used in this study.

Table 4.5: Fusion Types

Fusion Type	instances
Fusion 1 (F1)	Left Finger Vein (FV) and Left Palmprint
Fusion 2 (F2)	Right Finger Vein (FV) and Right Palmprint
Fusion 3 (F3)	Left Finger Vein (FV) and Right Palmprint
Fusion 4 (F4)	Right Finger Vein (FV) and Left Palmprint
Fusion 5 (F5)	Both Hands (Finger Veins and Palmprints)

The tables 4.6, 4.7, and 4.8 present a comprehensive comparison of the performance metrics for different fusion methods across three protocols (P1, P2, and P3), using three classifiers: Support Vector Machine (SVM), Random Forest (RF), and K-Nearest Neighbor (KNN). The metrics include precision (pr), recall (rec), F1-score (F1), rank-1 accuracy (R1), rank-5 accuracy (R5), and area under the curve (AUC).

Table 4.6: Performance Metrics for Protocol 1 using features level fusion.

Protocols	Fusion	Methods	Pr(%)	Rec(%)	F1(%)	R1(%)	R5(%)	AUC(%)
P1	Fusion 1	SVM	100	100	100	99.65	99.76	100
		RF	94.05	93.70	94	93.04	99.17	98.90
		KNN	100	100	100	99.76	99.88	99.94
	Fusion 2	SVM	99.29	99.53	99.37	99.53	100	100
		RF	99	99.50	99.30	99.06	100	99.99
		KNN	99	99	99	99.65	99.80	99.80
	Fusion 3	SVM	100	100	100	100	100	100
		RF	98	97.88	97.23	97.88	98.11	99.89
		KNN	99	99	99	99.41	100	100
	Fusion 4	SVM	100	100	100	99.60	100	100
		RF	97.17	98.11	97.47	98.11	99.06	99.98
		KNN	100	100	100	99.80	100	99.89
	Fusion 5	SVM	100	100	100	99.97	100	100
		RF	99	98	98	98.47	99.62	99
		KNN	100	100	100	100	100	100

Analyzing the performance metrics for Protocol 1, we observe that SVM achieves perfect scores 100% in precision, recall, and F1-score across all fusion methods, indicating exceptional performance. KNN also achieves perfect scores in these metrics, while RF maintains high accuracies above 94%. This suggests that Protocol 1 is highly effective for feature-level fusion, with SVM and KNN classifiers performing exceptionally well.

For Protocol 2, the performance metrics show greater variability. SVM maintains excellent performance with scores ranging from 99.21% to 100%, while KNN's pre-

Table 4.7: Performance Metrics for Protocol 2 using features level fusion.

Protocols	Fusion	Methods	Pr(%)	Rec(%)	F1(%)	R1(%)	R5(%)	AUC(%)
	Fusion 1	SVM	100	100	100	99.21	99.74	100
		RF	88	87	86	86.90	99	98
		KNN	97	96.75	96.64	96.46	99	99.94
	Fusion 2	SVM	99.21	99.53	99.37	99.53	100	100
		RF	86.34	83.49	82.55	83.96	96.17	99
		KNN	96.50	94.71	94.75	95.34	98.91	99
P2	Fusion 3	SVM	100	100	100	99.50	100	100
		RF	99.29	99.53	99.37	86.43	100	100
		KNN	95	94	93	94.61	98.11	99.80
	Fusion 4	SVM	99.53	99.53	99.37	99.53	100	100
		RF	82.62	81	82	81.08	97.50	98.57
		KNN	97	97	97	96.59	99	100
	Fusion 5	SVM	100	100	100	99.80	99.95	100
		RF	97	96	96	96.49	99.48	99
		KNN	100	100	100	99.86	99.95	99.97

Table 4.8: Performance Metrics for Protocol 3 using features level fusion.

Protocols	Fusion	Methods	Pr(%)	Rec(%)	F1(%)	R1(%)	R5(%)	AUC(%)
		SVM	100	100	100	100	100	100
	Fusion 1	RF	100	100	100	100	99.53	100
		KNN	100	100	100	100	100	100
		SVM	100	100	100	100	100	100
	Fusion 2	RF	97.88	98.58	98.11	98.58	99.53	100
		KNN	99.29	99.53	99.37	99.53	100	100
P3		SVM	100	100	100	100	100	100
	Fusion 3	RF	99.29	99.53	99.37	99.53	100	100
		KNN	100	100	100	100	100	100
		SVM	100	100	100	100	100	100
	Fusion 4	RF	93.79	95.75	94.42	95.75	100	99.98
		KNN	99	99	99	98.60	100	100
		SVM	99	100	99	99.75	100	100
	Fusion 5	RF	93.79	95.75	94.42	99.75	100	99.98
		KNN	99	100	99	99.75	100	100

cision drops significantly to 83–89%, indicating increased task complexity. RF also shows lower performance, with precision and recall values ranging from 81% to 99.53%. This suggests that Protocol 2 poses a greater challenge for classification, particularly for KNN.

Protocol 3 demonstrates the most robust performance across classifiers, particularly in Fusions 1–3, where even KNN achieves near-perfect scores 98–100%. SVM consistently achieves perfect scores 100% across all metrics and fusion methods, while RF maintains high accuracies above 93%. This indicates that Protocol 3 provides a highly consistent and effective environment for feature-level fusion.

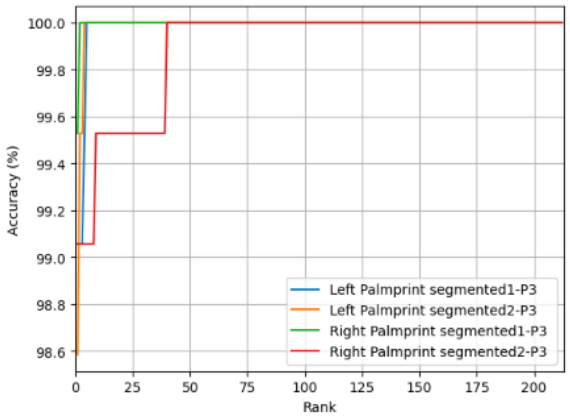


Figure 4.17: CMC curve for palmprint with preprocessing using P3.

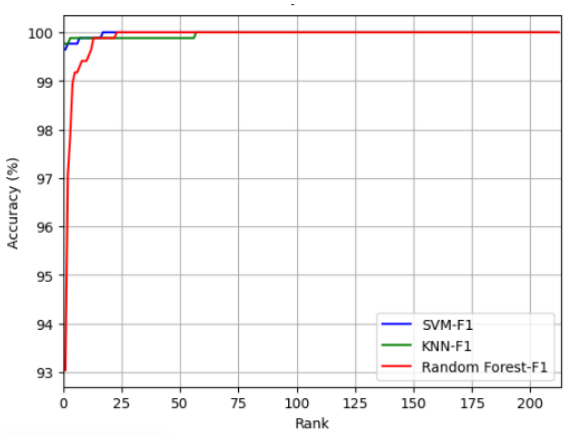


Figure 4.18: CMC curve for fusion level features 1 using P1.

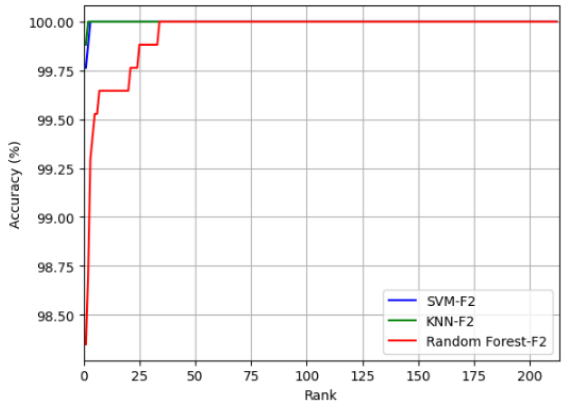


Figure 4.19: CMC curve for fusion level features 2 using P1.

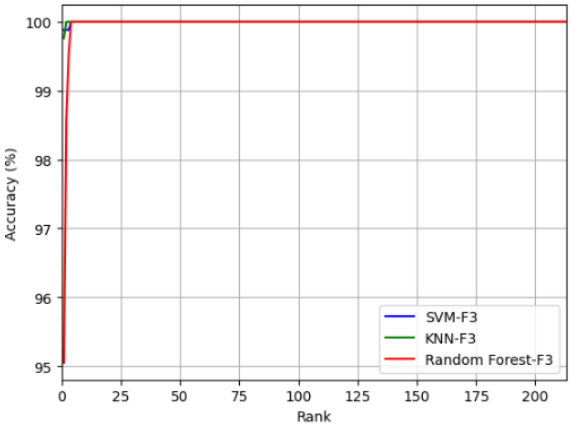


Figure 4.20: CMC curve for fusion level features 3 using P1.

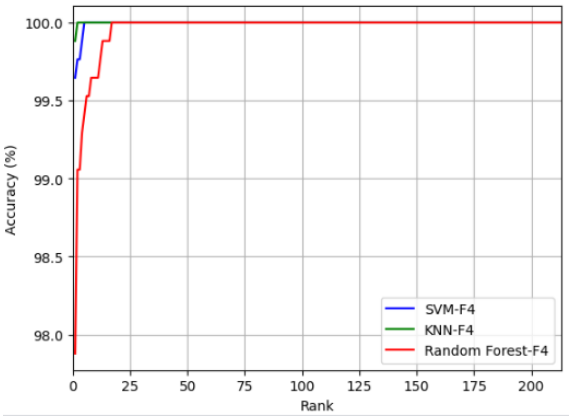


Figure 4.21: CMC curve for fusion level features 4 using P1.

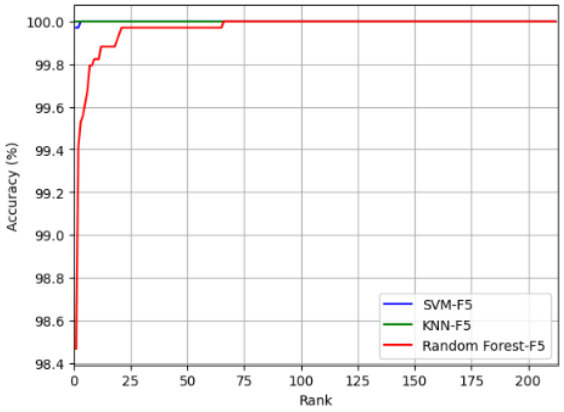


Figure 4.22: CMC curve for fusion level features 5 using P1.

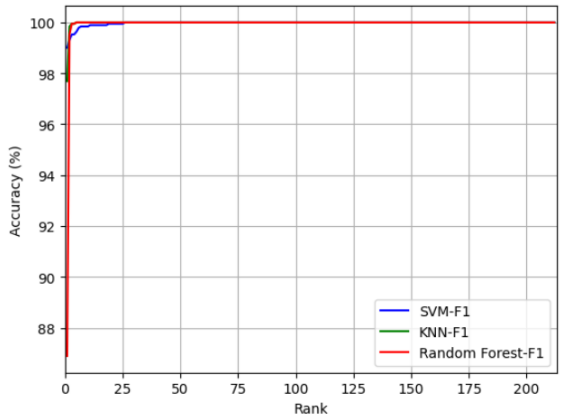


Figure 4.23: CMC curve for fusion level features 1 using P2.

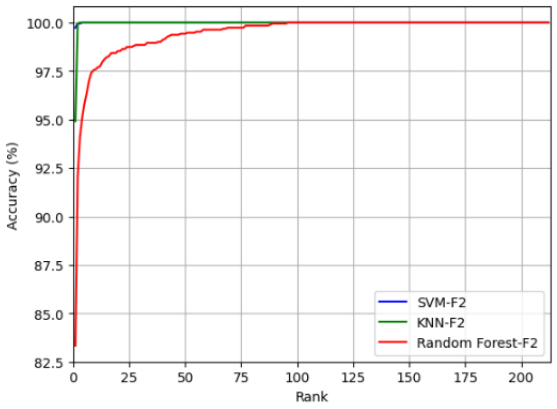


Figure 4.24: CMC curve for fusion level features 2 using P2.

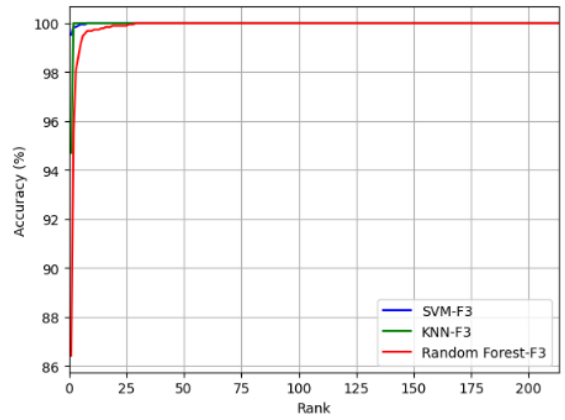


Figure 4.25: CMC curve for fusion level features 3 using P2.

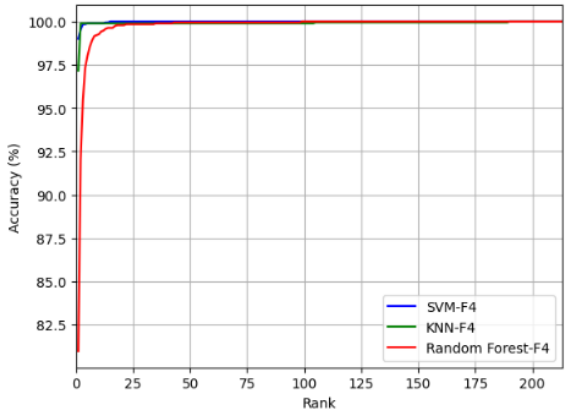


Figure 4.26: CMC curve for fusion level features 4 using P2.

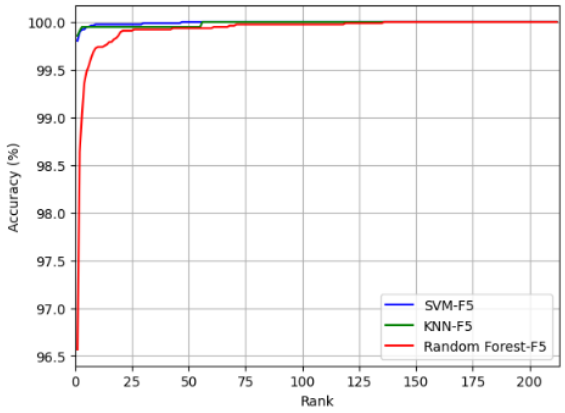


Figure 4.27: CMC curve for fusion level features 5 using P2.

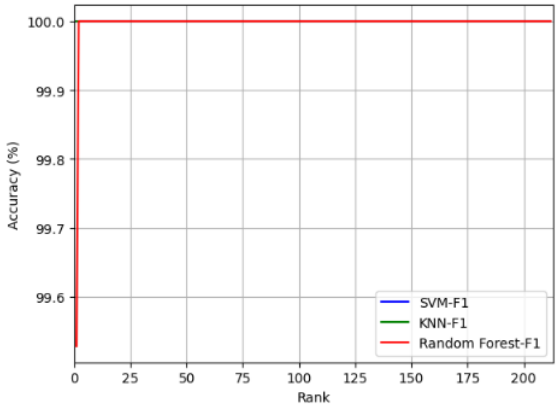


Figure 4.28: CMC curve for fusion level features 1 using P3.

4.5.3 Comparison of Unimodal and Multimodal Feature-Level Fusion

Table 4.9: Performance Metrics for Finger vein, Palmprint, and Fusion Approaches level features.

Datasets		Protocol	Pr(%)	Rc(%)	F1(%)	R1(%)	R5(%)	AUC(%)
Finger Vein	Left	P1	97.56	97.88	97.44	97.88	99.29	100
		P2	96.46	97.64	96.86	98.27	99.21	99.99
		P3	98.58	99.06	98.74	99.06	99.53	100
	Right	P1	98.27	97.41	97.23	97.41	98.58	99.97
		P2	97.23	96.23	96.08	96.23	98.27	99.94
		P3	98.71	98.27	98.24	97.64	98.58	100
Palmprint	Left	P1	97	98	97	98.35	99.20	99
		P2	97	96	96	96.70	99	98
		P3	98	99	98.74	99.06	100	100
	Right	P1	98	99	98.74	99.06	99	100
		P2	96	95.30	95	95.30	98.80	98
		P3	100	99	99	99.50	100	100
Fusion	Fusion 1	P1	100	100	100	99.76	99.88	99.94
		P2	100	100	100	99.21	99.74	100
		P3	100	100	100	100	100	100
	Fusion 2	P1	99	99	99	99.65	99.80	99.80
		P2	99.21	99.53	99.37	99.53	100	100
		P3	100	100	100	100	100	100
	Fusion 3	P1	100	100	100	100	100	100
		P2	100	100	100	100	99.50	100
		P3	100	100	100	100	100	100
	Fusion 4	P1	100	100	100	99.80	100	99.89
		P2	99.53	99.53	99.37	99.53	100	100
		P3	100	100	100	100	100	100
	Fusion 5	P1	100	100	100	100	100	100
		P2	100	100	100	99.86	99.95	99.97
		P3	100	100	100	100	100	100

Table 4.9 shows that fusion methods improve biometric system performance. However, some results are not perfect and need improvement. While most fusion techniques achieve very high scores, there are still some areas where fusion needs enhancement. To address this, we can explore better ways to combine features, employ improved machine learning methods, or optimize fusion level scores. Despite these minor issues, fusion remains a powerful method because it reduces errors and increases system reliability. By refining fusion techniques, we can further improve the accuracy and effectiveness of the system.

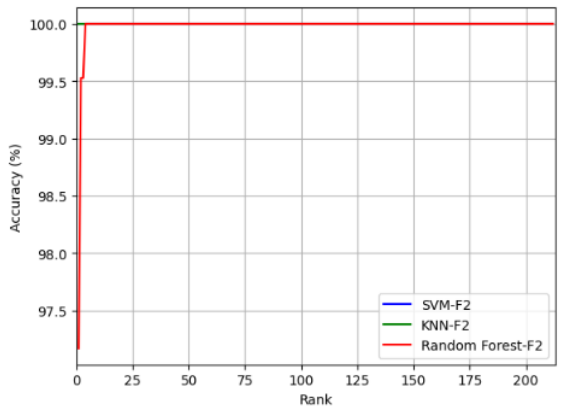


Figure 4.29: CMC curve for fusion level features 2 using P3.

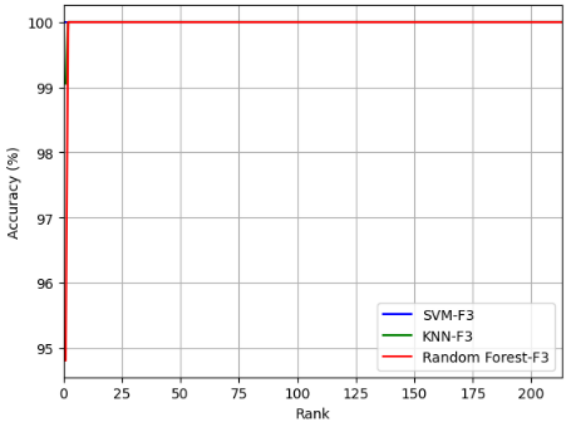


Figure 4.30: CMC curve for fusion level features 3 using P3.

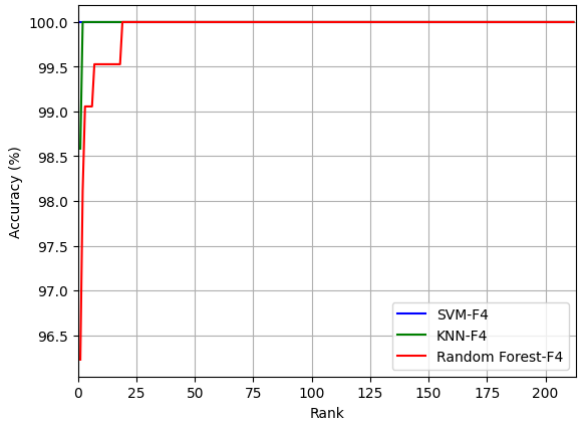


Figure 4.31: CMC curve for fusion level features 4 using P3.

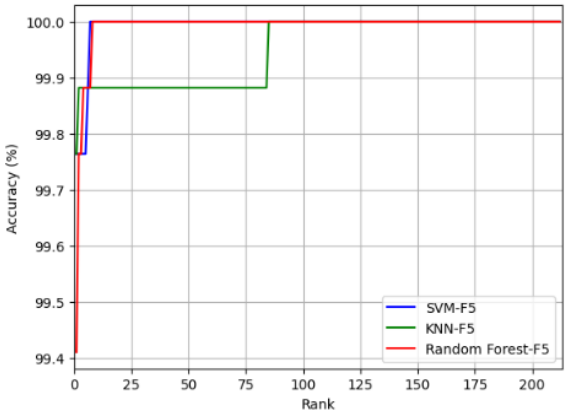


Figure 4.32: CMC curve for fusion level score F1 using P1.

4.6 Conclusion

This chapter introduced a multimodal biometric recognition framework using feature-level fusion of finger vein and palmprint modalities. We implemented ROI extraction, image enhancement with CLAHE, and feature extraction using MobileNetV2. The fused features were classified with SVM, RF, and KNN algorithms, yielding improved recognition performance. In the next chapter, we will explore score-level fusion to further enhance system accuracy and reliability.

Chapter 5

Score Level Fusion Methods for Finger Vein and Palmprint Biometrics

5.1 Introduction

In multimodal biometric systems, combining data from different sources early on can theoretically lead to better results. However, this approach, known as feature-level fusion, faces several challenges [130, 131]. These include:

- Potential incompatibility between feature-level data sets, which can arise from differences in data structure and quality.
- The risk of dimensional redundancy, where the size of the combined feature vectors increases significantly, leading to computational inefficiencies and reduced performance.

Because of these challenges, feature-level fusion is not widely used. Instead, most research focuses on score-level fusion, which combines data at a later stage. This chapter will explore the advancements and challenges in score-level fusion techniques, discussing how they address the limitations of feature-level fusion and improve the overall performance of multimodal biometric systems.

This chapter presents two experiments to evaluate score-level fusion methods in a multimodal biometric system that combines finger vein and palmprint data.

In the first experiment, we utilize the outputs of MobileNetV2 before feature-level fusion, specifically the Softmax outputs.

In the second experiment, we combine the output probabilities of several machine learning classifiers, including SVM, KNN, and Random Forest. We apply a range of score-level fusion techniques in both experiments, such as average, weighted sum, product, Borda count, max, harmonic, and Sugeno. This comprehensive exploration

of score-level fusion techniques enables us to identify the most effective combination strategy for our specific multimodal system.

5.2 Literature survey

Multimodal biometric systems have gained significant attention for their potential to enhance security and accuracy in identification processes. This survey examines recent advancements and challenges in score-level fusion techniques within these systems. The article [132] introduced a multimodal biometric framework that integrates fingerprint, palm print, and facial recognition through score-level fusion. The system employs two feature extraction methods: Accelerated Robust Features (SURF) and Scale-Invariant Feature Transform (SIFT). The weighted sum rule and fusion rules for score comparison enhance the system's performance. However, multimodal biometric systems may face challenges such as increased processing complexity, sensitivity to environmental conditions, and potential vulnerabilities to attacks. The study [133] explored the practical applications of soft computing techniques in biometric systems. The research proposed a hybrid optimization tool called the "Adaptive Fuzzy Genetic Algorithm (AFGA)," which can be adapted for both unimodal and multimodal biometric identification systems. This method effectively combines multimodal biometrics with fuzzy logic to achieve high verification accuracy. Key considerations include processing complexity, scalability, and potential attack vulnerabilities. In [134], the authors aimed to improve accuracy by fusing biometric features at the score level using fingerprint and iris data. Convolutional Neural Networks (CNN) were employed to extract features. The proposed biometric verification system, which integrates multiple biometric traits, outperforms single-trait systems. Challenges for multimodal biometric systems include data quality, scalability, and attack vulnerability. Kabir et al. [135] proposed fusing face profile and gait features at the matching score level. They extracted side face features (EFSI) and gait features (GEI) from video and fused them using three strategies. In another work, Kabir et al. [136] proposed matching-level fusion of fingerprint verification using minutiae and texture matching algorithms. Kim et al. [137] developed a deep CNN for finger-vein and finger shape modalities, capturing finger images with a near-infrared camera sensor. The matching distance scores of features were fused using weighted sum, product, and perceptron methods. AlWaisy et al. [138] proposed a deep belief network (DBN) for facial feature extraction and recognition, and a CNN for left and right iris feature extraction and classification. The scores were combined using score-level and rank-level fusion methods for multimodal biometric identification.

5.3 Proposed System

The proposed system presents two multimodal biometric score-level fusion strategies. Fusion score level 1 uses direct Softmax classifier output fusion using methods such as average, product, weighted sum, Borda count, max, harmonic, and Sugeno. In contrast, fusion scores level 2 utilizes an approach involving the outputs after feature concatenation from multiple classifiers (SVM, KNN, and Random Forest) before applying score-level fusion techniques. Figure 5.1 illustrates the proposed system.

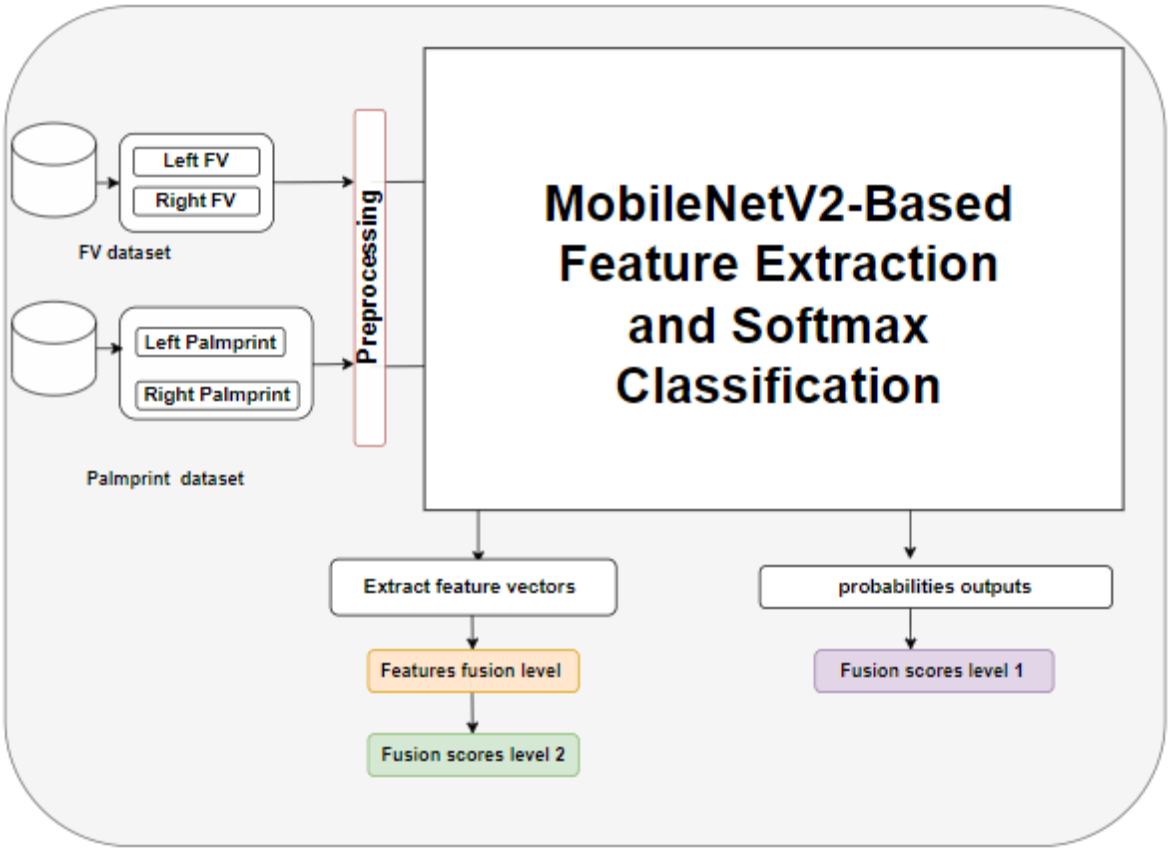


Figure 5.1: Proposed system.

The framework of fusion level scores is illustrated in Fig. 5.2.

After feature extraction using MobileNetV2, separate classifiers are used for each modality, generating probability scores for each class using the softmax function:

$$P(y|x) = \text{softmax}(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \tag{5.1}$$

where z_i represents the classifier output for class i, and K is the number of classes. These probability scores are then combined using the following fusion rules [139]:

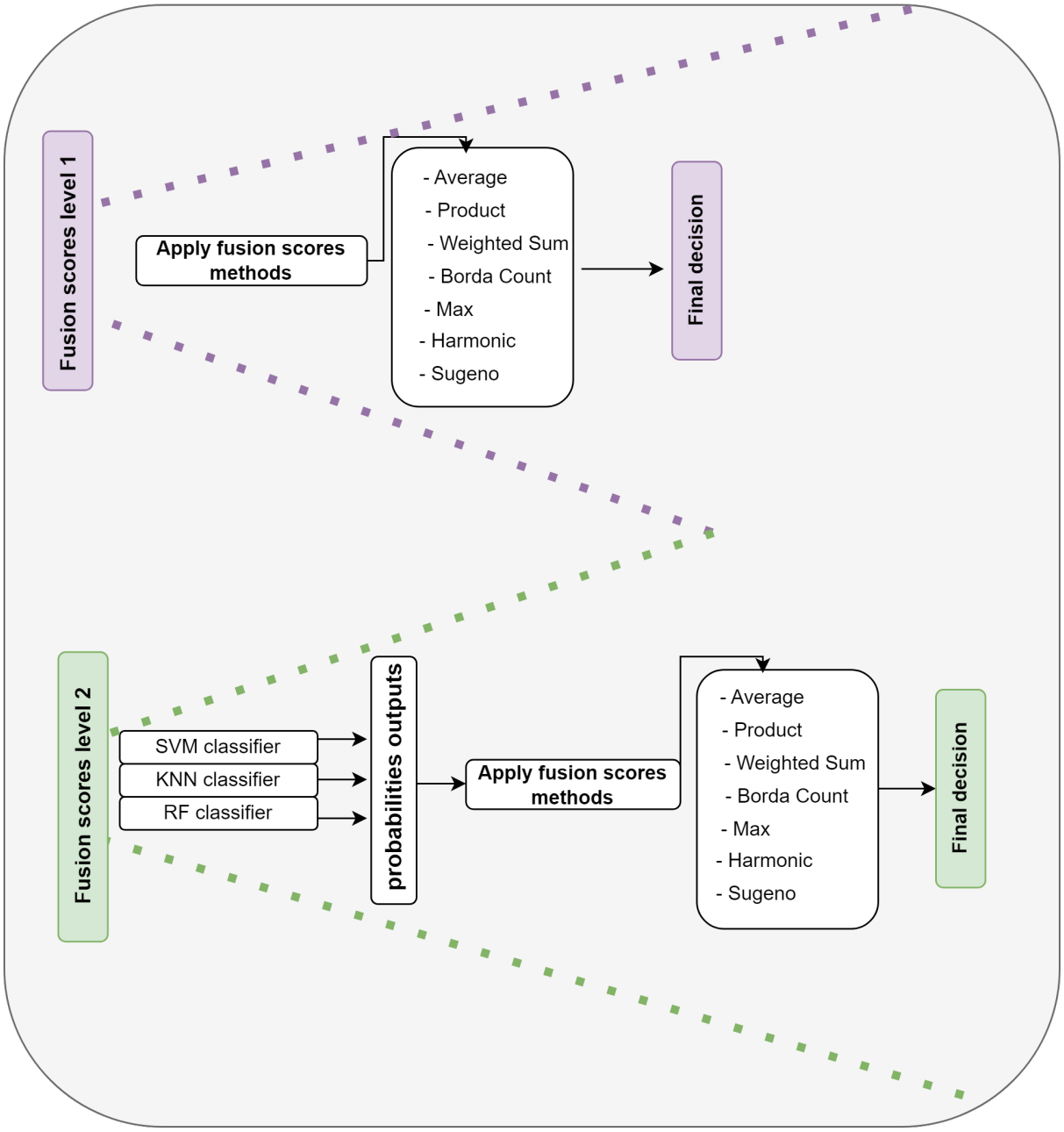


Figure 5.2: Proposed System for Score-Level Fusion.

5.3.1 Max Rule

This method selects the highest score among the models' predictions to make the final decision. The final class prediction is:

$$S_{\max} = \max\{P_1, P_2, \ldots, P_n\} \tag{5.2}$$

where P_i represents the prediction score from the i-th model.

5.3.2 Weighted Sum Rule

This method computes a weighted sum of the prediction scores, where the weights reflect the importance of each model. The final score is:

$$S_{\text{weighted}} = \sum_{i=1}^{n} w_i P_i \tag{5.3}$$

where w_i is the weight assigned to the i-th model, and $\sum_{i=1}^{n} w_i = 1$.

5.3.3 Product Rule

This method involves multiplying the prediction scores from each model together to make the final prediction. The final class prediction is:

$$S_{\text{product}} = \prod_{i=1}^{n} P_i \tag{5.4}$$

5.3.4 Average Rule

This method computes the average of the prediction scores across all models. The final score is:

$$S_{\text{avg}} = \frac{1}{n} \sum_{i=1}^{n} P_i \tag{5.5}$$

where n is the number of models.

5.3.5 Borda Count

This rule, typically used for rank-level fusion, is adapted here for score-level fusion. It assigns ranks to the predicted classes based on their probabilities for each model. The ranks are then summed across all models, and the class with the highest sum of ranks is selected [140]:

$$C_p = \sum_{j=1}^{N} rank_j(p) \tag{5.6}$$

where:

- C_p represents the combined Borda count for class p.
- N is the number of models.
- $rank_j(p)$ denotes the rank assigned to class p by model j.

5.3.6 Harmonic Mean Rule

This method calculates the harmonic mean of the prediction scores from each model. The harmonic mean is useful if the models' predictions are skewed. The final class prediction is:

$$S_{\text{harmonic}} = \left(\frac{1}{n} \sum_{i=1}^{n} \frac{1}{P_i} \right)^{-1} \tag{5.7}$$

5.3.7 Sugeno integral

The Sugeno integral [141–143] is defined by considering μ as a fuzzy measure on a set $A = \{a_1, a_2, \ldots, a_n\}$. The ranges of these elements are represented by a function r mapping A to the interval $[0, 1]$ such that $r(a_1) \leq r(a_2) \leq \cdots \leq r(a_n)$.

The Sugeno integral can be expressed as:

$$\int_{i=1}^{n} r(a)\, d\mu = \max_{1 \leq x \leq n} \left(\min \left(r(a_x), \mu(A_x) \right) \right) \tag{5.8}$$

The fuzzy measures of $\{a_i\}$, denoted as $\mu(\{a_i\})$, are primarily determined by domain experts for a specific statistical problem. These fuzzy measures assist in calculating $\mu(A_x)$ using the Sugeno λ-fuzzy measures, where $\lambda \in (-1, \infty)$, defined by:

$$1 + \lambda = \prod_{i=1}^{n} \left(\mu(\{a_i\})^{\lambda} + 1 \right) \tag{5}$$

After obtaining the values of λ from Eq. (5.9), the value of $\mu(A_x)$ is calculated using:

$$\mu(\{a_l, a_o\}) = \mu(\{a_l\}) + \mu(\{a_o\}) + \lambda \mu(\{a_l\}) \mu(\{a_o\}) \tag{5.10}$$

where $1 \leq l, o \leq n$. Algorithm [142, 144] provides a comprehensive explanation of

the Sugeno integral method.

Algorithm 5.1: Sugeno Integral Algorithm

Input: C, P, F

Output: Fused Score

$\lambda \leftarrow$ initialize by solving Equation (5.9);
predictions $\leftarrow$ initialize empty list of final predictions;
for *class_index* $\in \{1, 2, 3, \ldots, C\}$ **do**
    $P_\pi \leftarrow$ Sorted array P in descending order;
    $F_\pi \leftarrow$ Permutation of F corresponding to P_π;
    $f_{\text{prev}} \leftarrow F_\pi[0]$;
    pred $\leftarrow \min(P_\pi[0], F_\pi[0])$;
    for $n \in \{1, 2, 3, \ldots, C\}$ **do**
        $f_{\text{current}} \leftarrow f_{\text{prev}} + F_\pi[n] + \lambda \times F_\pi[n] \times f_{\text{prev}}$;
        pred $\leftarrow \max(\text{pred}, \min(P_\pi[n], f_{\text{current}}))$;
        $f_{\text{prev}} \leftarrow f_{\text{current}}$;
    end
    predictions$[i] \leftarrow$ pred;
end
class$(I) \leftarrow \arg\max(\text{predictions}[i])$;

5.4 Experimental Results and Discussion

5.4.1 First Exepriment

In this experiment, we evaluate the performance of various fusion techniques such as average, weighted sum, product, Borda count, max, harmonic, and Sugeno. The evaluation metrics include precision (PR), recall (Rec), F1-score (F1), rank-1 accuracy (R1), rank-5 accuracy (R5), and area under the curve (AUC). The results are presented in Tables 5.1, 5.2, and 5.3.we also present the CMC curves, limited to the top 10 ranks, to highlight the most critical performance range. These curves are displayed in Figures 5.3 to 5.17.

The first experiment, which utilized score-level fusion with softmax, demonstrated remarkably high performance across all protocols (P1, P2, and P3), the fusion methods including : Average, Product, Weighted Sum, Borda Count, Max, Harmonic, and Sugeno exhibit consistently high performance, with most metrics (precision, recall, F1 score, R1, R5, and AUC) approaching or achieving 100%. Notably, the Average, Product, and Max fusion methods consistently achieve perfect scores across all metrics for all protocols, indicating their robustness and reliability in this context. The Borda Count method, while slightly lower in performance compared to others, still

Table 5.1: Performance Metrics for Protocol 1 Using Score-Level Fusion with Softmax.

Fusion	Methods	pr(%)	rec(%)	F1(%)	R1(%)	R5(%)	AUC(%)
Fusion1	Average	100	100	100	100	100	100
	Product	100	100	100	100	100	100
	Weighted Sum	99.84	99.76	99.75	99.76	100	100
	Borda Count	99.53	99.29	99.25	99.29	99.76	99.99
	Max	100	100	100	100	100	100
	Harmonic	99.84	99.76	99.75	99.76	99.76	100
	Sugeno	99.53	99.29	99.25	99.06	100	100
Fusion2	Average	100	100	100	100	100	100
	Product	99.84	99.76	99.75	99.76	100	100
	Weighted Sum	100	100	100	100	100	100
	Borda Count	99.37	99.06	98.99	99.06	99.53	99.93
	Max	100	100	100	100	100	100
	Harmonic	99.69	99.53	99.50	99.53	99.76	99.99
	Sugeno	100	100	100	99.76	100	100
Fusion3	Average	99.84	99.76	99.75	99.76	100	100
	Product	100	100	100	100	100	100
	Weighted Sum	100	100	100	100	100	100
	Borda Count	99.84	99.76	99.75	99.76	100	100
	Max	99.84	99.76	99.75	99.76	100	100
	Harmonic	100	100	100	100	100	100
	Sugeno	100	100	100	100	100	100
Fusion4	Average	99.84	99.76	99.75	99.76	100	100
	Product	100	100	100	100	100	100
	Weighted Sum	99.53	99.29	99.25	99.29	100	100
	Borda Count	99.21	98.82	98.74	98.82	99.29	99.94
	Max	99.84	99.76	99.75	99.76	100	100
	Harmonic	99.84	99.76	99.75	99.76	100	100
	Sugeno	99.37	99.06	98.99	98.35	100	100
Fusion5	Average	99.91	99.88	99.88	99.88	99.88	100
	Product	100	100	100	100	100	100
	Weighted Sum	99.53	99.41	99.40	99.41	99.88	100
	Borda Count	99.91	99.88	99.88	99.88	100	100
	Max	99.91	99.88	99.88	99.88	99.88	100
	Harmonic	100	100	100	100	100	100
	Sugeno	99.62	99.53	99.50	99.41	99.88	100

maintains high accuracy, particularly in terms of AUC, which remains close to 100%.

5.4.2 Second Exempriment

In the experiment 4.5.2, we create a unified representation of feature vectors from finger vein and palmprint modalities using concatenation technique. The merged feature vector is then processed through machine learning classifiers to generate individual probability predictions. Subsequently, various score-level fusion techniques including average, weighted sum, product, Borda count, max, harmonic, and Sugeno methods are applied to integrate and optimize the classifier outputs. The results are presented in tables 5.4, 5.5, and 5.6, and the CMC curves are plotted in figures 5.18 to 5.32.

In the second experiment, which employed machine learning classifiers for score-level fusion, the performance metrics were also very high, though there were some variations compared to the softmax results. For instance, in Protocol 1 (P1), the Average and Weighted Sum methods achieved perfect scores (100%) for several metrics, while the Product method showed slightly lower performance, with an F1-score of

Table 5.2: Performance Metrics for P2 Using Score-Level Fusion with Softmax.

Fusion	Methods	pr(%)	rec(%)	F1(%)	R1(%)	R5(%)	AUC(%)
Fusion1	Average	100	100	100	100	100	100
	Product	99.88	99.84	99.84	99.84	100	100
	Weighted Sum	99.53	99.37	99.35	99.37	100	100
	Borda Count	99.22	98.90	98.88	98.90	99.69	99.99
	Max	100	100	100	100	100	100
	Harmonic	99.76	99.69	99.68	99.69	99.84	100
	Sugeno	99.34	99.06	99.05	99.37	100	100
Fusion2	Average	99.88	99.84	99.84	99.84	100	100
	Product	100	100	100	100	100	100
	Weighted Sum	99.88	99.84	99.84	99.84	100	100
	Borda Count	98.78	98.43	98.39	98.43	99.21	99.95
	Max	99.88	99.84	99.84	99.84	100	100
	Harmonic	99.29	99.06	99.03	99.06	99.53	99.99
	Sugeno	99.53	99.37	99.35	98.90	100	100
Fusion3	Average	100	100	100	100	100	100
	Product	100	100	100	100	100	100
	Weighted Sum	99.53	99.37	99.35	99.37	100	100
	Borda Count	99.76	99.69	99.68	99.69	100	100
	Max	100	100	100	100	100	100
	Harmonic	99.88	99.84	99.84	99.84	100	100
	Sugeno	99.69	99.53	99.53	99.53	100	100
Fusion4	Average	99.88	99.84	99.84	99.84	100	100
	Product	100	100	100	100	100	100
	Weighted Sum	99.88	99.84	99.84	99.84	100	100
	Borda Count	98.66	98.27	98.18	98.11	99.21	99.95
	Max	99.76	99.69	99.68	99.69	100	100
	Harmonic	99.65	99.53	99.51	99.53	100	100
	Sugeno	99.17	98.90	98.87	98.90	100	99.99
Fusion5	Average	99.87	99.84	99.84	99.84	100	100
	Product	100	100	100	100	100	100
	Weighted Sum	99.73	99.69	99.68	99.69	100	100
	Borda Count	99.25	99.14	99.13	99.37	99.84	99.99
	Max	99.80	99.76	99.76	99.76	100	100
	Harmonic	99.73	99.69	99.68	99.69	100	100
	Sugeno	99.41	99.29	99.28	99.14	100	100

around 99.27%. In Protocol 2 (P2), the performance was consistently high, with some methods achieving perfect scores and others close to 100 %. Similar to the first experiment, Protocol 3 again demonstrated the highest and most consistent performance across all fusion methods. This further highlights the effectiveness of the specific combination of modalities and classifiers used in Protocol 3. The CMC curves for this experiment likely mirrored the trends observed in the first experiment, with high recognition rates at lower ranks, indicating the strong performance of the fusion of finger vein and palmprint modalities in improving biometric recognition.

5.4.3 Comparison of Unimodal and Multimodal Fusion Level Scores

From table 5.7 When the two modalities are fused using the Softmax method, we can see that the performance metrics reach a perfect score of 100% across all protocols and evaluation metrics. This indicates that the Softmax approach to fusion is extremely effective at leveraging the complementary strengths of both finger vein and palmprint modalities. By normalizing and combining the scores, the Softmax fusion method suc-

Table 5.3: Performance Metrics for P3 Using Score-Level Fusion with Softmax.

Fusion	Methods	pr(%)	rec(%)	F1(%)	R1(%)	R5(%)	AUC(%)
	Average	100	100	100	100	100	100
	Product	100	100	100	100	100	100
	Weighted Sum	100	100	100	100	100	100
Fusion1	Borda Count	99.29	99.53	99.37	99.06	100	100
	Max	100	100	100	100	100	100
	Harmonic	100	100	100	100	100	100
	Sugeno	99.29	99.53	99.37	99.53	100	100
	Average	100	100	100	100	100	100
	Product	100	100	100	100	100	100
	Weighted Sum	100	100	100	100	100	100
Fusion2	Borda Count	97.88	98.58	98.11	98.58	100	100
	Max	100	100	100	100	100	100
	Harmonic	100	100	100	100	100	100
	Sugeno	98.58	99.06	98.74	99.53	100	100
	Average	100	100	100	100	100	100
	Product	100	100	100	100	100	100
	Weighted Sum	100	100	100	100	100	100
Fusion3	Borda Count	99.29	99.53	99.37	99.53	100	100
	Max	100	100	100	100	100	100
	Harmonic	100	100	100	100	100	100
	Sugeno	98.58	99.06	98.74	100	100	100
	Average	100	100	100	100	100	100
	Product	100	100	100	100	100	100
	Weighted Sum	100	100	100	100	100	100
Fusion4	Borda Count	99.29	99.53	99.37	99.53	100	100
	Max	100	100	100	100	100	100
	Harmonic	100	100	100	100	100	100
	Sugeno	99.29	99.53	99.37	99.53	100	100
	Average	100	100	100	100	100	100
	Product	100	100	100	100	100	100
	Weighted Sum	100	100	100	100	100	100
Fusion5	Borda Count	100	100	100	100	100	100
	Max	100	100	100	100	100	100
	Harmonic	100	100	100	100	100	100
	Sugeno	100	100	100	100	100	100

cessfully mitigates any small weaknesses that might be present when the modalities operate independently, resulting in flawless performance. also the fusion using a machine learning (ML) approach also yields impressive results, with most scores being in the high 99% range and several instances achieving 100%. Although the ML-based fusion occasionally produces scores slightly below perfect such as precision values in the range of 99.61% to 99.72% for certain protocols the overall performance remains nearly as high as that of the Softmax method. This suggests that ML-based fusion is also a highly effective strategy for combining the information from both modalities, although the Softmax method appears to edge it out by consistently delivering perfect scores.

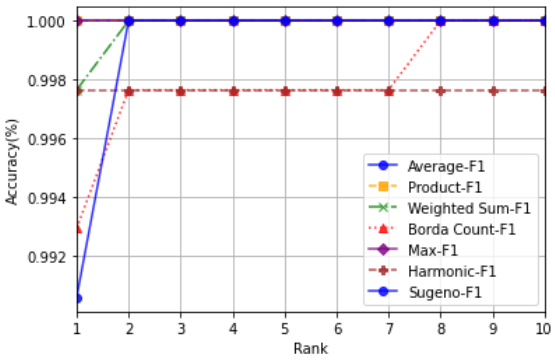


Figure 5.3: CMC Curves for F1-Level Scores Using P1 with Softmax Classifier.

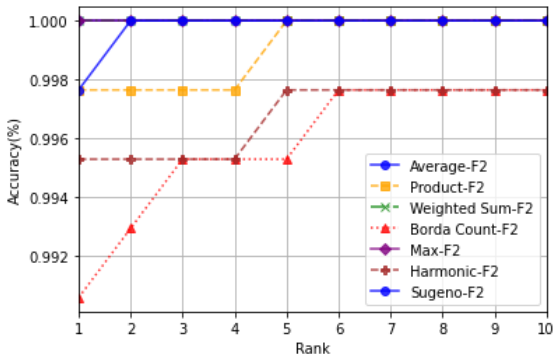


Figure 5.4: CMC Curves for F2-Level Scores Using P1 with Softmax Classifier.

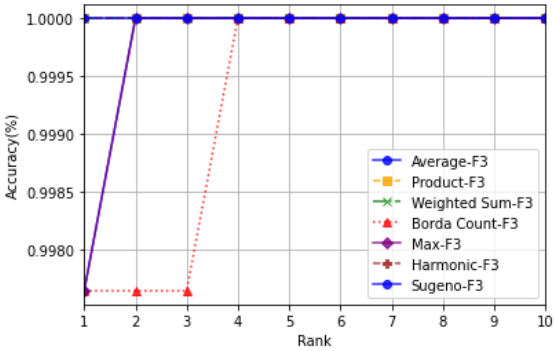


Figure 5.5: CMC Curves for F3-Level Scores Using P1 with Softmax Classifier.

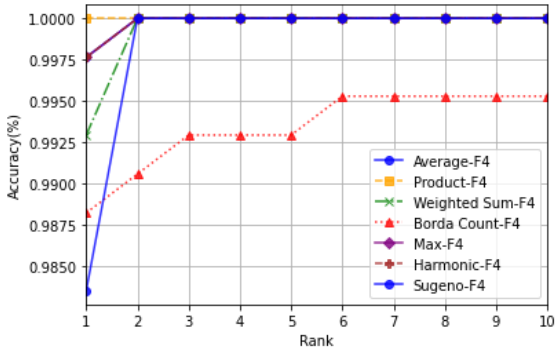


Figure 5.6: CMC Curves for F4-Level Scores Using P1 with Softmax Classifier.

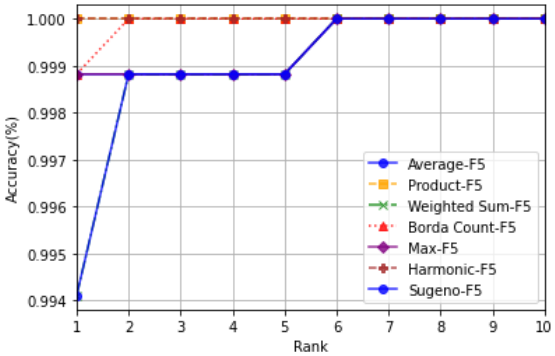


Figure 5.7: CMC Curves for F5-Level Scores Using P1 with Softmax Classifier.

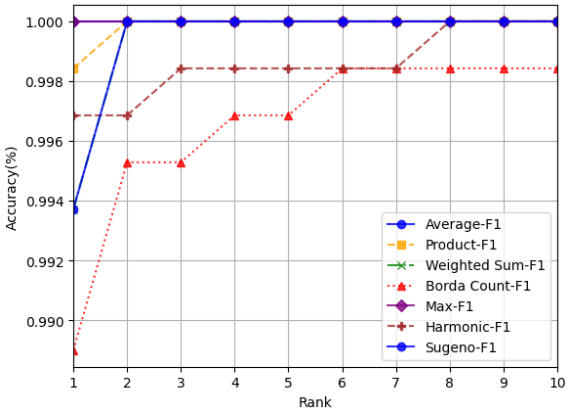


Figure 5.8: CMC Curves for F1-Level Scores Using P2 with Softmax Classifier.

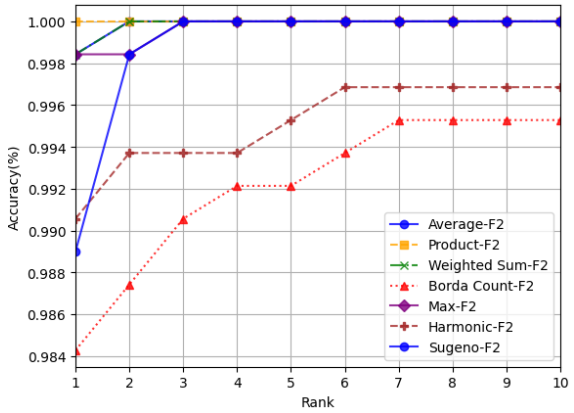


Figure 5.9: CMC Curves for F2-Level Scores Using P2 with Softmax Classifier.

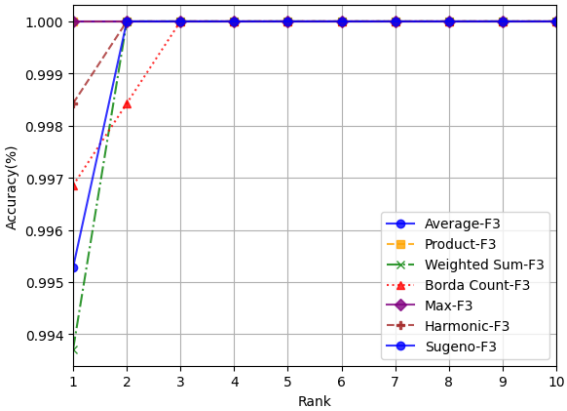


Figure 5.10: CMC Curves for F3-Level Scores Using P2 with Softmax Classifier.

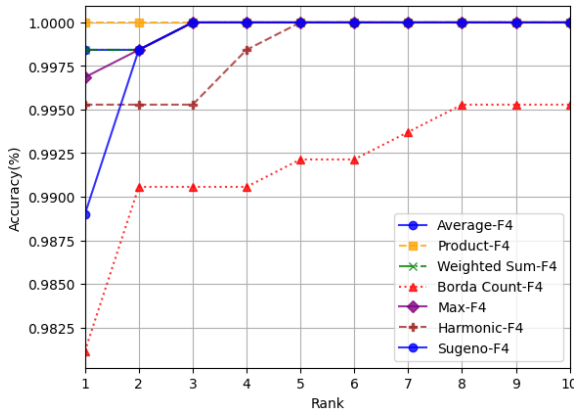


Figure 5.11: CMC Curves for F4-Level Scores Using P2 with Softmax Classifier.

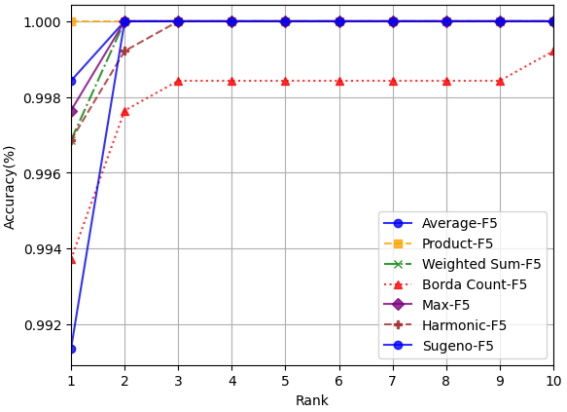


Figure 5.12: CMC Curves for F5-Level Scores Using P2 with Softmax Classifier.

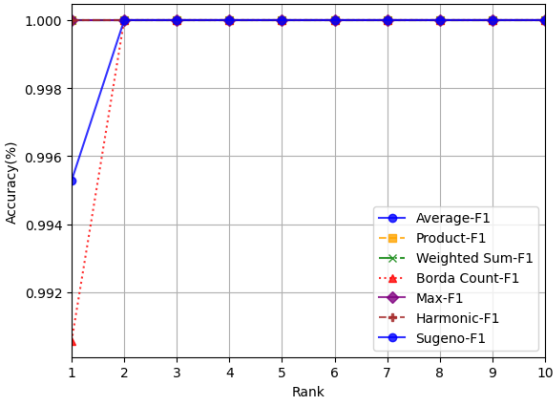


Figure 5.13: CMC Curves for F1-Level Scores Using P3 with Softmax Classifier.

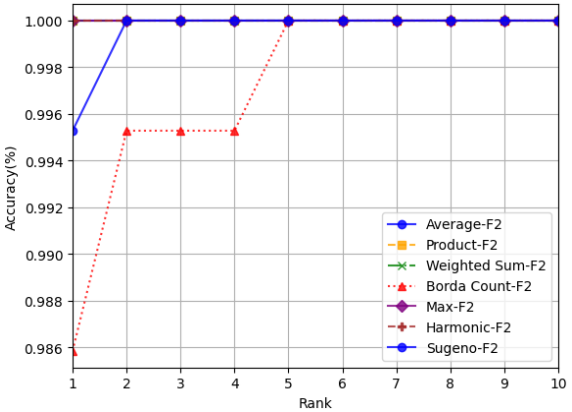


Figure 5.14: CMC Curves for F2-Level Scores Using P3 with Softmax Classifier.

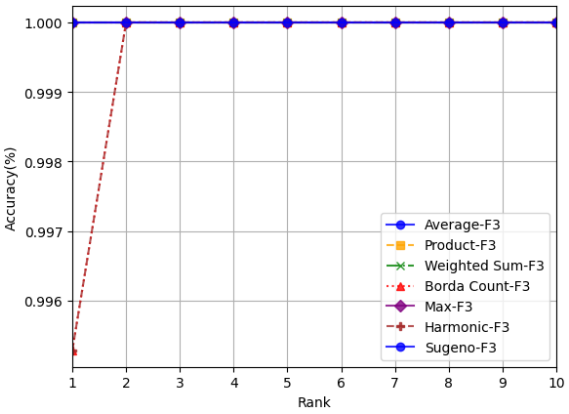


Figure 5.15: CMC Curves for F3-Level Scores Using P3 with Softmax Classifier.

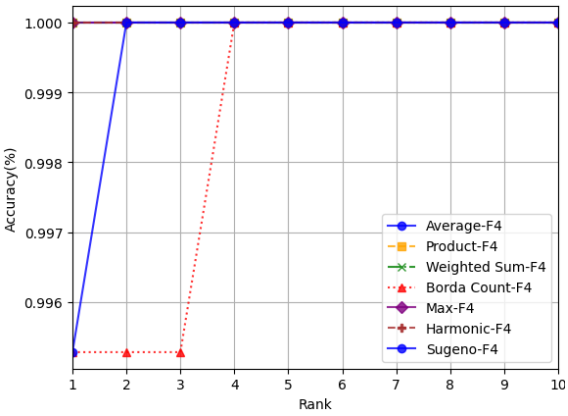


Figure 5.16: CMC Curves for F4 -Level Scores Using P3 with Softmax Classifier.

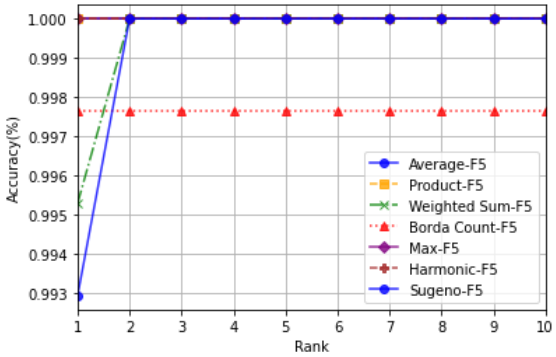


Figure 5.17: CMC Curves for F5-Level Scores Using P3 with Softmax Classifier.

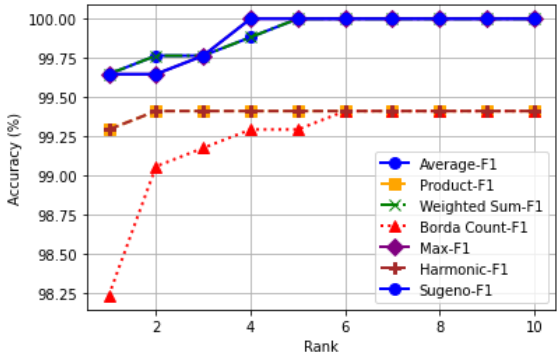


Figure 5.18: CMC curves for F5-level-scores using ML-P1.

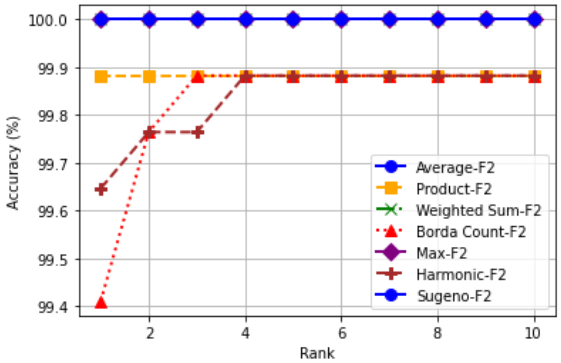


Figure 5.19: CMC curve for F2-level scores using P1 with ML.

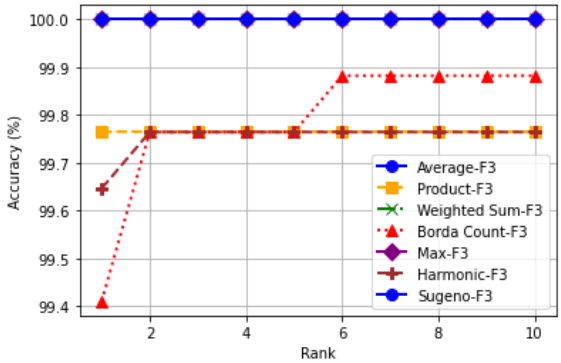


Figure 5.20: CMC curve for F3-level scores using P1 with ML.

Table 5.4: Performance Metrics for P1 Using Score-Level Fusion with ML Classifiers.

Fusion	Methods	pr(%)	rec(%)	F1(%)	R1(%)	R5(%)	AUC(%)
Fusion1	Average	99.72	99.65	99.62	99.65	100	100
	Product	99.58	99.29	99.27	99.29	99.41	100
	Weighted Sum	99.72	99.65	99.62	99.65	100	100
	Borda Count	98.89	98.47	98.37	98.23	99.29	100
	Max	99.72	99.65	99.62	99.65	100	100
	Harmonic	99.58	99.29	99.29	99.29	99.41	100
	Sugeno	99.72	99.65	99.62	99.65	100	100
Fusion2	Average	100	100	100	100	100	100
	Product	99.91	99.88	99.88	99.88	99.88	100
	Weighted Sum	100	100	100	100	100	100
	Borda Count	99.53	99.41	99.38	99.41	99.88	100
	Max	100	100	100	100	100	100
	Harmonic	99.72	99.65	99.64	99.65	99.88	100
	Sugeno	100	100	100	100	100	100
Fusion3	Average	100	100	100	100	100	100
	Product	99.84	99.76	99.77	99.76	99.76	100
	Weighted Sum	100	100	100	100	100	100
	Borda Count	99.56	99.41	99.39	99.41	99.76	100
	Max	100	100	100	100	100	100
	Harmonic	99.75	99.65	99.65	99.65	99.76	100
	Sugeno	100	100	100	100	100	100
Fusion4	Average	99.91	99.88	99.88	99.88	100	100
	Product	99.45	99.17	99.16	99.17	99.41	100
	Weighted Sum	99.91	99.88	99.88	99.88	100	100
	Borda Count	99.11	98.70	98.64	98.70	99.41	99.99
	Max	99.91	99.88	99.88	99.88	100	100
	Harmonic	99.39	99.06	99.05	99.06	99.41	100
	Sugeno	99.91	99.88	99.88	99.88	100	100
Fusion5	Average	100	100	100	100	100	100
	Product	100	100	100	100	100	100
	Weighted Sum	100	100	100	100	100	100
	Borda Count	99.97	99.97	99.97	99.97	100	100
	Max	100	100	100	100	100	100
	Harmonic	100	100	100	100	100	100
	Sugeno	100	100	100	100	100	100

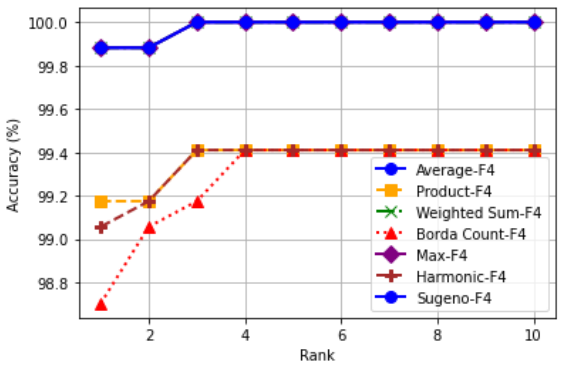


Figure 5.21: CMC curve for F4-level scores using P1 with ML.

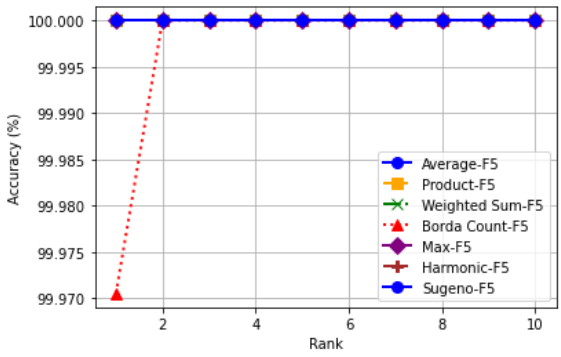


Figure 5.22: CMC curve for F5-level scores using P1 with ML.

Table 5.5: Performance Metrics for P2 Using Score-Level Fusion with ML Classifiers.

Fusion	Methods	pr(%)	rec(%)	F1(%)	R1(%)	R5(%)	AUC(%)
Fusion1	Average	99.63	99.58	99.57	99.58	99.95	99.99
	Product	99.55	99.48	99.41	99.48	99.95	99.97
	Weighted Sum	99.63	99.58	99.57	99.58	99.95	100
	Borda Count	99.50	99.42	99.36	99.42	99.84	100
	Max	99.68	99.63	99.63	99.63	99.95	99.99
	Harmonic	99.50	99.42	99.36	99.42	99.95	99.97
	Sugeno	95.43	93.66	93.31	92.61	99.95	99.98
Fusion2	Average	99.77	99.74	99.74	99.74	99.90	99.99
	Product	99.76	99.74	99.73	99.74	99.84	99.92
	Weighted Sum	99.77	99.74	99.74	99.74	99.90	99.99
	Borda Count	99.67	99.63	99.62	99.63	99.84	100
	Max	99.77	99.74	99.74	99.74	99.84	99.99
	Harmonic	99.72	99.69	99.67	99.69	99.84	99.92
	Sugeno	99.73	99.69	99.68	99.63	99.84	99.98
Fusion3	Average	99.95	99.95	99.95	99.95	100	100
	Product	99.95	99.95	99.95	99.95	99.95	99.97
	Weighted Sum	99.95	99.95	99.95	99.95	100	100
	Borda Count	99.95	99.95	99.95	99.95	99.95	100
	Max	99.95	99.95	99.95	99.95	100	100
	Harmonic	99.95	99.95	99.95	99.95	99.95	99.97
	Sugeno	99.87	99.84	99.84	99.90	100	100
Fusion4	Average	99.61	99.53	99.53	99.53	100	100
	Product	99.56	99.48	99.47	99.48	99.95	99.97
	Weighted Sum	99.53	99.42	99.43	99.42	100	100
	Borda Count	99.05	98.90	98.88	98.90	99.69	99.99
	Max	97.54	96.96	96.92	96.70	100	99.99
	Harmonic	99.26	99.11	99.10	99.11	99.90	99.97
	Sugeno	96.17	95.02	94.92	95.07	100	99.97
Fusion5	Average	99.91	99.91	99.91	99.91	99.99	100
	Product	99.86	99.86	99.86	99.86	99.92	100
	Weighted Sum	99.90	99.90	99.89	99.90	99.99	100
	Borda Count	99.74	99.72	99.72	99.72	99.97	100
	Max	99.90	99.90	99.89	99.90	99.99	100
	Harmonic	99.84	99.83	99.83	99.83	99.92	100
	Sugeno	99.85	99.84	99.84	99.87	99.99	100

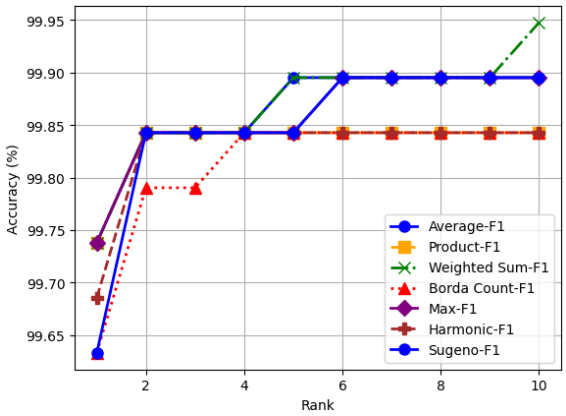


Figure 5.23: CMC curve for F1-level scores using P2 with ML.

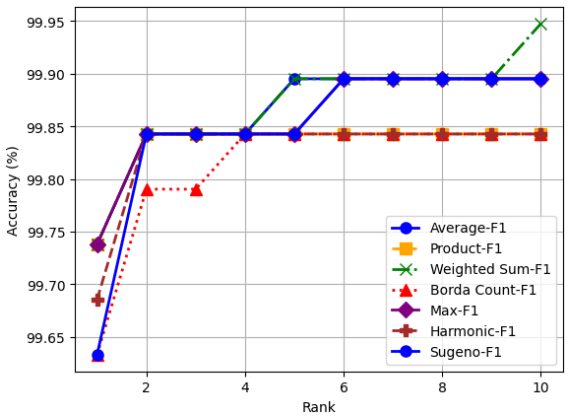


Figure 5.24: CMC curve for F2-level scores using P2 with ML.

Table 5.6: Performance Metrics for P3 Using Score-Level Fusion with ML Classifiers.

Fusion	Methods	pr(%)	rec(%)	F1(%)	R1(%)	R5(%)	AUC(%)
Fusion1	Average	100	100	100	100	100	100
	Product	99.76	99.53	99.37	99.53	99.53	99.76
	Weighted Sum	100	100	100	100	100	100
	Borda Count	99.76	99.53	99.37	99.53	99.53	100
	Max	100	100	100	100	100	100
	Harmonic	99.76	99.53	99.37	99.53	99.53	99.76
	Sugeno	99.76	99.53	99.37	99.53	100	100
Fusion2	Average	100	100	100	100	100	100
	Product	100	100	100	100	100	100
	Weighted Sum	100	100	100	100	100	100
	Borda Count	100	100	100	100	100	100
	Max	100	100	100	100	100	100
	Harmonic	100	100	100	100	100	100
	Sugeno	100	100	100	100	100	100
Fusion3	Average	100	100	100	100	100	100
	Product	100	100	100	100	100	100
	Weighted Sum	100	100	100	100	100	100
	Borda Count	100	100	100	100	100	100
	Max	100	100	100	100	100	100
	Harmonic	100	100	100	100	100	100
	Sugeno	100	100	100	100	100	100
Fusion4	Average	100	100	100	100	100	100
	Product	100	100	100	100	100	100
	Weighted Sum	100	100	100	100	100	100
	Borda Count	99.76	99.53	99.37	99.53	100	100
	Max	99.76	99.53	99.37	100	100	100
	Harmonic	100	100	100	100	100	100
	Sugeno	94.58	90.57	87.89	90.57	100	99.94
Fusion5	Average	100	100	100	100	100	100
	Product	100	100	100	100	100	100
	Weighted Sum	100	100	100	100	100	100
	Borda Count	100	100	100	100	100	100
	Max	100	100	100	100	100	100
	Harmonic	100	100	100	100	100	100
	Sugeno	100	100	100	100	100	100

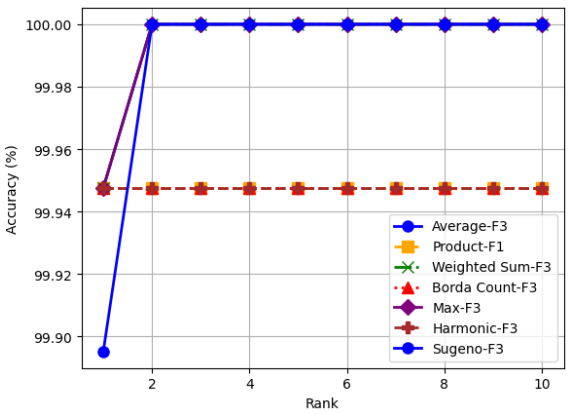


Figure 5.25: CMC curve for F3-level scores using P2 with ML.

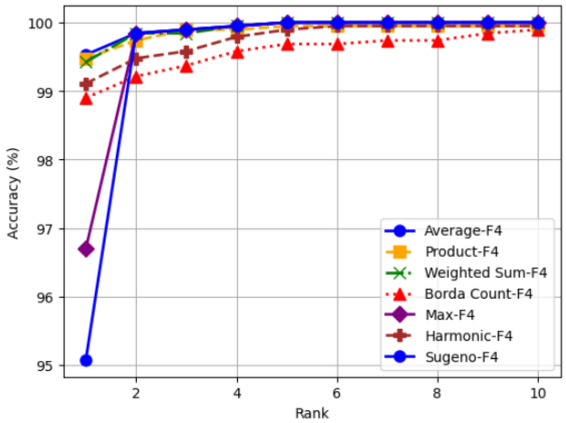


Figure 5.26: CMC curve for F4-level scores using P2 with ML.

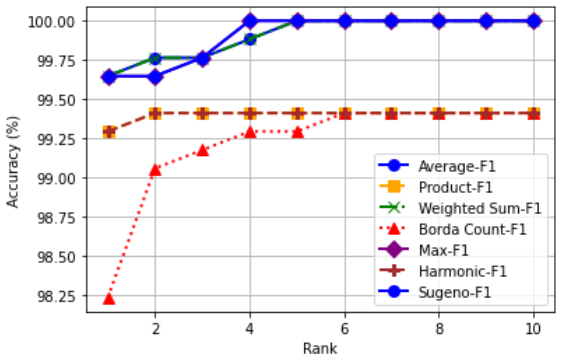


Figure 5.27: CMC curve for F5-level scores using P2 with ML.

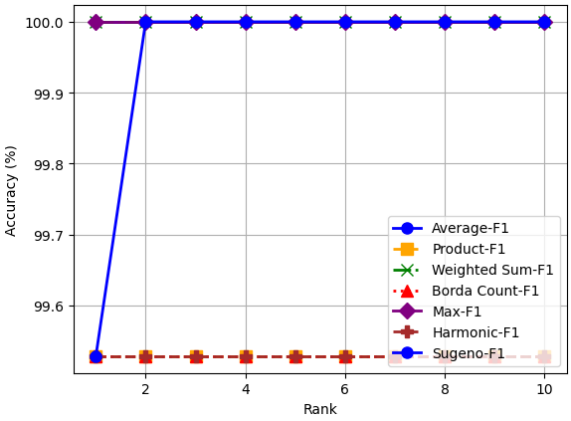


Figure 5.28: CMC curve for F1-level scores using P3 with ML.

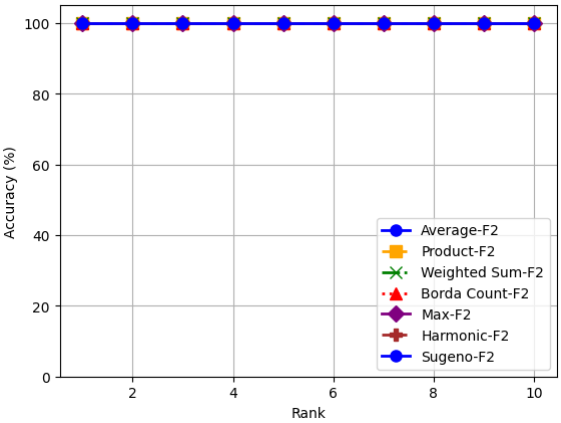


Figure 5.29: CMC curve for F2-level scores using P3 with ML.

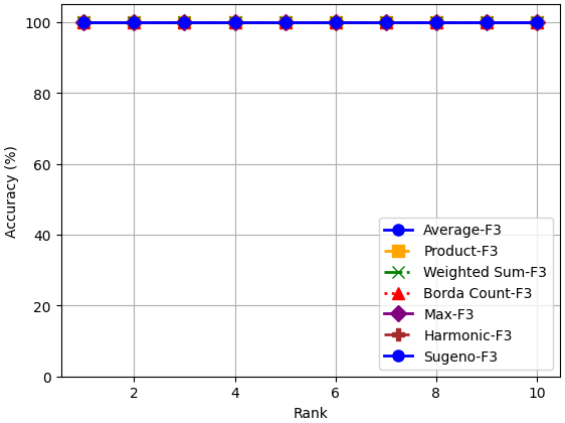


Figure 5.30: CMC curve for F3-level scores using P3 with ML.

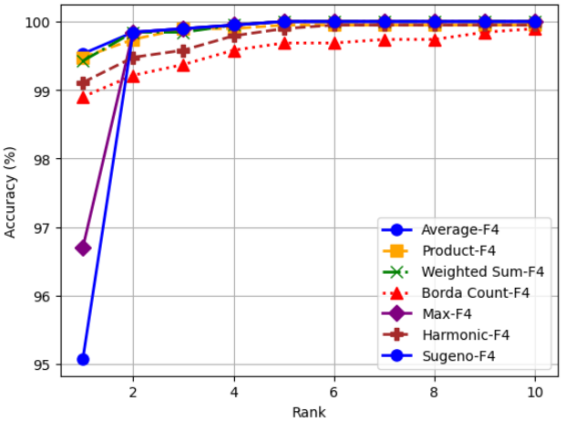


Figure 5.31: CMC curve for F4-level scores using P3 with ML.

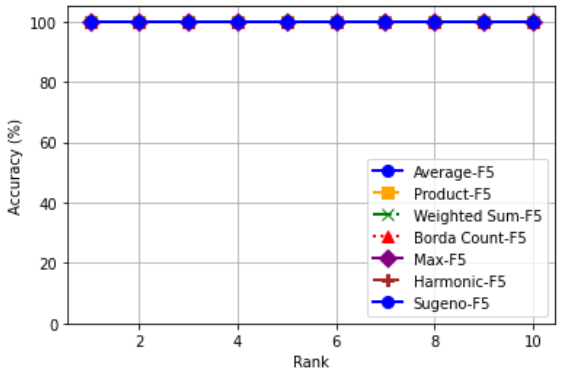


Figure 5.32: CMC curve for F5-level scores using P3 with ML.

Table 5.7: Performance Metrics for Finger vein, Palmprint, and Fusion Approaches level scores.

Datasets		Protocol	Pr(%)	Rc(%)	F1(%)	R1(%)	R5(%)	AUC(%)
Finger Vein	Left	P1	97.56	97.88	97.44	97.88	99.29	100
		P2	96.46	97.64	96.86	98.27	99.21	99.99
		P3	98.58	99.06	98.74	99.06	99.53	100
	Right	P1	98.27	97.41	97.23	97.41	98.58	99.97
		P2	97.23	96.23	96.08	96.23	98.27	99.94
		P3	98.71	98.27	98.24	97.64	98.58	100
Palmprint	Left	P1	97	98	97	98.35	99.20	99
		P2	97	96	96	96.70	99	98
		P3	98	99	98.74	99.06	100	100
	Right	P1	98	99	98.74	99.06	99	100
		P2	96	95.30	95	95.30	98.80	98
		P3	100	99	99	99.50	100	100
Fusion 1	Softmax	P1	100	100	100	100	100	100
		P2	100	100	100	100	100	100
		P3	100	100	100	100	100	100
	ML	P1	99.72	99.65	99.62	99.65	100	100
		P2	99.68	99.63	99.63	99.63	99.95	99.99
		P3	100	100	100	100	100	100
Fusion 2	Softmax	P1	100	100	100	100	100	100
		P2	100	100	100	100	100	100
		P3	100	100	100	100	100	100
	ML	P1	100	100	100	100	100	100
		P2	99.67	99.63	99.62	99.63	99.84	100
		P3	100	100	100	100	100	100
Fusion 3	Softmax	P1	100	100	100	100	100	100
		P2	100	100	100	100	100	100
		P3	100	100	100	100	100	100
	ML	P1	100	100	100	100	100	100
		P2	99.95	99.95	99.95	99.95	100	100
		P3	100	100	100	100	100	100
Fusion 4	Softmax	P1	100	100	100	100	100	100
		P2	100	100	100	100	100	100
		P3	100	100	100	100	100	100
	ML	P1	99.91	99.88	99.88	99.88	100	100
		P2	99.61	99.53	99.53	99.53	100	100
		P3	100	100	100	100	100	100
Fusion 5	Softmax	P1	100	100	100	100	100	100
		P2	100	100	100	100	100	100
		P3	100	100	100	100	100	100
	ML	P1	100	100	100	100	100	100
		P2	99.91	99.91	99.91	99.91	100	100
		P3	100	100	100	100	100	100

In summary, while the unimodal systems provide strong baseline performance, the application of feature-level fusion significantly boosts the system's accuracy by combining detailed modality-specific information. This is further refined by score-level fusion, which drives the performance to near or absolute perfection. Thus, the two-stage fusion approach first at the feature level and then at the score level results in a highly robust and reliable biometric recognition system.

5.5 Conclusion

The fusion of finger vein and palmprint modalities creates a more powerful and accurate biometric system, as evidenced by the improvements in key performance metrics when compared to unimodal systems. By combining FV and palmprint data, the fusion approach mitigates limitations inherent to a single modality, leading to greater robustness in identity verification tasks. Given these advantages, multimodal biometrics holds significant potential in applications requiring high security and reliability, such as border control, banking, and access control systems. Future studies could further investigate optimizing fusion techniques for real-time performance, exploring more diverse biometric modalities, and assessing scalability in larger, more varied populations.

Chapter 6

Conclusion

6.1 Summary of the Thesis Contributions

Biometrics plays an essential role in modern security systems by providing reliable, individual specific identification methods that enhance security. Unlike traditional security measures such as passwords or PINs, biometrics leverages unique physical or behavioral characteristics fingerprints, facial features, palmprint texture, finger vein, iris patterns, and voice, among others that are more difficult to replicate or falsify. In this research, we focused on finger vein and palmprint biometrics, two promising modalities known for their distinct and secure nature.

Finger vein recognition depends on the vein patterns within the finger, which are unique to each individual and located under the skin surface. This internal placement makes them highly resistant to forgery, as they are not easily visible or accessible for replication. Additionally, the veins' internal location means they are less affected by external factors such as dirt or skin conditions, ensuring reliable performance in various environments. Moreover, finger vein patterns exhibit significant stability over time, with only variations due to aging or health conditions, making them a reliable biometric trait for long term identification. However, finger vein recognition is hindered by variations in lighting and image quality.

Similarly, palmprint recognition uses the line patterns, texture, and ridge structures of the palm, which are highly individualistic and remain stable over time, making them a reliable biometric trait. However, palmprint recognition, particularly in contactless environments, is affected by changes in illumination and hand posture variations. Furthermore, factors such as dirt, scars, injuries, or skin deformations can introduce noise and may affect image quality and matching performance, compromising system robustness.

Both finger vein and palmprint recognition have proven effective as secure identification methods. By integrating them into a multimodal system, we can leverage their

complementary strengths, enhancing overall robustness, accuracy, and resistance to spoofing.

This thesis aimed to leverage deep learning techniques to develop an advanced multimodal biometric system that integrates both finger vein and palmprint recognition. To achieve this, we conducted multiple experiments, exploring various combinations of instances and features, and evaluating different fusion methods at multiple levels, as described below:

In Chapter Three, the research explored the potential and challenges of finger vein and palmprint recognition as highly secure biometric modalities. The study proposed a unified biometric framework that integrates deep learning-based feature extraction with machine learning classification.

The key contributions of this chapter include implementing two learning strategies single instance and multiple instance learning to enhance recognition robustness against physical anomalies. The researchers evaluated deep learning architectures, including VGG16, VGG19, and MobileNetV2, ultimately identifying MobileNetV2 as the most effective model. Preprocessing techniques were applied, such as histogram equalization for finger vein images and contrast-limited adaptive Histogram Equalization (CLAHE) for palmprint images to mitigate illumination and contrast issues. Additionally, the study compared machine learning classifiers (Support Vector Machine, K Nearest Neighbors, and Random Forest), with SVM demonstrating the best performance.

Experimental results revealed that fine tuning MobileNetV2 significantly improved feature extraction performance. The combination of deep learning and machine learning classifiers further enhanced recognition accuracy, with multiple instance learning outperforming single instance learning. This approach highlighted the advantage of incorporating multiple images per subject to address biometric variations.

In the fourth chapter, we explore the integration of finger vein and palmprint recognition systems into a unified multimodal framework to enhance biometric security. While unimodal systems, despite their successes, face limitations such as susceptibility to noise, non-universality, and spoofing risks, multimodal systems offer a promising solution by leveraging the complementary strengths of different biometric traits.

The proposed system begins with robust Region of Interest (ROI) extraction to focus on the most discriminative areas of the biometric samples. Notably, the ROI extracted from palmprint images in this work demonstrates superior results when compared to the dataset's original ROI, showing the effectiveness of our proposed approach.

To enhance the extracted features, contrast-limited adaptive Histogram Equalization (CLAHE) is applied, improving the contrast of both finger vein and palmprint images. Our findings from the previous chapter, where MobileNetV2 is classified as

the most efficient deep learning architecture for feature extraction, are further utilized in this multimodal context. Once features are extracted from both modalities, they are concatenated to create a comprehensive feature vector, which is then classified using machine learning algorithms such as Support Vector Machine (SVM), K Nearest Neighbors (KNN), and Random Forest (RF).

This chapter contributes to this thesis by advancing preprocessing techniques, feature extraction methods, and feature level fusion strategies that enhance the accuracy and reliability of biometric recognition systems. The improved ROI extraction, in particular, highlights the efficacy of our proposed methodology in addressing challenges such as environmental factors, hand positioning, and the uniqueness of individual biometric traits. These advancements lay the groundwork for more robust multimodal biometric systems that can offer enhanced security and resilience in real-world applications.

In Chapter Five, we examined the role of score-level fusion in multimodal biometric systems, highlighting its advantages over feature-level fusion in addressing challenges such as data incompatibility, dimensional redundancy, and computational inefficiencies.

Through two experiments, we explored the integration of finger vein and palmprint data at different stages of the biometric process, using MobileNetV2 outputs in the first experiment and classifier outputs in the second. The application of various score level fusion techniques such as average, weighted sum, product, Borda count, max, harmonic, and Sugeno has provided valuable insights into their effectiveness for improving system performance.

By combining the outputs of individual biometric traits and classifiers, score-level fusion offers a more flexible and computationally efficient alternative to feature-level fusion. This chapter's findings emphasize the potential of score-level fusion to enhance the overall accuracy and robustness of multimodal biometric systems, paving the way for more reliable identification processes in real-world applications.

In conclusion, this thesis demonstrates that deep learning can be effectively applied to multimodal biometrics by combining CNN-based feature extraction with feature and score-level fusion. This approach not only enhances recognition accuracy but also improves system robustness, making it more suitable for high-security applications. Our research contributes to the field of biometric identification by providing a framework that integrates finger vein and palmprint recognition, thereby advancing the capabilities of multimodal systems.

6.2 Future Work and Perspectives

Future research could expand upon this work by incorporating additional biometric modalities, such as iris or face recognition, or by exploring advanced fusion strategies, like decision level or hybrid fusion, to further optimize system performance. Furthermore, recent advancements in deep learning, such as transformer networks, offer exciting opportunities to enhance feature extraction and representation in multimodal biometrics. Through these continued advancements, biometric systems can become increasingly accurate, adaptable, and secure for real world applications in a wide range of security sensitive environments.

6.3 Author Publications

6.3.1 Journal Article

1.Title: A sequential combination of convolution neural network and machine learning for finger vein recognition system
Authors: Cheyma Nadir, Bilal Attallah, Youcef Brik.
Journal: Signal, Image and Video Processing
Volume: 18
Pages: 8267–8278
Year: 2024
DOI: https://doi.org/10.1007/s11760-024-03471-z

6.3.2 International Conference

1.Title: Finger Vein Based CNN Algorithms for Human Recognition
Authors: Cheyma Nadir, Bilal Attallah, Youcef Brik.
Conference: 2022 International Conference of Advanced Technology in Electronic and Electrical Engineering (ICATEEE)
Pages: 1–6
Publisher: IEEE

2.Title: Ear Recognition Using Ensemble of Deep Features and Machine Learning Classifiers
Authors: Bilal Attallah, Youcef Brik, Cheyma Nadir.
Conference: 2022 32nd International Conference on Computer Theory and Applications (ICCTA).
Date: 2022/12/17
Pages: 68–73

Publisher: IEEE

6.3.3 National Conference

1.Title: A Robust Convolutional Neural Network for Iris Recognition System .
Conference: The First National Conference for Applied Sciences and Engineering.
Date: November 17 and 18, 2024
Authors: Nadir Cheyma, Attallah Bilal, and Brik Youcef.

2.Title: Multiple-Instance Palmprint Recognition System Using Deep Neural Networks.
Conference: The 2nd National Conference on Electronics, Electrical Engineering, Telecommunications, and Computer Vision.
Date: November 24-25, 2024, Mila, Algeria
Authors: Nadir Cheyma, Sif Eddine Boudjellal, Attallah Bilal, and Brik Youcef.

3.Title: An Efficient Hybrid Conv-Transformer Architecture for Finger Vein Recognition Based on MobileViTv2 Model
Conference: The 2nd National Conference on Electronics, Electrical Engineering, Telecommunications, and Computer Vision.
Date: November 24-25, 2024, Mila, Algeria
Authors: Sif Eddine Boudjellal and Nadir Cheyma.

Bibliography

[1] Vishal Anand and Vijayalaxmi Kanhangad. Deeprespore: Cnn-based pore descriptor for high-resolution fingerprint detection. *Pattern Recognition Letters*, 2020.

[2] David D Zhang. *Automated biometrics: Technologies and systems*, volume 7. Springer Science & Business Media, 2013.

[3] Muhtahir O Oloyede and Gerhard P Hancke. Unimodal and multimodal biometric sensing systems: a review. *IEEE access*, 4:7532–7555, 2016.

[4] Anil Jain, Ruud Bolle, and Sharath Pankanti. *Introduction to biometrics*. Springer, 1996.

[5] Qingchen Zhang, Laurence T Yang, Zhikui Chen, and Peng Li. A survey on deep learning for big data. *Information Fusion*, 42:146–157, 2018.

[6] Chien Le and Raj Jain. A survey of biometrics security systems. *EEUU. Washington University in St. Louis*, 2009.

[7] Md Tahmid Hasan Fuad, Awal Ahmed Fime, Delowar Sikder, Md Akil Raihan Iftee, Jakaria Rabbi, Mabrook S Al-Rakhami, Abdu Gumaei, Ovishake Sen, Mohtasim Fuad, and Md Nazrul Islam. Recent advances in deep learning techniques for face recognition. *IEEE Access*, 9:99112–99142, 2021.

[8] A Kong, D Zhang, and M Kamel. A survey of palmprint recognition. *Pattern Recognition*, 42(7):1408–1418, 2009.

[9] Jasem Rahman Malgheet, Noridayu Bt Manshor, Lilly Suriani Affendey, and Alfian Bin Abdul Halin. Iris recognition development techniques: a comprehensive review. *Complexity*, 2021:1–32, 2021.

[10] Jin-Hyuk Hong, Jun-Ki Min, Ung-Keun Cho, and Sung-Bae Cho. Fingerprint classification using one-vs-all support vector machines dynamically ordered with naı¨ ve bayes classifiers. *Pattern Recognition*, 41(2):662–671, 2008.

[11] Marcos Faundez-Zanuy. Signature recognition state-of-the-art. *IEEE aerospace and electronic systems magazine*, 20(7):28–32, 2005.

[12] Changsheng Wan, Li Wang, and Vir V Phoha. A survey on gait recognition. *ACM Computing Surveys (CSUR)*, 51(5):1–35, 2018.

[13] David P Sidlauskas and Samir Tamer. Hand geometry recognition. In *Handbook of Biometrics*, pages 91–107. Springer, 2008.

[14] Md Mustafizur Rahman, Ajay Krishno Sarkar, Md Amzad Hossain, Md Selim Hossain, Md Rabiul Islam, Md Biplob Hossain, Julian MW Quinn, and Mohammad Ali Moni. Recognition of human emotions using eeg signals: A review. *Computers in Biology and Medicine*, 136:104696, 2021.

[15] Anil K Jain, Ruud Bolle, and Sharath Pankanti. *Biometrics: personal identification in networked society*, volume 479. Springer Science & Business Media, 2006.

[16] Shaveta Dargan and Munish Kumar. A comprehensive survey on the biometric recognition systems based on physiological and behavioral modalities. *Expert Systems with Applications*, 143:113114, 2020.

[17] JD Woodard Jr, NM Orlans, and PT Higgins. Biometrics: Identity assurance in the information age, 2003.

[18] Shimon K Modi. *Biometrics in identity management: Concepts to applications*. Artech House, 2011.

[19] Kamer Vishi and Sule Yildirim Yayilgan. Multimodal biometric authentication using fingerprint and iris recognition in identity management. In *2013 Ninth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pages 334–341. IEEE, 2013.

[20] Y Chen, H Vinck, D Gligoroski, and SJ Knapskog. An overview of the informationtheoretic perspective on biometric systems. *Norsk informasjonssikkerhetskonferanse-NISK*, 4(10):3142, 2011.

[21] Bir Bhanu and Venu Govindaraju. *Multibiometrics for human identification*. Cambridge University Press, 2011.

[22] Nicolas Morizet. Reconnaissance biométrique par fusion multimodale du visage et de l'iris. ..., 2009.

[23] Florent Perronnin and Jean-Luc Dugelay. Introduction à la biométrie authentification des individus par traitement audio-vidéo. *Traitement du signal*, 19(4), 2002.

[24] R Beveridge and M Kirby. Biometrics and face recognition. In *IS&T Colloquium*, page 25, 2005.

[25] Brian DeCann and Arun Ross. Relating roc and cmc curves via the biometric menagerie. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 6(4):1–8, 2013.

[26] Pushpa Singh, Narendra Singh, Krishna Kant Singh, and Akansha Singh. Diagnosing of disease using machine learning. In *Machine learning and the internet of medical things in healthcare*, pages 89–111. Elsevier, 2021.

[27] Leonardo Rundo and Carmelo Militello. Image biomarkers and explainable ai: handcrafted features versus deep learned features. *European Radiology Experimental*, 8(1):130, 2024.

[28] Rafael C Gonzalez. *Digital image processing*. Pearson education india, 2009.

[29] Anil K Jain. *Fundamentals of digital image processing*. Prentice-Hall, Inc., 1989.

[30] Robert M Haralick, Karthikeyan Shanmugam, and Its' Hak Dinstein. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, (6):610–621, 1973.

[31] Guillaume Thibault, Jesus Angulo, and Fernand Meyer. Advanced statistical matrices for texture characterization: application to cell classification. *IEEE Transactions on Biomedical Engineering*, 61(3):630–637, 2013.

[32] Chengjun Sun and William G Wee. Neighboring gray level dependence matrix for texture classification. *Computer Vision, Graphics, and Image Processing*, 23(3):341–352, 1983.

[33] Belur V Dasarathy and Edwin B Holder. Image characterizations based on joint gray level—run length distributions. *Pattern Recognition Letters*, 12(8):497–502, 1991.

[34] Mary M Galloway. Texture analysis using gray level run lengths. *Computer graphics and image processing*, 4(2):172–179, 1975.

[35] A Chu, Chandra M Sehgal, and James F Greenleaf. Use of gray value distribution of run lengths for texture analysis. *Pattern recognition letters*, 11(6):415–419, 1990.

[36] Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, volume 1, pages 886–893. Ieee, 2005.

[37] Timo Ojala, Matti Pietikäinen, and David Harwood. A comparative study of texture measures with classification based on featured distributions. *Pattern recognition*, 29(1):51–59, 1996.

[38] Dennis Gabor. Theory of communication. part 1: The analysis of information. *Journal of the Institution of Electrical Engineers-part III: radio and communication engineering*, 93(26):429–441, 1946.

[39] Stephane G Mallat. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE transactions on pattern analysis and machine intelligence*, 11(7):674–693, 1989.

[40] G. E. Hinton and R. R. Salakhutdinov. Reducing the dimensionality of data with neural networks. *Science*, 313(5786):504–507, 2006.

[41] S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 1997.

[42] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby. An image is worth 16x16 words: Transformers for image recognition at scale. In *Proceedings of the International Conference on Learning Representations (ICLR)*, 2021.

[43] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 1097–1105, 2012.

[44] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. In *Proceedings of the IEEE*, volume 86, pages 2278–2324, 1998.

[45] S. Ioffe and C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*, 2015.

[46] V. Nair and G. E. Hinton. Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th International Conference on Machine Learning (ICML)*, 2010.

[47] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15:1929–1958, 2014.

[48] Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.

[49] Yichuan Tang. Deep learning using linear support vector machines. *arXiv preprint arXiv:1306.0239*, 2013.

[50] Christopher J.C. Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2(2):121–167, 1998.

[51] Leo Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001.

[52] J. Ross Quinlan. Improved use of continuous attributes in c4.5. *Journal of Artificial Intelligence Research*, 4:77–90, 1996.

[53] A K Jain, A Ross, and S Prabhakar. An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1):4–20, 2004.

[54] A. Ross and A. K. Jain. Information fusion in biometrics. *Pattern Recognition Letters*, 24(13):2115–2125, 2003.

[55] A. K. Jain, K. Nandakumar, and A. Ross. Score normalization in multimodal biometric systems. *Pattern Recognition*, 38(12):2270–2285, 2007.

[56] Naoto Miura, Akio Nagasaka, and Takafumi Miyatake. Feature extraction of finger-vein patterns based on repeated line tracking and its application to personal identification. *Machine vision and applications*, 15:194–203, 2004.

[57] Naoto Miura, Akio Nagasaka, and Takafumi Miyatake. Extraction of finger-vein patterns using maximum curvature points in image profiles. *IEICE TRANSACTIONS on Information and Systems*, 90(8):1185–1194, 2007.

[58] Huafeng Qin, Lan Qin, and Chengbo Yu. Region growth–based feature extraction method for finger-vein recognition. *Optical Engineering*, 50(5):057208–057208, 2011.

[59] Jianfeng Zhang, Zhiying Lu, and Min Li. Active contour-based method for finger-vein image segmentation. *IEEE Transactions on Instrumentation and Measurement*, 69(11):8656–8665, 2020.

[60] Dongdong Zhao, Hui Ma, Zedong Yang, Jianian Li, and Wenbo Tian. Finger vein recognition based on lightweight cnn combining center loss and dynamic regularization. *Infrared Physics & Technology*, 105:103221, 2020.

[61] Changlong He, Zuojin Li, Liukui Chen, and Jun Peng. Identification of finger vein using neural network recognition research based on pca. In *2017 IEEE 16th International Conference on Cognitive Informatics & Cognitive Computing (ICCI* CC)*, pages 456–460. IEEE, 2017.

[62] Pengyang Zhao, Shuping Zhao, Jing-Hao Xue, Wenming Yang, and Qingmin Liao. The neglected background cues can facilitate finger vein recognition. *Pattern Recognition*, 136:109199, 2023.

[63] Chunxin Fang, Hui Ma, and Jianian Li. A finger vein authentication method based on the lightweight siamese network with the self-attention mechanism. *Infrared Physics & Technology*, 128:104483, 2023.

[64] Ismail Boucherit, Mohamed Ould Zmirli, Hamza Hentabli, and Bakhtiar Affendi Rosdi. Finger vein identification using deeply-fused convolutional neural network. *Journal of King Saud University-Computer and Information Sciences*, 34(3):646–656, 2022.

[65] Kashif Shaheed, Aihua Mao, Imran Qureshi, Munish Kumar, Sumaira Hussain, Inam Ullah, and Xingming Zhang. Ds-cnn: A pre-trained xception model based on depth-wise separable convolutional neural network for finger vein recognition. *Expert Systems with Applications*, 191:116288, 2022.

[66] D. Zhang and et al. Palmprint recognition: A review. *ACM Computing Surveys (CSUR)*, 45(3):1–40, 2012.

[67] C. Chen and et al. Palmprint identification based on principal lines. *International Conference on Image Processing*, pages 1–4, 2001.

[68] Y. Connie and et al. Handcrafted feature extraction for palmprint recognition. *International Conference on Image Processing*, pages 1–4, 2003.

[69] X. Wu and et al. Palmprint recognition using multi-scale texture features. *Proceedings of the International Conference on Document Analysis and Recognition*, pages 1–5, 2002.

[70] S. Minaee and R. Wang. Deep learning for palmprint recognition: A comprehensive review. In *Proceedings of the International Conference on Image Processing*, 2017.

[71] M. Samai and et al. Deep learning-based 2d and 3d palmprint recognition. In *Proceedings of the International Conference on Document Analysis and Recognition*, 2018.

[72] J. Zhong and et al. Siamese network for palmprint recognition. In *Proceedings of the International Conference on Document Analysis and Recognition*, 2018.

[73] A. Izadpanahkakhk and et al. Transfer learning for palmprint verification. In *Proceedings of the International Conference on Document Analysis and Recognition*, 2018.

[74] Mohd Shahrimie Mohd Asaari, Shahrel A Suandi, and Bakhtiar Affendi Rosdi. Fusion of band limited phase only correlation and width centroid contour distance for finger based biometrics. *Expert Systems with Applications*, 41(7):3367–3382, 2014.

[75] Yilong Yin, Lili Liu, and Xiwei Sun. Sdumla-hmt: A multimodal biometric database. In *Biometric Recognition: 6th Chinese Conference, CCBR 2011, Beijing, China, December 3-4, 2011. Proceedings 6*, pages 260–268. Springer, 2011.

[76] Zhenan Sun, Tieniu Tan, Yunhong Wang, and Stan Z Li. Ordinal palmprint represention for personal identification [represention read representation]. In *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, volume 1, pages 279–284. IEEE, 2005.

[77] Ajay Kumar. Incorporating cohort information for reliable palmprint authentication. In *2008 Sixth Indian conference on computer vision, graphics & image processing*, pages 583–590. IEEE, 2008.

[78] Heng-Da Cheng and XJ Shi. A simple and effective histogram equalization approach to image enhancement. *Digital signal processing*, 14(2):158–170, 2004.

[79] Garima Yadav, Saurabh Maheshwari, and Anjali Agarwal. Contrast limited adaptive histogram equalization based enhancement for real time video system. In *2014 international conference on advances in computing, communications and informatics (ICACCI)*, pages 2392–2397. IEEE, 2014.

[80] Jason Yosinski, Jeff Clune, Yoshua Bengio, and Hod Lipson. How transferable are features in deep neural networks? *Advances in neural information processing systems*, 27, 2014.

[81] Nima Tajbakhsh, Jae Y Shin, Suryakanth R Gurudu, R Todd Hurst, Christopher B Kendall, Michael B Gotway, and Jianming Liang. Convolutional neural networks for medical image analysis: Full training or fine tuning? *IEEE transactions on medical imaging*, 35(5):1299–1312, 2016.

[82] Maxime Oquab, Leon Bottou, Ivan Laptev, and Josef Sivic. Learning and transferring mid-level image representations using convolutional neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1717–1724, 2014.

[83] Simon Kornblith, Jonathon Shlens, and Quoc V Le. Do better imagenet models transfer better? In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2661–2671, 2019.

[84] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. Mobilenetv2: Inverted residuals and linear bottlenecks. *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4510–4520, 2018.

[85] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[86] Corinna Cortes. Support-vector networks. *Machine Learning*, 1995.

[87] Yihua Liao and V Rao Vemuri. Use of k-nearest neighbor classifier for intrusion detection. *Computers & security*, 21(5):439–448, 2002.

[88] Thomas Cover and Peter Hart. Nearest neighbor pattern classification. *IEEE transactions on information theory*, 13(1):21–27, 1967.

[89] Jaime Lynn Speiser, Michael E Miller, Janet Tooze, and Edward Ip. A comparison of random forest variable selection methods for classification prediction modeling. *Expert systems with applications*, 134:93–101, 2019.

[90] Adele Cutler, D Richard Cutler, and John R Stevens. Random forests. *Ensemble machine learning: Methods and applications*, pages 157–175, 2012.

[91] Abhishek Das, Harsh Agrawal, C Lawrence Zitnick, Devi Parikh, and Dhruv Batra. Convolutional neural networks for medical image analysis: Full training or fine tuning? *IEEE transactions on medical imaging*, 35(5):1299–1312, 2018.

[92] Lakpa Dorje Tamang and Byung Wook Kim. Fvr-net: finger vein recognition with convolutional neural network using hybrid pooling. *Applied Sciences*, 12(15):7538, 2022.

[93] Weiye Liu, Huimin Lu, Yifan Wang, Yupeng Li, Zhenshen Qu, and Yang Li. Mmran: A novel model for finger vein recognition based on a residual attention mechanism: Mmran: A novel finger vein recognition model. *Applied Intelligence*, 53(3):3273–3290, 2023.

[94] Hengyi Ren, Lijuan Sun, Jian Guo, Chong Han, and Ying Cao. A high compatibility finger vein image quality assessment system based on deep learning. *Expert Systems with Applications*, 196:116603, 2022.

[95] Chih-Hsien Hsia, Zi-Han Yang, Hong-Jyun Wang, and Kuei-Kuei Lai. A new enhancement edge detection of finger-vein identification for carputer system. *Applied Sciences*, 12(19):10127, 2022.

[96] Tingting Chai, Jiahui Li, Shitala Prasad, Qi Lu, and Zhaoxin Zhang. Shape-driven lightweight cnn for finger-vein biometrics. *Journal of Information Security and Applications*, 67:103211, 2022.

[97] Tingting Chai, Jiahui Li, Yanhong Wang, Guoying Sun, Changyong Guo, and Zhaoxin Zhang. Vascular enhancement analysis in lightweight deep feature space. *Neural Processing Letters*, 55(3):2305–2320, 2023.

[98] Dingzhong Feng, Shanyu He, Zihao Zhou, and Ye Zhang. A finger vein feature extraction method incorporating principal component analysis and locality preserving projections. *Sensors*, 22(10):3691, 2022.

[99] Shuping Zhao, Jigang Wu, Lunke Fei, Bob Zhang, and Pengyang Zhao. Double-cohesion learning based multiview and discriminant palmprint recognition. *Information Fusion*, 83:96–109, 2022.

[100] Tingting Chai, Shitala Prasad, and Shenghui Wang. Boosting palmprint identification with gender information using deepnet. *Future Generation Computer Systems*, 99:41–53, 2019.

[101] Wojciech Michal Matkowski, Tingting Chai, and Adams Wai Kin Kong. Palmprint recognition in uncontrolled and uncooperative environment. *IEEE Transactions on Information Forensics and Security*, 15:1601–1615, 2019.

[102] Nadia Amrouni, Amir Benzaoui, Rafik Bouaouina, Yacine Khaldi, Insaf Adjabi, and Ouahiba Bouglimina. Contactless palmprint recognition using binarized statistical image features-based multiresolution analysis. *Sensors*, 22(24):9814, 2022.

[103] N. Mokni, A. Benabdallah, and S. Chitroub. Fusion of principal lines and texture patterns for palmprint recognition. *Pattern Recognition Letters*, 87:24–31, 2017.

[104] Abdu Gumaei, Rachid Sammouda, Abdul Malik Al-Salman, and Ahmed Alsanad. An effective palmprint recognition approach for visible and multispectral sensor images. *Sensors*, 18(5):1575, 2018.

[105] Kaijun Zhou, Xiancheng Zhou, Lingli Yu, Lizhi Shen, and Shaoqian Yu. Double biologically inspired transform network for robust palmprint recognition. *Neurocomputing*, 337:24–45, 2019.

[106] Mohamed Hammami, Salma Ben Jemaa, and Hanene Ben-Abdallah. Selection of discriminative sub-regions for palmprint recognition. *Multimedia tools and applications*, 68:1023–1050, 2014.

[107] Y. Liang, Q. Zhao, and X. Luo. Multi-task progressive semi-relational learning for palmprint recognition. *Pattern Recognition*, 107:107494, 2020.

[108] Shanwen Zhang, Haoxiang Wang, Wenzhun Huang, and Chuanlei Zhang. Combining modified lbp and weighted src for palmprint recognition. *Signal, Image and Video Processing*, 12:1035–1042, 2018.

[109] Bilal Attallah, Amina Serir, and Youssef Chahir. Feature extraction in palmprint recognition using spiral of moment skewness and kurtosis algorithm. *Pattern Analysis and Applications*, 22:1197–1205, 2019.

[110] Gopal Chaudhary and Smriti Srivastava. A robust 2d-cochlear transform-based palmprint recognition. *Soft Computing*, 24(3):2311–2328, 2020.

[111] Mahdieh Izadpanahkakhk, Seyyed Mohammad Razavi, Mehran Taghipour-Gorjikolaie, Seyyed Hamid Zahiri, and Aurelio Uncini. Deep region of interest and feature extraction models for palmprint verification using convolutional neural networks transfer learning. *Applied Sciences*, 8(7):1210, 2018.

[112] Feng Yulin and Ajay Kumar. Best: Building evidences from scattered templates for accurate contactless palmprint recognition. *Pattern recognition*, 138:109422, 2023.

[113] Zhaoqun Li, Xu Liang, Dandan Fan, Jinxing Li, and David Zhang. Bpfnet: A unified framework for bimodal palmprint alignment and fusion. *arXiv preprint arXiv:2110.01179*, 2021.

[114] Chenyi Zhang, Jing Huang, Feng Yang, and Yaqin Liu. A hybrid fusion model of iris, palm vein and finger vein for multi-biometric recognition system. *Multimedia Tools and Applications*, 80:145–162, 2021.

[115] Anil Kumar, Rajesh Sharma, and Rajib Das. A hybrid multimodal biometric system using face and palmprint traits. *Pattern Recognition Letters*, 165:85–93, 2023.

[116] Sobhan Soleymani, Ali Dabouei, Hadi Kazemi, Jeremy Dawson, and Nasser M Nasrabadi. Multi-level feature abstraction from convolutional neural networks for multimodal biometric identification. In *2018 24th International Conference on Pattern Recognition (ICPR)*, pages 3469–3476. IEEE, 2018.

[117] Changxing Ding and Dacheng Tao. Robust face recognition via multimodal deep face representation. *IEEE transactions on Multimedia*, 17(11):2049–2058, 2015.

[118] Pedro H Silva, Eduardo Luz, Luiz A Zanlorensi, David Menotti, and Gladston Moreira. Multimodal feature level fusion based on particle swarm optimization with deep transfer learning. In *2018 IEEE Congress on Evolutionary Computation (CEC)*, pages 1–8. IEEE, 2018.

[119] Huda Moyasar Therar, Lect Dr Emad Ahmed Mohammed, and Ahmed Jadaan Ali. Multibiometric system for iris recognition based convolutional neural network and transfer learning. In *IOP Conference Series: Materials Science and Engineering*, volume 1105, page 012032. IOP Publishing, 2021.

[120] Sara Daas, Amira Yahi, Toufik Bakir, Mouna Sedhane, Mohamed Boughazi, and El-Bay Bourennane. Multimodal biometric recognition systems using deep learning based on the finger vein and finger knuckle print fusion. *IET Image Processing*, 14(15):3859–3868, 2020.

[121] Himanshu Purohit and Pawan K Ajmera. Optimal feature level fusion for secured human authentication in multimodal biometric system. *Machine Vision and Applications*, 32(1):24, 2021.

[122] Rasha O Mahmoud, Mazen M Selim, and Omar A Muhi. Fusion time reduction of a feature level based multimodal biometric authentication system. *International Journal of Sociotechnology and Knowledge Development (IJSKD)*, 12(1):67–83, 2020.

[123] K Padmavathi and K Thangadurai. Implementation of rgb and grayscale images in plant leaves disease detection–comparative study. *Indian Journal of Science and Technology*, 9(6):1–6, 2016.

[124] Robert A Hummel, B Kimia, and Steven W Zucker. Deblurring gaussian blur. *Computer Vision, Graphics, and Image Processing*, 38(1):66–80, 1987.

[125] Xin Wang. Laplacian operator-based edge detectors. *IEEE transactions on pattern analysis and machine intelligence*, 29(5):886–890, 2007.

[126] N Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66, 1979.

[127] J Serra. *Image Analysis and Mathematical Morphology*. Academic Press, 1982.

[128] Wee Lorn Jhinn, Michael Goh Kah Ong, Lau Siong Hoe, and Tee Connie. Contactless palm vein roi extraction using convex hull algorithm. In *Computational Science and Technology: 5th ICCST 2018, Kota Kinabalu, Malaysia, 29-30 August 2018*, pages 25–35. Springer, 2019.

[129] Dane Brown and Karen Bradshaw. Deep palmprint recognition with alignment and augmentation of limited training samples. *SN Computer Science*, 3(1):11, 2022.

[130] Xuekui Yan, Wenxiong Kang, Feiqi Deng, and Qiuxia Wu. Palm vein recognition based on multi-sampling and feature-level fusion. *Neurocomputing*, 151:798–807, 2015.

[131] Mohamed Soltane and B Mimen. Soft decision level fusion approach to a combined behavioral speech-signature biometrics verification. *International Journal of Signal Processing, Image Processing and Pattern Recognition–IJSIP*, 5(5), 2013.

[132] Rohit Srivastava, Dheeraj Singh, Ravi Tomar, and Sarishma. Three-layer multimodal biometric fusion using sift and surf descriptors for improved accuracy of authentication of human identity. *Innovative Trends in Computational Intelligence*, pages 119–142, 2022.

[133] N Malarvizhi, P Selvarani, and Pethuru Raj. Adaptive fuzzy genetic algorithm for multi biometric authentication. *Multimedia Tools and Applications*, 79(13):9131–9144, 2020.

[134] Poorti Sagar and Anamika Jain. Multinet: A multimodal approach for biometric verification. In *Computer Vision and Machine Intelligence: Proceedings of CVMI 2022*, pages 679–690. Springer, 2023.

[135] Waziha Kabir, M Omair Ahmad, and MNS Swamy. A new anchored normalization technique for score-level fusion in multimodal biometrie systems. In *2016 IEEE international symposium on circuits and systems (ISCAS)*, pages 93–96. IEEE, 2016.

[136] Waziha Kabir, M Omair Ahmad, and MNS Swamy. Score reliability based weighting technique for score-level fusion in multi-biometric systems. In *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 1–7. IEEE, 2016.

[137] Wan Kim, Jong Min Song, and Kang Ryoung Park. Multimodal biometric recognition based on convolutional neural network by the fusion of finger-vein and finger shape using near-infrared (nir) camera sensor. *Sensors*, 18.

[138] Alaa S Al-Waisy, Rami Qahwaji, Stanley Ipson, and Shumoos Al-Fahdawi. A multimodal biometrie system for personal identification based on deep learning approaches. In *2017 Seventh international conference on emerging security technologies (EST)*, pages 163–168. IEEE, 2017.

[139] Madasu Hanmandlu, Jyotsana Grover, Ankit Gureja, and Hari Mohan Gupta. Score level fusion of multimodal biometrics using triangular norms. *Pattern recognition letters*, 32(14):1843–1850, 2011.

[140] Jinfeng Yang, Yihua Shi, and Jinli Yang. Personal identification based on finger-vein features. *Computers in Human Behavior*, 27(5):1565–1570, 2011.

[141] Xiaofeng Liu, Lin Ma, and Joseph Mathew. Machinery fault diagnosis based on fuzzy measure and fuzzy integral data fusion techniques. *Mechanical Systems and Signal Processing*, 23(3):690–700, 2009.

[142] Rohit Kundu, Pawan Kumar Singh, Seyedali Mirjalili, and Ram Sarkar. Covid-19 detection from lung ct-scans using a fuzzy integral-based cnn ensemble. *Computers in Biology and Medicine*, 138:104895, 2021.

[143] Shang-Lin Wu, Yu-Ting Liu, Tsung-Yu Hsieh, Yang-Yin Lin, Chih-Yu Chen, Chun-Hsiang Chuang, and Chin-Teng Lin. Fuzzy integral with particle swarm optimization for a motor-imagery-based brain–computer interface. *IEEE Transactions on Fuzzy Systems*, 25(1):21–28, 2016.

[144] Oussama Bouguerra, Bilal Attallah, and Youcef Brik. Mri-based brain tumor ensemble classification using two stage score level fusion and cnn models. *Egyptian Informatics Journal*, 28:100565, 2024.