

Robust thermal face recognition for law enforcement using optimized deep features with new rough sets-based optimizer

Tarek Gaber^{a,b,*}, Mathew Nicho^c, Esraa Ahmed^d, Ahmed Hamed^e

^a School of Science, Engineering and Environment University of Salford, Manchester, UK

^b Faculty of Computers and Informatics, Suez Canal University, Ismailia, Egypt

^c Research and Innovation Centre, Rabdan Academy, Abu Dhabi, United Arab Emirates

^d Faculty of Science, Department of Mathematics, Suez Canal University, Ismailia, Egypt

^e Department of Computer Science, Faculty of Computers and Information, Damanhour University, Damanhour, 22511, Egypt

ARTICLE INFO

Keywords:

Thermal face images
Convolution Neural Networks
Multi-granulation rough set
Neighborhood relation
Particle swarm optimization
Charlotte-ThermalFace dataset

ABSTRACT

In the security domain, the growing need for reliable authentication methods highlights the importance of thermal face recognition for enhancing law enforcement surveillance and safety especially in IoT applications. Challenges like computational resources and alterations in facial appearance, e.g., plastic surgery could affect face recognition systems. This study presents a novel, robust thermal face recognition model tailored for law enforcement, leveraging thermal signatures from facial blood vessels using a new CNN architecture (Max and Average Pooling- MAP-CNN). This architecture addresses expression, illumination, and surgical invariance, providing a robust feature set critical for precise recognition in law enforcement and border control. Additionally, the model employs the NM-PSO algorithm, integrating neighborhood multi-granulation rough set (NMGRS) with particle swarm optimization (PSO), which efficiently handles both categorical and numerical data from multi-granulation perspectives, leading to a 57% reduction in feature dimensions while maintaining high classification accuracy outperforming ten contemporary models on the Charlotte-ThermalFace dataset by about 10% across key metrics. Rigorous statistical tests confirm NM-PSO's superiority, and further robustness testing of the face recognition model against image ambiguity and missing data demonstrated its consistent performance, enhancing its suitability for security-sensitive environments with 99% classification accuracy.

1. Introduction

Biometric authentication is being widely used as a tool for law enforcement due to its high accuracy, quick processing capabilities, and seamless integration with existing security systems, enhancing overall operational effectiveness and deterrence against criminal activities [1]. Face recognition is a well-known biometric authentication technique, owing to its significance in the commercial and governmental arenas. Criminal identification, airport face verification, smart home physical access control, and logging to mobile devices are some applications for face recognition. However, the accuracy of face recognition heavily depends on the quality of the facial image captured [2].

In this context, thermal imaging—a specialized subset of face-based biometric authentication—utilizes the heat patterns emitted by individuals to generate visual representations, offering several practical applications [2]. This technique effectively addresses many challenges inherent to the visible spectrum, including complications such as object shadows, clothing blending into the background, and variable lighting

conditions, such as low light environments [3]. However, it also faces its own set of challenges, such as blurred thermal signatures due to high ambient temperatures, surface reflections that can distort images, and a reduction in clarity at greater distances, all of which can degrade image quality. Widely adopted in military operations for border surveillance and by law enforcement agencies, thermal imaging facilitates the identification and authentication of individuals by analyzing their unique thermal signatures [4]. Additionally, thermal face authentication serves as a vital real-time tool for law enforcement officers, aiding in the detection and monitoring of individuals on “watchlists” who are sought by the police [5].

Generally, the face recognition process, including thermal infrared, consists of several stages such as image acquisition, face detection, feature extraction (FE), and matching [6]. Since face recognition relies on biometrics to identify people based on their traits, supervised learning is primarily used, which can leverage the success of deep learning (DL) [7].

* Corresponding author at: School of Science, Engineering and Environment University of Salford, Manchester, UK.

E-mail addresses: t.m.a.gaber@salford.ac.uk (T. Gaber), mnicho@ra.ac.ae (M. Nicho), esraa_ahmed@science.suez.edu.eg (E. Ahmed), ahmed_hamed@cis.dmu.edu.eg (A. Hamed).

<https://doi.org/10.1016/j.jisa.2024.103838>

There are two primary challenges inherent in face recognition systems, particularly affecting their integration into Internet of Things (IoT) applications like security and surveillance systems. Firstly, the process of face recognition, which involves face detection, FE, and matching, requires substantial computational resources, posing a challenge for IoT devices with limited capabilities [8]. This is particularly problematic in scenarios such as law enforcement using portable cameras, where high computational costs can hinder operational efficiency. Secondly, the effectiveness of face recognition systems, especially those using the visible spectrum, is heavily compromised by alterations in facial appearance, such as those resulting from plastic surgery. These changes can significantly disrupt the FE process, which is critical for accurate identification. Plastic surgery, while often used for legitimate medical reasons like correcting birth defects or injuries, can also be exploited to deliberately alter one's identity to evade law enforcement [9]. The increasing affordability and prevalence of cosmetic surgery further exacerbate this issue, challenging the reliability of facial recognition technologies.

To address these issues, a robust FE method is essential to ensure that the initial capture of facial characteristics is accurate and comprehensive. Additionally, feature selection (FS) [10] becomes critical in refining the feature set to manageable levels, reducing the computational load and enhancing the performance of classifiers [11]. This step is crucial for adapting face recognition technologies to environments with resource constraints, such as IoT devices, ensuring both efficiency and efficacy in real-world applications [12]. Therefore, the development of reliable FE and FS tools is fundamental for the success of thermal face recognition systems in security-sensitive environments. The study is guided by the following objectives.

1. Improving the FE performance to maintain robust facial extracted features.
2. Enhancing FS to mitigate the computation cost burden.

To achieve the first objective, this paper makes use of a new modified version of convolution neural networks (CNN) that uses the max/average pooling, referred to as MAP-CNN. While CNNs are employed for FS, they do not analyze features from a multigranulation perspective and struggle to directly handle heterogeneous data [13], which may compromise the quality of the feature subset. Consequently, to address the second objective, a novel version of particle swarm optimization (PSO) is proposed. This version is based on a new rough set model utilizing multigranulation theory with neighborhood relation. This approach guides the feature reduction task through a modified FS version of PSO, where the rough set model serves as a fitness function. PSO was selected as it has been argued that PSO is the elite algorithm in the swarm family and simple but yet employing a fewer number of parameters.

The contributions of the paper are as follows.

- Extracting deep features from thermal face images using MAP-CNN algorithm that is able to capitalize on the distinct thermal signatures from facial blood vessels, integrating the AVG-MAX Vector Pooling Block (VPB) to synergistically combines Max and Average Pooling methods for comprehensive feature capture. This approach ensures robust FE crucial for thermal face recognition, significantly enhancing computational efficiency and model stability for real-time law enforcement applications. The extracted features showed a better representation of the individuals than the features extracted using the traditional CNN.
- A novel version of the multi-granulation rough set called the Neighborhood multi-granulation rough set (NMGRS). It advances the MGRS theory by accommodating neighborhood relations. This innovation allows NMGRS to effectively manage heterogeneous datasets containing both categorical and numerical features. It excels in discerning the nuanced interrelations within data, thereby providing a more adaptive and accurate representation about

the data. By integrating distance measures, NMGRS extends the applicability of traditional MGRS to broader real-world scenarios where data inconsistency and diversity prevail.

- An NM-PSO based FS algorithm that integrates NMGRS with the PSO algorithm to select and reduce the high dimensionality of the features needed for face recognition. In this scenario, NMGRS serves as a fitness function to guide the PSO reduction of inconsistencies and heterogeneity in the dataset. It showed that it could select only 40% of the features while achieving more than 99% classification accuracy.
- The effectiveness of the suggested model was assessed through the utilization of several performance metrics. The metrics encompassed in this study are accuracy, precision, recall, and F1-Score. The metrics employed in this study demonstrated a full evaluation of our suggested model, showcasing its effectiveness and quality across various dimensions. Robustness analysis was conducted to assess the model's performance under varying degrees of ambiguity and missing data. The analysis demonstrates that the model exhibits considerable robustness, effectively managing ambiguity and incompleteness up to a threshold of 50%.
- The statistical analysis of the suggested model was conducted to demonstrate the significance of the various machine-learning approaches utilized in this work. This facilitates practitioners in making informed judgments by relying on robust evidence, hence decreasing the risk of making decisions based on arbitrary factors.

The remainder of the paper is organized as follows: Section 2 reviews the literature. Section 3 sheds light on the main concept and mathematical models needed for building our models. In Section 4, the proposed model is introduced in detail and is validated in Section 5. Finally, Section 6 concludes the paper.

2. Related work

Research on thermal face matching, using machine learning (ML) is limited in the literature, as searches on multiple databases did not yield many results. Methods used for the extraction of features in face matching can be accomplished using appearance-based methods, local descriptor methods, or CNN [6]. While thermal cameras can capture images in low-light conditions, the low information and depth texture quality are still considered weaknesses in infrared face recognition. However, CNN can extract more profound features even from low-information data using its convolutional layers [14]. In this respect, the DL classifier CNN has been proposed as the preferable option in most observed studies due to its performance and ability to process large amounts of data [6].

Using the Yale facial image dataset, Ramaiah et al. [15] deployed the CNN algorithm to get an average classification accuracy of 94% under varying illumination conditions. While this approach reinforces the use of CNN for thermal face recognition under varying lighting and facial expression conditions, the average accuracy of 94% and the inability to identify surgically altered faces may not meet the requirements of law enforcement agencies. In the same direction, Wu et al. [16] compared face recognition with traditional methods namely LBP, HOG, and moments invariant. They demonstrated that CNN achieves a higher recognition rate of 98%. While the experiment demonstrated the superiority of CNN in thermal face recognition under expression and illumination invariance, it is worth noting that the dataset did not include cases with surgically modified faces.

Yet, another direction is carried out by Fan et al. [17] who evaluated the efficacy of a fully convolutional network (FCN) in analyzing multi-face thermal images in unconstrained settings. The FCN achieved accuracies of 98%, 100%, and 100% for head rotation, expression, and illumination, respectively, showing optimal performance when faces are spaced apart, though its practicality for law enforcement is constrained by this requirement. Moreover, Grudzien et al. [6] presented

a thermal face verification method based on Siamese CNN, utilizing the ThermalFaceNet architecture to compare its performance with existing state-of-the-art CNN architectures. However, the research does not take care of either the effect of illumination or expression invariance. Furthermore, for better accuracy, Sayed et al. [18] deployed a pre-trained CNN to extract the features of the multiple convolution layers of the low-resolution thermal infrared images using a sample of 1500 resized thermal images, each with a resolution of 181×161 pixels. It achieved an impressive accuracy of 99% while significantly reducing processing and training times.

With the focus on different datasets, Manssour et al. [19] applied CNN to thermal and visual image datasets, attaining accuracies of 98% and 98% on the DHU and DHUFO databases, respectively, showcasing reduced training and processing times. Their approach also considered illumination and, to some extent, expression invariance but not surgical changes.

Another direction is carried out in the realm of hybrid algorithms. An example of this is the work carried out by Kakarwal et al. [20] who applied two ANN algorithms—backpropagation and Levenberg–Marquardt—across databases of visible, thermal, and fused images. The backpropagation algorithm achieved 92% accuracy, whereas Levenberg–Marquardt reached 83%. Despite CNNs being more proficient in image processing, these ANNs still show considerable performance.

More recently, using the thermal dataset USTC-NVIE, Mahouachi and Akloufi [21] proposed a Deep CNN architecture based on the FaceNet architecture and the Multi-task Cascaded Convolutional Networks (MTCNN) model for thermal face matching, demonstrating promising results. While the research demonstrated the promising application of infrared imaging, it mainly focuses on illumination invariance, placing less emphasis on expression invariance and without addressing surgical invariance. A similar attempt is carried out by Muller et al. [22] who demonstrated that established network architectures can be trained for the task of multiclass face analysis in thermal infrared focusing in specific areas of the face while irrelevant features such as glasses, masks and jewelry are not considered. While CNN provides promising results, the study primarily focuses on detecting health issues rather than face-matching; thus, it is not oriented toward law enforcement.

The DL-based approach proposed by Aji et al. [23] utilizes CNN to extract features from infrared images, employing a combination of Haar Cascade and Local Binary Pattern methods to delineate facial regions. The experiments demonstrate that CNN, combined with Haar Cascade, significantly enhances accuracy, achieving scores up to 95%. This promising experiment underscores the potential of thermal face recognition in law enforcement, showing high accuracy, precision, recall, and F1-scores. However, the study's focus on illumination invariance omits considerations for expression or surgical invariance. Similarly, Ashfaq and Akram [3] developed a framework based on the DL Inception v3 model using two datasets, Sleek thermal and FLIR, each including three classes: cat, car, and man. They achieved an accuracy of 98% for the FLIR dataset and 100% for the Sleek Thermal dataset using a CNN classifier. Since the dataset primarily consists of mixed images, the accuracy may not accurately reflect the recognition of human faces by law enforcement agencies.

A more recent attempt is carried out by Taspinar [24] who introduced a new CNN model, named LW-CNN (Light Weight CNN), which was compared with the pre-trained VGG16 model. LW-CNN achieves high classification accuracy on three different datasets with results reaching up to 98%, proving the model's effectiveness in handling thermal images without the need for pre-processing. Moreover, Tsai et al. [25] introduced a thermal imaging-based authentication system to verify employees wearing COVID masks against an organizational database, achieving a 94.1% accuracy rate with CNN. While impactful for organizational authentication, the relative accuracy limits its application in law enforcement.

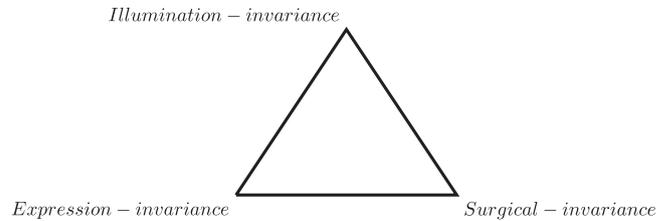


Fig. 1. Thermal Face Recognition Model (TFRT).

Last but not least, Bai et al. [26] developed a non-intrusive personal thermal comfort model that records and predicts interior inhabitants' thermal preferences using infrared facial recognition and ML. The method includes extracting facial temperatures from six locations of interest utilizing key point extraction techniques and infrared face identification. Using gradient boosting decision tree (GBDT) and random forest (RF) models, the feature relevance of these variables was determined. This method outperformed traditional deep neural networks in terms of training times, processing complexity, and prediction accuracy of 90%. However, predictions at different distances and surgical invariance were lacking. The literature review identified and demonstrated two major dimensions of thermal face recognition: illumination invariance and expression invariance, predominantly using the CNN classifier. However, this research proposes a third dimension necessary for law enforcement agencies, which is surgical invariance, and this dimension is lacking in all the research explored. Therefore, the proposed research is the first study in thermal face recognition to address illumination invariance, expression invariance, and surgical face invariance simultaneously through the TFRT model as per Fig. 1. This will aid in the detection of surgically altered faces, thus enhancing its applicability in law enforcement agencies. A summary of the literature review is given in Table 1.

3. Preliminaries

This section gives overviews of the techniques and algorithms used in the proposed thermal face detection model.

3.1. Multigranulation rough set

The Multigranulation Rough Set (MGRS) represents an advancement in classical rough set theory, tailored to concurrently manage diverse information systems and multiple granularities of knowledge. Its effectiveness is pronounced in contexts where information is plagued by vagueness, uncertainty, and incompleteness, which are commonplace in domains of knowledge representation and decision-making. Within the MGRS framework, the universe of discourse, denoted as U , is scrutinized through a collective of information systems or granulations, each delineating distinct attribute subsets, represented as C_i where $i = 1, 2, \dots, m$. These granulations coalesce into a structure denoted by $\mathcal{G} = \{C_1, C_2, \dots, C_m\}$. At the core of MGRS theory are the notions of lower and upper approximations of a set, which extend the classical rough set approximations to embrace a spectrum of granulations.

For a subset $X \subseteq U$, the lower approximation in the context of \mathcal{G} is articulated as:

$$\underline{\mathcal{G}}(X) = \bigcap_{i=1}^m C_i(X) = \{x \in U \mid \forall i \in \{1, \dots, m\}, \exists Y_i \in C_i : x \in Y_i \subseteq X\} \quad (1)$$

Conversely, the upper approximation is defined by:

$$\overline{\mathcal{G}}(X) = \bigcup_{i=1}^m \overline{C_i}(X) = \{x \in U \mid \exists i \in \{1, \dots, m\}, \exists Y_i \in C_i : x \in Y_i \cap X \neq \emptyset\} \quad (2)$$

Table 1
Summary of related work and identification of research gap.

Study, year	Illumination invariance	Expression invariance	Surgical invariance	Dealing with heterogeneity
Ramaiah et al. [15], 2015	✓	✗	✓	✗
Wu et al. [16], 2016	✓	✓	✗	✗
Fan et al. [17], 2017	✓	✗	✗	✗
Sayed et al. [18], 2018	✓	✓	✗	✗
Grudzien et al. [6], 2018	✗	✗	✓	✗
Manssor et al. [19], 2019	✓	✓	✗	✗
Mahouachi and Akloufi [21], 2021	✓	✗	✗	✗
Muller et al. [22], 2021	✗	✗	✓	✗
Kakarwal et al. [20], 2021	✓	✗	✗	✗
Aji et al. [23], 2022	✓	✗	✗	✗
Ashfaq and Akram [3], 2022	✓	✓	✗	✗
Tasi et al. [25], 2023	✗	✗	✓	✗
Bai et al. [26], 2024	✓	✓	✗	✗
Our model	✓	✓	✓	✓

The conventional MGRS model is primarily predicated on categorical attributes, a constraint that introduces unreliability in the approximation space when encountering numerical features. Addressing this limitation, we introduce an enhanced variant termed the Neighborhood Multigranulation Rough Set (NMGRS) model, which refines the approximation space computation by incorporating both categorical and numerical attributes.

An object u_i merits inclusion in the lower approximation of \mathcal{U}_{c_j} if it aligns with both categorical and numerical neighborhood criteria. This dual consideration precipitates a more precise approximation space, markedly beneficial for datasets that amalgamate categorical with numerical attributes. Building upon this refined model, we present the Particle Swarm Optimization NMGRS (NM-PSO) algorithm, wherein the NMGRS serves as the fitness function, pivotal in the iterative evaluation of feature subsets. To expedite convergence to the minimal feature subset, NMGRS systematically eliminates redundant features identified through successive iterations, ensuring a progressive refinement of the feature subset.

3.2. Convolutional neural network (CNN)

CNNs are an archetype of DL architectures, ingeniously designed to process data that adheres to a grid-like topology, for instance, images or sequential inputs. Their utility is extensively acknowledged in the realm of computer vision for tasks such as image classification, object detection, and image segmentation, to name a few. The underlying inspiration for CNNs can be traced to the cognitive processes of the human visual cortex, which they mimic to discern and analyze visual data [27]. A typical CNN comprises a stratified structure of convolution layers, non-linear activation functions like the Rectified Linear Activation Function (ReLU), pooling layers, fully connected layers, and regularization mechanisms such as dropout [28]. An exemplar CNN structure is depicted in Fig. 2.

In the CNN strata, the network employs a plethora of filters to parse the input, detecting distinctive, localized features within the data [29]. ReLU, as a non-linear activation function, is pivotal in capacitating the network to capture complex data relationships and patterns [30]. Subsequent to the convolutional layer, pooling layers perform down-sampling, thus attenuating the dimensionality of the feature maps while preserving essential features. This dimensionality reduction not only curtails the computational burden but also endows the network with translational robustness [31]. The fully connected layers, situated towards the end of the network, are tasked with synthesizing the high-level representations extracted from the input data and executing predictions [32]. To alleviate the risk of overfitting, dropout layers are incorporated, randomly deactivating a subset of neurons during training [28].

The CNN architecture is the most important factor of its performance and efficiency [33]. The arrangement of the layers, the elements used in each layer, and how they are constructed, all these factors affect the speed and accuracy at which various tasks can be carried out.

3.3. Max/average pooling CNN (MAP-CNN)

This section presents the MAP-CNN algorithm which was used to extract the features from the thermal face dataset. The improvement was done through the pooling layer by proposing a novel pooling, AVG-MAX pooling layer.

3.3.1. Vector pooling block

The vector pooling block is a novel component in the neural network that processes an input tensor I_{input} , with dimensions $C \times H \times W$, where C is the number of channels, H is the height, and W is the width of the tensor. This block is designed to include two distinct processing paths: one for horizontal pooling and the other for vertical pooling.

In the horizontal pooling path, the operation can be mathematically described as $P_{horizontal} = \max(I_{input}[c, h, :])$ for each channel c and height h . Similarly, the vertical pooling path operates along the height dimension. Each path then leads to a 1×1 convolutional layer, which can be represented as $F_{convolved} = K \cdot I_{input}$, where K is the 1×1 convolution kernel.

Following each convolution, a Rectified Linear Activation Function (ReLU) layer is integrated, described by $F_{ReLU} = \max(0, F_{convolved})$. This integration enhances the network's stability and accelerates its convergence.

The core feature of this pooling block is its ability to merge the feature vectors from both horizontal and vertical paths while maintaining dimensionality reduction benefits. This merging is achieved through an element-wise summation, $F_{combined} = F_{horizontal} + F_{vertical}$, ensuring a comprehensive feature vector representation. The resultant feature vector undergoes another ReLU layer, formulated as $F_{final} = \max(0, F_{combined})$, further aiding in rapid model convergence.

The vector pooling block's overall design is a testament to its effectiveness in FE. It leverages parallel pooling, strategic 1×1 convolutions, element-wise vector summation, and multiple ReLU activations to optimize the FE process. The operational mechanics and structure of the Vector pooling block are elucidated in Fig. 3.

3.3.2. MAX-AVG VBP pooling layer

The AVG-MAX Vector Pooling Block (VPB) in CNN architectures unites the two principal pooling methods: Max-Pooling and Average Pooling. Max-Pooling, denoted as $\text{MaxPool}(I_{input})$, selects the maximum element from a defined pooling region in the input tensor I_{input} . Conversely, Average Pooling, represented as $\text{AvgPool}(I_{input})$, calculates the mean of elements within the pooling region.

The AVG-MAX VPB ingeniously combines these techniques, extracting salient features by integrating the robustness of Max-Pooling with the comprehensive approach of Average Pooling. This is achieved through four parallel pathways: Max-horizontal pooling (MaxPool_H), Max-vertical pooling (MaxPool_V), Average horizontal pooling (AvgPool_H), and Average vertical pooling (AvgPool_V). Each pathway

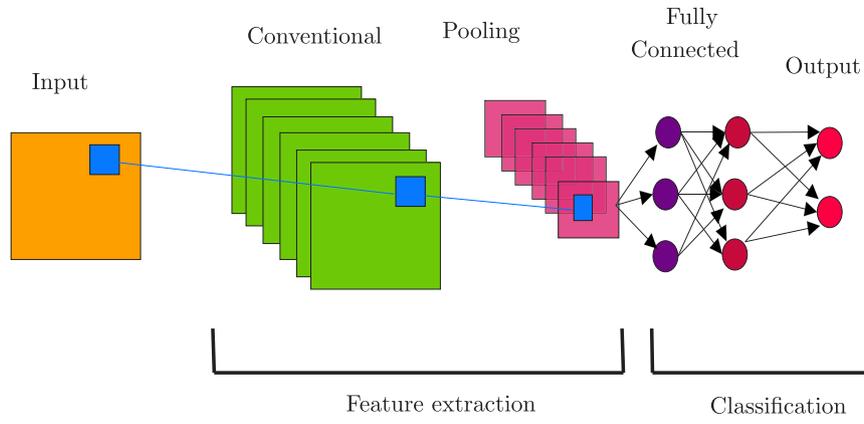


Fig. 2. Example of CNN.

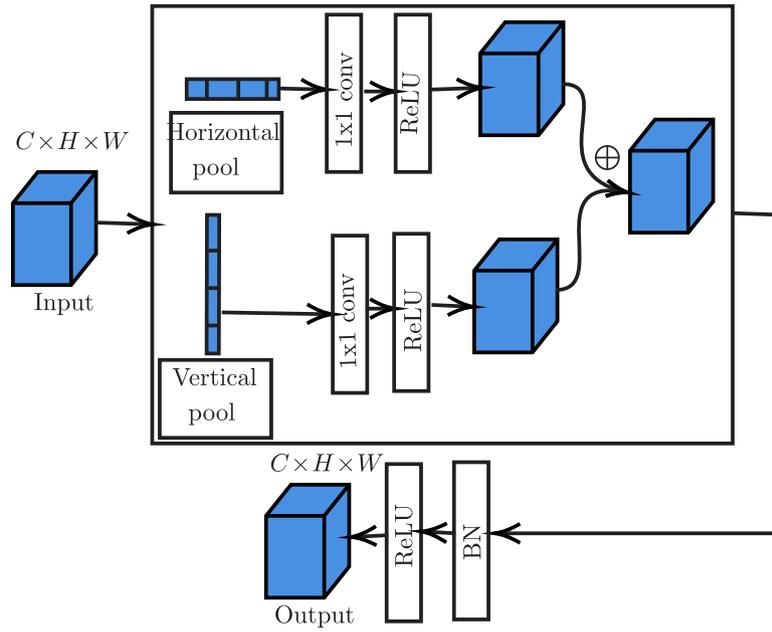


Fig. 3. Vector pooling block [28].

concludes with a 1×1 convolution ($F_{conv} = K \cdot I_{input}$), where K is the convolution kernel.

After convolution, the AVG-MAX VPB incorporates a Rectified Linear Activation Function (ReLU) layer ($F_{ReLU} = \max(0, F_{conv})$), promoting stability and faster convergence. The fusion of feature vectors from each ReLU layer is performed through element-wise summation ($F_{combined} = F_{ReLU1} + F_{ReLU2} + F_{ReLU3} + F_{ReLU4}$), harmonizing the insights from different pooling strategies.

This operation is followed by a batch normalization layer ($F_{BN} = \text{BatchNorm}(F_{combined})$), standardizing the feature set for enhanced network stability and model generalizability. The AVG-MAX VPB culminates with an additional ReLU layer post-batch normalization ($F_{final} = \max(0, F_{BN})$), further accelerating convergence.

Overall, the AVG-MAX VPB effectively combines concurrent pooling paths, 1×1 convolutions, and element-wise summation, along with batch normalization and dual ReLU activations, establishing a robust architecture for feature amalgamation. This design promotes network stability and expedites convergence, crucial for efficient CNN operations. The structure and functionality of the AVG-MAX VPB are detailed in Fig. 4.

4. Proposed robust thermal face recognition model

The proposed model combined the power of the DL and ML approaches. It used the CNN to extract deep features while the ML was used in the classification thus making our model computationally efficient. The following sections give overviews of the model components (i.e., FE, FS (a novel model was proposed), and the classification).

4.1. Feature extraction process

A CNN can be used as a powerful feature extractor in various ML tasks [34]. By leveraging the learned representations from CNNs, we can extract rich and meaningful features from input data. In CNNs, the initial layers perform local FE by applying convolutional filters across the input data. These filters capture low-level features such as edges, textures, or color gradients. As the network progresses through deeper layers, it learns to detect more complex and higher-level features, including object parts and semantic structures. The advantage of using a CNN as a feature extractor lies in its ability to automatically learn hierarchical and discriminative features from raw data. CNNs have to be highly effective in capturing complex patterns and structures in images, making them particularly well-suited for computer vision tasks [35].

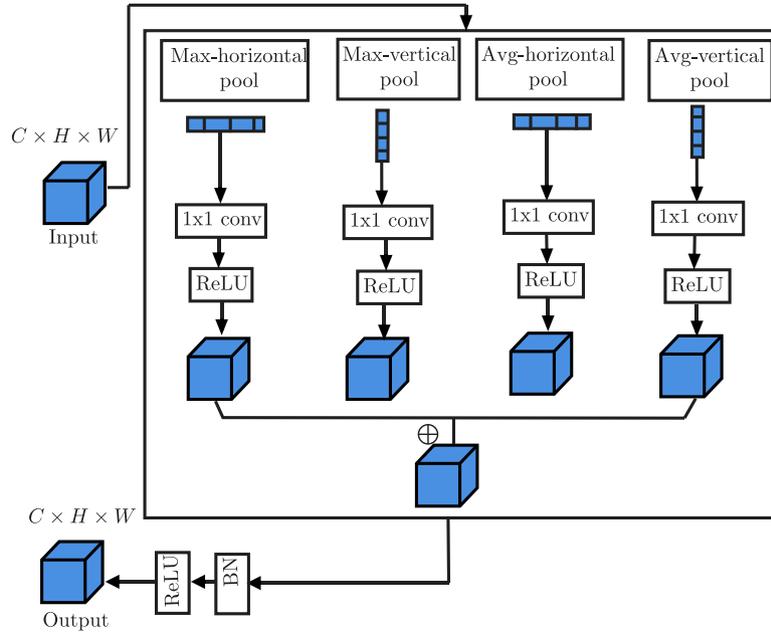


Fig. 4. AVG-MAX Vector Pooling Block (VPB) combining Max-Pooling and Average Pooling through four parallel paths (Max-horizontal Pool, Max-vertical Pool, Avg-horizontal Pool, Avg-vertical Pool), followed by 1×1 convolutions and ReLU activation. The outputs are merged via element-wise summation (\oplus), followed by batch normalization (BN), and a final ReLU activation [28].

Furthermore, transfer learning can be applied to leverage pre-trained CNN models. Pre-trained CNNs, which have been trained on large-scale datasets like ImageNet, have already learned a rich set of general features. By using these pre-trained models, we can extract features that are generic and transferable across different tasks. This approach is especially useful when the target dataset is small, as it allows us to benefit from the knowledge learned from large-scale datasets [36]. Therefore, CNNs can serve as powerful feature extractors by learning hierarchical representations of input data. These extracted features can be utilized in various downstream tasks, enabling us to leverage the impressive capabilities of CNNs for a wide range of machine-learning applications [37].

4.1.1. Feature extraction using AVG-MAX-VPB-based CNN

The novel MAP-CNN, presented in Section 3.3.2, could be a better FE technique than the traditional CNN. This is because the former is based on the combination of two pooling techniques: average pooling and max pooling with the concept of VPB. The traditional pooling layer uses a fixed square kernel size, which is not as effective in extracting a diverse set of features. In contrast, the AVG-MAX-VPB methodology takes advantage of the different characteristics of each pooling method, resulting in a more significant feature set [28].

In AVG-MAX-VPB, average pooling is used to compress the input image, as the output of average pooling is a smaller matrix containing the overall features of an image, with less emphasis on specific details. However, the max pooling method is used to identify the most prominent features of the image, as the output of max pooling is a smaller matrix with the most significant features of the image intact. While the VPB is used to extract both local and global features, it consists of two pathways that extract features along the horizontal and vertical orientations, respectively [28]. By using long and narrow pooling kernels along both pathways, the VPB is able to collect global features that span across large regions of the image, as well as local features that are specific to smaller regions. The combination of these pooling techniques with the concept results in a feature set that is more diverse and robust compared to using any of these pooling methods independently. This diverse feature set could be used for a variety of different tasks such as image classification, object detection, and object recognition. The steps of using CNN as a feature extractor are as follows:

1. Start by selecting a pre-trained CNN model that has been trained on a large-scale dataset.
2. Exchange each pooling layer in the pre-trained CNN with AVG-MAX-VPB.
3. Remove the Fully Connected Layers as these layers are responsible for mapping the extracted features to specific classes and are not needed for FE.
4. Extract Features from the Convolutional Layers.
5. Flatten or Pool the Feature Maps depending on the specific task and requirements.
6. Use the extracted features for various downstream tasks, such as image classification, object detection, or image retrieval.

4.2. Neighborhood multi-granulation rough set (NMGRS)

As shown above, MGRS has unreliable approximation space. To address this problem, a new variant of MGRS was introduced in this paper. This new variant used the neighborhood relation to design neighborhood MGRS (NMGRS). The description of this design process is given below.

In the context of NMGRS, each object u_i in the universe \mathcal{U} is associated with three distinct neighborhood classes, which are predicated upon the sets \mathcal{C} , \mathcal{N} , and \mathcal{A}^- , respectively. These classes are formally defined as follows:

$$\theta_{\mathcal{C},u_i} = \{u_j \mid \forall a_k \in \mathcal{C}, v_{u_i,a_k} = v_{u_j,a_k}\}, \quad (3a)$$

$$\theta_{\mathcal{N},u_i} = \{u_j \mid l_{\mathcal{N}}(u_i, u_j) \leq \epsilon\}, \quad (3b)$$

$$\theta_{\mathcal{A}^-,u_i} = \theta_{\mathcal{C},u_i} \cap \theta_{\mathcal{N},u_i}, \quad (3c)$$

where $\epsilon \in (0, 1]$ is a user-defined threshold, and $l_{\mathcal{M}}(u_i, u_j)$ denotes the Euclidean distance between objects u_i and u_j with respect to the attribute set \mathcal{M} , calculated as:

$$l_{\mathcal{M}}(u_i, u_j) = \sqrt{\sum_{k=1}^{|\mathcal{M}|} (v_{u_i,a_k} - v_{u_j,a_k})^2}.$$

Using these neighborhood classes, the NMGRS approximation space for a given decision subset $\mathcal{L}_{\mathcal{E}_j}$ is computed as follows. The lower

approximation $\underline{\Theta}_{\mathcal{U}_{\zeta_j}}$ is determined by:

$$\underline{\Theta}_{\mathcal{U}_{\zeta_j}} = \{u_i \in \mathcal{U} \mid \theta_{\mathcal{A}^-, u_i} \cap \mathcal{U}_{\zeta_j} \geq \tau |\mathcal{U}_{\zeta_j}|\}, \quad (4)$$

while the upper approximation $\overline{\Theta}_{\mathcal{U}_{\zeta_j}}$ is defined as:

$$\overline{\Theta}_{\mathcal{U}_{\zeta_j}} = \{u_i \in \mathcal{U} \mid \theta_{\mathcal{A}^-, u_i} \cap \mathcal{U}_{\zeta_j} \neq \emptyset\}.$$

The NMGRS model, through these calculations, promises an enhancement in computing the approximation space over the MGRS model, particularly by accommodating both categorical and numerical features in the data. The forthcoming theorem elucidates the efficiency of the NMGRS model relative to its MGRS predecessor.

Theorem 1. *The lower approximation, denoted as $\underline{\Theta}_{\mathcal{U}_{\zeta_j}}$, which is derived from the NMGRS model, serves to enhance the approximation yielded by the MGRS model. The refinement process ensures that the lower approximation in NMGRS encapsulates a more precise subset of objects that are unequivocally within the decision subset \mathcal{U}_{ζ_j} . This precision is achieved by considering both categorical and numerical features to determine the neighborhood classes, thereby leading to a more accurate and reliable approximation space.*

Proof. Consider an object u_i belonging to the lower approximation $\underline{\Theta}_{\mathcal{U}_{\zeta_j}}$ as computed by the NMGRS model, indicated by (4). This entails that the neighborhood class $\theta_{\mathcal{A}^-, u_i}$ is a subset of the decision subset \mathcal{U}_{ζ_j} . According to the Eq. (4), it implies that $\theta_{\mathcal{A}^-, u_i}$ does not encompass any objects associated with a disparate decision subset \mathcal{U}_{ζ_k} . Hence, the inconsistency measure $I(\theta_{\mathcal{A}^-, u_i}, \mathcal{U}_{\zeta_j})$ equals zero, leading to the conclusion that $I(\theta_{\mathcal{A}^-, u_i}, \mathcal{U}_{\zeta_k}) \geq I(\theta_{\mathcal{A}^-, u_i}, \mathcal{U}_{\zeta_j})$. As a result, u_i is also a member of the lower approximation $\underline{\Theta}_{\mathcal{U}_{\zeta_k}}$ when computed using the MGRS model as described by (4).

Conversely, if u_i is part of $\underline{\Theta}_{\mathcal{U}_{\zeta_j}}$ as derived through the MGRS model, indicated by (4), then $I(\theta_{\mathcal{A}^-, u_i}, \mathcal{U}_{\zeta_k}) \geq I(\theta_{\mathcal{A}^-, u_i}, \mathcal{U}_{\zeta_j})$. However, this inequality does not infer that $I(\theta_{\mathcal{A}^-, u_i}, \mathcal{U}_{\zeta_j})$ is null. Consequently, we cannot claim that $\theta_{\mathcal{A}^-, u_i}$ is entirely contained within \mathcal{U}_{ζ_j} . Thus, the lower approximation determined by the NMGRS model is inclusive or equal to that ascertained by the MGRS model. ■

Definition 1 (Fitness Function). The NMGRS based fitness function \mathcal{F} expresses the amount of goodness of the decision label d , of a dataset, on the set \mathcal{M} and is given by

$$\mathcal{F}_{\mathcal{M}}(d) = \frac{\sum_{j=1}^r |\underline{\Theta}_{\mathcal{U}_{\zeta_j}}|}{\sum_{j=1}^r |\overline{\Theta}_{\mathcal{U}_{\zeta_j}}|}. \quad (5)$$

The value of the fitness function, which is in $[0, 1]$, is called the goodness value. ■

It should be noted that the fitness function is used to differentiate between the necessary and unnecessary features. This procedure is employed as per Definition 2.

Definition 2 (Clipping the Data). Clipping a dataset involves removing features that do not contribute to determining its decision label d . This process focuses on eliminating irrelevant or redundant features, and streamlining the dataset for more effective analysis. By discarding these superfluous features, the dataset becomes more manageable and better suited for precise decision-making tasks. ■

In our research, the Neighborhood Multigranulation Rough Set (NMGRS) based fitness function, as outlined in Eq. (5), plays a pivotal role in discerning the indispensability of features in a dataset. This approach involves a strategic process to prune a set $\mathbb{A} \subseteq \mathcal{A}^-$ by iteratively removing individual features and evaluating the impact on the dataset's fitness value.

The procedure is as follows: for each feature a_i in \mathcal{A}^- , it is removed from the set \mathbb{A} , and the fitness value \mathcal{F} is calculated both before and after this removal. If the fitness value decreases subsequent to the removal of the feature, it implies that the feature is integral to the dataset, thereby necessitating its reinstatement into \mathbb{A} . Conversely, if the fitness value remains constant, it indicates the feature's redundancy, warranting its permanent exclusion. Formally, a feature a_i is necessary if $\mathcal{F}_{\mathcal{A}^- - \{a_i\}}(d) < \mathcal{F}_{\mathcal{A}^-}(d)$, and unnecessary otherwise.

The culmination of this iterative process yields a minimal subset \mathbb{A} of \mathcal{A}^- that encapsulates only the essential features. This subset \mathbb{A} is defined as minimal with respect to the decision d if $\mathcal{F}_{\mathbb{A}}(d) \geq \mathcal{F}_{\mathcal{A}^-}(d)$ and for each feature a_i within \mathbb{A} , $\mathcal{F}_{\mathbb{A} - \{a_i\}}(d) < \mathcal{F}_{\mathbb{A}}(d)$.

4.3. Improved particle swarm optimization using NMGRS for feature selection

Particle Swarm Optimization (PSO) is an evolutionary computational technique that addresses optimization problems by iteratively refining candidate solutions in accordance with a specific quality metric. The algorithm is inspired by the collective behavior observed in natural systems, such as birds flocking or fish schooling, particularly in their food-finding strategies.

To elaborate, consider a flock of birds searching for food. Within this flock, certain birds exhibit a higher proficiency in locating food sources. These proficient birds can be analogized as leaders who guide the direction of the search, while the rest of the flock, akin to followers, adapt their movements based on the leaders' cues. Periodically, the leaders disseminate their knowledge about the food's location, facilitating the followers in recalibrating their trajectories and enhancing the overall efficacy of the search. This dynamic interplay allows for the continuous exchange of roles where a follower, upon improved performance, may ascend to a leader's status, while a leader may transition to a follower's role if outperformed.

The PSO algorithm captures this essence by assigning each member of the swarm, represented as particles in the algorithm, a unique identity. These particles, akin to birds in a flock, navigate the solution space, with their positions symbolizing potential solutions to the optimization problem. The algorithm iteratively updates the positions of these particles, influenced by both their own best-known positions and the swarm's collective best position. Through this process, the swarm collectively converges towards an optimal solution, with each particle contributing to and benefitting from the shared knowledge of the group's experiences. These adept birds can be perceived as leaders, providing directional insights, while the others can be viewed as followers, adapting their positions based on the information acquired from the leaders. Periodically, the leaders share their knowledge about the food location, allowing the followers to adjust their trajectories and enhance their search efficacy. This dynamic allows for the potential elevation of a follower to a leader, based on its performance in locating the food, and a leader may become a follower if outperformed. Consider a flock of 10 birds, with only the top 3 efficient ones deemed as leaders. The flock adjusts its position every 5 minutes to evaluate their proximity to food and share insights. The leading trio in each round disseminates the positional information contributing to their successful forage, potentially altering the leader-follower dynamic in subsequent rounds.

There are numerous ways to deploy PSO. We plan to create a streamlined version, tailored to efficiently and precisely solve FS problems. The algorithm will emulate actual flocking behavior to pinpoint the optimal feature subset in a dataset. Each bird is assigned a unique number i , $i = 1, 2, \dots, N$, and a role, either a leader or a follower. These unique numbers are static, but the roles can interchange based on each iteration's performance. Initially, all birds assume random positions, representing potential solutions, and commence the search. A position, in this scenario, is a vector—known as a particle—that corresponds to a subset of the dataset features. After each iteration, the

PSO algorithm designates the top-performing $n < N$ birds as leaders allowing them to share their particle data with the followers to refine the subsequent search. Thus, the PSO algorithm resolves the FS tasks by a two-step process: an exploration step and an exploitation step, continually iterating until the optimal solution is identified.

4.3.1. NM-PSO exploration step

Particle $i \geq 1$ possesses in iteration $k \geq 1$ a position vector $\mathbf{X}_{i_k} = [x_1, x_2, \dots, x_n]$, $x_j \in \{0, 1\}$, continually undergoing updates in consecutive iterations. This vector specifies a subset of the features of the Data under review, where x_j highlights the features considered. For instance, $\mathbf{X}_{3_2} = [1, 0, 1, 0]$ signifies that particle 3 in iteration 2 is concentrating on features a_1 and a_3 .

Definition 3 (Vector-Set Correspondence). For any position vector $\mathbf{X} = [x_1, x_2, \dots, x_n]$, the equivalent feature set \mathbb{B} consists only of the features with non-zero entries in \mathbf{X} . Conversely, for any feature set \mathbb{B} , the matching position vector \mathbf{X} is the vector with a 1 for each element corresponding to a feature in \mathbb{B} . This implies $\mathbf{X} \Leftrightarrow \mathbb{B}$ if $x_j = 1 \Leftrightarrow b_j \in \mathbb{B}$. ■

As an illustration, with feature set $\mathbb{B} = \{a_1, a_3\}$, the related position vector is $\mathbf{X} = [1, 0, 1, 0]$, and reciprocally.

$$\mathbf{X} = [1, 0, 1, 0] \Leftrightarrow \mathbb{B} = \{a_1, a_3\}$$

Each particle in the Particle Swarm Optimization (PSO) algorithm commences its journey in iteration one by randomly selecting a subset of the initial n conditional features. This forms the initial position vector, which is subject to iterative adjustments influenced by the PSO algorithm. As the process progresses to iteration $k \geq 1$, the algorithm performs several key steps. The position vectors of particles, particularly those not in leadership roles, are modified. Concurrently, the fitness value δ of each particle's position vector is computed, necessitating its transformation into a corresponding feature set as per the defined criteria. The algorithm then identifies the top p particles with the highest fitness values as leaders, indicating their proximity to the optimal feature subset. This progression involves adjusting non-leader position vectors, assessing the fitness of these vectors, refining corresponding feature sets based on fitness evaluation, generating new position vectors from these refined sets, and promoting the particles with the highest fitness values to leadership roles in the next iteration. The iterative procedure persists until either a predetermined number of iterations T is reached or a feature subset is identified that matches or exceeds the fitness value of the original set.

4.3.2. NM-PSO exploitation step

In PSO, the exploitation step is crucial for particles to refine their positions in the search space based on their personal best positions and the global best position found by the swarm. Let S_k represent the swarm at iteration k , consisting of particles i , each having a position vector \mathbf{X}_{i_k} in the search space.

At each iteration $k \geq 1$, every particle i updates its velocity $\mathbf{V}_{i_{k+1}}$ and position $\mathbf{X}_{i_{k+1}}$ as follows:

$$\mathbf{V}_{i_{k+1}} = w\mathbf{V}_{i_k} + c_1r_1(\mathbf{P}_{i_k} - \mathbf{X}_{i_k}) + c_2r_2(\mathbf{G}_k - \mathbf{X}_{i_k})$$

$$\mathbf{X}_{i_{k+1}} = \mathbf{X}_{i_k} + \mathbf{V}_{i_{k+1}}$$

where:

- w is the inertia weight, controlling the impact of the previous velocity.
- c_1 and c_2 are cognitive and social coefficients, respectively.
- r_1 and r_2 are random numbers uniformly distributed in $[0, 1]$.
- \mathbf{P}_{i_k} represents the personal best position of particle i up to iteration k .
- \mathbf{G}_k represents the global best position found by the swarm up to iteration k .

Once the new positions are computed, each particle evaluates the objective function at its new position, and the personal best positions and the global best position are updated accordingly. This process of exploitation helps the particles converge towards the optimal solution, leveraging both individual and collective experiences. NM-PSO uses the above mentioned strategies as per Algorithm 1 and is depicted graphically in Fig. 5.

5. Experimental results and discussion

To validate the proposed NM-PSO capabilities as a FS tool, we coded it in Python and run on a PC with Windows 11, Core i7 and 8 GB RAM. The objective of this experiment is to prove that using the proposed NMGRS as a fitness function to guide PSO reduction task is of major significance. Throughout the experiments, the dataset is extracted from Charlotte-ThermalFace dataset [38] about thermal face recognition. Two methodologies (traditional CNN and improved one) were used to extract the features. Therefore, we have in hand two datasets referred to as Alex-net (extracted from CNN) and Alex-net-AVG-MAX (extracted from MAP-CNN). For the sake of clarity, NM-PSO performance (number of selected feature subset size, convergence speed and classification performance as in [39]) is compared against ten related FS algorithms.

5.1. Dataset description

Charlotte-ThermalFace dataset [38] was used to evaluate the proposed thermal face recognition model. This dataset is the optimal choice for evaluating thermal face recognition in law enforcement applications as it contains images covering illumination and expression invariance. Additionally, since the images are thermal, the dataset also addresses surgical face invariance. The dataset comprises a large number of high-resolution thermal infrared face images which are collected in different environments, including indoor and outdoor settings, with varying lighting conditions, at different distances and diverse backgrounds. This diversity ensures that the dataset covers a wide range of real-world scenarios, making it highly suitable for law enforcement applications. One notable feature of the Charlotte-ThermalFace dataset is its inclusion of face images captured at various distances. The variation in camera-to-subject distance is crucial for evaluating the performance of face recognition systems in scenarios where individuals may be close to or far from the camera, as is often the case in law enforcement.

5.2. Scenario I: NM-PSO performance

In this scenario, we conducted a comprehensive evaluation of the newly developed NM-PSO FS algorithm, alongside testing our proposed FE method (CNN-AVG-MAX). The FE process was executed using both the conventional CNN model and the enhanced CNN-AVG-MAX version, resulting in two distinct datasets: CNN-data and MAP-CNN-data. Subsequently, the NM-PSO-based FS algorithm was applied to these datasets and benchmarked against ten contemporary algorithms, including PSO [40], SFO [41], CO [42], WOA [43], MEHHO [44], BHHO [45], IHHO [46], GA [47], DE [48], and GWNO [49]. The comparative analysis focused on the size of the selected feature subsets and the computational time required for FS.

The primary comparative analysis involved assessing the number of features selected by the 11 algorithms (10 competitors and NM-PSO) using both CNN-data and MAP-CNN-data. Table 2 displays the results, highlighting that NM-PSO (bold in the last row) consistently selected the smallest feature subset. Specifically, for the CNN-data, NM-PSO selected 2032 features, achieving a reduction rate of approximately 50%, outperforming its closest competitor, IHHO, which selected 3247 features (21% reduction rate). This result indicates that NM-PSO surpasses its best competitor by nearly 30% in terms of CNN-data. In the case of MAP-CNN-data, NM-PSO selected 1458 features, corresponding

Algorithm 1: Neighborhood multigranulation rough set particle swarm optimization (NM-PSO) algorithm

```

Input : ( $\mathcal{U}, \mathcal{A}, \mathcal{V}$ ) // Dataset
           $N$  // # of employed particles
           $n$  // # of best performing particles
           $T$  // Maximum number of iterations
Output:  $\mathbb{N}$  // Reduced subset of the dataset
1 Calculate  $\mathcal{F}_{\mathcal{A}^-}(d)$ , as per (5).
  //Setup phase:
2  $t = 1$ 
3  $\mathcal{F} = []$  //list to keep the particles' fitness; initially empty
4 for  $i = 1$  to  $N$  do
5   Create the pivot vector  $X_{ik}$  consisting of elements  $[y_1, y_2, \dots, y_n]$ .
6   Formulate the corresponding set  $\mathbb{A}_i$  in accordance with Definition 3.
7   Compute the fitness value  $\mathcal{F}_{\mathbb{A}_i}(d)$ , Eq. (5), and add it to the list  $\mathcal{F}$ .
8   Clip  $\mathbb{A}_i$ , as per Definition 2 //Removing superfluous features.
9   Reassemble the vector  $X_{ik}$  to reflect  $\mathbb{A}_i$  as per Definition 3.
10 end
11 Formulate the set  $\mathbb{L}_k$  as  $i_1, i_2, \dots, i_m$ , s.t.  $i_j$  represents the index of top  $m$  values in  $\mathcal{F}$ , i.e designating the most effective particles as leaders.
  //Iterative phase:
12 do
13    $t = t + 1$ 
14    $\mathcal{F} = []$ 
15   for  $i = 1$  to  $N$  do
16     if  $i \in \mathbb{L}_{k-1}$  then
17       //For a leading particle  $i$  do
18       Compute the fitness measure  $\mathcal{F}_{\mathbb{A}_i}(d)$  in accordance with Eq. (5) and include this measure in the fitness list  $\mathcal{F}$ .
19     end
20     else
21       //For a non leading particle  $i$  do
22       Derive the new position vector  $Y_{ik}$  based on the previous iteration's vector  $Y_{i_{k-1}}$ 
23       Based on Definition 3, develop the relevant feature set  $\mathbb{A}_i$  that aligns with the updated position vector  $Y_{ik}$ .
24       Compute the fitness value  $\mathcal{F}_{\mathbb{A}_i}(d)$  for the set  $\mathbb{A}_i$ , in accordance with Eq. (5), and then add this fitness score to the list  $\mathcal{F}$ .
25       Clip  $\mathbb{A}_i$  in accordance with Definition 2 //Effectively removing any superfluous features from  $\mathbb{A}_i$ .
26       After pruning, recreate the pivot vector  $X_{ik}$  so that it corresponds to the refined feature set  $\mathbb{A}_i$ , following the guidelines
27       outlined in Definition 3.
28     end
29   end
30   Form the leader set  $\mathbb{L}_k$  by selecting the particles with indices  $i_1, i_2, \dots, i_m$ , corresponding to the highest  $m$  fitness values in the set  $\mathcal{F}$ .
31   //This step designates the top-performing particles as leaders for the next iteration.
32   Allocate the highest fitness value,  $\mathcal{F}_{\mathbb{A}_i}(d)$ , to the variable  $Hdep$ . //This step ensures that the highest fitness value achieved in the
33   current iteration is recorded for subsequent comparison and analysis.
34 while ( $Hdep < \mathcal{F}_{\mathcal{A}^-}(d)$  AND  $t < T$ );

```

to a 64% reduction rate, while the best competitor, PSO, selected 3054 features (25% reduction rate). Thus, NM-PSO demonstrated a 39% improvement over its predecessor. These findings lead to two key observations: (a) NM-PSO emerges as the superior FS algorithm for both thermal face datasets, and (b) the performance of all algorithms enhanced with the MAP-CNN-data, compared to the traditional CNN. This suggests that the MAP-CNN-AVG-MAX model is more effective in extracting representative features from thermal face images than the traditional CNN model, underlining its efficacy. Remarkably, NM-PSO showed a significant 16% improvement in reduction rate with the MAP-CNN-data compared to the 12% improvement demonstrated by the next best algorithms, PSO and CO.

The NM-PSO algorithm's exceptional performance can be attributed to several key design elements that collectively enhance its efficiency and accuracy. Firstly, the precise computation of the lower approximation in the NMGRS, as detailed in Eq. (4), plays a pivotal role. This calculation ensures an equitable balance between the numerical and categorical aspects of the data, thus preventing any one type from overshadowing the other. Essentially, this means that objects are included in the lower approximation based on a comprehensive analysis of both their numerical and categorical characteristics. Secondly, the algorithm's process of feature elimination is meticulously designed to consider the significance of each feature in relation to the decision class, aligning with Definition 2 (data clipping). This approach is distinct from traditional methods as it determines the relevance of features within a more sophisticated framework, taking into account the interplay between lower and upper approximations. Additionally, the NM-PSO algorithm is structured such that once a feature is deemed irrelevant and removed in a particular iteration, it is not reconsidered in subsequent iterations. This strategy ensures that the elimination of a

Table 2

Comparative analysis of the feature subset sizes selected by the NM-PSO algorithm and its ten competing counterparts, focusing on both standard CNN-data and the enhanced CNN-data.

FS algorithm	# selected features with respect to CNN-data	Percentage of improvement with respect to CNN-data	# selected features with respect to MAP-CNN-data	Percentage of improvement with respect to MAP-CNN-data
SFO	3448	15.82	3361	17.94
GWNO	3165	22.73	3151	23.07
CO	3619	11.65	3120	23.83
MEHHO	3480	15.04	3002	26.71
BHHO	3490	14.79	3231	21.12
GA	3410	16.75	3109	24.10
IHHO	3247	20.73	3007	26.59
WOA	3567	12.92	3153	23.02
PSO	3533	13.75	3054	25.44
DE	3266	20.26	3090	24.56
NM-PSO	2032	50.39	1458	64.40

feature does not compromise the distinctiveness of the objects, as they continue to be effectively differentiated using the remaining features.

In the second phase of these experiments, we focused on evaluating the time efficiency of the algorithm in locating the optimal feature subset. As illustrated in Fig. 6, it is apparent that the NM-PSO algorithm (represented by the red bar) consistently outperforms others in terms of time efficiency for both datasets. This implies that NM-PSO demonstrates remarkable speed in dimensionality reduction, underscoring its effectiveness in swiftly identifying the most compact and relevant feature subset. This high level of time efficiency can be primarily attributed to the sophisticated design of the NM-PSO's fitness

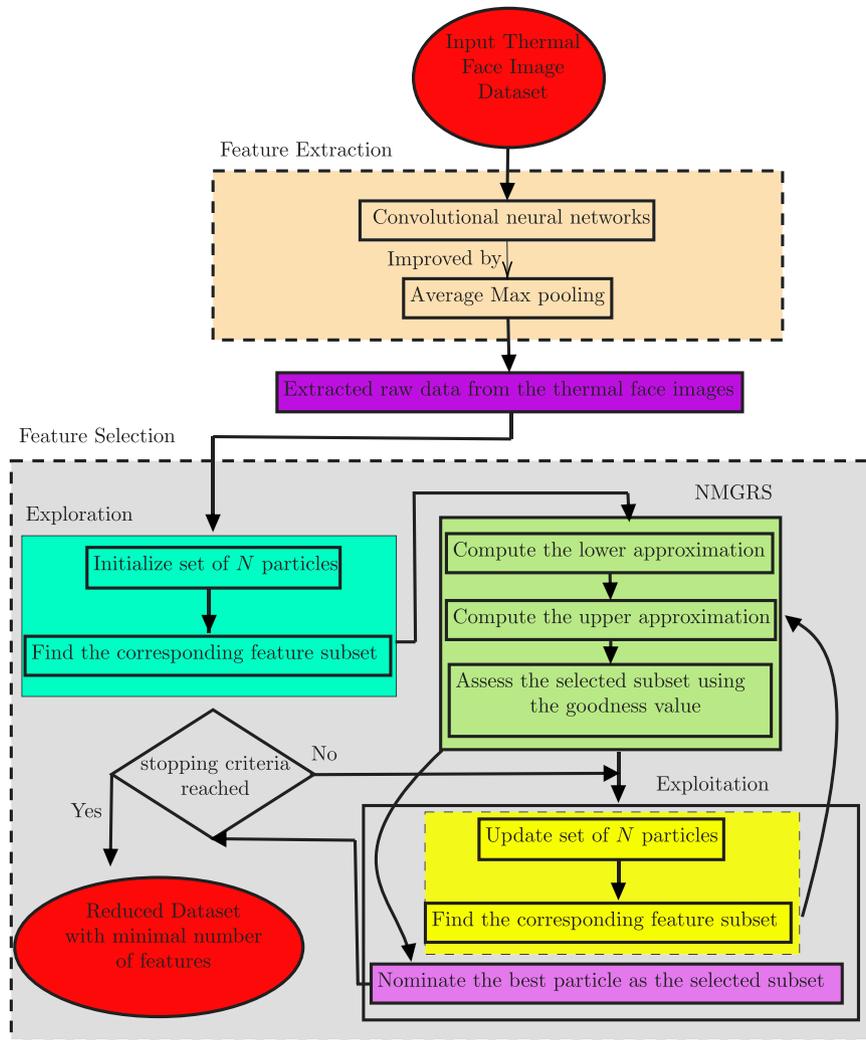


Fig. 5. The flowchart of the proposed model. The first block represents the feature extraction task for the thermal face images in which the CNN is improved using the average max pooling. The resulting equivalent raw data for the thermal face images are fed to the second block which represents the FS task implemented using PSO while improved using NMGRS.

function, which is integrated with the NMGRS model. This integration is crucial as it guarantees that once a feature is deemed irrelevant and excluded from the subset, it is not reconsidered in future iterations. This approach significantly streamlines the FS process, ensuring rapid convergence towards the minimal feature subset, thus saving valuable computational time and resources.

5.3. Scenario II: NM-PSO convergence

The second scenario examines the convergence speed (i.e., the algorithm’s behavior across iterations until the desired feature subset is identified). Evaluating this convergence speed is pivotal in bio-inspired FS algorithms. This is because for an FS algorithm to assuredly converge to the optimal minimal feature subset, it should consistently exhibit high fitness values across iterations. The investigation of the convergence offers an insightful perspective on the efficacy of our proposed NM-PSO FS algorithm comparing with the other 10 algorithms during the FS operation. The results of this scenario are summarized in Fig. 7. As illustrated in this figure, the fitness trajectory of the NM-PSO consistently surpasses that of the competing algorithms, indicating the robust and sustained superiority of NM-PSO throughout the process.

Table 3 Comparison of the standard deviation of fitness values for NM-PSO and its counterparts using CNN-data and MAP-CNN-data.

FS algorithm	Standard deviation with respect to CNN-data	Standard deviation with respect to MAP-CNN-data
SFO	0.20	0.26
GWNO	0.18	0.21
CO	0.19	0.25
MEHHO	0.24	0.16
BHHO	0.19	0.24
GA	0.25	0.21
IHHO	0.18	0.28
WOA	0.27	0.15
PSO	0.24	0.16
DE	0.22	0.18
NM-PSO	0.13	0.12

To rigorously evaluate the consistency and precision of the fitness values generated by the NM-PSO algorithm in comparison with competing algorithms, an analysis of the standard deviation (σ) of these fitness values was conducted. This analysis provides an insight into the variability of the fitness values throughout the \mathbb{T} iterations, supplementing

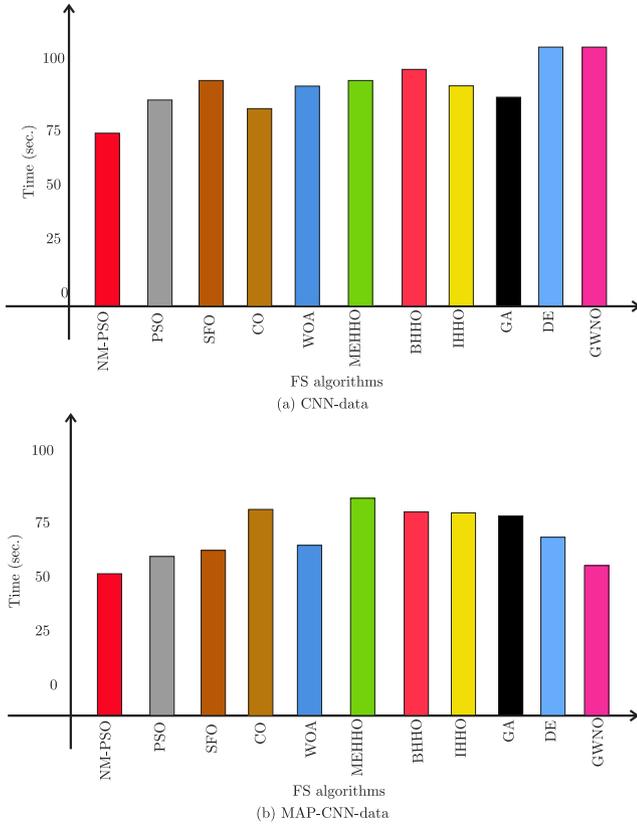


Fig. 6. The time taken by the 11 FS algorithms to reduce the data dimensionality. Clearly, the red bars, for the NM-PSO reduced data, are lower than those of the competitors confirming its applicability. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the visual interpretation presented in Fig. 7. The standard deviation is calculated using the formula:

$$\sigma = \sqrt{\frac{1}{T-1} \sum_{i=1}^T (\mathcal{M}_{\mathbb{A}_i} - \mu)^2},$$

where $\mathcal{M}_{\mathbb{A}_i}$ represents the fitness value of the feature subset \mathbb{A}_i selected in iteration i , and μ denotes the average of these T values, computed as follows:

$$\mu = \frac{1}{T} \sum_{i=1}^T \mathcal{M}_{\mathbb{A}_i}.$$

Table 3 reveals that NM-PSO consistently exhibits the lowest standard deviation compared to other algorithms in both datasets. This highlights the stability of NM-PSO's fitness values, indicating its efficient convergence.

The observations suggest that NM-PSO consistently yields dependable outcomes. Detailed scrutiny of the results indicates minimal fluctuation in the fitness function's value across iterations. This is corroborated by Fig. 7, highlighting that the most significant fluctuations in fitness value transpire during the initial iterations, subsequently approaching a quasi-stable state. This underscores NM-PSO's proficiency in rapidly identifying the targeted subgroup. The efficacy can be attributed to its advanced NMGRS architecture, which prioritizes the elimination of redundant and irrelevant features. Essentially, the harmonious equilibrium between exploration and exploitation processes ensures that NM-PSO, as previously delineated, consistently identifies the desired feature subset in every execution.

5.4. Scenario III: Impact of NM-PSO feature selector on thermal face recognition performance

To evaluate the impact of the quality of the features selected by NM-PSO in representing the original data classes (i.e., faces), a comparative analysis was conducted on the classification performance using both CNN-data and MAP-CNN-data. This analysis leveraged four renowned classifiers: random forest (RF), support vector machine (SVM), k nearest neighbors (k NN), and logistic regression (LR). The evaluation metrics encompassed precision, recall, accuracy, and F1-Score. It is crucial to note that, although the dataset comprised three decision classes, a binary classification approach was adopted. This was facilitated using the widely-recognized one vs. rest (OvR) method, alternatively termed one vs. all (OvA) [50], as recently demonstrated in [39].

Tables 4 and 5 present the classification outcomes for the subsets chosen by our proposed algorithm (NM-PSO) and its ten counterparts, utilizing four distinct classifiers. A meticulous evaluation underscores the preeminence of the subset selected by NM-PSO (highlighted in the bold column). Specifically, the subset delineated by NM-PSO consistently surpasses the subsets chosen by competing algorithms in terms of classification accuracy. Detailed scrutiny suggests that the NM-PSO-selected subset exhibits superior performance irrespective of the classifier or dataset in use. This level of consistency is not mirrored by other competitors. For instance, while IHHO emerged as the top performer with CNN-data, WOA took the lead with MAP-CNN-data. This further reinforces NM-PSO robustness and adaptability across different datasets. Moreover, it is noteworthy that NM-PSO when paired with the RF method, achieves an impeccable accuracy of 100%. Such an outcome suggests the potential for linear separability of the utilized data produced by NM-PSO.

Although we contend that the MAP-CNN-data is of superior quality compared to the traditional dataset, it is noteworthy that the classification accuracy for the RF classifier in conjunction with the NM-PSO algorithm reaches 100%. This phenomenon can be attributed to the inherent data-pruning capabilities of RF, which efficiently eliminates superfluous features, thereby facilitating a streamlined FS process. Consequently, this additional layer of selection by RF is instrumental in achieving such high accuracy levels. In contrast, competitor algorithms do not exhibit the same level of accuracy with these datasets. This discrepancy highlights the capability of NM-PSO to identify a near-optimal feature subset that can be effectively fine-tuned using RF, underscoring its robustness even when the dataset quality might not be optimal.

5.5. Scenario IV: Statistical analysis

For our study, NM-PSO was subjected to a series of experimental comparisons. Based on the results from the above three scenarios, NM-PSO exhibited superior performance over ten alternative algorithms in identifying a minimal feature subset while demonstrating rapid convergence. NM-PSO has also helped the classifiers to produce excellent accuracy for face recognition as illustrated in Scenario III. To bolster this observation, a supplementary statistical analysis was carried out on the fitness values attained post the 20th iteration. This comparative analysis was bifurcated: the first phase aimed to ascertain whether there was a statistical deviation between our method and the ensemble of ten competitor algorithms. We commenced with the null hypothesis positing that the algorithmic performances were equivalent, with the objective of discerning whether this hypothesis could be refuted. Should the null hypothesis be invalidated in this preliminary phase, it would necessitate the activation of the subsequent phase. In this segment, ten discrete comparisons were conducted, each juxtaposing our algorithm against an individual competitor from the group.

To determine which hypotheses were negated, we scrutinized the null hypothesis. This hypothesis posited equivalence between our method and each individual competitor's approach, enabling us to

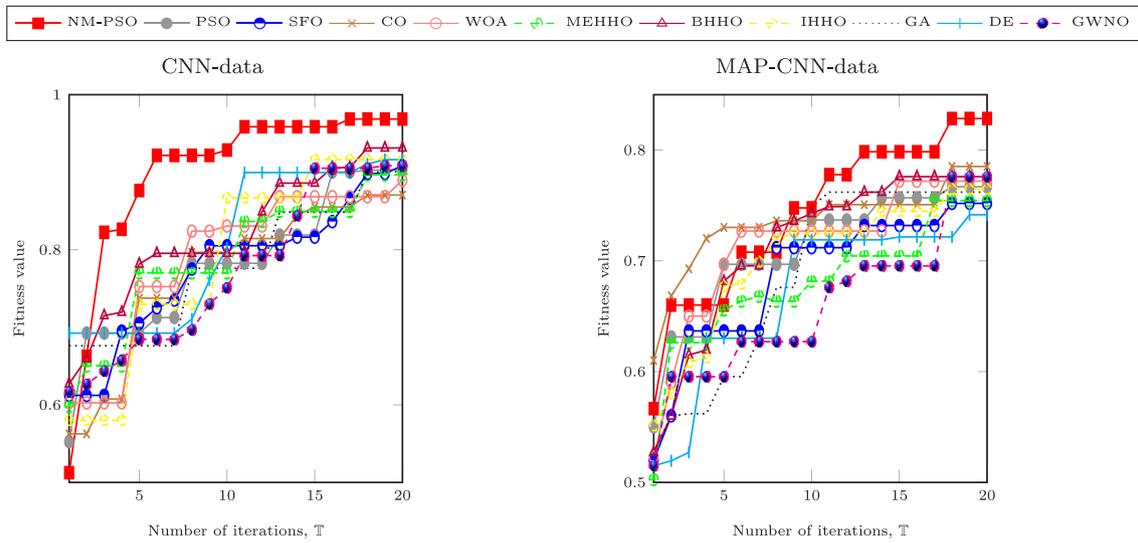


Fig. 7. Fitness value versus the number of iterations. NM-PSO curve (red), tops the other curves which reflects NM-PSO’s ability to find a reliable minimal feature subset. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

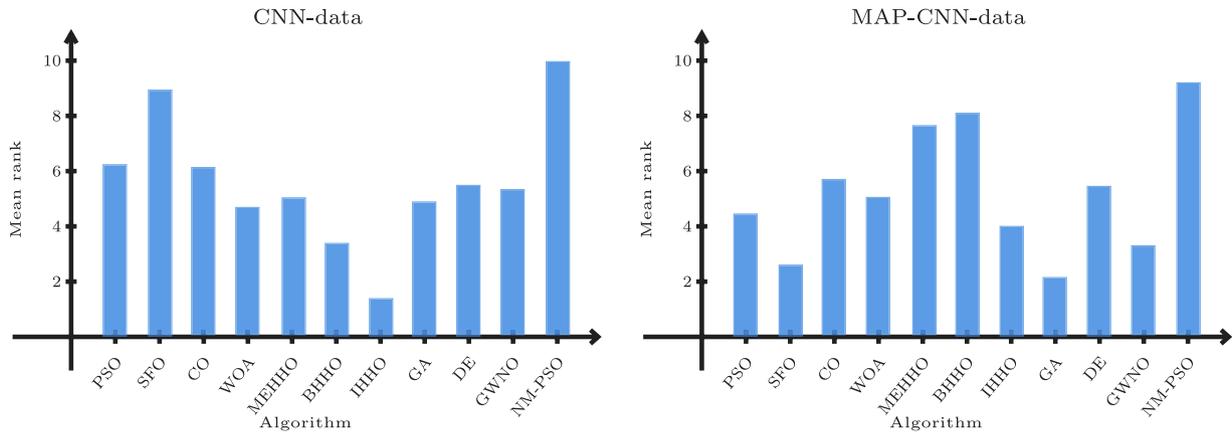


Fig. 8. The mean accuracy rank of the Friedman test for NM-PSO (our algorithm is shown in the last column) and ten competitor algorithms. Clearly, the NM-PSO bar tops the other bars, confirming its significant difference from the competitor algorithms.

Table 4

Comparative analysis of five distinct classification algorithms, gauging their performance through four key metrics over CNN data. These algorithms were applied to feature subsets selected by NM-PSO and ten other competing algorithms.

Classification algorithm	Performance metric	Original Data	SFO	GWNO	CO	MEHHO	IHHO	WOA	GA	BHHO	PSO	DE	NM-PSO
RF	Accuracy	100.0	91.33	82.11	78.99	81.24	82.11	91.55	79.11	84.32	73.12	77.44	100.0
	Precision	100.0	85.24	89.12	85.84	84.56	85.14	86.24	82.22	82.41	26.14	75.55	100.0
	Recall	100.0	83.58	84.25	84.12	79.96	80.25	82.16	86.36	75.52	65.55	90.63	100.0
	F1-Score	100.0	76.71	84.62	82.14	83.34	77.63	82.27	84.63	81.78	76.87	90.66	100.0
SVM	Accuracy	99.60	91.46	78.64	79.17	85.35	79.66	81.89	82.37	76.91	61.97	89.34	99.54
	Precision	99.60	89.52	76.14	78.85	89.14	76.37	74.61	76.96	89.65	74.50	75.39	99.54
	Recall	99.60	86.64	79.15	78.90	77.17	82.89	88.63	90.35	77.23	29.78	87.20	99.54
	F1-Score	99.60	83.19	76.16	86.33	85.84	86.99	88.54	88.58	87.21	20.91	78.55	99.54
KNN (K = 5)	Accuracy	99.16	83.83	76.94	80.24	79.52	82.61	81.22	74.61	74.24	79.11	80.63	99.13
	Precision	99.16	73.64	85.41	84.51	79.66	79.23	81.14	89.23	90.15	19.67	81.94	99.13
	Recall	99.16	80.27	85.15	76.66	83.37	74.57	78.15	85.25	88.11	48.88	76.60	99.13
	F1-Score	99.16	78.22	89.66	83.84	88.98	83.89	78.14	80.11	78.20	48.52	86.60	99.13
LR	Accuracy	98.63	82.87	89.38	78.67	78.52	76.94	76.14	80.25	81.30	54.14	80.31	98.61
	Precision	98.64	76.41	84.97	88.11	87.44	78.14	81.17	76.25	76.25	35.17	84.22	98.61
	Recall	98.63	80.42	79.58	76.41	74.25	79.52	74.71	87.24	74.6	27.88	79.30	98.61
	F1-Score	98.63	85.16	73.47	85.52	80.88	80.11	90.77	82.27	90.47	90.74	86.14	98.61

Table 5

Comparative analysis of five distinct classification algorithms, gauging their performance through four key metrics over MAP-CNN data. These algorithms were applied to feature subsets selected by NM-PSO and ten other competing algorithms.

Classification algorithm	Performance metric	Original improved CNN-Data	PSO	SFO	CO	WOA	MEHHO	BHHO	IHHO	GA	DE	GWNO	NM-PSO
RF	Accuracy	100.0	80.87	79.23	83.36	82.36	79.32	79.36	91.36	82.35	85.97	78.01	100.0
	Precision	100.0	65.87	86.25	80.39	89.66	78.25	76.96	74.98	78.25	79.45	86.25	100.0
	Recall	100.0	85.22	87.54	86.87	77.55	86.14	87.35	73.97	82.87	77.23	84.63	100.0
	F1-Score	100.0	93.15	83.77	84.47	88.41	81.42	80.14	80.58	90.46	86.69	76.84	100.0
SVM	Accuracy	99.97	62.69	90.89	90.75	74.78	81.36	89.75	81.27	83.52	80.84	83.75	99.97
	Precision	99.97	21.11	77.67	78.12	78.96	81.95	88.22	82.25	90.29	90.12	80.66	99.97
	Recall	99.97	60.97	76.54	79.12	76.65	76.84	75.25	73.85	86.66	83.15	87.35	99.97
	F1-Score	99.97	42.24	73.12	81.35	73.23	91.52	90.58	78.68	86.37	91.63	89.64	99.97
KNN (K = 5)	Accuracy	99.41	67.13	80.15	75.54	86.37	76.57	81.95	81.99	76.89	89.37	75.25	99.43
	Precision	99.41	47.28	90.36	77.44	84.95	74.22	88.67	73.40	74.54	84.85	89.63	99.43
	Recall	99.41	35.33	81.98	79.58	76.77	81.24	78.77	75.23	83.54	81.41	73.99	99.43
	F1-Score	99.41	4.62	78.74	87.65	73.98	86.24	78.10.2	78.51	80.52	77.52	91.74	99.43
LR	Accuracy	99.90	68.63	86.16	82.37	90.74	77.28	84.36	76.20	90.27	87.37	73.14	99.85
	Precision	99.90	65.03	86.52	91.47	90.56	77.36	91.55	82.01	89.89	78.88	88.25	99.85
	Recall	99.90	37.87	91.22	91.58	87.68	76.24	77.68	87.54	79.22	85.74	78.88	99.85
	F1-Score	99.90	1.53	79.37	82.62	87.32	83.25	88.99	87.00	89.28	86.99	85.85	99.85

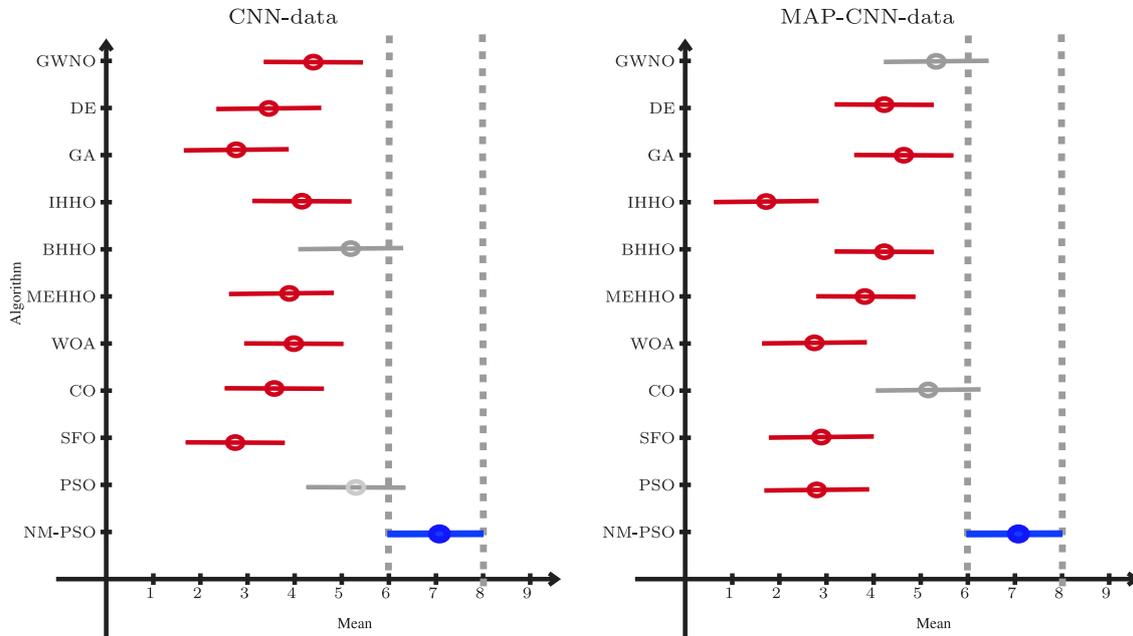


Fig. 9. Bonferroni post-hoc test for the mean classification accuracy value of NM-PSO (bold blue line) and ten FS competitor algorithms for CNN-data and MAP-CNN-data. Clearly, there is little overlap with NM-PSO. When the bars do not overlap with the blue line, this means that there is a significant difference from NM-PSO, and the null hypothesis is rejected. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

identify instances where the hypothesis was rejected. Given that the fitness value data, as depicted in Fig. 7, lacks a predefined distribution, our analysis leveraged the established non-parametric Friedman test [51]. For enhanced visualization, this dataset was bifurcated into two categories: CNN-data and its improved variant.

For statistical analyses presented within this paper, Matlab was employed as the tool of choice. It facilitated the computation of a p -value for the aggregate dataset, which subsequently became an input for a multi-comparison function executing a post-hoc test. In our scenario, the Bonferroni test served this purpose. It is paramount to note that when the derived p -value is below a specified threshold α , the standing null hypothesis is repudiated. Consistently, the threshold $\alpha = 0.05$ was adopted in our analyses. Upon executing the Friedman test on the comprehensive fitness value dataset on two separate occasions, the resultant p -values were discerned as $1.158e - 9$ and $2.125e - 10$. Notably, in both instances, the p -value was inferior to the α threshold, leading to the rejection of the null hypothesis. Furthermore, Fig. 8

illustrates the mean accuracy of the algorithms using the Friedman test. To reinforce the validity of these outcomes, the post-hoc test was deployed. The ensuing results from the Bonferroni post-hoc test are illustrated in Fig. 9. A meticulous examination underscores minimal intersections with NM-PSO in either case, fortifying the premise of NM-PSO's statistical distinctiveness from the ten competing algorithms. This offers a rigorous mathematical affirmation of NM-PSO's supremacy and its distinctiveness from the other contenders.

5.6. Scenario V: Robustness analysis of the proposed model

Robustness in face recognition refers to the system's ability to accurately identify individuals across a range of varying conditions such as different lighting, angles, facial expressions, occlusions, and across demographic changes. The Charlotte-ThermalFace dataset is ideally suited for evaluating robustness in our thermal face recognition systems

used in law enforcement. It encompasses a diverse array of high-resolution thermal infrared images, captured under varying conditions such as different lighting, expressions, and distances, as well as indoor and outdoor settings. This variety ensures the dataset effectively tests the system's performance in real-world scenarios. Further, we designed two other scenarios (ambiguity and missing data) to demonstrate the robustness of the proposed methods:

1. Case I (Increased Ambiguity Level): Thermal cameras, commonly used for security purposes, may produce ambiguous images under certain conditions, including but not limited to:
 - (a) High Ambient Temperatures [52]: In environments with elevated ambient temperatures, such as in industrial settings or outdoors on sunny days, the background heat can obscure or blur thermal signatures of human beings, complicating clear identification.
 - (b) Reflections [53]: Thermal cameras can capture infrared radiation reflected off shiny surfaces like metal, glass, or water. This reflection can distort thermal images, showing altered temperatures that can mislead interpretations about the sources and nature of heat signatures.
 - (c) Distance and Focal Adjustments [54]: The clarity of thermal images diminishes with distance. Humans located far from the camera may appear blurred or indistinct. Moreover, incorrect focal adjustments can lead to images lacking sharpness and detail, which complicates the interpretation of thermal data.

The challenge of these three scenarios lies in ensuring that the NM-PSO algorithm continues to function effectively even when ambiguity reaches 50%.

2. Case II (Increased Missing Values, *): Errors in thermal camera measurements can result in missing facial feature values [55]. The algorithm must leverage existing data to compensate for these missing values. The objective is to maintain algorithmic performance even when the proportion of missing values surpasses the existing values, approaching a 50% threshold of incompleteness.

In this experiment, attention is exclusively focused on data refined by the MAP-CNN due to its superior performance in prior tests. In scenarios involving either an increased level of ambiguity or a higher percentage of missing data, it is expected that the performance of the NM-PSO algorithm will naturally diminish. However, the essential aim of this investigation is not merely to confirm that performance deteriorates under these conditions, but rather to assess the nature of this degradation.

The question now, therefore, revolves around how gracefully the NM-PSO algorithm manages this decline. The findings from this series of tests reveal that the NM-PSO's degradation in performance, in terms of both feature subset size and accuracy, occurs in a controlled and gradual manner. This graceful degradation highlights the algorithm's robustness, demonstrating its ability to maintain a reliable level of performance even as the complexity or ambiguity of the dataset increases.

In the first case, the experiment quantitatively assesses the performance of the NM-PSO algorithm relative to its competitors (see in Table 6) under varying levels of feature ambiguity. This is systematically tested by artificially modifying the feature values on an object-by-object basis before executing the FS, with increments ranging from 10% to 50% in steps of 10%. Consequently, five distinct dataset versions are created, and both NM-PSO and its competitors are evaluated independently on each dataset. As the feature values are altered, the dependency of the labels on the conditional features shifts, leading to changes in the optimal feature subset. The variations in feature subset size are detailed in Table 6, and the corresponding changes in classification performance across four classifiers are depicted in Fig. 10.

Table 6

Comparative analysis of the feature subset sizes selected by the NM-PSO algorithm and its ten competing counterparts, with respect to varying ambiguity percentage.

FS algorithm	10%	20%	30%	40%	50%
SFO	3620	3655	3689	3724	3758
GWNO	3323	3355	3387	3418	3450
CO	3800	3836	3872	3909	3945
MEHHO	3654	3689	3724	3758	3793
BHHO	3665	3699	3734	3769	3804
GA	3581	3615	3649	3683	3717
IHHO	3409	3442	3474	3507	3539
WOA	3745	3781	3817	3852	3888
PSO	3710	3745	3780	3816	3851
DE	3429	3462	3495	3527	3560
NM-PSO	2134	2154	2174	2195	2215

Table 7

Feature subset size selected by NM-PSO with varying missing value percentage over MAP-CNN-data.

Missing value percentage	Feature subset size
10	2241
20	2375
30	2541
40	2744
50	2991

A detailed examination of Table 6 reveals that with increasing levels of ambiguity, the size of the feature subset tends to expand for all algorithms, including NM-PSO, due to the diminished discriminative power of the features. Notably, the feature subsets determined by NM-PSO consistently remain smaller compared to its competitors. Observing the classification performance depicted in Fig. 10, the NM-PSO curve outperforms those of other algorithms. Although performance degradation is a common trend among all algorithms as the ambiguity level rises, the decline in performance is notably milder for NM-PSO, underscoring its robustness against high degrees of ambiguity. This resilience can be attributed to the sophisticated design of its fitness function, which effectively guides the feature reduction process within PSO. Specifically, NM-PSO's ability to analyze data from a multigranular perspective, treating categorical and numerical features distinctly, plays a crucial role. Hence, even with up to 50% alterations in facial feature values, NM-PSO maintains a commendable accuracy in human identification.

In the subsequent experiment concerning Case II, we introduce missing values into the dataset by artificially inserting null entries (denoted by '*') for various objects. It is important to emphasize that NM-PSO is readily adaptable to accommodate these missing values. This adaptation involves a modification of the categorical neighbor function, Eq. (3a), as delineated in Eq. (6). This modification allows NM-PSO to effectively handle incomplete datasets, ensuring that the presence of missing data does not impede the algorithm's performance.

$$\theta_{\mathcal{E}, u_i} = \{u_j \mid \forall a_k \in \mathcal{C}, v_{u_i, a_k} = v_{u_j, a_k} \vee v_{u_i, a_k} = * \wedge v_{u_j, a_k} \neq *\}, \quad (6)$$

Furthermore, for numerical features, the approach to calculating the difference between two feature values is adapted to accommodate missing data, as delineated below:

1. If both values are present, the difference is computed through traditional subtraction.
2. If one value is missing, the difference is treated as the value of the existing feature. This is effectively equivalent to imputing a zero for the missing value.

On the contrary, competitor algorithms employ traditional classifiers as their fitness functions, which inherently lack the capability to directly address missing values without resorting to preliminary imputation strategies. This limitation necessitates additional preprocessing

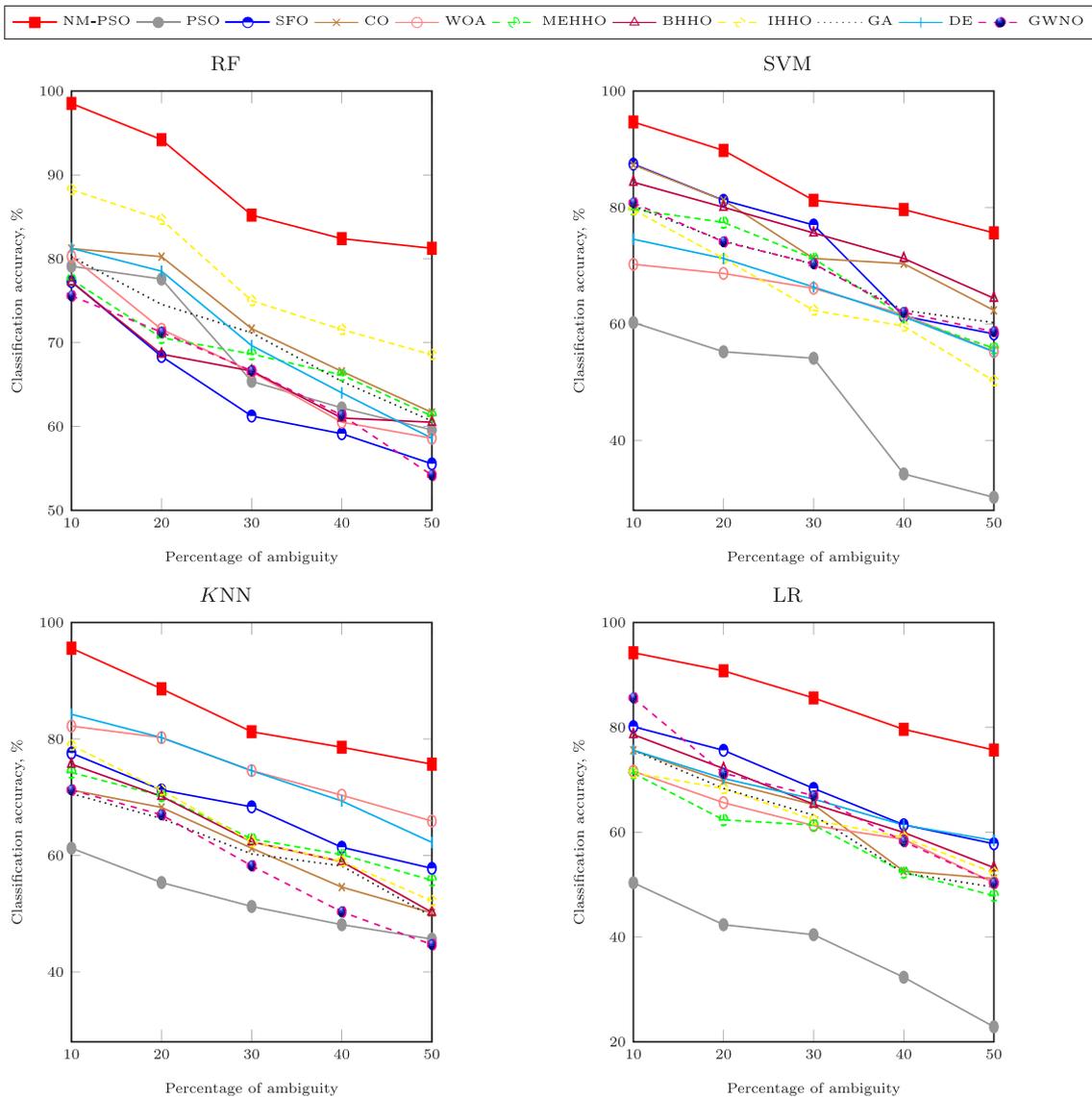


Fig. 10. Classification accuracy of trained by the FS selected by NM-PSO and its competitors versus the ambiguity percentage.

steps, effectively causing these algorithms to revert to methodologies akin to those applied in Case I, thus rendering their application somewhat redundant for our purposes. Therefore, for Case II, we exclusively assess the performance of NM-PSO. Mirroring the experimental approach of Case I, we introduced missing values artificially to the dataset, denoted by “?”. This manipulation was conducted on an object basis, progressively increasing the proportion of missing values from 10% to 50%, in increments of 10%. Consequently, we generated five distinct datasets, each subjected to an independent execution of NM-PSO. The resultant sizes of the feature subsets are documented in Table 7. Additionally, the impacts of these variations on classification performance, assessed through four different classifiers, are illustrated in Fig. 11, providing a comprehensive view of the robustness and adaptability of NM-PSO in handling incomplete data, thus making our thermal face recognition model robust.

Please note that for the classification task, we re-imputed the data following the FS task, which is justified as the imputation is now performed post-FS. Upon closer examination of the results, it is evident that the degradation rate remains relatively low until the missing values constitute 40% of the dataset. Beyond this point, the degradation rate increases marginally until it reaches 50%. This resilience can be attributed to the use of RST notions by the NMGRS algorithm,

which addresses missing values without resorting to imputation or elimination, unlike other common algorithms.

5.7. Comparison with related work

End-to-end CNN-based image classification methods have demonstrated significant success in various image recognition tasks, including thermal face recognition. However, these methods often encounter challenges when addressing complex scenarios involving illumination, expression, and surgical invariance. Our proposed MAP-CNN model, integrated with the NM-PSO algorithm, offers several critical improvements in these areas.

Studies by Ramaiah et al. [15], Wu et al. [16], Fan et al. [17], Sayed et al. [18], Kakarwal et al. [20], Aji et al. [23], and Ashfaq and Akram [3] have achieved commendable recognition rates under varying lighting conditions, focusing primarily on illumination invariance. Despite this, these approaches often fail to maintain accuracy when combined with factors such as expression and surgical changes. The maximum accuracy achieved by these methods was 95% by Aji et al. [23], with the minimal feature subset consisting of 3541 features, as demonstrated by Ashfaq and Akram [3]. These results highlight the

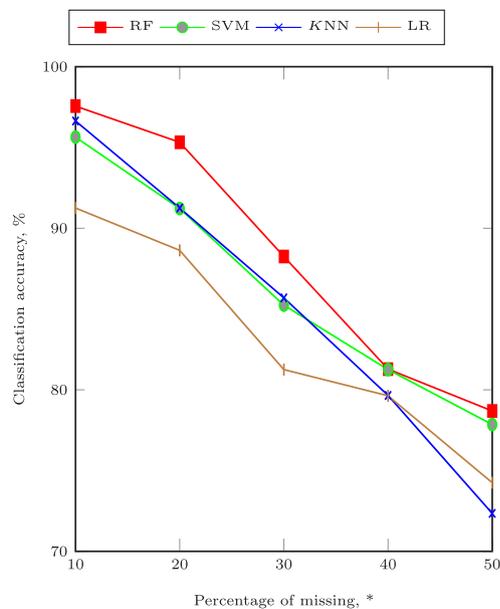


Fig. 11. Classification accuracy of trained by the FS selected by NM-PSO versus the missing value percentage.

superiority of our model (achieving 99%) over traditional CNN-based algorithms in handling illumination invariance.

Similarly, Manssour et al. [19] and Bai et al. [26] achieved 98% and 90% accuracy, respectively, by considering both illumination and expression invariance, but not surgical invariance. Additionally, our model demonstrates superiority over the work of Muller et al. [22], who focused on health-related face analysis post-surgical changes, achieving 93% accuracy where our proposed face recognition achieved 99%. The integration of NM-PSO with MAP-CNN results in a significant reduction in feature dimensions (by 57%), enhancing computational efficiency while maintaining high classification accuracy. This integration allows our model to effectively address illumination, expression, and surgical invariances, as well as handle heterogeneous data. This is a notable improvement over traditional end-to-end CNN methods, which typically require extensive computational resources and time for training and inference. Furthermore, unlike conventional CNN-based models that cannot directly handle missing values, our model is designed to manage such scenarios, providing a crucial advantage.

Therefore, the MAP-CNN model combined with NM-PSO not only surpasses existing methods in terms of accuracy and feature efficiency but also offers enhanced robustness and adaptability to complex real-world scenarios, making it a superior choice for thermal face recognition tasks.

6. Conclusions

Thermal imaging technology has been deployed in multiple domains for identification and authentication, utilizing one or at most two dimensions, namely illumination invariance and/or expression invariance. However, this is the first known research that has triangulated illumination, expression, and surgical invariance, thus providing a robust recognition method for law enforcement agencies. This was achieved by extracting robust facial features using the new MAP-CNN architectures. Then, a novel FS algorithm, NM-PSO, was introduced to select the most discriminative features reducing the computational cost but keep the accuracy high. This innovative approach amalgamated the robustness of PSO with the principles of neighborhood multi-granulation rough set (NMGRS). An exhaustive comparative analysis of NM-PSO against ten contemporaneous algorithms elucidated its

preeminence, both in terms of reduced feature subset dimensions and rapid convergence thus making the research methodically novel as well. Moreover, when benchmarked with four renowned classifiers, the feature subset delineated by NM-PSO consistently outperformed subsets derived from alternative algorithms in classification efficacy. A subsequent statistical evaluation, deploying the Friedman/Bonferroni test, corroborated the distinct advantage of NM-PSO over its contemporaries. Robustness testing of the face recognition model against image ambiguity and missing data further demonstrated its consistent performance. The model was able to provide reliable results with up to 50% missing data and 50% image ambiguity, showcasing its resilience and effectiveness even under challenging conditions. This highlights the model's practical applicability in real-world scenarios where data imperfections are inevitable. As the NM-PSO proved its applicability, we plan to develop a distributed version of it in the near future either over MapReduce or the YARN framework. This makes it possible to extend the algorithm benefit to multi-disciplines.

CRedit authorship contribution statement

Tarek Gaber: Writing – review & editing, Writing – original draft, Supervision, Software, Resources, Project administration, Methodology, Formal analysis, Conceptualization. **Mathew Nicho:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Conceptualization. **Esraa Ahmed:** Writing – review & editing, Writing – original draft, Software, Methodology, Conceptualization. **Ahmed Hamed:** Writing – review & editing, Writing – original draft, Software, Methodology, Conceptualization.

Declaration of competing interest

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

Data availability

The data is publicly available.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT in order to improve the language and readability of the paper. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References

- [1] Sharma S, Saini A, Chaudhury S. A survey on biometric cryptosystems and their applications. *Comput Secur* 2023;103458.
- [2] Awad AA, Ali AF, Gaber T. An improved long short term memory network for intrusion detection. *PLoS One* 2023;18(8):e0284795.
- [3] Ashfaq Q, Akram MU. Convolutional neural network based thermal image classification. In: 2022 2nd international conference on digital futures and transformative technologies (iCoDT2). IEEE; 2022, p. 1–6.
- [4] del Rio JS, Moctezuma D, Conde C, de Diego IM, Cabello E. Automated border control e-gates and facial recognition systems. *Comput Secur* 2016;62:49–72.
- [5] Singh M, Baranwal N, Singh K, Singh A, Zhou H. Deep learning-based biometric image feature extraction for securing medical images through data hiding and joint encryption-compression. *J Inf Secur Appl* 2023;79:103628.
- [6] Grudzien A, Kowalski M, Palka N. Face re-identification in thermal infrared spectrum based on ThermalFaceNet neural network. In: 2018 22nd international microwave and radar conference. MICON, IEEE; 2018, p. 179–80.
- [7] Budiarsa R, Wardoyo R, Mudholifah A. Face recognition for occluded face with mask region convolutional neural network and fully convolutional network: a literature review. *Int J Electr Comput Eng* (2088-8708) 2023;13(5).
- [8] Ibrahim A, Tharwat A, Gaber T, Hassanien AE. Optimized superpixel and AdaBoost classifier for human thermal face recognition. *Signal Image Video Process* 2018;12:711–9.

- [9] Rusia MK, Singh DK. A comprehensive survey on techniques to handle face identity threats: challenges and opportunities. *Multimedia Tools Appl* 2023;82(2):1669–748.
- [10] Hamed A, Mohamed MF. A feature selection framework for anxiety disorder analysis using a novel multiview harris hawk optimization algorithm. *Artif Intell Med* 2023;102605.
- [11] Salem OA, Liu F, Chen Y-PP, Hamed A, Chen X. Effective fuzzy joint mutual information feature selection based on uncertainty region for classification problem. *Knowl-Based Syst* 2022;257:109885.
- [12] Pavlidis I, Symosek P. The imaging issue in an automatic face/disguise detection system. In: *Proceedings IEEE workshop on computer vision beyond the visible spectrum: methods and applications* (cat. no. PR00640). IEEE; 2000, p. 15–24.
- [13] Liu X, Peng Y, Lu Z, Li W, Yu J, Ge D, Xiang W. Feature-fusion segmentation network for landslide detection using high-resolution remote sensing images and digital elevation model data. *IEEE Trans Geosci Remote Sens* 2023;61:1–14.
- [14] Hassan RJ, Abdulazeez AM. Deep learning convolutional neural network for face recognition: A review. *Int J Sci Bus* 2021;5(2):114–27.
- [15] Ramaiah NP, Ijjina EP, Mohan CK. Illumination invariant face recognition using convolutional neural networks. In: *2015 IEEE international conference on signal processing, informatics, communication and energy systems. SPICES, IEEE*; 2015, p. 1–4.
- [16] Wu Z, Peng M, Chen T. Thermal face recognition using convolutional neural network. In: *2016 international conference on optoelectronics and image processing. ICVIP, IEEE*; 2016, p. 6–9.
- [17] Fan Y, Zhai G, Wang J, Hu M, Liu J. Multiple thermal face detection in unconstrained environments using fully convolutional networks. In: *Pacific rim conference on multimedia*. Springer; 2017, p. 24–33.
- [18] Sayed M, Baker F. Thermal face authentication with convolutional neural network. *J Comput Sci* 2018;14(12):1627–37.
- [19] Manssor SA, Sun S. TIRFaceNet: thermal IR facial recognition. In: *2019 12th international congress on image and signal processing, bioMedical engineering and informatics (CISP-BMED)*. IEEE; 2019, p. 1–7.
- [20] Kakarwal S, Chaudhari K, Deshmukh R, Patil R. Thermal face recognition using artificial neural network. In: *2020 international conference on smart innovations in design, environment, management, planning and computing. ICSIDEMPC, IEEE*; 2020, p. 300–4.
- [21] Mahouachi D, Akhloufi MA. Adaptive deep convolutional neural network for thermal face recognition. In: *Thermosense: thermal infrared applications XLIII*. Vol. 11743, SPIE; 2021, p. 15–22.
- [22] Müller D, Ehlen A, Valeske B. Convolutional neural networks for semantic segmentation as a tool for multiclass face analysis in thermal infrared. *J Nondestruct Eval* 2021;40(1):9.
- [23] Aji MES, Alfanz R, Prakasa E. Infrared image analysis for human face recognition. In: *2022 5th international conference on computing and informatics. ICICI, IEEE*; 2022, p. 157–62.
- [24] Taspinar YS. Light weight convolutional neural network and low-dimensional images transformation approach for classification of thermal images. *Case Stud Therm Eng* 2023;41:102670.
- [25] Tsai T-H, Lu J-X, Chou X-Y, Wang C-Y. Joint masked face recognition and temperature measurement system using convolutional neural networks. *Sensors* 2023;23(6):2901.
- [26] Bai Y, Liu L, Liu K, Yu S, Shen Y, Sun D. Non-intrusive personal thermal comfort modeling: A machine learning approach using infrared face recognition. *Build Environ* 2024;247:111033.
- [27] Mohamed EA, Rashed EA, Gaber T, Karam O. Deep learning model for fully automated breast cancer detection system from thermograms. *PLoS One* 2022;17(1):e0262349.
- [28] A. Mohamed E, Gaber T, Karam O, Rashed EA. A Novel CNN pooling layer for breast cancer segmentation and classification from thermograms. *PLoS One* 2022;17(10):e0276523.
- [29] Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Adv Neural Inf Process Syst* 2012;25.
- [30] Nair V, Hinton GE. Rectified linear units improve restricted boltzmann machines. In: *Proceedings of the 27th international conference on machine learning (ICML-10)*. 2010, p. 807–14.
- [31] Sermanet P, Eigen D, Zhang X, Mathieu M, Fergus R, LeCun Y. Overfeat: Integrated recognition, localization and detection using convolutional networks. 2013, arXiv preprint arXiv:1312.6229.
- [32] LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;521(7553):436–44.
- [33] Houssein EH, Emam MM, Ali AA, Suganthan PN. Deep and machine learning techniques for medical imaging-based breast cancer: A comprehensive review. *Expert Syst Appl* 2021;167:114161.
- [34] Zhou B, Khosla A, Lapedriza A, Oliva A, Torralba A. Learning deep features for discriminative localization. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, p. 2921–9.
- [35] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, p. 770–8.
- [36] Shin H-C, Roth HR, Gao M, Lu L, Xu Z, Nogues I, Yao J, Mollura D, Summers RM. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Trans Med Imaging* 2016;35(5):1285–98.
- [37] Yosinski J, Clune J, Bengio Y, Lipson H. How transferable are features in deep neural networks? *Adv Neural Inf Process Syst* 2014;27.
- [38] Ashrafi R, Azarbayjani M, Tabkhi H. Charlotte-ThermalFace: A fully annotated thermal infrared face dataset with various environmental conditions and distances. *Infrared Phys Technol* 2022;124.
- [39] Hamed A, Tahoun M, Nassar H. KNNHI: Resilient K NN algorithm for heterogeneous incomplete data classification and K identification using rough set theory. *J Inf Sci* 2022;1–25.
- [40] Al-Tashi Q, Kadir SJA, Rais HM, Mirjalili S, Alhussain H. Binary optimization using hybrid grey wolf optimization for feature selection. *Ieee Access* 2019;7:39496–508.
- [41] Gomes GF, da Cunha SS, Ancelotti AC. A sunflower optimization (SFO) algorithm applied to damage identification on laminated composite plates. *Eng Comput* 2019;35(2):619–26.
- [42] Abdel-Basset M, El-Shahat D, El-henawy I, de Albuquerque VHC, Mirjalili S. A new fusion of grey wolf optimizer algorithm with a two-phase mutation for feature selection. *Expert Syst Appl* 2020;139:112824.
- [43] Lu Z, Wang C, Guo J. A hybrid of fish swarm algorithm and shuffled frog leaping algorithm for attribute reduction. In: *2018 13th world congress on intelligent control and automation. WCICA, IEEE*; 2018, p. 1482–7.
- [44] Too J, Liang G, Chen H. Memory-based Harris hawk optimization with learning agents: a feature selection approach. *Eng Comput* 2021;1–22.
- [45] Thaher T, Heidari AA, Mafarja M, Dong JS, Mirjalili S. Binary Harris Hawks optimizer for high-dimensional, low sample size feature selection. In: *Evolutionary machine learning techniques*. Springer; 2020, p. 251–72.
- [46] Zhang Y, Liu R, Wang X, Chen H, Li C. Boosted binary Harris hawks optimizer and feature selection. *Eng Comput* 2021;37(4):3741–70.
- [47] Katoch S, Chauhan SS, Kumar V. A review on genetic algorithm: past, present, and future. *Multimedia Tools Appl* 2021;80:8091–126.
- [48] Price K, Storn RM, Lampinen JA. *Differential evolution: a practical approach to global optimization*. Springer Science & Business Media; 2006.
- [49] Hamed A, Nassar H. Efficient feature selection for inconsistent heterogeneous information systems based on a grey wolf optimizer and rough set theory. *Soft Comput* 2021;25(24):15115–30.
- [50] Gao X, He Y, Zhang M, Diao X, Jing X, Ren B, Ji W. A multiclass classification using one-versus-all approach with the differential partition sampling ensemble. *Eng Appl Artif Intell* 2021;97:104034.
- [51] Hamed A, Tahoun M, Nassar H. K n n h i: Resilient k n n algorithm for heterogeneous incomplete data classification and k identification using rough set theory. *J Inf Sci* 2023;49(6):1631–55.
- [52] Molineux J, Lee T, Kim KJ, Kang K-S, Lyons NP, Nishant A, Kleine TS, Durfee SW, Pyun J, Norwood RA. Fabrication of plastic optics from chalcogenide hybrid inorganic/organic polymers for infrared thermal imaging. *Adv Opt Mater* 2024;12(7):2301971.
- [53] Xue S, Huang G, Chen Q, Wang X, Fan J, Shou D. Personal thermal management by radiative cooling and heating. *Nano-Micro Lett* 2024;16(1):153.
- [54] Zhong S, Zhao X, Liu D, Su H, Xie Z, Fan B. High-resolution, lightweight remote sensing via harmonic diffractive optical imaging systems and deep denoiser prior image restoration. *IEEE Trans Geosci Remote Sens* 2024.
- [55] Morresi N, Cipollone V, Casaccia S, Revel GM. Measuring thermal comfort using wearable technology in transient conditions during office activities. *Measurement* 2024;224:113897.