
Sparse Neural Network for horse face detection in a Smart Riding Club Biometric System

Islem Jarraya¹ · Wael Ouarda¹ · Fatma

BenSaid¹ · Adel M. Alimi¹

Received: date / Accepted: date

Abstract Horses and breeders need to be safe on the farm and the riding club. On account of the great value of the horse, the breeder needs to protect it from theft and disease. In this context, it is important to detect and to recognize the identity of each horse for security reasons. In fact, this paper proposes a Smart Riding Club Biometric System (SRCBS) consisting in automatically detecting and recognizing horses as well as humans. The proposed system is based on the facial biometrics for a horse and the gait biometrics for a human due to their simplicity and intuitiveness in an uncontrolled environment. The present work focuses mainly on horse face detection and recognition. Animal face detection is still extremely difficult given the fact that face textures and shapes are grossly diverse. In addition, recent detectors require a huge dataset for training and represent a huge number of parameters and layers, leading to so much training time. For re-

Islem Jarraya
E-mail: islem.jarraya@enis.tn

Wael Ouarda
E-mail: wael.ouarda@enis.tn

Fatma BenSaid
E-mail: fatma.bensaid.tn@ieee.org

Adel M. Alimi
E-mail: adel.alimi@enis.tn

¹REGIM-Lab.: REsearch Groups in Intelligent Machines, University of Sfax, National Engineering School of Sfax (ENIS), BP 1173, Sfax, 3038, Tunisia

solving these problems and also for a useful detection system, this paper proposes a Sparse Neural Network (SNN) based on sparse features for horse face detection. Different global and local features were performed to identify horses and humans for the recognition process. Due to the unavailability of horse databases, this paper presents a new benchmark for horse detection and recognition in order to evaluate our proposed system. This system achieved an average precision equal to 90% for horse face detection and a recognition rate equal to 99.89% for horse face identification.

Keywords Horse Face Detection · Identification · Neural Network · THDD · THoDBRL

1 Introduction

Object detection and identity recognition have been widely employed in recent years for several security reasons. Animals, just as humans, need to be safe. Farm animals control is required for verification of the source, the identity and the production process as well as for the livestock health surveillance. In addition, the outbreak of diseases such as bovine spongiform encephalopathy (BSE), foot-and-mouth disease (FMD) and classical swine fever (CSF), and the importance of export markets for domestic producers, requires the implementation of animal detection and identification programs that allow farmers to trace cattle from birth. Hence, the regulatory provisions (Orientation Law 1999) [13] made the equine identification compulsory to ensure the origin of each horse and certify their identities for the purpose of fighting against theft and fraudulent horses as well as making sure of healthy monitoring.

In this context, assuring the security of animals and humans seems extremely interesting in a riding club. Owing to its great value, it is very necessary to help essentially the horse as well as its owner so as to be controlled without obstacles or difficulties in the barn, the box or the race track.

Well-known traditional methods of animal identification use plastic ear tags, tattoos, freeze branding, hot-iron branding or RFID electronic ear tags. However, these techniques may threaten the well-being of animals [6] and require effort to identify the animal. Because of these problems, another detection and identification way should be used for effective control by means of a stressed robust biometric marker that is invariant, fraud-proof, fast, accurate, inexpensive and non-invasive to capture. Researchers have resorted to other identification methods using different physiological biometric modalities such as muzzle patterns [24, 23] and retina pattern [2]. Although these patterns are effective, they require an effort to get the identity of the animal. In fact, these methods necessitate direct contact between the animal and the camera to capture the muzzle or retina patterns. Thus, a biometric detection and recognition system needs to be created for the comfort of both animals and humans taking into account the subject in motion in order to be well-being and lack the direct contact between the subject and the sensor.

Gait acquisition does not require contact with the sensor, just as the face biometrics, and is less likely to be exposed to obscurity than other biometrics. In fact, gait identification is easier to use and more secure than the other biometrics. The subject can be identified in any view and at any location. In a riding club, the gait biometrics are very useful for human identification but not for horses. In the stable, the horses remain in their boxes without walking and only the face appears. Therefore, the facial biometrics are the most suitable biometric traits for horse identification. In fact, the most suitable biometric traits that can be useful for a smart riding club security are the facial biometrics for horses and the gait biometrics for humans due to their simplicity and intuitiveness in an uncontrolled environment.

Face recognition is one of the most promising modalities for horse identification [16, 17, 34] despite the scarcity of research in this field. According to the 5th edition of the FEI ¹, the description of the horse face is particularly different. In fact, the

¹ Federation Equestre Internationale (2000), 5 th edition (Identification of Horses with the narrative and the diagram), url:<https://www.fei.org/>

horse hair color is very varied and contains different texture patterns such as the blaze, the strip, the snip, the star and the lip marking on the head. The variation in direction and shape of these texture patterns is different from one horse to another.

Indeed, the use of the face pattern is more effective for the horse identification.

This paper suggests renewing the Smart Riding Club Biometric System (SRCBS) proposed in [34] using other modalities, features, methods and techniques to achieve a more efficient automatic system with an improved performance. The proposed SRCBS is very important for breeders to ensure the safety of the horse due to its great value. Renewing the horse face detection and recognition systems is the main interest in this work. Regretfully, the face detection task is still very difficult mainly because of the large intra-class variation, the illumination change, the variable pose, the complex background, the partial occlusion and the uncontrolled environment. Face textures and shapes are grossly diverse, which makes the animal face detection extremely difficult. This is probably the reason why the number of related work approaches for animals is small. Despite these difficulties, recent research has achieved significant progress to resolve the interesting detection problems. The detection rate has reached nearly 90% of the human face using boosting-based and CNN-based (Convolutional Neural Network) approaches [18, 44]. The fact that Convolutional Neural Networks (CNNs) are strong and fascinating classification tools is among the reasons why Deep Learning is immensely popular and widely used for computer vision tasks. Despite the rapid development of CNN, there are different challenges during its training. First, CNN requires a huge dataset for training [19]. Second, pooling layers eliminate a great deal of information and ignore the realationship between image parts according to Hinton [39]. Finally, CNNs represent a big number of parameters and layers, leading to much training time and high computational complexity.

According to the above-listed problems, the challenging questions at this step are:

- How to create a neural network as a backbone detector with a small number of parameters and layers in order to minimize the training time and the computational complexity.
- How to determine the powerful features and effectively detect horse faces using a small database for training.
- How to recognize horses using the simplest method for training and testing.

The challenging issue consists in establishing a safer smart riding club and proposing a new application to detect and recognize the horse face and human gait with an easy and fast way without the need for a huge data for training. The different contributions in this paper are as follows:

- Renewing the proposed SRCBS in [34] using other modalities and other methods.
- Making a Sparse Neural Network for horse face detection (SNN) using the smallest number of parameters and without the need for a huge horse data for training.
- Making a system for horse face detection based on the proposed SNN.
- Employing a sparse feature selection method for face detection. The feature selection field has been well adapted resorting to many learning methods for pattern recognition applications. However, these are not devoted to object detection applications.
- Using the proposed sparse feature selection method as a training optimizer for a sparse hidden layer of SNN instead of such traditional algorithms as ADAM and SGD.
- Making a new horse database called Tunisian Horse Detection Database (THDD), which could contribute to the research community of the animal biometrics. To the best of our knowledge, this is the only dataset of public face image that is available for research on horse detection.
- Proving that our proposed SNN could get better performance than the CNNs of the traditional detectors. In the experiments, the efficiency of SNN was

demonstrated for face classification compared to the other Convolutional Neural Networks using the smallest number of parameters and without the need for a huge dataset.

- Showing through experimental studies that the proposed recognition system for horses got a better performance than the related works.

The rest of this paper is organized as follows: Section 2 presents the related works on animal face detection and recognition. However, the proposed SRCBS and its sub-systems are described in Section 3. Section 4 is devoted to the presentation of THoDBRL’2015 and THDD databases. Section 5 focuses on the experimental study, whereas section 6 shows the limitations of the system. The conclusion of this paper which also presents some possible future work is eventually drawn in Section 7.

2 Related works

2.1 Animal face detection

The number of works in this area is very limited due to the complication of the animal face detection and recognition tasks. The existing related works in this field are as follows:

Zhang *et al.* [53] proposed a set of Haar of Oriented Gradients (HOOG) to capture the texture and shape features on the animal head (such as cats, tigers, pandas, foxes and cheetahs). They used SVM for classification and decision calculation. Using the Cat Database, they found a precision equal to 95% and a recall equivalent to 99.8%.

Yamada *et al.* [50] proposed detecting dog and cat heads using edge-based features. They selected four directional features (Horizontal, Vertical, Upper Right and Upper Left) to detect the facial characteristics. They used a multi-layer classifier for features classification. Yamada *et al.* performed their method on a set of

cat and dog images from the web. The recall rate was equal to 85% on the cat set and 90% on the dog set.

Mukai *et al.* [32] focused on the cat and dog face detection. They used the same Viola-Jones method and employed both the Haar and the HOG descriptors for feature extraction. Using 58 images from the Cat Database for the test, they found a recall equal to 96.6% and a precision equivalent to 75.7%. However, they achieved a recall equal to 98.3% and a precision equivalent to 90.8% using 60 images from the Stanford Dogs Dataset.

These traditional animal face detectors with handcrafted features, have been replaced in the recent works by deep convolutional neural networks with the ability to extract discriminating face features.

Vlachynska *et al.* [49] used the faster R-CNN proposed in [38] with ResNet-101 for dog face detection. They found an Average Precision equal to 98% on the Columbia Dogs Dataset.

Tureckova *et al.* [46] who used the YOLOv3 detector with DarkNet-53 for dog face detection noticed an Average Precision equivalent to 92% on the Columbia Dogs Dataset and the Oxford-IIIT Pet Dataset.

The proposed detectors [46, 49] have already taken on known CNN architectures (ResNet and DarkNet) as backbones. In addition, other detectors for animal detection [20] based on CNN were propounded. However, Convolutional Neural Networks (CNNs) ignore the relationship between image parts, represent a big number of parameters and layers and require a huge dataset for training. This leads to much training time and high computational complexity.

To deal with the above-listed problems, this paper introduces the Sparse Neural Network for animal face detection SNN. The proposed network effectively exploits the animal face characterization to obtain fast classification and detection using the smallest number of parameters and without the need for a huge dataset.

2.2 Animal face recognition

In the last decades, the development of facial recognition systems has been achieved using manually-noted databases in order to locate the facial area in the image. Overall, facial recognition systems have not been automated by facial detection systems [16, 17, 34, 42, 22, 40, 31]. However, these methods allow high recognition rates but their systems lack automatic face detection, which is why the animal face detection system is important to ensure safety and security.

Jarraya *et al.* [16] benefited from the horse face properties and proposed an approach for horse identity recognition using frontal facial features of 47 horses. They used the Gabor filters and the LDA for features extraction. However, the Euclidian (Euc) distance and MahCosine (MC) distance were employed for classification. They validated our previous system using the THoDBRL'2015 and they achieved a recognition rate equal to 95.74%.

Jarraya *et al.* [17] proposed a multi-view horse face recognition using the THoDBRL'2015 that contains 47 horses. They used the Gabor filters for face characterization, the Stacked Auto-encoders (SA) to reduce the size of the feature vector and SVM for classification. Using 9 images for training, they obtained a recognition rate equal to 94.22% on the frontal images.

Ouarda *et al.* [34] proposed a new feature descriptor (RNGLBP) based on the Gabor and LBP features. They tested the proposed approach on the THoDBRL'2015 using the SVM classifier. They reached a recognition rate equal to 98.77%.

Shi *et al.* [42] propounded the Residual InterSpecies Equivariant Network (RiseNet) for deep cross-species feature learning. The features of the lower and the upper halves of faces were learned separately. They merged these features as additional information to improve the performance of the proposed RiseNet. Shi *et al.* performed the suggested method on the THoDBRL'2015. They found a recognition rate equal to 82.56%.

Face recognition was performed on other animals such as cattle by [22, 40] and dogs by [31]. Kumar *et al.* [22] used the SURF and the LBP for feature extraction and the Chi-Square for classification. The performance was about 92,5% with a database of 40 cattle.

Salama *et al.* [40] used the Bayesian optimization to automatically set the parameters for the AlexNet convolutional neural network. They performed their method on a database of 52 sheep. They achieved a recognition rate of 98%.

Mougeot *et al.* [31] proposed a deep learning approach based on a deep convolutional neural network for the face recognition and verification of dogs. They found a recognition rate equal to 88% on a database of 48 dogs.

The previously-mentioned related works conducted their proposed methods on small datasets that contained on average between 40 and 50 subjects with no more than 10 images for each subject. This number is not sufficient for a deep neural network training. This explains the high recognition rate for works [16, 17, 34, 22] that used handcrafted features and the low rate for works [31, 42] which used deep neural network approaches. Hence, the proposed solution in this paper is to avoid the use of a deep neural network which represents a big number of parameters.

Owing to the importance of studying the horse, some works have been carried out recently and have given new computer solution for control and security objectives. In fact, according to North [33], the interaction between the horse and the computer is more substancial.

Hummel *et al.* [15] and Li *et al.* [27] proposed using the face pattern as it is rich in information about the life of horses such as pain, disease and feelings. The objective of Hummel *et al.* [15] was to recognize the pain in equines. They suggested employing the HOG features and SVM for pose estimation and the SIFT, LBP, HOG and VGG16 features as well as SVM for pain recognition. Using their own equine dataset, Hummel *et al.* found the F1 score to be equal to 0.89 for pose estimation and 0.53-0.87 for pain estimation.

The objective of Li *et al.* [27] was to detect EquiFACS units automatically from

the horse face images. They proposed to test the DRML and AlexNet for horse facial AU recognition. They found an accuracy between 54.0% and 58.1% using DRML and an accuracy between 52.8% and 57.0% with AlexNet on their own dataset.

Bragança *et al.* [5] propounded to improve the gait classification of horses using data generated by the Inertial Measurement Unit (IMU). They built a dataset of 120 horses which included 7.576 strides of 8 different gaits. Their gait classification model based on the LSTM network achieved 97% of accuracy.

3 The proposed Smart Riding Club Biometric System (SRCBS)

In this section, the proposed Smart Riding Club Biometric System (SRCBS) which involves three sub-systems for horse and human recognition is introduced. The objective of the first sub-system (HFD-SF system: Horse Face Detection using Sparse Features) is to detect the horse face. As for the second sub-system (HIR-FB system: Horse Identity Recognition based on Face Biometrics), its objective is to develop a contactless solution for horse recognition using facial features. The objective of the third sub-system (HIR-GB system: Human Identity Recognition based on Gait Biometrics) is to recognize the person from a specific distance using the gait modality. In fact, two Camera positions were proposed for human gait and horse face detection and recognition. The first camera was placed in the barn near the horses for the facial biometrics capture. The second one could be placed anywhere in the club for the capture of the human gait biometrics from a distance.

Fig. 1 illustrates the proposed architecture of the new SRCBS system.

3.1 HFD-SF system: Horse Face Detection based Sparse Features

The difficult question, at this stage, is how to determine the powerful features and the most useful method to effectively detect horse faces using a small number of parameters and without the need for a huge set of data.

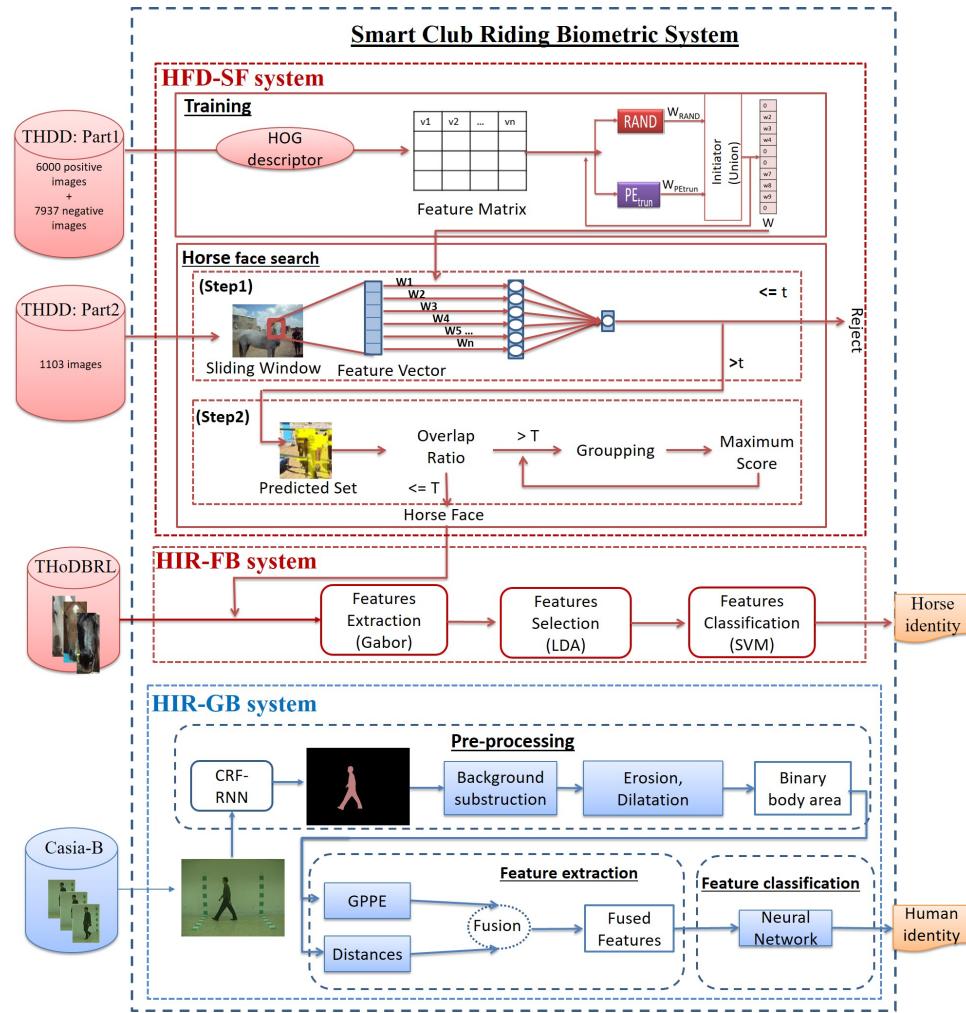


Fig. 1 The proposed Smart Riding Club Biometric System

Due to the large diversity of horse head textures, it would be a sophisticated task to develop a face detector. Although horses have distinctive ears, this characteristic cannot be focused on because the horse moves them frequently and changes their shape and position in a fascinating way. Therefore, our work concentrates on detecting the horse face without considering its ears. On the other hand, we have observed that all horses have distinctive head forms, nose and profile eyes. The horse faces have a globally similar shape, but locally variant colors and textures. For this reason, finding out how to effectively use the shape features has

been our focus in order to create a robust horse face detector. Based on this idea, a detection method that focuses on the most expressive oriented gradient features has been proposed. In fact, a Sparse Neural Network for horse face detection using sparse gradient features called (SNN) has been suggested.

3.1.1 Gradient features for horse face description

Since gradient features have performed well in several high-level computer vision tasks such as object detection [37], testing their performance on horse face detection is proposed in the present work. This descriptor describes the apparent objects and shapes by estimating the direction of the edges or the intensity distribution [7]. The description was carried out by dividing the image into small adjacent regions, called cells, and by calculating the gradient directions for each cell in the histogram. The magnitude (Mag) and direction in pixel (x, y) were calculated according to the following equations where $I(x, y)$ is the brightness value of the image in (x,y) (Eq. 1-4).

$$d_y = I(x, y - 1) - I(x, y + 1) \quad (1)$$

$$d_x = I(x - 1, y) - I(x + 1, y) \quad (2)$$

$$Mag(x, y) = \sqrt{d_y^2 + d_x^2} \quad (3)$$

$$\theta(x, y) = \tan^{-1}\left(\frac{d_y}{d_x}\right) \quad (4)$$

For each cell, an oriented gradient feature vector was constructed by quantizing θ into K orientation bins weighted by the gradient magnitude [7]. The overlapped cells were grouped and normalized in order to form a wider spatial region (block). The concatenation of the block histograms formed the gradient descriptor [7]. There were significant challenges in adapting gradient features for horse face detection and selecting the effective ones of the complicated face textures.

3.1.2 Sparse Neural Network (SNN) architecture

To deal with the above-listed problems (in the Introduction), the smallest network with a small number of parameters was suggested for horse face detection. In fact, the proposed Sparse Neural Network is composed of three layers; an input layer (gradient features), a sparse hidden layer and an output layer for feature classification as in the MLP network. The input layer represents the gradient feature vector. The sparse hidden layer had a size equal to the input layer. Each neuron in the input layer had a unique relationship with the opposite neuron in the hidden layer, which reduced the number of parameters. This layer was trained using a proposed sparse feature selection method that will be explained in the next section. In fact, this method produced a weight vector W through the input gradient features. W was used as the weight vector of the hidden layer. The weight vector contained a big number of zeros and therefore all the values corresponding to zero would be falling in the hidden layer giving rise only to the pertinent features that represented 10% of the input vector. The linear activation function was used for the hidden layer (Eq. 5) where G is the gradient feature vector, n refers to the number of neurons in the input and in the hidden layer, w represents a weight value in the hidden layer and g stands for a gradient feature value in the input layer. The output layer contained only one neural because the classification is binary. This layer was trained by the Stochastic Gradient Descent (SGD) optimizer with momentum using a sigmoid transfer function. The NN-Sparse architecture is presented in fig. 2.

$$f(WG) = \begin{pmatrix} w_1 \\ w_2 \\ \dots \\ w_n \end{pmatrix} \times \begin{pmatrix} g_1 \\ g_2 \\ \dots \\ g_n \end{pmatrix} = \begin{pmatrix} wg_1 \\ wg_2 \\ \dots \\ wg_n \end{pmatrix} \quad (5)$$

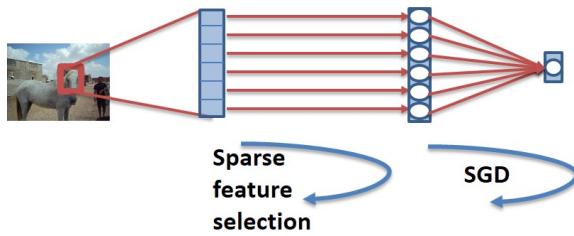


Fig. 2 SNN architecture

3.1.3 Sparse hidden layer training based on sparse feature selection method

In other research works, the feature selection field has been well adapted resorting to many learning methods for pattern recognition applications. However, these are not devoted to object detection applications. The main challenge of feature selection methods is how to reach accurate results using a small number of active features. Despite the efficiency of these methods, they are not accurate enough to process real world data using a small number of features. In this paper, a feature selection method that treats this limitation and integrates an automated negotiation process between the PE_{trun} and $RAND$ algorithms has been proposed for binary classification. The input pattern sequence was (x_t, y_t) where x is the input gradient features of d dimension, $t = 1, \dots, T$ refers to the number of iteration and y is the desired output. The sparse selection method required a classifier W_t which contained at most B non-zero elements (where $B > 0$ is a predefined constant). Thus, the classification of x_t depended only on B features and was made by the function: $sgn(W_t x_t)$. The classifier W_t would be updated in each trial t and the learner would classify the instance x_t . This scenario was repeated until $t = T$. It has been assumed that the learner was provided with full inputs of every training instance. The $RAND$ algorithm which was described and used in [3] randomly selects and picks B features in a learning task. The PE_{trun} algorithm which was described and used in [3] is a perceptron modified by a simple truncation. Both of these algorithms should respect the following 6 steps:

- Step 1: Initializing the weight vector W by zero values

- Step 2: Applying the prediction function: $f_t = W' * x_t$
- Step 3: The feedback y_t ; the real class
- Step 4: Calculating the error: if $y_t * f_t \leq 0$, $err_{count} = err_{count} + 1$
- Step 5: Updating the weight vector: $W = W + y_t * x_t$
- Step 6: Selecting relevant features

Since our aim is to select only relevant features, unnecessary ones were reduced to zero. In the first step, all input features were considered to be irrelevant. Therefore, the weight vectors of the two algorithms were initialized by zeros. The role of *RAND* and *PE_{trun}* algorithms was to keep zeros for irrelevant features and increase the weights of useful features. The first five steps were common between the two *RAND* and *PE_{trun}* algorithms. The only difference was within the feature selection manner, which was processed in the final step. The *RAND* and *PE_{trun}* algorithms participated both in the negotiation process and tried to select the best features. In fact, our key contribution was to incorporate the automated negotiation between the learning algorithms to improve the classification performance. The error rate was considered as the utility function of each negotiator. Negotiation, in this sense, involves the minimum number of mistakes to select the relevant features with the minimum execution time. The *RAND* participant sends W_{RAND} while the *PE_{trun}* sends $W_{PE_{trun}}$ to the initiator. The initiator creates the union of W_{RAND} and $W_{PE_{trun}}$ into a W . The same scenario will be repeated and the participants will take the newly vector W in each iteration (Fig. 3). Algorithm 1 describes the different steps of the initiator upon receiving the W_{RAND} and $W_{PE_{trun}}$ proposals. Algorithm 2 and 3 represent the negotiation of the sparse feature selection method.

3.2 HIR-FB system: Horse Identity Recognition based on Facial Biometrics

According to [16, 17, 34] it can be concluded that the Gabor descriptor is effective for horse face characterization. Gabor descriptor has proven its sufficiency in many recent works for face recognition [26, 4, 30]. Jarraya *et al.* [16] demonstrated that

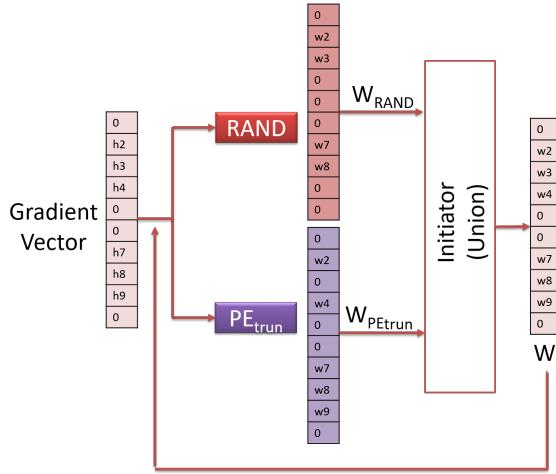


Fig. 3 The proposed feature selection method

Algorithm 1 $W = \text{initiator}(W_{\text{RAND}}, W_{\text{PE}_{\text{trun}}})$

Input

W_{RAND} : The output of the RAND algorithm

$W_{\text{PE}_{\text{trun}}}$: The output of the PE_{trun} algorithm

Initialization $W = 0$

NumFeatures: initial number of features

```

1: for i=1, 2, ..., NumFeatures do
2:   if  $W_{\text{RAND}}(i) \simeq 0$  then
3:      $W(i) = W_{\text{RAND}}(i)$ 
4:   else
5:     if  $W_{\text{PE}_{\text{trun}}}(i) \simeq 0$  then
6:        $W(i) = W_{\text{PE}_{\text{trun}}}(i)$ 

```

the LDA was better than the PCA for the selection of Gabor features. Multi-class SVM effectiveness for horse face feature classification was proven in [34]. Thus, it would be interesting to use these techniques and propose an identification system using the Gabor filters for feature extraction, the LDA for feature selection and multi-class SVM for classification.

Algorithm 2 Negotiation based on *RAND* method**Input***B*: The number of selected features**Initialization** $W_{PE_{trun}}$ $W_1 = 0$ $W_{RAND} = 0$

- 1: **for** $t=1, 2, \dots, T$ **do**
- 2: Receive x_t
- 3: Make prediction $sgn(x_t^T W_t)$
- 4: Receive y_t
- 5: **if** $y_t x_t^T W_t \leq 0$ **then**
- 6: $\tilde{W}_{t+1} = W_t + y_t x_t$
- 7: $V = Zeros(size(\tilde{W}_{t+1}, 1), 1)$
- 8: $perm_t = randperm(size(\tilde{W}_{t+1}, 1))$
- 9: $C_t = perm_t(1 : B)$
- 10: $V(C_t) = 1$
- 11: $W_{RAND} = \tilde{W}_{t+1} V$
- 12: $\tilde{W}_{t+1} = initiator(W_{RAND}, W_{PE_{trun}})$
- 13: Send W_{RAND}

Algorithm 3 Negotiation based on *PE_{trun}* method**Input***B*: The number of selected features**Initialization** W_{RAND} $W_1 = 0$ $W_{PE_{trun}} = 0$

- 1: **for** $t=1, 2, \dots, T$ **do**
- 2: Receive x_t
- 3: Make prediction $sgn(x_t^T W_t)$
- 4: Receive y_t
- 5: **if** $y_t x_t^T W_t \leq 0$ **then**
- 6: $\tilde{W}_{t+1} = W_t + y_t x_t$
- 7: $W_{PE_{trun}} = Truncate(\tilde{W}_{t+1}, B)$
- 8: $\tilde{W}_{t+1} = initiator(W_{RAND}, W_{PE_{trun}})$
- 9: Send $W_{PE_{trun}}$

3.3 HIR-GB system: Human (worker and rider) Identity Recognition based on

Gait Biometrics

3.3.1 Pre-processing (CRF-RNN for worker and rider detection)

Since CNNs have achieved a great success in natural image analysis and the CRF outperformed other existing solutions in structural learning, the Conditional Ran-

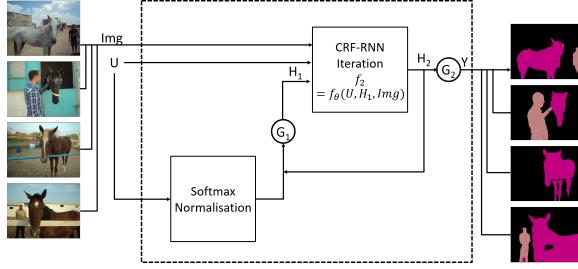


Fig. 4 The CRF-RNN Network

dom Fields as Recurrent Neural Networks (CRF-RNN) [54] segmentation method has been employed for the human body detection. To describe the deep learning system for semantic image segmentation, it is necessary to understand how repeated iterations are organized as an RNN. One iteration of the algorithm could be formulated as a stack of CNN layers. The transformation done by one CRF-RNN iteration was denoted by f_θ using an image Img , pixel-wise unary potential values U and a marginal probabilities estimation H from the previous iteration. The next estimation of marginal distributions after one iteration was given by $f_\theta(U, H, Img)$. Eq. 6-8 represent the behaviour of the network where T is the number of iterations while the θ vector represents the CRF parameters [54].

$$H_1(t) = \begin{cases} \text{Softmax}(U), & t = 0 \\ H_2(t-1), & 0 < t \leq T \end{cases} \quad (6)$$

$$H_2(t) = f_\theta(U, H_1(t), Img), \quad 0 \leq t \leq T, \quad (7)$$

$$Y(t) = \begin{cases} 0, & 0 \leq t < T \\ H_2(t)), & t = T \end{cases} \quad (8)$$

Using CRF-RNN, the output is a number of regions with different classes. The region reference map was obtained and only the human references have been selected. As shown in Fig. 4, the CRF-RNN of semantic image segmentation was benefited from in order to detect the human boundary area represented by the light pink color.

3.3.2 Worker and rider identity recognition

The main advantage of the gait biometrics is that it offers a great potential for recognition at a low resolution and from a distance. Using the proposed solution in [11], the gait signature was represented by the fusion of two types of silhouette characterization. The Gait Pal and Pal Entropy (GPPE) features were merged with four proposed distances from the silhouette. The proposed distances were d_1 , d_2 , d_3 and the slope of axis of least inertia of the silhouette area θ which were detailed in [11]. The proposed features were performed on the Casia-B dataset using the MLP classifier. Fig. 1 shows the human gait recognition solution for a Smart Riding Club Biometric System.

4 Databases

To the best of our knowledge, there is no public horse face image pattern database that can be used for the evaluation of detection algorithms. Therefore, the THDD was prepared for the horse face detection system while the THoDBRL'2015 was used for the horse face recognition system.

4.1 THDD

The THDD includes a set of frontal horse images. The digital images were taken at different distances ranging from 1 to 2 meter from horses. The capture was performed by a video camera. Two digital cameras were used. The first is of 10.1 Mega Pixels and at a resolution of 640 * 480 pixels, whereas the second is of 15 Mega Pixels and at a resolution of 12080 * 720 pixels. The collected data set consists of 1103 horse images and 6000 cropped face images for 60 Barbaro, Arabian and hybrid horses (Table 1). In fact, there were two sets: The first one was for horse face classification that includes 6000 positive images and 7937 negative images, the second one was for horse face detection that includes 1103 images. The first set was used as a training set and the second set was used as a test set in

this paper. Most of these animal images had near frontal view. Fig. 5 shows some sample images from the training set of the database. Fig. 6 shows some sample images from the test set of the database.



Fig. 5 Horse images from the training set of THDD



Fig. 6 Horse images from the test set of THDD

4.2 THoDBRL'2015

Our THoDBRL'2015 is a multi-view horse face database. The digital images were taken at a distance of about 1 meter from the horses when they were in the barn. The capture was achieved by the video camera using a digital still camera of 10.1 Mega Pixels and at a resolution of 640 * 480 pixels. The horse face data were captured from 3 views: frontal view, right view profile and left view profile of the horse. Our THoFDRL'2015 contains 470 frontal face images for 47 Barbaro, Arabian and hybrid horses (Table 1), 470 left profile images and 470 right profile images. In fact, following the same database construction process of the most related works

[21, 45, 22, 1], 10 frontal face images. Therefore, this database contains 10 left profile images and 10 right profile images were taken per horse.

In the THoDBRL'2015 database, there are three sets: The first one contains the captured videos, the second one includes selected images, the third one has the cropped images.

Table 1 Some Details about the used Databases

Database	Subjects	Views	breeds	Cite
THDD	60	Frontal	Barbaro	-
			Arabian	
			Hybrid (Barbaro-Arabian)	
THoDBRL'2015	47	Frontal	Barbaro	37
		Right	Arabian	
		Left	Hybrid (Barbaro-Arabian)	

4.3 Capture condition

4.3.1 Horses in motion:

The animal is not sane and it is impossible to fix its head and keep its stability. It may change its place and position for any time. These changes of place spoil the distance of 1 meter (for THoDBRL'2015) and 1-2 meters (for THDD) between the camera and the animal. In order to fix the head of the animal as much as possible to obtain adequate images, the capture of horse videos was done when they were in barns. The movements of the horse were reduced but did not disappear. Thus, the distance between the horse face and the camera slightly changed on the two databases .

4.3.2 Natural conditions:

Videos were taken from four equestrian centers in Sfax (a city in the east coast of Tunisia) at daylight. The camera was hand-held and positioned in front of the

photographer eyes at a distance of about 10 cm from his face. In order to guarantee a maximum fixation of the animal head and to obtain adequate images, the videos were taken when the horses were in barns in natural conditions and without any pressure.

There was a change of the lighting in the horse face according to the position of the animal head and the sun. In addition, there were shadows such as the shadow of the walls or the leaves (Fig. 7). The background also varied in each of the captured videos (Fig. 8).



Fig. 7 An illustration of luminance variation in three horses from THoDBRL'2015 and THDD

4.4 Data collection procedure

As the horse cannot be fixed or kept stable, a capture of a video for each horse of about 50 seconds was opted for in order to obtain adequate poses. Based on human observation, the best photo was selected with different poses, background and luminance. Selected images were neither 100% front nor 100% profile. The view could be inclined to the right or to the left. Consequently, the facial images of our database had almost the same size and resolution. The difference was not



Fig. 8 An illustration of background variation in some horses from THoDBRL'2015 and THDD

very large.

In the THDD, each face area was manually detected with four ROIs (Xmin, Ymin, Width, Height) of the selected image (fig. 9).

In the THoDBRL'2015, each face area was manually cropped of the selected image for three views (fig. 10).

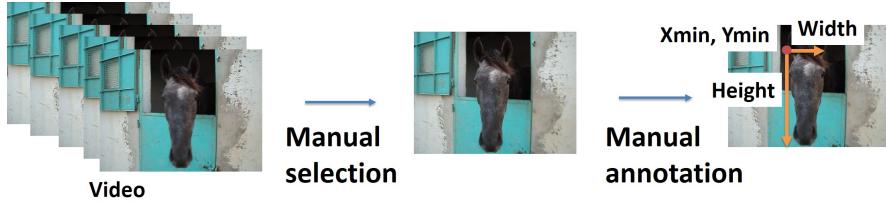


Fig. 9 THDD collection procedure

5 Experimental study of SRBCS

5.1 Implementation details

Gradient features: According to the experiment, nine unsigned orientations, a cell size equal to 8×8 and a grouped block including 2×2 cells were used.

Sparse hidden layer of SNN: A small fraction of features equal to 10% of the feature dimensions was selected as proposed in [3]. This fraction is enough to find

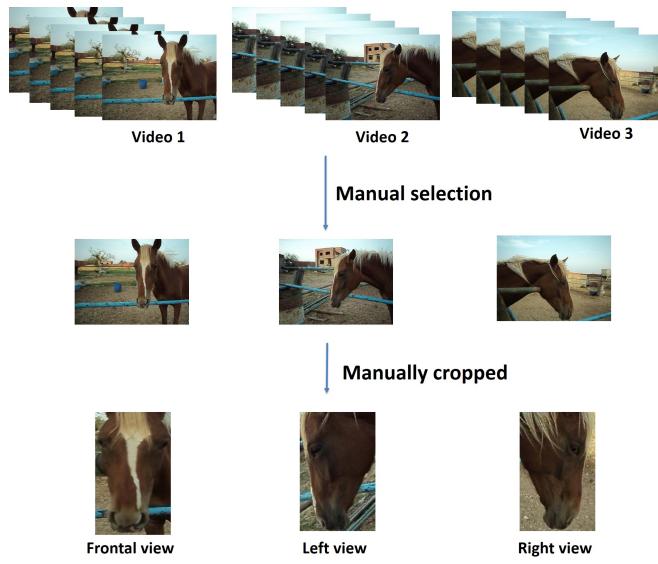


Fig. 10 THoDBRL'2015 collection procedure

favorable prediction results [3]. Indeed, 90% of the weight vector W of the hidden layer were equal to zero.

Intersection over Union (IoU): Following the Pascal challenge [12], A detected bounding box was considered as a true positive detection only when the Intersection over Union (IoU) ratio was equal to or larger than 50%. For a more accurate evaluation, the metrics proposed by COCO² challenge were used. 10 different IoU thresholds were considered from 0.5 to 0.95 in steps of 0.05. The average precision was calculated over 10 IoU.

Block generation for detection process: A very popular searching strategy has been proposed in [48] to detect face instances in the image. A sequential scan of all possible regions in the image was done by a sliding window. The highly accurate real-time human frontal face detector presented by [47] and [48] used the sliding window strategy. This technique is still used for its sufficiency in many recent detection works [14, 25, 36]. In this paper, this detection strategy was chosen owing to its interesting performance.

The digital images of the database were taken at distances ranging from 1 to 2

² <https://competitions.codalab.org/competitions/12061>

meter from the horses. Hence, the camera was considerably not too far from the animal and, consequently, the horse face region was a bit wide in the image. Thus, the horse face size in the image varied between $80 * 155$ to $360 * 640$. Taking this characteristic into account, the first stage consisted in scanning the entire image with a resized window (from $80 * 155$ to $360 * 640$) with 8 pixels in width, 20 pixels in height and 10 pixels stride. To control the presence of the horse face in each window , the output value of the SNN was monitored. Each window classified as face, would be kept to collect all predicted windows. The predicted windows were filtered applying the Non-Maximum Suppression algorithm (NMS).

Gabor descriptor: The same parameters used in [16, 17] were employed.

Multi-class SVM: The linear kernel of the multi-class SVM was used.

5.2 Experimental results

In the literature, the detection procedure usually includes three steps: block generation (multi-scale sliding windows or region proposals), face classification (in the backbone of the detector) and post-processing (non-maximum suppression and bounding box regression). In fact, the performance of face detectors is mainly influenced by the face classification network also known as the backbone. Duan *et al.* [10] discovered that the detector and the classifier of the general object detection have comparable performances using the same backbone. This explains that the designed backbone for the classification dataset is applied easily to the general object detection which gets an excellent mAP (mean Average Precision) score. Therefore, it is necessary to evaluate the proposed SNN and the HFD-SF system for horse face classification and detection on the THDD.

Using the gait modality, the CRF-RNN for human body detection and the proposed solution by [11], the recognition rates equal to 95.41% on Casia-B database were close to the recognition rates of the last SRCBS system in [34] using face biometrics even though the capture distances in Casia-B were larger than the distances in the databases used in [34]. The recognition result was the same as [11]

because CRF-RNN detected all human bodies of the test set. Then, it would be more efficient to replace the old system with ours.

5.2.1 HFD-SF system

Classification evaluation: For a good evaluation, the proposed SNN was compared with other CNNs such as MobileNetV2 [41] and GoogLeNet [43] which have represented the smallest number of parameters as well as ResNet-50 [51] which has been widely used for human face detection [29, 52]. In fact, MobileNetV2 contains 3.4 million pareameters, GoogLeNet contains 5 million parameters and ResNet-50 contains 25 million parameters. CNNs need a huge data for training to perform at its best. During the training of ResNet-50, GoogleNet and MobileNetV2, three types of data augmentation were used. Reflection, X and Y axis translation, and random scaling were applied to the images of the THDD training set. The training was done with a mini-batch size equal to 10. A transfer learning of the three pre-trained CNNs (on ImageNet) was made for classification on the training set of the THDD. The last Fully Connected (FC) layer in these CNNs was replaced with another FC layer having two outputs (face/non-face). The classification was made on the test set of the THDD which contains 1124 horse faces and 10000 negative images randmoly selected from the background.

The comparison includes the accuracy of classification and the number of parameters of each network. Fig. 11 shows the accuracy of the classifications. It is very obvious that the accuracy of SNN was very large compared to GoogLeNet and ResNet-50. The accuracy of MobileNetV2 and NN-Sparse is almost the same (Table 2). SNN presents competitive results with a small number of parameters equal to 76.928.

Detection evaluation: Since human and horse faces share similar structures (two eyes, nose and mouth), starting with the existing human frontal face detection approaches has been proposed. Unfortunately, applying these approaches directly on horses met some obvious difficulties. In fact, the horse faces have large ap-

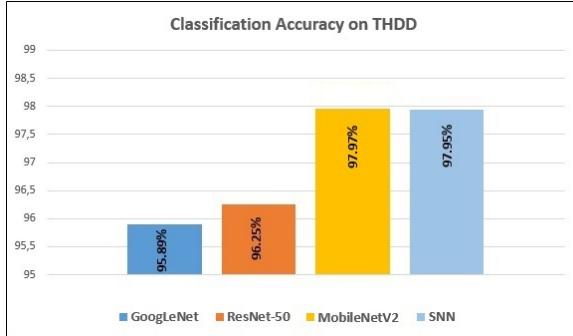


Fig. 11 Classification results on THDD using SNN, GoogLeNet, MobileNetV2 and ResNet-50

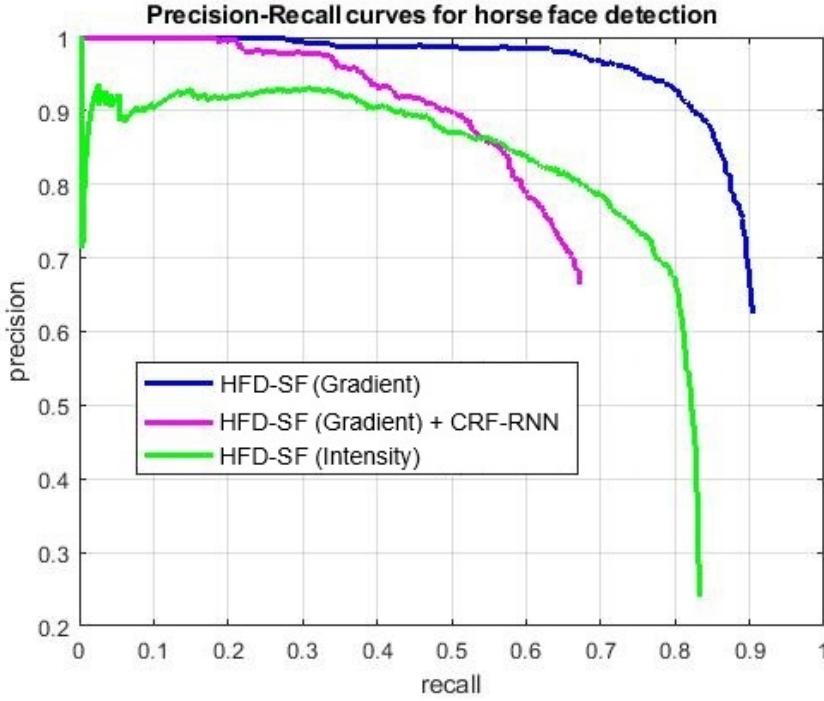
Table 2 Classification Rates of different networks on THDD

CNN	ACC	PPV	TPR	SPC	NPV	F1
ResNet-50	96.25%	95.61%	65.93%	99.66%	96.30%	78.04%
GoogLeNet	95.89%	97.99%	60.59%	99.86%	95.75%	74.88%
MobileNetV2	97.97%	90.35%	89.15%	98.96%	98.78%	89.88%
SNN	97.95%	89.02%	90.93%	98.74%	98.98%	89.96%

pearance variations and intricate textures compared to the human face. Due to the small amount of data in our database, we could not use CNN human face detection methods as they require a large number of images for effective training. Since there are no related works to compare the results, our approach was compared with the Viola-Jones detector that used for cat face detection by [32]. Our detector was found to achieve a higher performance. The Average Precision of the two detectors (HFD-SF and Viola-Jones) was 90.00% and 70.00% respectively. The detection results were compared (shown in Table 3) using two kinds of features, gradient features and intensity features (using LBP features). The average precision and recall of the two descriptors (gradient, intensity) using the proposed detection system were (90%, 90.39%) and (71.40%, 83.27%) respectively. Fig. 12 shows precision-recall curves on the THDD. This figure represents the performance of the proposed detector HFD-SF on the THDD and reports that the proposed method had the biggest critical region. The Haar (Fig. 14) and weighted LBP features (Fig. 12) gave the poorest performance because of the large texture variations and shapes of the horse head. With the help of the proposed SNN, we have

Table 3 A comparison between the results of the gradient features and the intensity features

Features	AP	Recall	Precision	F1
Viola-Jones(HOG)	70.00%	-	-	-
HFD-SF(Intensity)	71.40%	83.27%	-	-
HFD-SF(Gradient)	90.00%	90.39%	63.50%	84.87%

**Fig. 12** Precision-recall curves on THDD

improved the performance (Fig. 15). Using the evaluation metrics of the COCO challenge, Fig. 13 displays precision-recall curves calculated at 10 different IoU thresholds. Using the proposed SNN, the average precision varied between 84% and 90% for the first four IoU thresholds.

To reduce the search time for the horse faces in the image, the CRF-RNN was applied for regions selection to select the horse bodies. The face search was restricted to the selected area only. The experiment showed that the search with Sliding Window (SW) gave better results (Fig. 12). The proposed SNN proved its performance on the last experiments for classification and detection. Fig 11

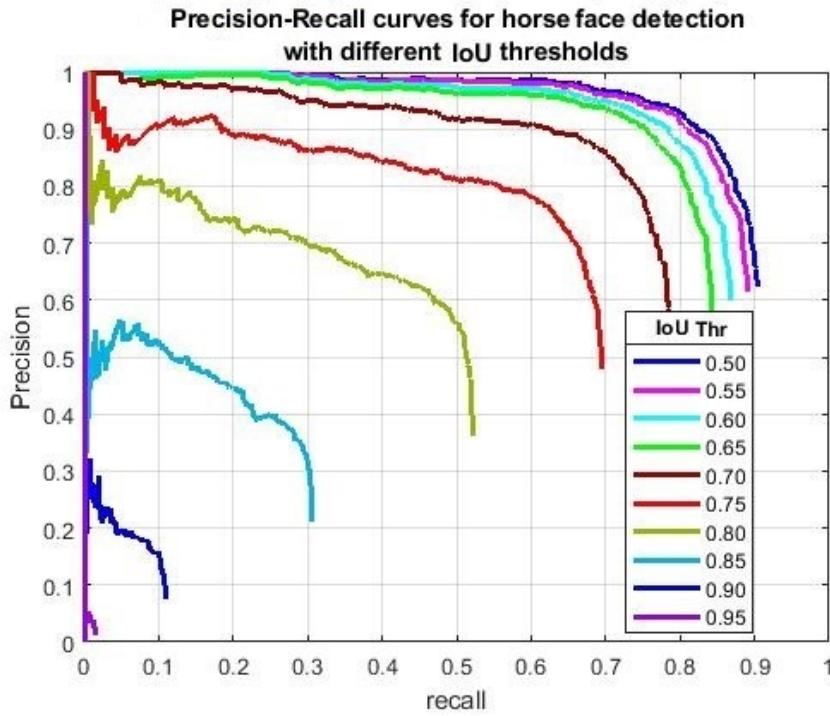


Fig. 13 Precision-Recall curves calculated at various IoU thresholds according to the COCO challenge

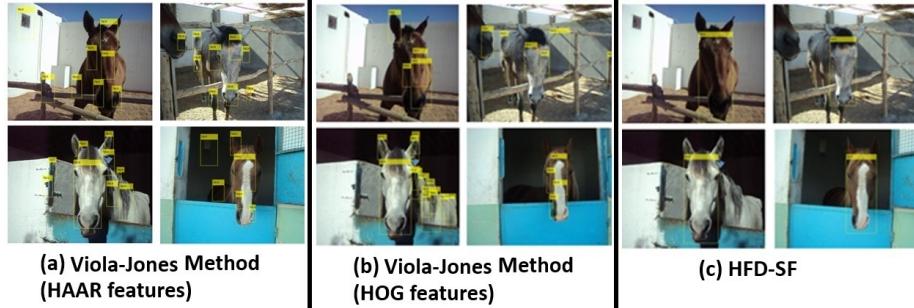


Fig. 14 Comparison between the detection results of Viola-Jones method [47] and the proposed HFD-SF

presents the classification accuracy on the THDD. It is very obvious that the SNN gave competitive results compared to the other networks. In fact, the SNN overcame the accuracy of the other CNNs by about 1.25% on the THDD (Table 2). The same thing was noticed for the detection process. HFD-SF based on SNN gave encouraging results with an Average Precision equal to 90%.

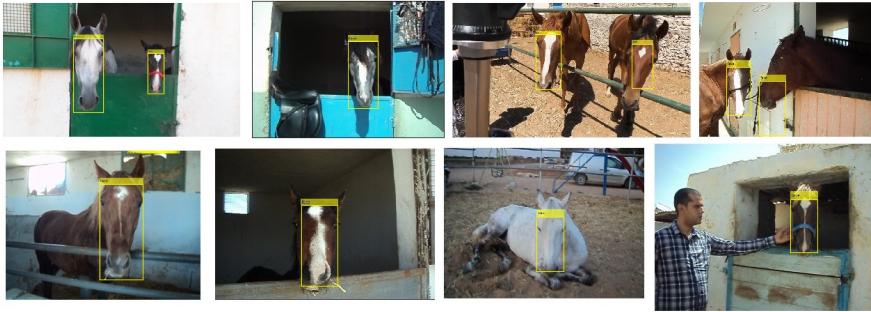


Fig. 15 Detection results on the THDD

The use of a sparse feature selection method as a learning algorithm enhanced the information transmitted to the output layer. In fact, the proposed sparse hidden layer and training methodology contributed to a proper distinction between true and false detection (face/non-face). SNN extracted the relevant features using the sparse hidden layer and then classified the candidate bloc using the output layers. Unlike the other networks, SNN kept as much information as possible by minimizing the number of operations and parameters. Consequently, the sparse hidden layer positively influenced decisions and brought detection closer to reality.

5.2.2 HIR-FB system

Among the THoDBRL'2015 images, only the frontal face images were interesting and seven face images were chosen per horse (329 images in total) for the training step of the facial recognition system. To better verify the performance and the efficiency of the HFD-SF system in SRCBS and also to achieve end-to-end recognition system, an investigation on horse identity recognition has been integrated based on our detection algorithm results. Using the proposed HFD-SF system, faces were detected on 3 images from the THDD for each horse in the THoDBRL'2015. These images were used for evaluation on the HIR-FB system. In fact, three face images per horse (141 images in total) were used as test images.

In the identification of applications, the camera is fixed and the background was

static. It is easy to take a background model and identify objects of interest by detecting changes in the background. In our system, the case was different. In fact, the capture was done in natural conditions with varied background. Because of these reasons and the variety of luminance in our database, the facial area of the horse was cropped with Matlab from the original images of the database and resized (the same image size 160 * 380) as shown in Fig. 16.

Table 4 shows that the proposed recognition system gave the highest performance



Fig. 16 Samples of frontal face images from the training set of four horses; (a) first horse, (b) second horse, (c) third horse, (d) fourth horse.

in two processes: Verification and Identification with interesting results compared to the results of the related works. Fig. 17 of ROC curves and Fig. 18 of CMC curves disclose that the HIR-FB had the biggest critical region compared to the others. HIR-FB recorded an 8.5% and a 6.62% improvement in the identification and verification rates respectively. It is worth reminding that there has already been a suitable horse face recognition system in [16, 34, 17]. However, their previously published results were less than the present results.

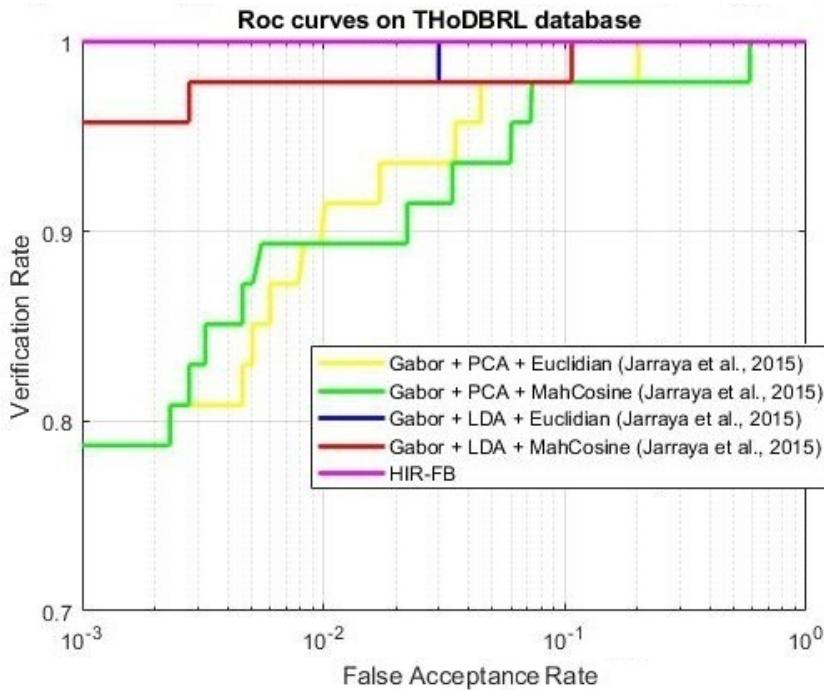
For more evaluation, the proposed approaches were compared with some famous and standard systems of face recognition including VGG-Face [35], SphereFace [28] and ArcFace [9] and RiseNet [42].

VGG-Face [35] is a convolutional neural network composed of 37 layers including input, convolutional, relu, max pooling and softmax layers.

SphereFace [28] is based on a convolutional neural network of 64 layers including convolutional, fully connected and softmax layers with an angular softmax loss

Table 4 Evaluation of the performance of HIR-FB system on THoDBRL'2015.

Approach	Rank-1	VR at	VR at	VR at
	RR	FAR=0.001	FAR=0.01	FAR=0.1
Gabor + PCA + Euc [16]	82.98%	77.19%	92.10%	99.08%
Gabor + PCA + MC [16]	80.85%	77.19%	89.14%	99.09%
Gabor + LDA + Euc [16]	95.74%	95.19%	99.08%	99.08%
Gabor + LDA + MC [16]	95.74%	95.19%	99.08%	99.08%
Gabor + SA [17]	94.22%	-	-	-
RNGLBP + SVM [34] (Old SRCBS)	98.77%	99.08%	100%	100%
HIR-FB	99.89%	100%	100%	100%

**Fig. 17** ROC curves of HIR-FB system and the approaches of [16] on THoDBRL'2015

(A-Softmax) that enables the CNN to learn angularly discriminative features.

ArcFace (Additive Angular Margin Loss) was proposed by [9] to get highly discriminative features for human face recognition. ArcFace is incorporated within CNN architectures such as ResNet-50 and ResNet-100.

RiseNet proposed by [42] is an animal face recognition framework .

The above-listed systems are based on pre-trained networks on the ImageNet [8].

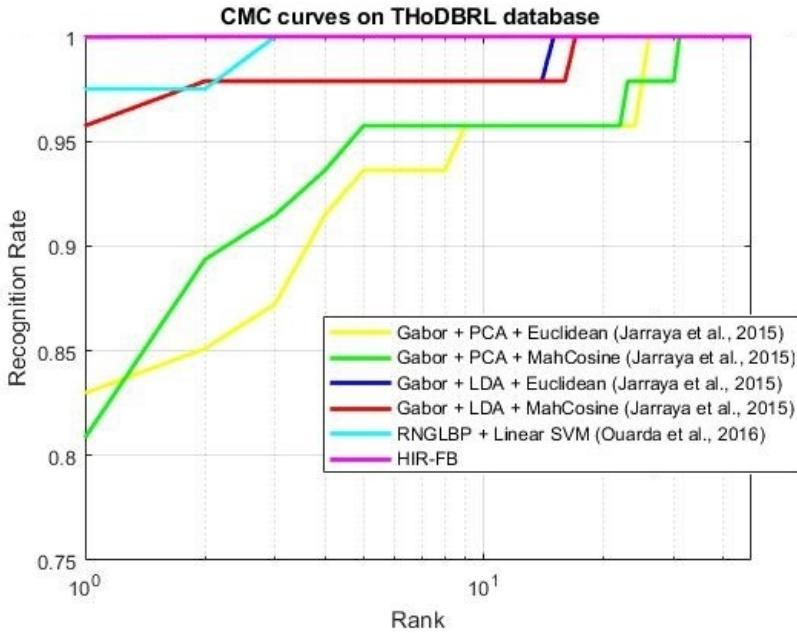


Fig. 18 CMC curves of HIR-FB system and the approaches of [16, 34] on THoDBRL'2015

A transfer learning on the THoDBRL'2015 obtained the results shown in Table 5. This table proves the superiority of the proposed approach based on Gabor features, LDA and SVM. The proposed HIR-FB proved its performance on the

Table 5 Comparison between the horse face recognition rates of the proposed approaches and the Related works

Ref	Methods	Accuracy
[35]	VGG Face	72.62%
[28]	SphererFace	74.98%
[9]	ArcFace	75.32%
[42]	Fine-tuning of Arcface	81.79%
[42]	RiseNet	82.56%
HIR-FB	Gabor+LDA+SVM	99.89%

experimental part for the classification and verification processes. In fact, the results were very encouraging with a recognition rate equal to 99.89%. The proposed system improved the recognition rate by 1.12% compared to the old SRCBS [34] and 22.44% in comparison with the standard systems for face recognition

[35, 28, 9, 42]. According to these results, facial pattern can be considered as a well biometric marker for horse identification.

The multi-class SVM was used to separate the data in n classes by optimal hyperplanes. It uses different kernels for feature transformation in a new helpful representation to facilitate the optimization of the margin by reducing feature complexity. The choice of the best and simplest kernel was proposed in this work to easily separate the feature classes. The use of the multi-class SVM based on a linear kernel is justified by the linearity of data using the Linear Discriminant Analysis (LDA) which represents the features in a new space based on the projection of eigenvectors. In addition, the Gabor features and LDA proved their efficiency for horse face recognition.

6 Limitations

Owing to the photos which were taken close to the pets in the used databases, the faces of the animals were not very small and the system easily detected and recognized them. Nevertheless, the proposed SRCBS cannot detect very small faces when the animal present in the photo is very far away. This was due to the poor resolution of the facial area as well as the lack of the important details and information. However, the performance of the sparse feature selection method decreased as the information was reduced. The more information there is, the fairer the SNN does the classification. The problem of detecting small faces is actually a challenge in the backbone of the most popular detectors. In an open space, when the horse is not in the stable, the animal will move around and sometimes the camera will only be able to capture a part of its body. The proposed SRCBS cannot in this case identify the horse and the face biometrics become useless. Therefore, the addition of the gait biometrics to the recognition system is effective. The construction of a multimodal identification system based on the face and the gait biometrics is an urgent necessity.

7 Conclusion

In this paper, a new biometric system for smart riding club security called SRCBS was introduced. This system is based on two main sub-systems to detect and recognize both horses and humans.

The proposed HFD-SF system for horse face detection is based on the proposed network SNN. The SNN gave competitive classification results compared to the others networks. In fact, the SNN overcame the accuracy of the other CNNs by about 1.25%. This detection system achieved encouraging results. The experiments on the THDD proved that our system is efficient by reaching a useful detection rate equal to 90%. The use of a sparse feature selection method as a learning algorithm enhanced the information transmitted to the output layer. In fact, the proposed sparse hidden layer and training methodology contributed to proper distinction between true and false detections (face/non-face). Unlike the other networks, SNN kept as much information as possible using a smaller number of parameters equal to 76928.

The proposed HIR-FB for horse face recognition proved their performance by a recognition rate equal to 99.89% on the THoDBRL'2015. In fact, HIR-FB enhanced the recognition rate by 1.12% compared to the old SRCBS [34] and 22.44% in comparison with the standard systems for face recognition. Thus, the facial pattern can be considered as a good biometric marker for horse identification according the found results.

In order to obtain a more efficient SRCBS system, with supplementary services for humans, our paper has proposed to use our previously developed approach in [11]. Using the gait modality, the CRF-RNN for human body detection and the proposed solution by [11], the recognition rates equal to 95.41% on Casia-B database were close to the recognition rates of the last SRCBS system in [34] using face biometrics.

Our future perspectives can be summarized as follows:

- Developing the SRCBS system for horse and human detection and identification at real time.
- Expanding the proposed HFD-SF system in two directions. First, it is important to improve its performance and effectiveness by designing more discriminant features. Second, extending the HFD-SF system to other animals is among our plans for the future.

Acknowledgment

The research leading to these results has received funding from the Tunisian Ministry of Higher Education and Scientific Research under the grant agreement number LR11ES48.

The authors would like to acknowledge that this research was supported by four Riding Clubs in Sfax, Tunisia: Equestrian Clubs of road Mahdia, road Tunis in Sakit Ezzit, road Ain Km 17, and road Saltnia Km 17.

References

1. Awad AI, Zawbaa HM, Mahmoud HA, Nabi EHHA, Fayed RH, Hassanien AE (2013) A robust cattle identification scheme using muzzle print images. In: Federated Conference on Computer Science and Information Systems, pp 529–534
2. Barron UG, Corkery G, Barry B, Butler F, McDonnell K, SWard (2008) Assessment of retinal recognition technology as a biometric method for sheep identification. International Journal of Computer Vision 60:156–166
3. BenSaid F, Alimi AM (2021) Online feature selection system for big data classification based on multi-objective automated negotiation. Pattern Recognition 110:107–629

4. Boughida A, Kouahla MN, Lafifi Y (2021) A novel approach for facial expression recognition based on gabor filters and genetic algorithm. *Evolving Systems*
5. Bragança FS, Broomé S, et al MR (2020) Improving gait classification in horses by using inertial measurement unit (imu) generated data and machine learning. *Scientific Reports* 11:1–3
6. D S Edwards DUP A M Johnston (2001) A comparison of commonly used ear tags on the ear damage of sheep. *Animal Welfare* 10(2):141–151
7. Dalal N, Triggs B (2005) Histograms of oriented gradients for human detection. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp 886–893
8. Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L (2009) Imagenet: A large-scale hierarchical image database. In: IEEE Conference on Computer Vision and Pattern Recognition, pp 248–255
9. Deng J, Guo J, Xue N, Zafeiriou S (2019) Arcface: Additive angular margin loss for deep face recognition. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp 4685–4694
10. Duan J, Liao S, Zhou S, ZLi S (2016) Face classification: A specialized benchmark study. In: Chinese Conference on Biometric Recognition, pp 22–29
11. Ehaimir ME, Jarraja I, Ouarda W, Alimi AM (2017) Human gait identity recognition system based on gait pal and pal entropy (gppe) and distances features fusion. In: Sudan Conference on Computer Science and Information Technology (SCCSIT), pp 1–5
12. Everingham M, Gool LV, Williams CKI, Winn J, Zisserman A (2010) The pascal visual object classes (voc) challenge. *International Journal of Computer Vision* 88:303–338
13. France (1999) Loi n 99-574 du 9 juillet 1999 d'orientation agricole. JORF n°0158 <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000000395813/>

14. Hang Z, Wang Y, Huang S (2021) Accurate heavy hitter detection with sliding window approach. In: IEEE International Conference on Computer and Communication Systems (ICCCS), pp 512–516
15. Hummel HI, Pessanha F, Salah AA, van Loon TJP, Veltkamp RC (2020) Automatic pain detection on horse and donkey faces. In: IEEE International Conference on Automatic Face and Gesture Recognition, pp 793–800
16. Jarraya I, Ouarda W, Alimi AM (2015) A preliminary investigation on horses recognition using facial texture features. In: IEEE International Conference on Systems, Man, and Cybernetics, pp 2803–2808
17. Jarraya I, Ouarda W, Alimi AM (2017) Deep neural network features for horses identity recognition using multiview horses' face pattern. In: ICMV, pp 52–56
18. Jiang H, Learned-Miller E (2017) Face detection with the faster r-cnn. In: IEEE International Conference on Automatic Face Gesture Recognition, pp 650–657
19. Joshi S, Verma DK, Saxena G, Paraye A (2019) Issues in training a convolutional neural network model for image classification. In: International Conference on Advances in Computing and Data Sciences, pp 282–293
20. Kishore T, Jha A, Kumar S, Bhattacharya S, Sultana M (2021) Deep cnn based automatic detection and identification of bengal tigers. In: International Conference Computational Intelligence in Communications and Business Analytics, pp 503–507
21. Kumar S, Singh S (2014) Biometric recognition for pet animal. Journal of Software Engineering and Applications 07:470–482
22. Kumar S, Tiwari S, Singh SK (2015) Face recognition for cattle. In: International Conference on Image Information Processing (ICIIP), pp 65–72
23. Kumar S, Pandey A, Satwik KSR, Kumar S, Singh SK, Singh AK, Mohan A (2017) Deep learning framework for recognition of cattle using muzzle point image pattern. Measurement 116:1–17

24. Kusakunniran W, Wiratsudakul A, Chuachan U, Kanchanapreechakorn S, Imaromkul T (2018) Automatic cattle identification based on fusion of texture features extracted from muzzle images. In: IEEE International Conference on Industrial Technology (ICIT), pp 1484–1489
25. L Kulanuwat CC, et al (2021) Anomaly detection using a sliding window technique and data imputation with machine learning for hydrological time series. Water 13
26. Li C, Huang Y, Huang W, Qin F (2021) Learning features from covariance matrix of gabor wavelet for face recognition under adverse conditions. Pattern Recognition 119:108085
27. Li Z, Broomé S, Andersen PH, Kjellström H (2020) Automated detection of equine facial action units. arXiv:210208983
28. Liu W, Wen Y, Yu Z, Li M, Raj B, Song L (2017) Sphereface: Deep hypersphere embedding for face recognition. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp 6738–6746
29. Liu W, Hasan I, Liao S (2020) Center and scale prediction: A box-free approach for pedestrian and face detection. arXiv:190402948v3
30. M MaryHelta Daisy PK (2021) Investigation of rotated local gabor features in face recognition using fusion techniques. Journal of Ambient Intelligence and Humanized Computing 12:5895–5908
31. Mougeot G, Li D, Jia S (2019) A deep learning approach for dog face verification and recognition. In: Pacific Rim International Conference on Artificial Intelligence, pp 418–430
32. Mukai N, Zhang Y, Chang Y (2018) Pet face detection. In: Nicograph International, pp 52–57
33. North S (2016) Do androids dream of electric steeds? the allure of horse-computer interaction. interactions 23:50–53
34. Ouarda W, Trichili H, Alimi MA, Solaiman B (2016) Towards a novel biometric system for smart riding club. JIAS 11:201–213

35. Parkhi OM, Vedaldi A, Zisserman A (2015) Deep face recognition. In: Proceedings of the British Machine Vision Conference (BMVC), pp 41.1–41.12
36. R Li KW B Qu, et al (2021) Identification of two methylated fragments of an sdc2 cpg island using a sliding window technique for early detection of colorectal cancer. Water 11:1941–1952
37. Rathor S, Kumari S, Singh R, Gupta P (2021) Two layers machine learning architecture for animal classification using hog and lbp. In: Proceedings of International Conference on Communication and Artificial Intelligence, pp 445–453
38. Ren S, He K, Girshick R, Sun J (2015) Faster r-cnn: Towards real-time object detection with region proposal networks. In: International Conference in Neural Information Processing Systems, pp 91–99
39. Sabour S, Frosst N, Hinton G (2017) Dynamic routing between capsules. In: International Conference on Neural Information Processing Systems, p 3859–3869
40. Salama A, Hassaniene AE, Fahmy A (2019) Sheep identification using a hybrid deep learning and bayesian optimization approach. IEEE Access 7:31681–31687
41. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L (2018) Mobilenetv2: Inverted residuals and linear bottlenecks. In: IEEE Conference on Computer Vision and Pattern Recognition, pp 4510–4520
42. Shi X, Yang C, Xia X, Chai X (2020) Deep cross-species feature learning for animal face recognition via residual interspecies equivariant network. In: European Conference on Computer Vision, pp 667–682
43. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z (2016) Rethinking the inception architecture for computer vision. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp 2818–2826
44. Taunk P, Jayasri G, Priya J P, Kumar NS (2020) Face detection using viola jones with haar cascade. Test Engineering and Management 83

45. Tharwat A, Gaber T, Hassanien AE (2014) Cattle identification based on muzzle images using gabor features and svm classifier. In: International Conference on Advanced Machine Learning Technologies and Applications, pp 236–247
46. Tureckova A, Holik T, Oplatkova ZK (2020) Dog face detection using yolo network. MENDEL Soft Computing Journal 26:2571–3701
47. Viola P, Jones M (2001) Robust real-time face detection. In: IEEE International Conference on Computer Vision, pp 747–747
48. Viola P, Jones MJ (2004) Robust real-time face detection. International journal of computer vision 57(2):137–154
49. Vlachynska A, Oplatkova ZK, Turecek T (2018) Dogface detection and localization of dogface's landmarks. CSOC 764:465–476
50. Yamada A, Kojima K, Kiyama J, Okamoto M, Murata H (2011) Directional edge-based dog and cat face detection method for digital camera. In: IEEE International Conference on Consumer Electronics (ICCE), pp 87–88
51. Zagoruyko S, Komodakis N (2017) Wide residual networks. arXiv:160507146v4
52. Zhang B, Li J, Wang Y, Tai Y, Wang C, Li J, Huang F, Xia Y, Pei W, Ji R (2020) Asfd: Automatic and scalable face detector. arXiv:200311228
53. Zhang W, Sun J, Tang X (2011) From tiger to panda: Animal head detection. IEEE Transactions on Image Processing 20:1696–1708
54. Zheng S, Jayasumana S, Romera-Paredes B, Vineet V, Su Z, Du D, Huang C, Torr PHS (2015) Conditional random fields as recurrent neural networks. In: IEEE International Conference on Computer Vision, pp 1529–1537