Research Article

Jinmao Tong, Zhongwang Cao, and Wen Fu*

Motion vector steganography algorithm of sports training video integrating with artificial bee colony algorithm and human-centered AI for web applications

https://doi.org/10.1515/jisys-2022-0093
received February 16, 2022; accepted December 12, 2022

Abstract: In multimedia correspondence, steganography schemes are commonly applied. To reduce storage capacity, multimedia files, including images, are always compressed. Most steganographic video schemes are, therefore, not compression tolerant. In the frame sequences, the video includes extra hidden space. Artificial intelligence (AI) creates a digital world of real-time information for athletes, sponsors, and broadcasters. AI is reshaping business, and although it has already produced a significant impact on other sectors, the sports industry is the newest and most receptive one. Human-centered AI for web applications has substantially influenced audience participation, strategic plan execution, and other aspects of the sports industry that have traditionally relied heavily on statistics. Thus, this study presents the motion vector steganography of sports training video integrating with the artificial bee colony algorithm (MVS-ABC). The motion vector stenography detects the hidden information from the motion vectors in the sports training video bitstreams. Artificial bee colony (ABC) algorithm optimizes the block assignment to inject a hidden message into a host video, in which the block assignment is considered a combinatorial optimization problem. The experimental analysis evaluates the data embedding performance using steganographic technology compared with existing embedding technologies, using the ABC algorithm compared with other genetic algorithms. The findings show that the proposed model can give the highest performance in terms of embedding capacity and the least error rate of video steganography compared with the existing models.

Keywords: steganography, artificial bee colony, motion vector, sports training video, artificial intelligence, human-centered, web applications

1 Prologue study

Today, computers and internet technology applications are growing, and the dissemination of new media is growing every day. Digital information can be updated and intercepted quickly by users on any open platform [1]. As a result, data protection is critical, and several scientists suggested many works to improve information security [2]. Unknown individuals or hackers on the internet can quickly intercept confidential data when transmitted. This makes it easier to encrypt information exchanged over the web application...
using steganography techniques. Steganography characterizes confidential communication design [3]. The message is sent to the public domain through cryptography in an encrypted manner. The message’s true contents are revealed to an outsider who understands there must be communication [4]. The main purpose of steganography, unlike cryptography, is to conceal existence of the message to the observer so that the observer does not know that it communicates a hidden message [5]. A cover object is identified and received after the embedding is considered a stego object. The object is used to mask the hidden message [6].

“Steganalysis” applies in this sense to methods intended to separate cover-objects and stego-objects. It must be remembered that differentiation must be made without knowing the secret key and often without knowing the particular algorithm which can be used to embed the secret message [7]. The object can be killed, and its delivery to the destination can be aborted until the message is identified. Thus, Steganalysis determines only the presence in the cover object of a hidden message. Applications of Steganalysis are as described in multimedia [8]. Multimedia files such as image frames, motion pictures (videos), and audio are ideal object covers for steganography. Embedding based on motion vectors seems to be a common video steganography branch, where motion vectors are altered to embed hidden passages [9]. Web application-based methods pick the primary relevant motion vectors based on certain selection rules and adjust the chosen motion vectors for inclusion. Video files can handle a comparatively large payload [10]. Successive video frames as a series of still pictures are modeled by a subset of current video steganography algorithms and are extended into the video chain. Such approaches that yield sub-optimal solutions as video time movements are not considered. Time complexity should be considered to overcome this problem [11,12]. Different methods have been suggested for steganography, which gains videos such as compression, scale, etc.

The steganography video is very productive and effective because it is more powerful than an image. In steganographic video domains like MP4, Moving Picture Experts Community (MPEG), Audio Recording Interleave, or several different video file formats can be used (AVI) [13,14]. Steganography can be typically categorized into two major categories of uncompressed and compressed images in the video. In general, information enclosed in the video is equivalent to information obscured from an image in which data are hidden in multiple video frames [15]. Steganography techniques can therefore be implemented on video in an image. The most common technique is the Discrete Cosine Transform [16]. Video steganalysis has been comparatively less vigilant until now than picture steganalysis. The method steganalysis uses the collusion attack to locate the cover video on stego video, using kurtosis and entropy features for classification [17]. Human-centered AI for web applications suggested to use the similarity discrepancies from frame to frame to statistically distinguish the components between the stego video and the cover video.

Motion capture technology of moving targets based on human-centered AI for web applications in sports training is used to intellectualize present professional sports training. Sports skills are a cognitive capacity and a tacit knowledge from an educational psychology perspective. Although sports technology, language, and image play an auxiliary role in sports training, they are not the primary teaching technique.

Therefore, this study introduces the motion vector steganography incorporating the artificial bee colony algorithm for sports training video (MVS-ABC). The motion artificial bee colony algorithm (ABC) optimizes the block task to insert a secret message in a hosted video in sports simulation video streams. The blocked task is an optimizing combination problem. In addition, AI is increasingly becoming a part of the mainstream sports industry. Digital competition has ushered in a new age in professional sports, and the reliance on players’ and coaches’ skills and knowledge has begun to diminish. In human-centered AI for web applications, quantitative evaluation of athletic performance is done by computing sports characteristics, motion recognition, and the division of the hitting stage in football, among other things.

The rest of the article is distributed as follows: Section 2 reports the major research innovations related to the given context. The design of the proposed MVS-ABC system has been discussed in Section 3. Next the experimental analysis and findings have been discussed in Section 4. Section 5 includes the conclusion and the future scope of the proposed system.
2 Existing works

2.1 Video steganography

Yao and Yu [18] suggested a video steganography technique for payload allocation based on a study of distortion variations in motion vectors. They studied the distortion of motion vector adjustment induced by data integration. A rate-distortion model was then derived, representing the proliferation of residue variance in succeeding inter-coded frames. The residual distinction weight was determined for each inter-coded frame. The strategy for allocating interframe payload was finally developed to limit the propagation of residue variations.

The new MPEG-2 video compression method based on movement vector steganography is proposed by Rana et al. [19]. The secret bit was hidden in the video by modifying the motion vector in the uniform areas. Higher motion vectors were used for embedding purposes to increase statistical undetectability. An efficient search window suggested the candidates’ macroblocks for embedding inhomogeneous regions. To adjust the motion vectors with the candidate macro-block motion vectors for effective embedding with a greater probability of non-detection in re-encoding steganalysis methods, their novel framework was implemented.

Cao et al. [20] proposed secure video steganography based on intra prediction model (IPM), to boost our algorithm’s efficiency as a new integration technique. Unlike adaptive steganography, changes to the current block were automatically made before the next block intra prediction such that changes in the existing frame would not impact the next. Furthermore, the remainder of the values were revised to allow optimum utilization of the changed IPM after updating the IPM. Finally, a cover selection rule was applied to maintain the balance between capability and video quality, where only eligible blocks are combined with a hidden message.

The integer wavelet transform based on the Singular value decomposition (IWT-SVD) scheme was suggested by Pilania and Gupta [21] to protect watermarks in a video cover file. The cover of video file concealed watermarks allows a greater capacity to camouflage. The watermark concealment in the sub-bands HH and LL was demonstrated to achieve high perceptive efficiency, greater robustness, and lower device costs.

2.2 Intelligence algorithm in steganography

Georges and Magdi [22] explored numerous artificial intelligence methods used to hit the greatest possible level of concealed knowledge, and picture quality since the low-quality distortion signs indicate hidden element inclusion. They discussed the implementation scope of the ABC algorithm as a part of their detailed study.

Zebari et al. [23] reviewed various swarm intelligence algorithms based on pictorial steganography, and they illustrated the significance of using swarm intelligence algorithms in the steganography of images. The swarm intelligence algorithm demonstrated its value and approval with valid claims in the proposed work.

Kaur et al. [24] proposed a novel approach to image steganography for securing secret data using a hybrid hiding model that secures the transmission of an image from attackers. They added Image Hiding Encryption and Decryption (IHED). Besides, a novel mid-search Africa buffalo model describes the encoding mechanism on mid-frequence values. Validated by many attacks such as a novel WFSA, RS steganalysis, chi-square attack and visual attack, the proposed model’s utility was verified.

Behbahani [25] suggested a new blink steganalysis approach based on features to detect stego images with JPEG format from the cover (clean). They applied a selection technique based on an improved ABC in this regard. The approach proposed depends upon the wrapper-based approaches, classification efficiency, and the chosen function vector size. Two big datasets of JPEG images were used to carry out the experiments.

The particle swarm optimization (PSO), bi-orthogonal wavelet transform (BWT), and genetic algorithm were combined by Pramanik et al. [26]. The upgraded version of the host photos was included with the PSO,
and the host images became clearer and brighter in an upgraded version. The selective host image sub-bands were chosen by BWT. The most suitable hidden image was chosen from various hidden images generated after mutation–genetic algorithm. An advanced technique creates a coded password in the hidden picture. A revolutionary technique was applied to image steganography to transfer sensitive information in a cover image.

Denysova et al. [27] suggested to examine the issue of deciding on cloud technologies (CT) for remote education for experts in physical culture and sport. The writers analyzed distance learning systems from both local and international countries to solve the problem (e-Learning Management System, LMS). Because of the comparative functional properties of LMS systems and cloud services, the future implementation of distant learning for physical education experts shows promise due to CT.

The category-wise literature survey showed that the motion vector steganography approaches performed well, and swarm intelligence optimization (SIO) algorithms’ scope in steganography was significant. The research gap for integrative motion approaches is discovered in these research reports. Therefore, this study presents the MVS-ABC. The motion vector stenography transmits secret information from motion vectors in sports simulation videos. ABC optimizes the block task for injecting a hidden message in the video hosting, which considers the block task to be an optimizing combination issue.

### 3 The proposed methodology

In this section, the video steganographic algorithm using motion vectors has been discussed with significant description. Following this, the ABC algorithm and human-centered AI for web applications integration in motion vector steganography has been explained. The following shows the motion vector steganography procedure in detail.

Figure 1 describes the transmission and retrieval parts of the motion vector video steganography. Figure 1 shows the selected cover video of sports training from the human metabolome database (HMDB) dataset [28]. Each frame was measured using a motion detector process to locate a motion vector. The motion estimation took place with 8 × 8 pixel blocks, which are non-overlapping from left to right and from top to bottom. The motion vector detection and the frame selection perform motion compensation. Cumulative absolute difference (CAD) between the present and preceding frames determines the motion vectors. In human-centered AI for web applications used to remove the cover/hide a secret message, the blocks with the motion vector can be used. The macro-block seems to have the relevant magnitude values for the hiding place. Each frame is determined using modified entropy until all motion vectors have been collected. The secret message encoding performs least significant bit (LSB) substitution approach enhanced with ABC. The ABC algorithm performs the block assignment. Figure 1(b) illustrates how the secret message is extracted. Figure 1 shows that the stego video included with stego frames received from the source at the destination through the public (unsecure) communication channel is passed to the extraction module that performs wavelet transformation to extract the secret message covered by the stego video. The message decoding section includes frame extraction, frame identification, motion vector identification, digital cosine transformation, etc. The following section details the proposed system model statistically. Steganographic video is more powerful and effective than pictorial in its high capacity using human-centered AI for web applications. In steganography videos, in the format like the MPEG, MP4, and AVI, etc., can be used. Here the golf.avi video has been selected for generating the stego video. The input video usually is categorized as uncompressed and compressed videos by video steganography. In essence, the information concealed in the video is like information hidden in a picture where the information is hidden in various video systems.
3.1 Statistical model for motion vector-based video steganography

Consequently, image steganography technique can be applied to video steganography. In video steganography, LSB technology, in particular, was used to ensure high protection and visual consistency. The first and the most significant step in video steganography is the frame selection. The mathematical model for this has to be considered with the following.

Figure 2 illustrates the imaginary visualization of a single frame with a stream of bits. Let \( Y(t) \) be the training video for the period \( t = 1,2,3,\ldots, T \), where \( T \) is the maximum duration of the inputted video. The input video consists of \( N \) non-overlapping frames of \( 8 \times 8 \) pixels and can be represented as seen in equation (1a).

\[
Y(t) = \{Y_1(t), Y_2(t), Y_3(t), \ldots, Y_N(t)\}, \quad (1a)
\]

\[
Y_n(t) = \begin{bmatrix}
y_n(1 \times 1) & \cdots & y_n(1 \times 8) \\
\vdots & \ddots & \vdots \\
y_n(8 \times 1) & \cdots & y_n(8 \times 8)
\end{bmatrix}, \quad (1b)
\]
\[ Y_t = y_{jk} \text{ for } 0 \leq j \leq A, \quad 0 \leq k \leq B. \]  

Equation (1b) represents a feature matrix of the \( n \)th frame of the given video \( Y(t) \). Since this proposed system uses \( 8 \times 8 \) image matrix representation, the above equations express the general representation structure of every frame. The frame selection from the set of frames in a video based on the modified entropy performs the motion vector estimation and motion compensation. Since the message embedding uses the blocks that have a motion vector, the following equations determines the magnitude of the motion vector:

\[ |M_v| = \sqrt{h^2(j) + v^2(j)}, \]  
\[ |SM_v| = \sqrt{Sh^2(j) + Sv^2(j)}, \]  
\[ Sh(j) = h(j) + \eta^h(j); \quad Sv(j) = v(j) + \eta^v(j); \quad \text{for } j = 1, 2, 3, ..., M \].

In equation (2a), the \( h(j) \) and \( v(j) \) denote the \( j \)th macroblocks' horizontal vector and vertical vector, respectively. The \( | \cdot | \) operator eliminates the negative sign and obtains the magnitude of the motion vector to make them scalar values. In view of the motion estimation algorithm, the modulation of motion vectors using to-installed data can be done in a motion estimation process. The motion vector for each macroblock applies with the LSB substitution method. Equation (2b) measures the motion vector of the stego video after embedding data, and equation (2c) evaluates the respective horizontal vector \( (Sh(j)) \) and vertical vector \( (Sv(j)) \) using the random variables \( \eta^h(j) \) and \( \eta^v(j) \), for all \( j \) from 1 to \( M \), where \( M \) is the maximum number of macroblocks. The random number variables \( \eta^h(j) \) and \( \eta^v(j) \) can take any positive or negative values of integers to model the stego frame under the constraints, as given below:

\[ P(m = 0) = 1 - P(m = A_m), \quad \text{where } P(m = A_m) = \frac{E_R}{2}. \]

Equation (3) determines the probability mass function \( (P(m)) \) of the embedding macroblock, where \( m \) represents a macroblock number among \( M \) macroblocks, that was updated with the secret bits (data). In equation (3), the variable \( A_m \) denotes the amplitude of the stego signal after the motion vector modification. The parameter \( E_R \) measures the data embedding rate. Normally \( A_m = 1 \) shows the lowest effect on compression and the highest potential of the \( E_R = 1 \) embedding.
\[ Sh(j) = h(j) + \zeta^h(j) \cdot \eta^h(j); \quad \text{for} \quad j = 1, 2, 3, \ldots, M. \]

\[ Sv(j) = v(j) + \zeta^v(j) \cdot \eta^v(j); \]

The distortion control can be assured by selecting the suitable values for control variables named selective factors \( \zeta^h(j) \) and \( \zeta^v(j) \) respective to horizontal vector and vertical vector of the \( j \)th macroblock. These selective factors can be in the range of \((0,1)\). Equation (4) denotes the modified embedding motion vectors (horizontal and vertical) derived from the following equation:

\[
\zeta^h(j) = \begin{cases} 
1; \sqrt{Sh^2(j) + Sv^2(j)} > T \quad \text{and} \quad \arctan\left(\frac{Sv^2(j)}{Sh^2(j)}\right) = \phi^A, \\
0; \quad \text{Otherwise}
\end{cases}
\]

(5a)

\[
\zeta^v(j) = \begin{cases} 
1; \sqrt{Sh^2(j) + Sv^2(j)} > T \quad \text{and} \quad \arctan\left(\frac{Sv^2(j)}{Sh^2(j)}\right) = \phi^O, \\
0; \quad \text{Otherwise}
\end{cases}
\]

(5b)

\[
\arctan\left(\frac{Sv^2(j)}{Sh^2(j)}\right) = \tan^{-1}\left(\frac{Sv^2(j)}{Sh^2(j)}\right).
\]

(5c)

Equations (5a) and (5b) help to determine the stego video motion vectors using equation (4), with the detailed constraints as expressed in each. When the stego motion vector resultant exceeds the magnitude threshold defined in \( T \), and \( \arctan\left(\frac{Sh^2(j)}{Sh^2(j)}\right) \) equals to the acute angle \( \phi^A \), the selective variable \( \zeta^h(j) \) takes the value 1, and 0 for otherwise. Similarly, when the stego motion vector resultant exceeds the magnitude expressed in \( T \), and \( \arctan\left(\frac{Sh^2(j)}{Sh^2(j)}\right) \) equals to the obtuse angle \( \phi^O \), the selective variable \( \zeta^v(j) \) takes value 1, and 0 for otherwise. The \( \arctan\left(\frac{Sh^2(j)}{Sh^2(j)}\right) \) determines the inverse tan function of the resultant \( \frac{Sv^2(j)}{Sh^2(j)} \), as expressed in equation (5c).

\[
e' = - \sum_{j=1}^{M} P(j) \cdot \log2(P(j)) + P(j) \cdot \exp(1 - P(j)).
\]

(6)

After selecting the stego frames, entropy updation has been performed, and the modified entropy \( e' \) can be expressed, as seen in equation (6). The \( P(j) \) determines the probability of choosing the \( j \)th macroblock (detectable macroblock), whereas \( (1 - P(j)) \) gives the non-detectable macroblock. These macroblocks within the selected frames have been applied with LSB substitution method.

The following describes the general LSB substitution method.

Figure 3 gives the example of LSB substitution approach for video steganography. The secret message (text, audio, video, or image) to be transmitted has been converted into bitstreams of 8-bit blocks, as shown in the input section of Figure 3. Similarly, the cover video has to be converted to bitstreams (8-bit blocks) and then identifies the least significant bit among each block that refers to the motion vectors. The above examples differentiate each data blocks using four different colors. The blue color represents the LSB from the motion vector block that can be modified. With this knowledge and the secret keys among senders and receivers, the message can be easily extracted without losing information and in a secure manner.

\[
SM = \{sm_1, sm_2, sm_3, \ldots, sm_n\}; \quad sm_j \in \{0, 1, \ldots, 2^{b-1}\} \quad \text{and} \quad j = \{1, 2, 3, \ldots, n\},
\]

(7a)

\[
sm_j' = \sum_{i=0}^{b-1} (sm_{j+b-i} \cdot 2^{b-1-i}); \quad SM' = \{sm_j'\},
\]

(7b)

Assume, the hidden \( n \)-bit message \( SM \) has been embedded into the \( b \)-rightmost LSBs of the cover-image \( M \), and the \( SM \) is revamped to get a \( b \)-bit virtual \( SM' \) image conceptually, as shown in the above equation.
Equations (7a) and (7b) represent the secret message embedded using LSB substitution and the embedded bits in the macroblock of each key frame with embedded data by motion vectors of the cover video. The SM represents the secret message with substreams of \( \{ sm_1, sm_2, sm_3, \ldots, sm_n \} \) in which each \( sm \) lies in the value range of \([0,1, \ldots, 2^{b-1}]\). Equation (7b) gives the embedded bits in the secret message \( SM' = \{ sm'_1 \} \) in which \( b \) denotes the total number of bits in the secret message.

### 3.2 ABC algorithm in the proposed video steganography

The proposed MVS-ABC system uses the ABC algorithm for improved block assignment for video steganography applied with LSB substitution method. The ordinary ABC algorithm is structured to solve computational problems of optimization. In MVS-ABC, the algorithm uses the general equations to find the food source and the scout bee position to develop a solution in the continuous area of operations. These equations’ solutions are discrete, so they cannot be used to directly solve problems using combinatorial optimization. Even so, with necessary changes, the algorithm can be extended for this problem type. The MVS-ABC sets up the block assignment list to find the macroblock for secret message embedding as a food source in the general ABC algorithm. Every food source dimension (feasible solution space) is the assignment of a secret image to a host image by one block. Each block will appear only once in a block assignment list. Each source of food thus has \( m \) dimensions that fit \( m \) blocks for the \( m \) block of the hidden image. Before explaining the improved ABC model, it is to be needed to brief the general ABC algorithm.

\[
F_{SP} = OF(y_j); y_j \in \mathbb{R}^D \quad \text{and} \quad j = \{1, 2, 3, \ldots, N_F\}.
\] (8)

Equation (8) represents the general process of the ABC algorithm that generates the randomly distributed food source, where the \( F_{SP} \) represents the food source position distribution process, \( OF(y_j) \) denotes the objective function of determining how perfect the combinatorial solution concerning to \( y_j \) is. The variable \( y_j \) gives the \( D \)-dimensional vector for the food source position for all number of food sources from 1 to \( N_F \). The colony is often broken into bees and bee types within the ABC algorithm. Every other search space solution consists of a set of parameters for optimization that further indicate a food source position. The number of bees working equals the number of sources of food. For each food supply, there is one bee working.

\[
P_{jk} = y_{jk} + \delta_{jk}(y_{jk} - y_{l_k}),
\] (9a)

\[
p(j) = \frac{F_j}{\sum_{m=1}^{N_F} F_m}.
\] (9b)

The population is exposed to repetitive loops of improving feasible solutions, preferring feasible solutions, and rejecting unsuccessful solutions after initialization. All working bees pick a new position for the
food source to upgrade viable solutions. The preference depends on the region of the food source chosen. The updated food source position can be expressed as equation (9a). The variable $P_{jk}$ denotes the updated feasible solution which changes by comparison with the randomly chosen location of its neighboring solution $y_k$ with its previous solution value $y_{jk}$. Equation (9b) gives the probability of selecting a proposed source of food, where $F_j$ denotes the fitness value for $j$th region.

The random integer $\partial_{jk}$ takes the value in the interval $[-1,1]$ for adjusting the preceding solution at the next iteration, to transform it into a new solution. The indexes $i$, $j$, and $k$ must follow the constraints as follows: $i \epsilon \{1, 2, 3, \ldots, N_f\}$, $j \epsilon \{1, 2, 3, \ldots, D\}$, and $k \epsilon \{1, 2, 3, \ldots, D\}$. The difference in position $(y_{jk} - y_k)$ gives the motion object identification. If the new position has a better fitness score, the old food source position in the bee’s memory has been substituted by the new candidate food source, as shown in Figure 4. Working bees can further move back to their hive and exchange their fresh food sources’ fitness value with the spectators. Each bee selects one of the food sources according to the importance of fitness from the bees used in the next step.

$$y_{jk} = y^\text{min}_j + r[0, 1] \times (y^\text{max}_j - y^\text{min}_j).$$  \hspace{1cm} (10)

If the food source’s fitness value increases, the likelihood of the food source being preferred by the viewer’s increases. Following the food source selection, onlooker bees will visit the food supply chosen and pick a new applicant food source location in its vicinity. Equation (7a) expresses and quantifies the latest candidate food source using human-centered AI for web applications. The third move is to discontinue and replace every food source position not increasing the fitness benefit with a new position arbitrarily decided by a scout bee. The above allows discouraging inadequate remedies. Equation (10) determines the current random location selected by the scout bee. In the above equation, $y^\text{min}_j$ and $y^\text{max}_j$ represent the lower and upper bounds of the food source position or feasible solution space in $j$th dimension, correspondingly. The function $r[0,1]$ is the random number generator and generates the value between 0 and 1. For the regulation of the number of iterations, the maximum number of cycles is used, and a termination criterion. This method has been replicated till its number of iterations reaches the maximum number of cycles. As noted previously, it retains a list of blocks allocated to the ABC algorithm as a food source. Figure 5a provides an example of an 8-block message embedded frames’ block assignment list with a 16-block cover frame.

Figure 5a provides an example of an 8-block message embedded frames’ block assignment list with a 16-block cover frame. SB0, SB1, ..., SB7 represents the 8-block secret message whereas CB0, CB1, ..., CB15 denotes the 16-block cover frame selected based on the motion vectors. Figure 5b illustrates the cover frame with embedded secret message blocks. Therefore, Figure 5b gives the feasible solution for selecting the stego frame construction similar to getting the best position for the food source within the bee colony.

$$N_{CB} = \frac{CB_{\text{len}}}{l}; \quad N_{SB} = \frac{SB_{\text{len}}}{l}.$$  \hspace{1cm} (11)
The number of blocks for the cover frame and secret message can be determined using equation (11). 

\[ N_{CB} \] gives the number blocks for the cover frame array and \( N_{SB} \) is the total number of blocks in the array of secret message segments to be added in the selected cover frame. The \( C_{Blen} \) and \( S_{Blen} \) are the length of the cover block array and the secret block array, respectively. The parameter \( l \) is the block length.

Figure 6 shows the sports evaluation using a web application. Athletes are required to pick a learning goal from the library’s reference collection. Finally, they choose a standard test of a certain complexity. The difficulty level of all workouts is set to zero, and their skill levels are pre-established. An athlete’s skill level, goal level, content, and indicator relevance are all factored into the system’s ability to choose the best workout group and degree of difficulty for them. The system will identify fresh exercise content for athletes if the real-time performance of athletes under a specific exercise content achieves expectations after updating the difficulty progress and associated skill indicators in the exercise \( Z(m)X(n) \) group’s exercise content.

\[
Z(m)X(n) = P^{(m)} - (z^{(k)} + y^{R}R^{(k)}) - R_{d}(e)R_{k}(g) + (y^{(m)} + z^{S}g^{(m)}).
\]  

(12)

Using a human-centered AI in a web 4.0 system, \( P^{(m)} \), a command will be selected based on \( z^{(k)} \), its likeness to the present real-time performance \( y^{R}R^{(k)} \). And athletes will be given feedback \( R_{d}(e) \) based on the command’s influence \( e \) on their performance. Interactive training ends \( (g) \) when athletes’ skill levels are not significantly different from the reference levels, the athletes can \( y^{(m)} \) calibrate different skill indicator values using standard tests \( z^{S}g^{(m)} \) if they believe that the system updates their skill levels inaccurately or creates a
change in skill levels during self-practice using equation (12). The system will then proceed with interactive training depending on the new skill levels discovered.

### 3.3 Secret message extraction

The secret message can be retrieved through the communication channel at the receiver’s side. On the receiver side, the received stego video has been classified into a set of frames and then identifies the keyframes based on the motion vectors. Many machine learning and deep learning algorithms can be applied for stego frames identification. For each selected frame, the macroblocks of the hidden message co-ordinates are indicated by reference to the motion analysis of the database. The block is then transformed by $8 \times 8$ wavelet transformation for each selected block. Certain rules can be used to extract the message. Every first pair of factor coefficients is smaller than the second, the message bit is set to 1. Since the transmitting stego video has been encrypted with various encryption technologies, the respective decryption technologies can be used for decrypting the stego video and then message bits are rearranged to get the message at the receiver.

\[
\begin{align*}
    y'_l &= y_l - y_{lj} \mod 2^b + sm'_l, \quad (13a) \\
    sm'_l &= y'_l \mod 2^b. \quad (13b)
\end{align*}
\]

The $n'$ pixel subset \{y_{l1}, y_{l2}, ..., y_{ln'}\} is selected in a predefined sequence from the cover image $M$. Since the process of the secret message embedding has been successfully completed with a $b$-bit LSB replacement of $y_l$ by $sm'_l$. In algebra, the $y_l$ value of the pixel chosen is modified to form the stego-pixel $y'_l$ for storing the $b$-bit message $sm'_l$, and shown in equation (13a). The embedded messages can easily be extracted in the extraction process, given the $SM$ stego frame, without reference to the original cover. The pixels that store

\[\]

Figure 6: Sports evaluation using web application.
the bits of the secret message are chosen from the stego frame with the same sequence as in the embedding process. For secret bits of the message, the b-bit LSBs of the selected pixels are extracted and lined up. The embedded message bits can be retrieved mathematically by equation (13b). The modulus operator returns the LSB of the respective macroblock in the stego frame and retrieves the original message.

4 Experimental analysis and findings

This section evaluates the proposed MVS-ABC model on video steganography compared with some existing models included in the literature survey (Section 2). For the better understanding, approach/model referenced in a previous study [18] was named VST-PA, in another study [19] as MVC-MVS, and in the study [20] as SVS-IPM. Another existing model IWT-SVD [21] was kept as such. The experiment was implemented in the software MATLAB 2020a that run on a PC with a 10th Generation Intel®Core™ i5-10300H (8MB Cache, up to 4.5 GHz, four cores) Windows 10 NVIDIA® GeForce GTX® 1650 Ti 4GB GDDR6 with memory 8GB, onboard, DDR4, 2933 MHz. The number of iterations (minimum number of cycles) was set to 25. For the purpose, the number of employed bees, and the number of onlooker bees for the proposed ABC system, was set to 12. The entire model was assessed based on the performance of the video steganography and its perfectness. Each evaluation parameter was analyzed for two cover videos from two data repositories; they are golf.avi from HMDB dataset [28] and football.avi from Video Trace Library [29] and two different messages were embedded (Message 1 [64-bit] and Message 2 [128-bit]). The parameters used for the evaluation are as follows.

4.1 Secret message embedding capacity

This parameter identifies the message embedding capacity (MEC) in each stego frames in the cover video. The MEC can be determined by computing the embedded information bits per frame. The overall embedding capacity of the stego video was evaluated by taking the average, as expressed in the following equation.

$$\text{MEC} = \left( \frac{N_b}{S_c} + T_f \right) \cdot 100.$$  (14)

Equation (14) computes the MEC for each stego frames, where $N_b$ and $S_c$ represent the number of bits in the secret message and the size of the cover frame or stego frame is selected based on motion vectors. The $T_f$ measures the total number of stego frames selected. Figure 7 illustrates how well the proposed approach performs for the selected sports videos compared to the existing models.

Figure 7(a) and (b) shows that the football.avi can embed 1 bit per pixel with 100% accuracy for the proposed MVS-ABC framework and human-centered AI for web applications for message 1 and message 2, respectively. In contrast, the least capacity model for video steganography using sports training video was observed as VST-PA. Since the football training video encloses more number of motion vectors, this steganography approach performed well in this dataset. And for both the message, the proposed MVS-ABC framework resulted in significant performance.

4.2 Peak signal-to-noise ratio (PSNR)

PSNR determines how much distortion has occurred for the stego frame by computing the invisibility of the stego video. The PSNR with 50 and above were considered excellent for steganography performance. The following equations help to achieve the results.
\[
\text{PSNR} = 20 \log_{10} \left( \frac{255}{\text{RMSE}} \right), \quad (15a)
\]

\[
\text{Root mean square error (RMSE)} = \sqrt{\frac{1}{A \times B} \sum_{j=1}^{A} \sum_{k=1}^{B} (P_{\text{cf}}(j, k) - P_{\text{sf}}(j, k))^2}. \quad (15b)
\]

Equation \((15a)\) gives the invisibility in the video steganography using the proposed MVS-ABC technique, where the RMSE can be determined by equation \((15b)\). The \(P_{\text{cf}}\) and \(P_{\text{sf}}\) are the pixel intensities of cover frame and the stego frame, respectively. The frame dimension can be represented by \(A \times B\), where \(A\) denotes the rows and \(B\) gives the columns.

The PSNR for both the messages in the preceding description was evaluated, and Figure 8a gives the same for message 1 of 64 bit, and Figure 8b for the second message of 128 bit. The experimental analysis resulted in the highest PSNR for the MVS-ABC of 54 for football.avi while embedding the message 1, whereas the lowest PSNR was observed with 49.7 in football.avi for the same message embedding by VST-PA. In contrast, the highest PSNR was observed for the MVS-ABC (53.25) in golf.avi and the least value
of 50.37 in golf.avi for VST-PA while embedding message 2. The obtained results of PSNR in message embedding can assure the better performance of the proposed MVS-ABC in video steganography using moving vectors of the sports training video.

4.3 Similarity index using normalized cross-correlation (NCC)

The similarity index using NCC determines the closeness between the cover frame and the stego frame. The overall closeness among multiple frames was measured by taking the average of the NCC obtained by each frame evaluation.

\[
NCC = \frac{\sum_{j=1}^{A}\sum_{k=1}^{B} P_{c}(j, k) \times P_{s}(j, k)}{P_{c}(j, k)^2}. \tag{16}
\]

Equation (16) computes the NCC for individual stego frame similarity with their respective cover frame. Using the result obtained, the average NCC can be evaluated.
The above chart compares the results of various video steganography models, along with the proposed framework. Figure 9 illustrates the PSNR value per frame for three distinct video hosts is compared here and Figure 8b includes the results of NCC for hiding message 2 within the sports training video (cover video). Here the golf.avi gives the highest value of NCC, which always lies between −1 and 1. The NCC for MVS-ABC was the highest with 1 compared with other video steganography models.

![Chart comparing PSNR values for different video steganography models](chart.png)

**Figure 9:** The PSNR value per frame for three distinct video hosts is compared here.

The above chart compares the results of various video steganography models, along with the proposed framework. Figure 9 illustrates the PSNR value per frame for three distinct video hosts is compared here and Figure 8b includes the results of NCC for hiding message 2 within the sports training video (cover video). Here the golf.avi gives the highest value of NCC, which always lies between −1 and 1. The NCC for MVS-ABC was the highest with 1 compared with other video steganography models.

![Chart comparing NCC values for different video steganography models](chart_ncc.png)

**Figure 10:** NAE for (a) Message 1 and (b) Message 2.
4.4 Error rate

This section evaluates the error rate using normalized absolute error (NAE), one of the most common error rate evaluation parameters. The normalized absolute error can be defined as the absolute error rate between the cover frame and the stego frame where the data get normalized related to the stego frame.

\[
\text{NAE} = \frac{\sum_{j=1}^{A} \sum_{k=1}^{B} P_{c}(j, k) \times P_{s}(j, k)}{P_{s}(j, k)^{2}}.
\]

Equation (17) determines the performance of the proposed framework with the least error rate compared with the existing video steganography models. The following figures mark the observations.

Figure 10a and b illustrates the achieved error rates (NAE) for the transmitted secret message for 64-bit information and 128-bit information, respectively. The above-plotted graph gives that the proposed MVS-ABC framework for video steganography with the golf.avi scores better than that with football.avi. The lowest error rate measured was about 0.000073, which was for the golf.avi.

Figure 11 shows the accuracy ratio. Sports recognition systems can be evaluated using several datasets. The analysis is built on two pillars: uncertainty and perceived consistency. Specific information on the fluctuating security price is provided for each division identification. The percentage of test data components were properly identified from all test data elements. Because of the precise depth and body data,
human-centered AI for web applications can effectively measure sports activity behaviors. Finger and back pull identification is more accurate than left and right waves, and motion detection accuracy is the same. The accuracy ratio is given in equation (18).

\[ G_s - F_s = \psi(X_sV_{sf} - G_{s+1}V_{fh} + d_h) - \psi(XV_{sf} + g_h). \]  

(18)

As shown in equation (18), \( X_f \) found that users are more inclined to \( G_{s+1} \) identify vertical vibrations than \( V_{sf} \) horizontal ones. A simple detection \( d_h \) and tracking result directly influence more complicated \( g_h \) gesture recognition. There have been several settings and \( \psi \) locations where sports performed demanding motions in this investigation. The suggested approach, which uses this article’s motion vector steganography, has an average accuracy of 97.5%.

Finally, all the above parameters promised the proposed video steganography’s applicability for real-time secret communication with highest performance compared to the existing models. Compared to current techniques, the suggested methodology has more integration potential and approximates the same PSNR. The LSB will cover the details of the matches that are not identified. The above findings from the approach suggest that the MVS-ABC method provides high-quality stego frames and retrieves the hidden message even though stego frames are invaded with noise.

5 Conclusion and future scope

The research introduced integrative sports training footage from the ABC algorithm (MVS-ABC). In the sports simulation video bitstreams, the motion vector stenography detected secret information from the motion vectors. The ABC optimized the block assignment for a secret message to be inserted into a hosted video, where the block assignment was treated as a problem with the combinatorial optimization. A virtual training environment can be created using human-centered AI for web applications, which many athletes use regularly. Our study directly influences people’s daily lives because of advances in AI. The experimental study evaluated the data embedding outputs compared to the current using steganographic technology and the error rate using ABC algorithms instead of other genetic algorithms. The MVS-ABC was the best with the highest embedding capacity and least error rate. In the future, this study plans to integrate the Convolution Neural Network prediction model for predicting the stego frames at the earliest.

Author contributions: Conception and design of study: Wen Fu, Jinmao Tong; Acquisition of data: Wen Fu; Analysis and/or interpretation of data: Wen Fu, Zhongwang Cao.

Conflict of interest: The authors declare that there is no conflict of interest in this article.

Data availability statement: The authors declare that there is no data availability statement in this article.

References


Shankar A, Jaisankar N. Dynamicity of the scout bee phase for an artificial bee colony for optimized cluster head and network parameters for energy efficient sensor routing. Simulation. 2018;94(9):835–47.


