Co-Design Secure Control Based on Image Attack Detection and Data Compensation for Networked Visual Control Systems

Dajun Du, Lang Wu, Changda Zhang, Zixiang Fei, Lisi Yang, Minrui Fei, and Huiyu Zhou

***Abstract*—The incomplete and untrue data caused by cyber attacks (e.g., image information leakage and tampering) will affect control performance and even lead to system instability. To address this problem, a novel co-design secure control method based on image attack detection and data compensation for networked visual control systems (NVCSs) is proposed. Firstly, the existing problems of NVCSs under image attacks are analyzed, and a co-design secure control method including image encryption, watermarking-based attack detection and online data compensation is presented. Then, a detector based on double-layer detection mechanism of timeout and digital watermarking is designed for real-time, integrity and authenticity discrimination of the image. Furthermore, according to the detection results, an online compensation scheme based on cubic spline interpolation and post-prediction update is proposed to reduce the effect of cumulative errors and improve control performance. Finally, the online compensation scheme is optimized by considering the characters of networked inverted pendulum visual control systems, and experimental results demonstrate the feasibility and effectiveness of the proposed detection and control method.**

***Index Terms*—Networked visual control systems, cyber attacks, attack detection, online data compensation, secure control.**

I. Introduction

Vision-based instrumentation and measurement systems [1], [2] have been applied for acquiring and processing signals, also known as non-contact, non-invasive, or non-destructive inspection. The technologies of visual sensing and image processing are widely used in variety of industrial automation fields, e.g. [3]–[5], robot control, unmanned driving and unmanned aircraft control. This leads to the rapid development of networked visual control systems (NVCSs) [6]. While bringing convenience, vision-based measurement faces large challenges of image security (i.e., preservation of image information). Since a large number of images need to be transmitted in real time, data integrity and authenticity cannot be guaranteed especially under an open network environment, facing the risk of image leakage, tampering, and so forth. It in turn leads to the incomplete and untrue image data, and ultimately affects the stability of NVCSs [7], [8].

Such incomplete and untrue data mainly derive from two categories: 1) network inherent factors such as data packet losses, network-induced delay and so on [9]–[12]; 2) cyber attacks [13]–[17] such as denial-of-service (DoS) attacks, crop attacks and noise attacks, etc., leading to information forgery or even loss [18]–[21]. These factors pose huge challenges to security control of NVCSs.

These problems have stimulated some research works by considering incomplete and untrue data caused by cyber attacks. For example, DoS attacks decline system performance by blocking data transmission [22], and a compensation mechanism using the latest received data packets is designed to alleviate the influence of DoS attacks [23]. Deception attacks such as replay attacks and false data injection attacks (FDIAs) destroy data authenticity [24], and a distributed observer combined with attack detection algorithm is designed to resist random or intermittent replay attacks [25]. To reduce the oscillation caused by FDIAs, a terminal integral adaptive sliding mode control algorithm using the estimation error as adaptive factor is proposed [26]. Moreover, more different methods on attack detection, state estimation and security control under cyber attacks are summarized in [27], [28]. However, these researches have not considered cyber attacks against the images.

In NVCSs, the mechanism of image attacks is more complex than non-image attacks, because image attacks will damage the quality of the transmitted images, and lead to being unable to extract complete and true state information. To explore secure control methods of NVCSs under image attacks, the existing researches are basically aimed at image information leakage and tampering. To protect the security of the images, chaos theory is employed to design some image encryption techniques based on image pixels [29], [30]. Furthermore, by using cyclic generation of confrontation network, the encryption and decryption of images are achieved based on deep learning [31]. To detect image attacks, the fragile image watermarking methods [32]–[34] are proposed to find image tampering and its locations. Some approaches based on image attacks detection are studied. For example, an image restoration method is presented by integrating non-local self-similarity and global structure sparsity [35]. A diagonal mapping algorithm is proposed to guarantee the image tampered content recovery [36], and a flexible deep framework is proposed based on discriminant convolutional neural network for various image

The work of D. Du, L. Wu, C. Zhang, Z. Fei, L. Yang, and M. Fei was supported in part by the National Science Foundation of China under Grant Nos. 92067106, 61773253, 61803252, and 61833011, 111 Project under Grant No. D18003, and Project of Science and Technology Commission of Shanghai Municipality under Grant Nos. 20JC1414000, 19510750300, 21190780300. (Corresponding Author: Lang Wu; Changda Zhang).

D. Du, L. Wu, C. Zhang, Z. Fei, L. Yang, and M. Fei are with Shanghai Key Laboratory of Power Station Automation Technology, School of Mechatronic Engineering and Automation, Shanghai University, Shanghai 200444, China (e-mail: ddj@i.shu.edu.cn; m71997@shu.edu.cn; changdazhang@shu.edu.cn; zxfei@shu.edu.cn; lisiyangfive@163.com; mrfei@staff.shu.edu.cn).

H. Zhou is with School of Computing and Mathematical Sciences, University of Leicester, Leicester LE1 7RH, U.K. (e-mail: hz143@leicester.ac.uk).

restoration tasks [37].

However, these existing methods are difficult to simultaneously satisfy real-time and image security requirements in NVCSs. There are the following challenging problems and difficulties:

(1) The existing image security algorithms have usually high complexity and relatively low efficiency, leading to their inapplicability to high real-time environments. How to design high efficiency image security algorithm for NVCSs is the first challenge.
(2) Image security approach can only achieve the corresponding information confidentiality functions, but how to design image security detection is the second difficulty.
(3) Most existing studies have been performed based on known assumption of data loss or tampering characteristic, but the assumption that the attacks obey a specific probability distribution is not always consistent with actual diverse and random image attacks. How to further guarantee stability of NVCSs based on attack detection is the third challenge.

Therefore, considering real-time and accuracy requirements of NVCSs under image attacks, a novel co-design secure control method is proposed. The main contributions of this paper are summarized as follows:

(1) The existing problems of traditional NVCSs under image attacks are analyzed, and the overall framework of co-design secure control method including image encryption, watermarking-based attack detection and online data compensation is elaborated.
(2) A detector based on double-layer detection mechanism of timeout strategy and digital watermarking is designed for real-time and integrity discrimination of the images, which can judge wether the image is valid and provide the detection result by a trigger signal.
(3) Considering detection result and cumulative errors, a cubic spline interpolation online compensation scheme based on post-prediction update is proposed to improve control performance of NVCSs.

The remainder of this paper is organized as follows. In Section II, we have analysed vision-based measurement of NVCSs and the necessity of new secure control method. In Section III, co-design secure control method for NVCSs is proposed, including double-layer detection and online compensation method. Section IV optimizes the online compensation scheme, and discusses different experimental results for the NVCSs. The conclusions and future researches are given in Section V.

II. Problem Formulation

A. Vision-based Measurement of NVCSs

Vision-based measurement of NVCSs are shown in Fig. 1, which includes the controlled plant, industrial camera (i.e., visual sensor), remote control terminal (with image information extraction unit and controller) and the actuator. Firstly, the real-time images of the controlled plant are captured by industrial cameras, which are then transmitted to remote control terminal via network. Then, states x_k of the controlled

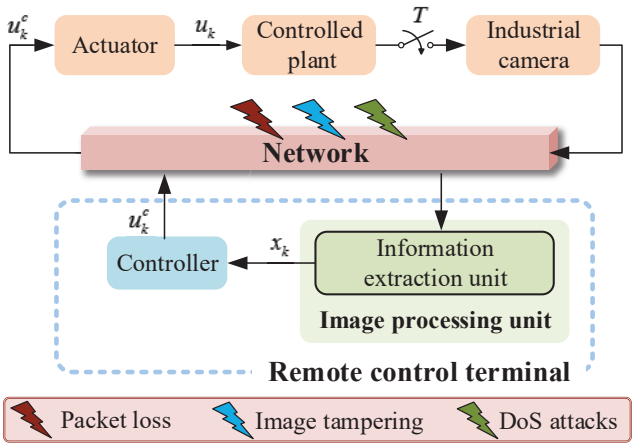


Fig. 1. Vision-based measurement of NVCSs.

plant are extracted from the received images in the image information extraction unit. Furthermore, according to x_k, the control signals u_k^c will be calculated in the controller and transmitted to the actuator via network. Finally, the actuator derives the controlled plant to keep stability.

To well construct the above NVCSs, three basics of vision-based measurement [2] is discussed by taking the networked inverted pendulum visual control system [38] as example:

1) *Visual sensors:* There are three main aspects on visual sensors: i) The visual sensor can be a visible-light camera, an infrared camera, a laser scanner, an x-ray scanner, or any other sensor, e.g., a visible-light charge-coupled device (CCD) Aca640-120gm monochrome industrial camera with the adjusted frame rate is used in our experiments. ii) The lighting conditions and the parallelism of the planes of the camera and the controlled plant directly affect the quality of the captured image. To solve these two problems, the five light emitting diode (LED) stroboscopic fluorescent lamps with adjustable illumination levels as the light sources and a level meter are used in our experiments. iii) After having installed the industrial camera, an appropriate sampling mechanism should be set, e.g., a time-triggered sampling mechanism is designed to improve the efficiency of image processing and the real-time performance of the system.
2) *Image pre-processing:* The image acquired by the visual sensor could have deficiencies such as noise and data redundancy [2]. For instance, to reduce data redundancy, the regions containing the cart and the pendulum motion areas are found from the original image, and then the standard Canny edge detection algorithm is used to detect the edge information of the cart and the pendulum.
3) *Image analysis:* The purpose of this stage is to analyze the image and extract the necessary information in remote control terminal. For instance, the cart position is obtained by applying the translation between the pixel coordinate system and the world coordinate system since the camera and the moving plane are fixed, and the pendulum angle is measured in the pixel coordinate

system by using the edge detection technique based on the Hough Transformation.

For now, with the well-established vision-based measurement, to achieve the integrity and authenticity of the state information, secure transmission of images has become the primary issue of NVCSs.

B. Problem Analysis of Image Attacks Detection

The states x_k depends importantly on the images, when the images suffers from cyber attacks, the integrity and authenticity of the image are destroyed, leading to incomplete and untrue state and control information. The specific reasons are listed as follows:

1) It is difficult to accurately detect the types and areas of attacks in real time from the transmitted images. Image attacks usually include copy-move, slicing, noise attacks, geometric attacks, etc., which need to be identified through image target detection or forensics. However, for control systems with high real-time requirements, the large amount of image data will bring a great image processing delay.
2) It is tough to describe attacks quantitatively from the extracted information. Due to the characteristic of image attacks, the impact of image tampering on the extracted information is indirect, which different from direct attacks on traditional sensing acquisition data. The extracted information from image may involve only part of the area, and the impact of the attack on the extracted information cannot be estimated.
3) NVCSs take data availability as the highest priority. Assuming that image attacks conform to a specific emergence pattern does not satisfy the real random attack situation. Unilateral consideration from attack detection or data compensation cannot guarantee the stability of the system under cyber attacks.

Remark 1: Although accurate detection of image attacks can identify attacks, the system cannot afford the delay cost. Even if the type and area of the attack are detected, the impact of the attack on the extracted information is difficult to quantify. Assuming that the attacked image is directly discarded, there will be greater pressure on the system information compensation, which is only suitable for a good network environment. Therefore, it is necessary to weigh the attack recognition accuracy and recognition speed. Moreover, to guarantee secure control of NVCSs, attack detection and data compensation need to be co-designed urgently.

III. Co-Design Secure Control Method for NVCSs

The above has presented the framework of traditional NVCSs and analysed the corresponding drawbacks. To cope with the drawbacks, a new co-design secure control method for NVCSs is fully designed.

A. Framework of Co-Design Secure Control

To achieve secure control of NVCSs under image attacks, an co-design security control method is proposed, and the corresponding framework is shown in Fig. 2. Unlike traditional NVCSs shown in Fig. 1, extra four units are added, i.e., image pre-processing, attack detector, data compensation and buffer units. These four units are analysed as follows:

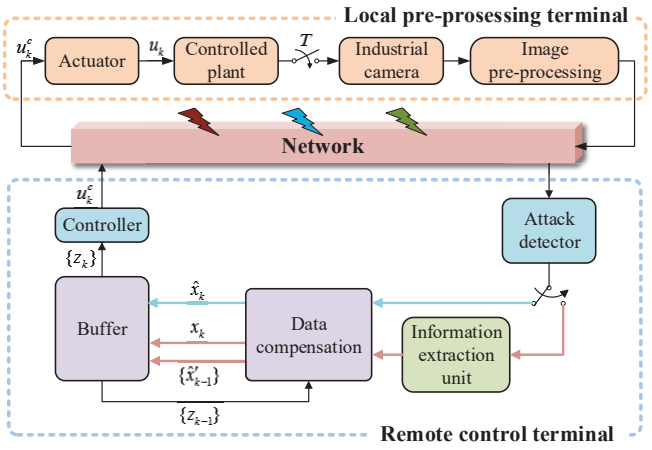


Fig. 2. Framework of co-design secure control method for NVCSs.

1) *Image pre-processing:* The controlled plant is sampled periodically by industrial camera, and these captured real-time images are preprocessed by watermarking encryption. They are then transmitted to remote control terminal via network.
2) *Attack detector:* Since the image may be lost or corrupted from attacks, a image detector is deployed in remote control terminal. When the image arrives at attack detector beyond the maximum allowable time, it will be discarded, going directly to data compensation unit. Otherwise, the image tampering detection is performed, and the whole process will be analysed by the following Section III. B (Double-level attack detector).
3) *Data compensation:* When the image is judged as "invalid" by attack detector, the lost state information will be online compensated by the following Section III. C (Online compensation based on cubic spline interpolation).
4) *Buffer:* The states x_k of the controlled plant are extracted and stored in the buffer to support controller design and data compensation.

For convenience, a flag γ_k is used to indicate whether data transmission is normal. $\gamma_k = 1$ represents the normal transmission of data, which means that true state information can be extracted from the received image. $\gamma_k = 0$ represents the invalidity of the data, which means that the predicted value should be used to compensate. Then the controller input signal can be expressed as

$$z_k = \gamma_k x_k + (1 - \gamma_k)\hat{x}_k, \tag{1}$$

where x_k represents true state, $\hat{x}_k$ represents the predicted value of x_k, and z_k represents the controller input.

Remark 2: In comparison with the traditional NVCSs, the induced four new units can achieve image encryption, watermarking-based attack detection and co-design of online data compensation with buffer.

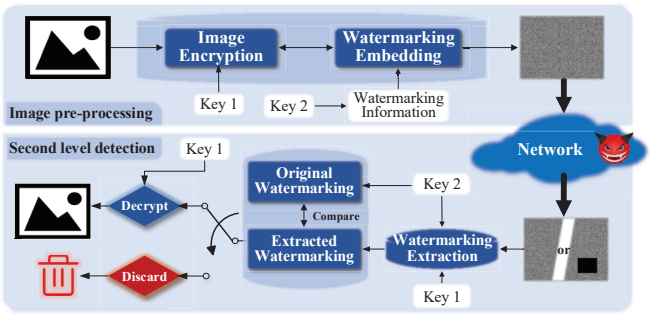


Fig. 3. Image security detection method based on image encryption and fragile watermarking.

B. Double-level Attack Detector for the Image

Let us begin achieving co-design secure control from attack detector. To make real-time judgement for image integrity and authenticity, a double-level attack detector is shown in Section I. A of the supplementary materials, which will be detailed presented in the following.

1) First-Level Detection, i.e., Time-Based Detection: For real-time availability of the images, time-based policy is treated as first level detection. According to the time from the timer, if the image reaches attack detector timeout, it will be judged as "invalid" and then directly enter the data compensation unit, otherwise it is "valid" image. Moreover, the timer will be reset after each image arrival or time-out.

2) Second-Level Detection, i.e., Image Security Detection: To guarantee the security of images and detect image attacks, some advanced image encryption and watermarking algorithms have been proposed. However, these approaches mainly solve the security problems of still images, which usually have high computational complexity.

Remark 3: To verify whether existing advanced encryption or watermarking approaches meet the real-time requirement of NVCSs, they have operated on networked inverted pendulum visual control system (NIPVCS) experimental platform [38]. As analyzed in Section I. B of the supplementary materials, it is revealed that they directly destroy the system stability due to high time.

To satisfy the real-time and detection requirements, a new image security detection method as shown in Fig. 3 is designed as the second level detection. Unlike the watermarking-based security detection method directly based on state information [39], it combines image encryption with random fragile watermarking to reduce both the sacrifice of image information and computational complexity of the algorithm.

At image pre-processing unit of the local pre-processing terminal, two pseudo-random sequences are generated firstly by iterating the logistic mapping:

$$\lambda_{n+1}^{en} = \mu^{en}\lambda_n^{en}(1-\lambda_n^{en}), \lambda_{n+1}^{wm} = \mu^{wm}\lambda_n^{wm}(1-\lambda_n^{wm}), \quad (2)$$

where $\lambda_n^{en}, \lambda_n^{wm} \in (0,1)$, $\mu^{en}, \mu^{wm} \in (0,4)$ and $n \in \mathbb{Z}$. The relevant parameters of the logistic mapping are set as Key 1 (*i.e.*, λ_0^{en}, μ^{en}) and Key 2 (*i.e.*, λ_0^{wm}, μ^{wm}).

Then, the original image I_o is encrypted by Key 1, where the pixel of image at the i^{th} row and j^{th} column is denoted by $P_{i,j}$. The specific encryption steps are as follows:

i) *Keys updating:* Getting the image time-stamp T_n, offset the Keys 1 and 2 by considering T_n as a disturbance:

$$\begin{aligned} T_n &\leftarrow mod(T_n, 1000)/1000 \\ \lambda_0^{en} &\leftarrow |\lambda_0^{en} - T_n|, \lambda_0^{wm} \leftarrow |\lambda_0^{wm} - T_n| \\ \mu^{en} &\leftarrow 3.9 + |\mu^{en} - 3.9 - T_n| \\ \mu^{wm} &\leftarrow 3.9 + |\mu^{wm} - 3.9 - T_n| \end{aligned} \quad (3)$$

where "$\leftarrow$" represents the assignment operation.

ii) *Recorders initialization:* Get the row number N and column number M of the image to ensure that each row and column is encrypted only once. The un-encrypted number of rows and columns are set as $\mathcal{N}_{row} = N, \mathcal{N}_{col} = M$, the row and column recorder of the image are set as $\mathcal{R}[i] = i, i = 0, \ldots, N-1$ and $\mathcal{C}[j] = j, j = 0, \ldots, M-1$ respectively.

iii) *Rows and columns selection:* Firstly, set $\epsilon_{row}^{en}[l]$ and $\epsilon_{col}^{en}[l]$, $l = 0, 1, 2, 3$ to record the iteration values for Logistic map of row and column selection. Then, iterate (2a) once to get $\epsilon_{row}^{en}[l]$ by Key 1, and the row number $\mathcal{P}_{row}[l]$ or $\mathcal{S}_{row}[l]$ in row recorder is selected by (4a) and (4b). Moreover, overwriting the selected sequence number with the last unselected sequence number stored in the recorder and update $\mathcal{N}_{row}$ by (4c) and (4d). After that, the un-encrypted row numbers are stored in the first $\mathcal{N}_{row}$ positions of the recorder. Finally, repeat these steps four times to select four rows.

$$\begin{aligned} &\mathcal{P}_{row}[l] \leftarrow \lfloor \epsilon_{row}^{en}[l]\mathcal{N}_{row} \rfloor, \mathcal{S}_{row}[l] \leftarrow \mathcal{R}[\mathcal{P}_{row}[l]], \\ &\mathcal{R}[\mathcal{P}_{row}[l]] \leftarrow \mathcal{R}[\mathcal{N}_{row}], \mathcal{N}_{row} \leftarrow \mathcal{N}_{row} - 1. \end{aligned} \quad (4)$$

So does columns selection, *i.e.*,

$$\begin{aligned} &\mathcal{P}_{col}[l] \leftarrow \lfloor \epsilon_{col}^{en}[l]\mathcal{N}_{col} \rfloor, \mathcal{S}_{col}[l] \leftarrow \mathcal{C}[\mathcal{P}_{col}[l]], \\ &\mathcal{C}[\mathcal{P}_{col}[l]] \leftarrow \mathcal{C}[\mathcal{N}_{col}], \mathcal{N}_{col} \leftarrow \mathcal{N}_{col} - 1. \end{aligned} \quad (5)$$

where "$\lfloor * \rfloor$" represents a rounding down operation.

iv) *Encryption parameters generation:* Generate the encryption parameters $\beta_{row}[l]$ and $\beta_{col}[l]$ as follows for $l = 0, 1, 2, 3$:

$$\begin{aligned} \beta_{row}[l] &= mod(\lfloor \epsilon_{col}^{en}[3-l] \times 10000 \rfloor, 256), \\ \beta_{col}[l] &= mod(\lfloor \epsilon_{row}^{en}[3-l] \times 10000 \rfloor, 256). \end{aligned} \quad (6)$$

v) *Pixels encryption and rank swapping:* Using the encryption parameters generated by (6) to encrypt the pixel values of the selected rows and columns, i.e.,

$$\begin{aligned} &P_{\mathcal{S}_{row}[l],j} \leftarrow P_{\mathcal{S}_{row}[l],j} \oplus (\beta_{row}[l] >> (mod(j,4) << 1)), \\ &P_{i,\mathcal{S}_{col}[l]} \leftarrow P_{i,\mathcal{S}_{col}[l]} \oplus (\beta_{col}[l] >> (mod(i,4) << 1)). \end{aligned} \quad (7)$$

where "$>>$" and "$<<$" represent rotate right and left operation, and $\oplus$ represents XOR operation. Then, the rows and columns are swapped in pairs respectively.

$$\begin{aligned} &P_{\mathcal{S}_{row}[l],j} \leftrightarrow P_{\mathcal{S}_{row}[l+1],j}, P_{i,\mathcal{S}_{col}[l]} \leftrightarrow P_{i,\mathcal{S}_{col}[l+1]}, \\ &if\ l\ is\ even \end{aligned} \quad (8)$$

where "$\leftrightarrow$" represents the replacement operator.

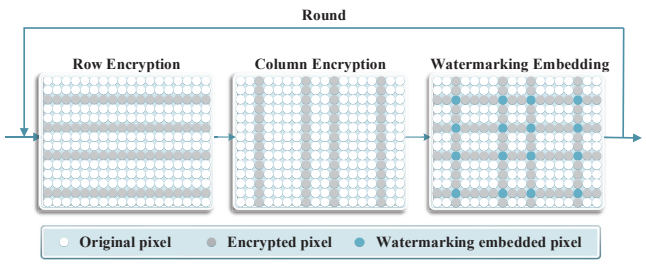


Fig. 4. Watermarking Embedding.

vi) Repeating (iii)-(v) until $\mathcal{N}_{row} = 0$ or $\mathcal{N}_{col} = 0$.

During the image encryption process, the cross iteration of rows and columns is used, while the intersection of rows and columns (i.e., $P_{\mathcal{S}_{row}[l_r],\mathcal{S}_{col}[l_c]}$, $l_r = 0, 1, 2, 3$, $l_c = 0, 1, 2, 3$) is selected as the watermarking embedding location, as shown in Fig. 4. Here, the specific watermarking embedding method is as follows:

i) *Watermarking parameters generation:* Iterating (2b) once to get $\zeta^{wm}[s], s = 0, 1, 2, ..., 15$ by Key 2 in (3), and set $g[s] = mod(\lfloor \zeta^{wm}[s] \times 1000 \rfloor, 2^4)$. Then, the watermarking parameters is split to get the watermarking information $wm_L[s] = g[s]\&3$ and pixel shift distance $wm_H[s] = (g[s]\&12) >> 1$.

ii) *Watermarkings embedding:* Embed watermarking information in the lower two bits of the pixel, i.e.,

$$\begin{aligned} P_{\mathcal{S}_{row}[l_r],\mathcal{S}_{col}[l_c]} &\leftarrow P_{\mathcal{S}_{row}[l_r],\mathcal{S}_{col}[l_c]}\&252 \oplus wm_L[s], \\ P_{\mathcal{S}_{row}[l_r],\mathcal{S}_{col}[l_c]} &\leftarrow P_{\mathcal{S}_{row}[l_r],\mathcal{S}_{col}[l_c]} >> wm_H[s]. \end{aligned} \tag{9}$$

iii) Repeating (i)-(ii) until all intersections are embedded with watermarking.

At the second level detection, after the encrypted image is received, the embedded watermarking is extracted from the same position according to Keys 1 and 2 before decryption. In comparison with the extracted watermarking and original watermarking, it will be detected whether the encrypted image has suffered tampering based on its intensity. If the intensity of the detected tampering exceeds attack detection threshold, the image is judged as "invalid" and discarded. Otherwise, the image can be decrypted by Key 1, which is an inverse process of image encryption.

The pseudo code of the image pre-processing (encryption and watermarking embedding) is Algorithm 1.

The result from double level attack detector will be represented by a trigger signal γ_k. If the image is judged as "invalid" (i.e., $\gamma_k = 0$), then the image is discarded and next process will directly enter data compensation unit. Otherwise, $\gamma_k = 1$, the "valid" image will be decrypted by Key 1 and the corresponding available state information x_k are extracted at information extraction unit.

For NVCSs, the complete and true image is the cornerstones of system stability. Moreover, the large-granularity-based row-column cross selection method greatly improves the detection efficiency of algorithm. In comparison with authenticity detection of state information extracted from the image, the proposed double-level attack detection has a higher real-time capability. The former needs to extract information from the image before detection, while the latter detects firstly the validity of the image, only "valid" image is processed to extract state information. Therefore, under the latter case, state information extraction maybe skipped if the detector judges the image as "invalid", which avoid the meaningless time-consuming of image processing.

Algorithm 1 Image Security Detection Method

Input: The image captured by industrial camera: I_o, The pixels of the image: $P_{i,j}$, The image time-stamp: T_n; The image encryption key (Key 1): λ_0^{en} and μ^{en}, The fragile watermarking key (Key 2): λ_0^{wm} and μ^{wm};

Output: The encrypted image I_w,

```
Keys updating (3),
Recorders initialization,
repeat
    Rows and columns selection (4) and (5),
    Encryption parameters generation (6),
    Pixels encryption (7) and rank swapping (8),
    repeat
        Watermarking parameters generation,
        Watermarkings embedding (9),
    until All intersections are embedded with watermarkings.
until N_row = 0 or N_col = 0.
```

Remark 4: In comparison with the retransmission mechanism [40], the proposed double-level detection mechanism (combing timeout strategy, image encryption and watermarking) enhances the attack resistance of images and can detect tampering locations on images, which improves the efficiency. For instance, if 1% of 1000 frames of images (i.e., 10 frames) have timed out, and 2% (i.e., 20 frames) are judged as 'invalid" under image attacks, then more 30-times retransmission will be produced based on the retransmission mechanism under single-layer detection, which cannot guarantee system stability because the retransmitted image may not reach the remote control terminal in time. However, the proposed double-level detection mechanism does not need the retransmission, while directly compensates the "invalid" image. Therefore, the proposed double-level detection mechanism can improve efficiency, because the computational time of compensation is far less than retransmission time.

C. Analysis of Real Time and Computational Burden

The process control systems do have less strict requirements on the control period, which is usually on the seconds or minutes level [41], [42]. However, the motion control systems with fast-changing characteristics, and its control period is generally on the millisecond level [43]–[45]. As an ideal class of the motion control platform, NIPVCS can only maintain stability if meeting high real-time requirements.

High real-time performance can be achieved with low computational burden of the algorithm. The computational complexity of our proposed algorithm based on the logistic map iteration is $O((M + N) + (MN)/2)$, which is much

smaller than $O(2MN)$, $O(3MN)$ or $O(24MN)$ of some advanced pixel-based algorithms. That is, our algorithm reduces time-consuming from generating encryption as well as watermarking parameters and converting floating point numbers with the same size of images. It is worth noting that the time-consuming of the decryption is about half of the entire algorithm. If images can be pre-judged as invalid after being attacked, this will avoid nearly half of the unnecessary image decryption time. Experimental results of the proposed method are analysed in Section I. B of the supplementary materials, which meets the real-time requirement of NVCSs.

D. Online Compensation Based on Cubic Spline Interpolation

If the images are judged as "invalid" and discarded by the above double-level attack detector, data compensation must be designed to guarantee the stability of NVCSs. Due to good characteristics such as continuity of derivatives and interpolation and low computational burden, etc., cubic spline interpolation algorithm [46] is adopted, as shown in Section I. C of the supplementary materials. Then, online compensation is described as follow.

1) Cubic Spline Interpolation Algorithm: To compensate state information caused from "invalid" images, the predicted values in the k^{th} sampling period can be calculated by spline interpolation. Firstly, an interval $[t_1, t_n]$ is divided into $n-1$ intervals, *i.e.*, $t_1 < t_2, \cdots, t_{n-1} < t_n$. Then, the function $S(t)$ is a cubic polynomial on each interval $[t_i, t_{i+1}]$. Given $y_i = f(t_i), i = 1, \cdots, n$ on the node t_i and $S(t_i) = t_i$, $S(t)$ is called as spline function on the nodes $t_1, \cdots, t_n$.

To solve function coefficients of spline interpolation, the following conditions need be satisfied:

i) In each interval $[t_i, t_{i+1}]$, $S_i(t)$ is a cubic polynomial and $S_i(t) = a_i + b_i(t-t_i) + c_i(t-t_i)^2 + d_i(t-t_i)^3$, where $a_i, b_i, c_i, d_i, i = 1, \cdots, n-1$. Thus, there are $4(n-1)$ unknown coefficients.
ii) Zero error at the node is satisfied, i.e., $S_i(t_i) = y_i$.
iii) The curve $S(t)$ is smooth. The first and second-order derivative $S'(t)$ and $S''(t)$ are all continuous in $[t_1, t_n]$, i.e., $S'_i(t_{i+1}) = S'_{i+1}(t_{i+1})$, $S''_i(t_{i+1}) = S''_{i+1}(t_{i+1})$.

Letting $h_i = t_{i+1} - t_i$, $m_i = 2c_i$, the above (i)-(iii) leads to $h_i m_i + 2(h_i + h_{i+1})m_{i+1} + h_{i+1}m_{i+2} = 6\left[\frac{y_{i+2}-y_{i+1}}{h_{i+1}} - \frac{y_{i+1}-y_i}{h_i}\right]$. Hence, the parameters in $S_i(t)$ are $a_i = y_i$, $b_i = \frac{y_{i+1}-y_i}{h_i} - \frac{h_i}{2}m_i - \frac{h_i}{6}(m_{i+1}-m_i)$, $c_i = \frac{1}{2}m_i$, $d_i = \frac{m_{i+1}-m_i}{6h_i}$.

A non-node boundary (Not-A-Knot) is used to add extra restrictions on the differentiation between the endpoints x_1 and x_n, i.e., $S'''_1(t_2) = S'''_2(t_2)$, $S'''_{n-2}(t_{n-1}) = S'''_{n-1}(t_{n-1})$. Then, the above restrictive conditions becomes $h_2(m_2 - m_1) = h_2(m_3 - m_2)$, $h_{n-1}(m_{n-1} - m_{n-2}) = h_{n-2}(m_n - m_{n-1})$. Therefore, the corresponding piecewise trinomial polynomial function curve coefficients a_i, b_i, c_i, d_i can be obtained by the solution of m_i.

Remark 5: The accuracy of the predicted data generated by interpolation gradually decreases as the number of "invalid" images increases. From the interpolation strategy, when the historical compensation data becomes historical data, they will have an impact on future data compensation. Therefore, only data prediction cannot fulfill stable operation requirement for high real-time control systems. It is necessary to improve the existing algorithm.

2) Design of Online Compensation scheme: It is found by the experiments that when only predictions are made in the above cubic spline interpolation algorithm, there exist the following drawbacks:

i) If the current control signal u_k is correlated with the current state x_k and buffer $\{z_{k-1}\}$ including the previous states, the historical prediction error will decline control performance when x_k is not lost and x_{k-1} is lost.
ii) Under poor network environments, the historical data $\hat{x}_{k-1}$ will be used to predict the current lost data, but the prediction error will adversely affect the current prediction.

Due to the existence of prediction errors described above, the cumulative errors will lead to an excessive accumulation of errors after a period of operation, which affects the stability of NVCSs. To solve the above problems, we propose an online compensation strategy based on cubic spline interpolation, which is mainly divided into data prediction phase and re-updating phase of historical prediction data. The data from the prediction phase will be transmitted to the controller to calculate control signals, while re-updating phase of historical prediction data reduces the accumulated errors to provide more accurate historical data for next prediction by improving the accuracy of historical prediction data.

To achieve online compensation strategy, a buffer is firstly deployed to record historical data $\{z_{k-1}\}$ for supporting data compensation of invalid images. When invalid images are discarded, the lost data are replaced by $\hat{x}_k$, which is predicted in data compensation unit by $\{z_{k-1}\}$ from the buffer. Then, the received first valid states x_k after discarding data will be used to update historical compensation data and be transferred to the buffer.

Specifically, three cases (i.e., the current data is invalid, the current data is valid and previous data was invalid, and both the current data and previous data are valid) are processed:

i) When the current data x_k is judged as invalid, taking the previous data $\{z_{k-1}\}$ as known data, the current data is predicted by cubic spline external interpolation algorithm, and $\{z_k\}$ will be updated by the predicted value $\hat{x}_k$. It is treated as the prediction phase.
ii) When x_k is valid but x_{k-1} is judged as invalid, the latest historical predicted data $\hat{x}_{k-1}, \cdots, \hat{x}_{k-\tau}$ will be re-updated to $\hat{x}'_{k-1}, \cdots, \hat{x}'_{k-\tau}$ by x_k. Then, the buffers $\Lambda \triangleq \{z_{k-\tau}, \cdots, z_{k-1}, z_k\}$ will be updated by $\hat{x}'_{k-\tau}, \cdots, \hat{x}'_{k-1}$ and x_k. It is treated as the update phase.
iii) When both x_k and x_{k-1} are valid, the buffer $\{z_k\}$ is updated directly.

Considering continuous character of attacks, τ_k is used to record the continuous invalidation at the k^{th} instant, representing the number of equivalent continuous packet losses. If $\gamma_k = 1$, set $\tau_k = 0$; if $\gamma_k = 0$, set $\tau_k = \tau_{k-1} + 1$.

Therefore, the pseudo-code of online compensation strategy is summarized in Algorithm 2.

Algorithm 2 Online Compensation Strategy Based on Cubic Spline Interpolation

Input: $\gamma_k, \tau_{k-1}, x_k,$

Output: $z_k, \tau_k,$

Initial $\tau_0 \leftarrow 0$ and $k \leftarrow 1$,
for each $k \in [1, \infty)$ **do**
  if $\gamma_k = 0$ **then**
    $\tau_k \leftarrow \tau_{k-1} + 1$,
    Using cubic spline interpolation prediction to obtain $\hat{x}_k$,
    $z_k \leftarrow \hat{x}_k$,
  else
    if $\tau_{k-1} \neq 0$ **then**
      Using cubic spline interpolation to update $\hat{x}'_{k-1}, \ldots, \hat{x}'_{k-\tau}$,
      Update historical values in the buffer $z_{k-1}, \ldots, z_{k-\tau}$,
    end if
    $\tau_k \leftarrow 0$,
    $z_k \leftarrow x_k$,
  end if
  $k \leftarrow k + 1$,
end for

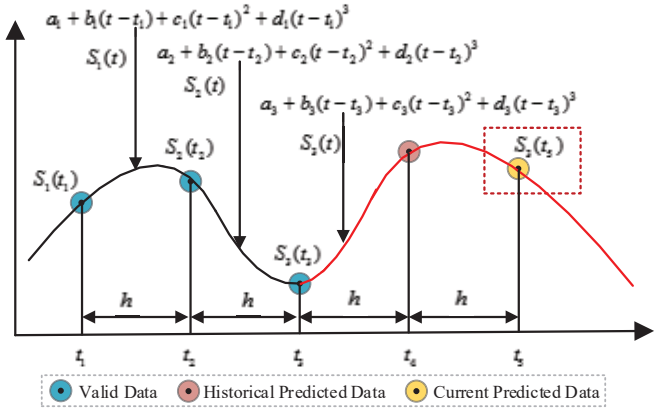


Fig. 5. An illustration of prediction phase.

As mentioned above, the compensation of invalid data is divided into two main phases: prediction and update.

The prediction phase uses a multi-step prediction approach. When invalid data occurs, the invalid data x_k is predicted by interpolation based on the known historical data $x_{k-1}, \ldots, x_{k-i}$ (i is the number of known historical data selected in the prediction phase). Each prediction is based on previous i steps, regardless of the presence of historical predicted data in previous i steps. Taking the prediction phase shown in Fig. 5 as an example. The coefficients of the segmented cubic polynomial are obtained by interpolating cubic splines with the known sampled values (which may contain historical prediction data, such as t_4) at $t_1, \ldots, t_4$. Then, the last segment function $S_3(t)$ is used as the motion trajectory of the controlled plant, and the corresponding function value obtained at the corresponding instant is the predicted value of the invalid sampling value at instant t_5.

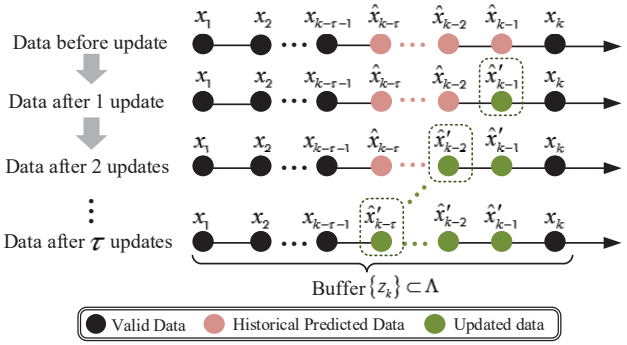


Fig. 6. An illustration of τ update of Multi-step update phase.

Remark 6: The predicted values $\hat{x}_k$ will be used in the calculation of control signal u_k^c, which has an effect on the motion of the controlled plant, so historical prediction data is also considered as known historical data for next invalid data in case of successive invalidation. It is worth noting that as the time interval between the prediction data and real valid data increases, the accuracy of the prediction gradually decreases in short-period sampling system.

To improve the accuracy of historical prediction data when next continuous invalid data occurs, an update phase after prediction is proposed to form pre-predict and post-update strategy. When the first valid data x_k appears after continuous data invalidation at τ_{k-1} instants, a multi-step update is used to interpolate historical prediction data $x_{k-1}, \ldots, x_{k-\tau}$ one by one based on the current real data x_k and historical data $x_{k-1}, \ldots, x_{k-j}$ ($j > \tau_{k-1}$, and j is the number of known data selected in the update phase) to ensure that the used historical data is closer to the real data when data invalidation occur again.

Remark 7: Multi-step update is defined as a process of τ update based on historical forecast data, and the corresponding updation process is shown in Fig. 6. According to the number of τ of historical prediction data, τ rounds of updations are performed, in the order of time interval between historical prediction data and the latest real data from near to far. Only one historical prediction data is updated in each round, and the buffer will be updated at the end.

Taking the update phase shown in Fig. 7 as an example, the data at t_4 and t_5 are historical prediction data, which means that the amount of historical prediction data is 2 and two rounds of updates are required. In the first round, the data at t_5 will be predicted again. In the second round, the data at t_4 will be predicted again. The segmentation function between two moments before and after the update moment is used as the motion curve. The first round of updation takes data at t_1, t_2, t_3, t_4, t_6 as known sampled data to obtain the coefficients of the segmented cubic polynomial. The segmentation function between t_4 and t_6 is used to take the updated data at the corresponding instant t_5. In the second round, the data at t_1, t_2, t_3, t_5, t_6 are taken as known sampled data, where the data at t_5 is the updated data. The corresponding segmentation function between t_3 and t_5 is used to take the updated data at t_4.

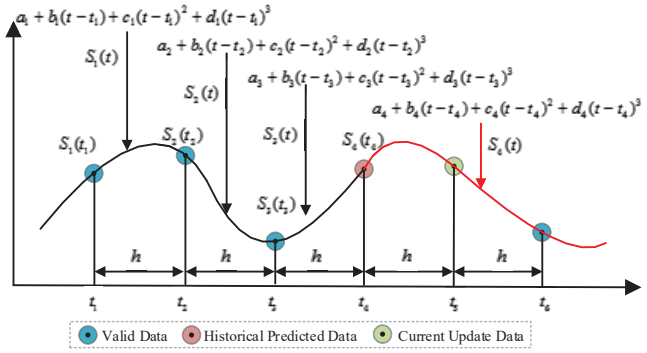


Fig. 7. An illustration of one update of Multi-step update phase.

This phase is based on the current valid data for reducing the impact of prediction errors. On the other hand next round of data prediction is operated to reduce the accumulation of prediction errors.

Remark 8: It is worth noting that the update phase requires multiple rounds of updations, which consumes a certain amount of computation time to update at the current instant. Therefore, it is considered to copy firstly the current buffered data, and then judge the necessity of updating based on whether this information is adopted or not, when next round of consecutive predictions appears. If the historical prediction data is adopted by next round of prediction, the time margin generated from skipping the image decryption and information extraction units at prediction phase is used to update, which will reduce unnecessary data updations.

IV. Experimental Analysis

A. Establishment of the Experimental Platform

With nonlinear and unstable characteristics, inverted pendulum control system is an ideal platform for experiment verification. Therefore, networked inverted pendulum visual control system (NIPVCS) is employed to verify the feasibility and effectiveness of the proposed secure control method, as shown in Fig. 8. NIPVCS is mainly composed of actuator, inverted pendulum, visual sensor (industrial camera), image pre-processing unit, double-layer detector, data compensation unit, image information extraction unit and controller. Visual sensor has the highest frame-rate up to 120fps and the highest resolution up to 659×492, which can satisfy the measurement resolution and real-time requirements of inverted pendulum control system.

Considering NIPVCS under time-triggered mechanism, the effective sampling period T_s needs to be selected to satisfy $T_s > \overline{T}$ ($\overline{T}$ represents the upper bound of system delay). In fact, system delay fluctuates due to the influence of compution processing ability, network environment and other factors, so there exist upper bound $\overline{T}$ and lower bound $\underline{T}$. Therefore, to select a suitable T_s, the fluctuation range of the delay needs to be analyzed of 2000 image samples is shown in Section I. D of the supplementary materials. $\overline{T}$ and $\underline{T}$ are obtained as:

$$\underline{T} = 26ms, \overline{T} = 34ms. \tag{10}$$

Therefore, $T_s = 35ms > \overline{T}$ is taken as the sampling period.

Fig. 8. Experimental platform of the networked inverted pendulum visual control system (NIPVCS).

NIPVCS with fast time variation has large sawtooth fluctuations in pendulum angle curve due to short control period, which is not conducive to long-term data prediction compensation. Therefore, its state is considered to be segmented to obtain the smoother state curve. From the pendulum angle curves at the $2k^{th}$ and $(2k+1)^{th}$ shown in Section I. E of the supplementary materials, it appears the gentler motion trend between the interval sampling points of the state compared to the original one, which makes it more suitable for data compensation. Therefore, the state are divided into two time series ($2k^{th}$ and $(2k+1)^{th}$) for compensation respectively.

However, the above division reduces the correlation of the prediction information. Therefore, to ensure the prediction accuracy, considering the correlation between x_k and x_{k-1}, the prediction results are corrected by the state information of the previous instant in the prediction phase.

$$\hat{x}_{k|k-1} = e_{pre}\hat{x}_k + (1 - e_{pre})z_{k-1} \tag{11}$$

where e_{pre} is the correction scale factor and $e_{pre} \in (0, 1)$, $\hat{x}_k$ is the predicted value of the invalid state x_k, z_{k-1} is the state of previous instant cached in the buffer, and $\hat{x}_{k|k-1}$ is the correction value of $\hat{x}_k$ which will be recorded as z_k.

In the update phase, assuming that the current new valid state belongs to the $2k^{th}$ series, then only the historical predicted state of the $2k^{th}$ series will be updated in the update phase, while the $(2k+1)^{th}$ series can be updated until next valid state appears. However, the calculation of control signal is highly correlated with the state of previous instant, so the state updation starts from a pre-update of the nearest neighboring historical prediction data. It follows that

$$\hat{x}'_{k-1|k} = e_{up}z_{k-1} + (1 - e_{up})x_k \tag{12}$$

where e_{up} is pre-update scale factor and $e_{up} \in (0, 1)$, x_k is new valid state, and $\hat{x}'_{k-1|k}$ is pre-update value of z_{k-1}. The correction scale factor $e_{pre} = 0.8$ and the pre-updated scale factor $e_{up} = 0.7$ are selected by several experiments.

B. Security Detection and Real-time Control Experiments

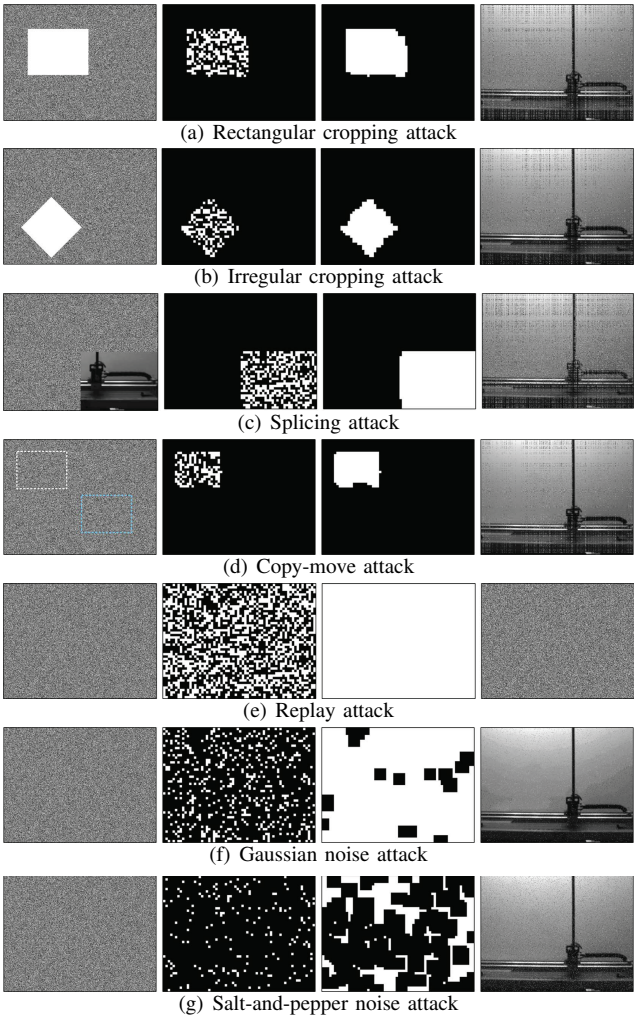
1) Analysis Method of Watermarking Detection: The watermarking embedded pixels of the proposed algorithm is random and scattered for the located. A block localization method is employed to support the detection, where 640×480 image will be divided into 10×10 blocks for initial tampering detection. If a watermarking embedded pixel is detected to be tampered, the corresponding block will be marked as be tampered. After extraction and comparison of all watermarkings, a 64×48 marked image is obtained. Morphological operation are then performed on this marked image to exclude small black holes, as shown in Section I. F of the supplementary materials, which can achieve high attack recognition accuracy.

2) Real-time Performance: The real-time performance of system with detector and system without detector are shown in Section I. G of the supplementary materials. It can be seen that the curve fluctuation of pendulum angle has slightly increased, but the system can still maintain stability after adding detector, which demonstrates that the proposed detector can satisfy real-time requirement.

3) Threshold Selection of Attack Detection: To analyze the effect of attack intensity on system performance, the experiments are operated by considering typical attacks including cropping attacks (rectangular cropping and irregular cropping), splicing attacks, copy-move attacks, replay attacks and noise attacks (Gaussian noise and salt-and-pepper noise) as shown in Fig. 9. Among them, from left to right are the encrypted images under attacks, initial marked images (initial tampering detection using block localization method), final marked images (after morphological operation of initial marked images) and decrypted images respectively; white part of the images is the tampered area detected; the range shown by blue dashed line in the image under copy-move attack is the copied area, and the range shown by white dashed line is the moved area.

As can be seen in Fig. 9, regardless of regular shape of the clipping attack, the cropping region can be detected and effectively localized. The splicing attack, copy-move attack and replay attack present the same results of tampering region as the cropping attack. Gaussian noise and salt-and-pepper noise are global tampering attacks, where initial marked images show global random distribution. It can be seen from the final marked images in Figs. 9(f) and 9(g) that further detection for the attack can more clearly reflect their global characteristics. Moreover, the decrypted image indicates that the information in the tampered area will not be completely lost under a certain intensity of attack, which is because the impact of the attack is dispersed to other regions of the image by encryption.

To guarantee the stability of NIPVCS, a suitable detection thresholds need to be set to achieve quantitative identification of images. Therefore, the tampering rates of typical attacks detected at different intensities are analyzed by appropriate threshold. The proposed block localization method is used to detect different types of attacks with different attack intensities, and the stability of system is observed without warning, as shown in Tables I and II. Here, attack intensity refers to the ratio of the number of pixels suffering from tampering to the whole image; "Tampering rate 1" represents the tamper rate obtained from the initial detection; "Tampering rate 2" represents the tamper rate obtained from the final detection; "$\surd$" represents the ability of the system to maintain stable operation under a continuous attack of the corresponding intensity, otherwise "$\times$".



(a) Rectangular cropping attack
(b) Irregular cropping attack
(c) Splicing attack
(d) Copy-move attack
(e) Replay attack
(f) Gaussian noise attack
(g) Salt-and-pepper noise attack

Fig. 9. Detecting and decrypting images under attack.

From Table I, it can be concluded that for local tampering (cropping, splicing and copy-move attacks), the tampering rate 2 can effectively represents the true tampering ratio and the ratio of tampering rate 1 to 2 is close to 1:2. Moreover, when the tampered area is greater than 17%, the inverted pendulum cannot maintain stability. For replay attack, the whole image is tampered and no valid information can be obtained to maintain system stability, but the ratio of tampering rate 1 to 2 is also close to 1:2. Therefore, these types of area-based tampering can be considered as having a ratio of "Tampering rate 1" to "Tampering rate 2" close to 1:2, and 17% can be selected as the threshold of "Tampering rate 2". The threshold of "Tampering rate 2" is 17%, which means that the area-type attacks with "Tampering rate 2" less than 17% can be resisted by the controller and the impact on image information can be ignored.

From Table II, it can be concluded that for global tam-

TABLE I
ATTACK DETECTION RESULTS UNDER DIFFERENT REGIONAL ATTACK TYPES

Attack type	Attack intensity	Tampering rate 1	Tampering rate 2	System stability
Rectangular cropping attack	4%	1.92%	4.13%	√
	8%	3.79%	7.77%	√
	16%	7.81%	15.59%	√
	25%	11.59%	24.66%	×
Irregular cropping attack	4%	1.91%	4.12%	√
	8%	3.86%	7.83%	√
	16%	7.92%	16.27%	√
	25%	12.35%	25.18%	×
Splicing attack	4%	1.86%	3.87%	√
	8%	3.91%	7.79%	√
	16%	7.71%	15.79%	√
	25%	11.97%	24.85%	×
Copy-move attack	4%	1.92%	4.09%	√
	8%	3.81%	7.89%	√
	16%	7.93%	15.87%	√
	25%	11.46%	24.80%	×
Replay attack	1	45.67%	100.00%	×

TABLE II
ATTACK DETECTION RESULTS UNDER DIFFERENT NOISE ATTACK TYPES

Attack type	Attack intensity	Tampering rate 1	Tampering rate 2	System stability
Gaussian noise attack	$\mu = 0, \sigma = 1$	8.04%	38.38%	√
	$\mu = 0, \sigma = 2$	15.71%	94.37%	√
	$\mu = 0, \sigma = 4$	24.10%	97.73%	√
	$\mu = 0, \sigma = 6$	28.48%	98.70%	×
Salt-and-pepper noise attack	4%	1.99%	4.46%	√
	10%	7.58%	33.20%	√
	20%	14.48%	62.66%	√
	30%	17.76%	86.32%	√
	40%	27.67%	92.97%	×
	80%	45.57%	98.83%	×

pering (Gaussian noise and Salt-and-pepper noise), the ratio of "Tampering rate 2" to "Tampering rate 1" increases and then decreases as the noise intensity increases, and the ratio of two tamper rates is three times or more at moderate intensity. Unlike the regional tampering, inverted pendulum still remains stable when "Tampering rate 2" is greater than 17%. The system is not stable when "Tampering rate 1" of Gaussian noise is greater than 25%, while Salt-and-pepper noise is smaller than Gaussian noise. Therefore, to distinguish medium-intensity noise attack, we can set a second threshold to determine whether medium-intensity noise attack is a resistible attack, based on the judgment that "Tampering rate 2" is three times or more than "Tampering rate 1". To be conservative, "Tampering rate 1" equal to 18% is selected as the second threshold, which indicates that the encrypted image under noise attack with "Tampering rate 1" less than 18% is considered to be recoverable.

Remark 9: An appropriate threshold of tampering rate of image attacks for double-layer detector is the key to judge the validity of images. Most of the existing threshold selection methods generally include theoretical derivation [47], statistical analysis [48] and machine learning [49], etc. According to the statistical results of real-world experiments, the thresholds of image attack tampering rate are determined. For the NIPVCS shown in Fig. 8, the double thresholds achieve the appropriate division of the damage on image information under different attack intensities. The "tampering rate 2" is equal to 17% as the first threshold. If it is less than 17%, the images are valid; if it is greater than 17%, we then analyse the ratio of "tampering rate 1" and "tampering rate 2". If "tampering rate 2" is 3 times or more than "tampering rate 1", and "tampering rate 1" is less than 18%, the image is valid, otherwise invalid. From Tables I and II, the selection of other thresholds is too conservative, which will affect the validity judgment of images.

Remark 10: When the system still remains stability under noise attacks, the tampering rate detected by Gaussian noise and Salt-and-pepper noise demonstrates a large difference, so a uniform threshold cannot be obtained to classify them. That is due to the global characteristics of Gaussian noise, which will add different intensities of Gaussian interfere to each pixel, while the watermarking information is embedded in each bit of the pixel that it has a better detection effect. On the other hand, Salt-and-pepper noise is selected a certain percentage of pixels to change its value to 0 or 255, and the watermarking information is also embedded at intervals, naturally the detection probability of it will be lower. Therefore, the smaller tampering rate among them is selected as the second threshold value conservatively.

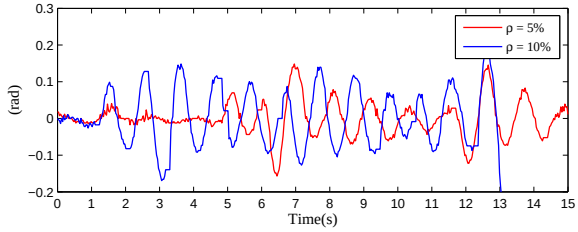
C. Co-Design Compensation and Real-Time Control Experiments

Experimental analysis is performed on NIPVCS to compare the proposed online compensation schemes step by step, and they are experimented under different network environments to analyze the factors affecting their performance.

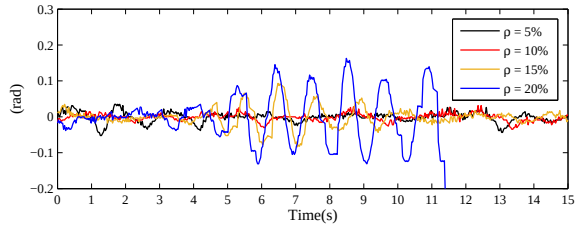
1) Control Performance Comparison under Different Compensation Schemes: Considering the invalidation of data be equated to data loss, and the frequency of data invalidation will be expressed as the loss rate ρ. To verify the effectiveness of the prediction phase (without the update phase), the traditional method (deferring the previous data), single-step prediction (based on cubic spline interpolation) and multi-step prediction are compared under the number of consecutive loss $\tau = 4$. The performance of the above three prediction compensation methods is tested by gradually increasing ρ, and the performance of the above three prediction methods is compared under the same ρ, as shown in Figs. 10 and 11.

From Fig. 10, with $\tau = 4$, it can be obtained that the tolerance to the loss rate is multi-step prediction $>$ single-step prediction $>$ traditional method. From Fig. 11, under $\rho = 15\%$, it can be obtained that the control curve fluctuation situation is multi-step prediction $<$ single-step prediction $<$ traditional method. Multi-step prediction performs best in both tolerance of loss rate and control stability.

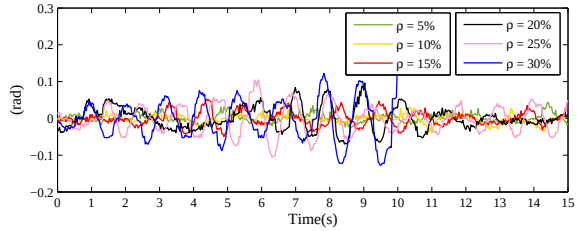
Furthermore, under $\tau = 4$ and with multi-step prediction phase, three update methods of non-update, single-step update and multi-step update were compared, and their performance are tested by gradually increasing ρ, as shown in Figs. 10(c) and 12. It can be obtained that the tolerance to the loss rate is multi-step update $>$ single-step update $>$ non-update.



(a) Traditional Method



(b) Single-step Prediction



(c) Multi-step Prediction

Fig. 10. Prediction compensation methods in different ρ.

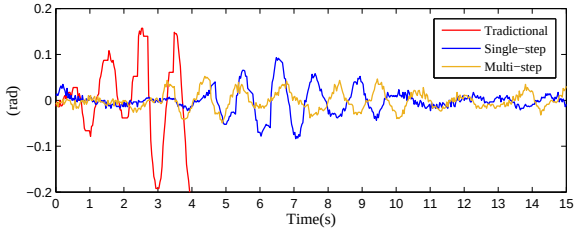
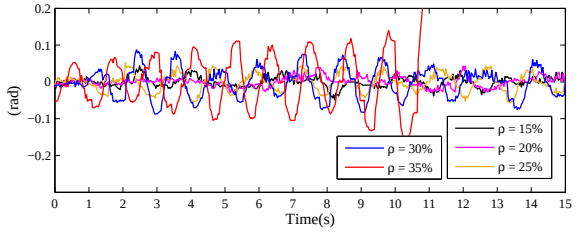


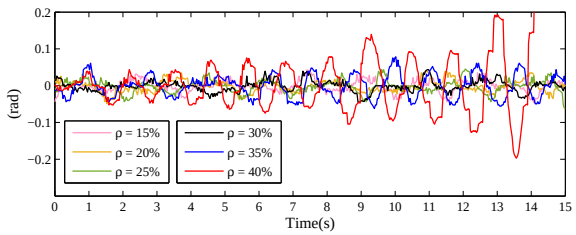
Fig. 11. Comparison of prediction compensation methods in $\rho = 15\%$.

The prediction error based on cubic spline interpolation will accumulate excessively after a period of time, which will affect the stability of NVCSs. Therefore, two updating methods (single-step update and multi-step update) are used to reduce the prediction error in the real-time experiments, which provides more accurate historical data for the next prediction. The performance of the above five data compensation methods is compared under the same ρ and τ, as shown in Fig. 13. It can be obtained that the fluctuation of the control curve is multi-step update (with multi-step prediction) $<$ single-step update (with multi-step prediction) $<$ multi-step prediction $<$ single-step prediction $<$ traditional prediction, and the multi-step update method based on historical prediction data has the best performance in loss rate tolerance and control stability.

The prediction accuracy is evaluated by the root mean squared errors between historical prediction data and real data, i.e., $PA = \sqrt{(1/N_{inv}) \sum_{n=1}^{N_{inv}} (err(n))^2}$, where N_{inv} represents the total number of invalid data and $err(n)$ represents the n^{th} error between the historical prediction and real data. When PA is smaller, it indicates that the prediction accuracy is higher, i.e., the compensation performance is better. As shown in Fig. 14, under the same $\rho = 30\%$ and $\tau = 4$, the prediction error fluctuation based on multi-step update (with multi-step prediction) is smaller. Moreover, the prediction accuracies of five methods are calculated respectively, as shown in Section I. H of the supplementary materials. Experimental results show that the proposed multi-step update with multi-step prediction has higher prediction accuracy, i.e., the historical prediction data are closer to real value.



(a) Single-step Update with Multi-step Prediction



(b) Multi-step Update with Multi-step Prediction

Fig. 12. Update compensation methods in different ρ.

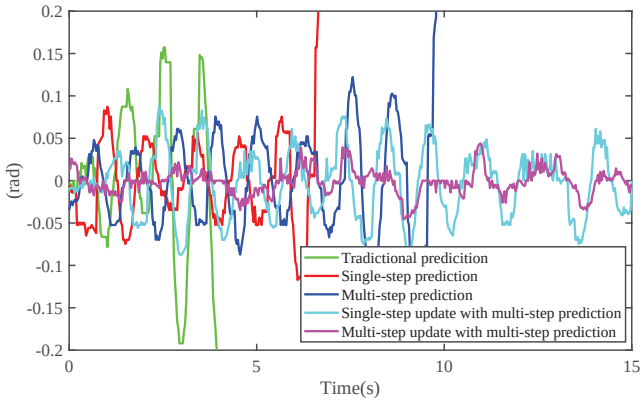


Fig. 13. Comparison of compensation methods in $\rho = 30\%$.

To summarize, multi-step prediction compensation works best in the prediction stage, and multi-step update works best in the update stage. The online compensation strategy based on multi-step prediction and multi-step update can effectively compensate for the invalid data under network uncertainty of the control system.

2) Control Performance Comparison under Different Network Environments: In fact, the variable network environment makes data loss random (e.g., random number of consecutive

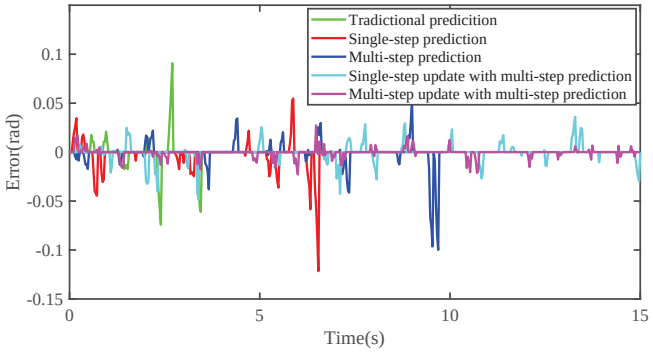


Fig. 14. Comparison of errors under different methods ($\rho = 30\%$).

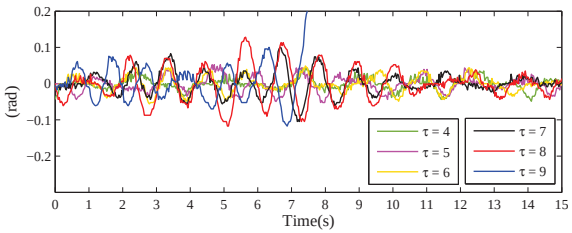


Fig. 15. Comparison under different τ in $\rho = 15\%$.

loss, random loss frequency), which affects the online date compensation and thus leads to different compensation effects of the proposed strategy. Therefore, the control performance under different number of consecutive loss τ and loss rate ρ will be experimentally analyzed to further explore the impact of network environment on the proposed strategy.

Firstly, the effect on the system control is observed by increasing τ, as shown in Fig. 15. It can be seen that for the same ρ, the fluctuation of the pendulum angle keeps increasing as τ increases. Within a certain range of τ, the control system at the lower ρ fluctuates for a period of time after data loss, but it still can reach a stable state. Moreover, the inverted pendulum control system starts to lose stability with $\tau > 8$.

The absolute value of angular error is calculated for $\tau = 4, \ldots, 9$ under $\rho = 15\%$, as shown in Fig. 16. It can be seen that the prediction error increases with the increase of τ at the same ρ. The more distant the prediction data is from the valid data, the lower the prediction accuracy is, which leads to the decrease of control performance and the increase of deviation between the current state of the controlled plant and the stable state. As the control system is in a fluctuating state, it takes a long time to regain a relatively stable state after obtaining valid data, resulting in a long oscillation period. What's more, with τ increases further, the system oscillates too much and exceeds the controllable boundary before the valid data is obtained, which leads to system destabilization.

The RMSEs of the data prediction phase for continuous lost data at different τ and ρ are counted respectively, as shown in Section I. I of the supplementary materials. With the same τ, the prediction error increases as ρ increases. Although updating the predicted data can reduce the cumulative error, the effect of the error still requires some time for the normal

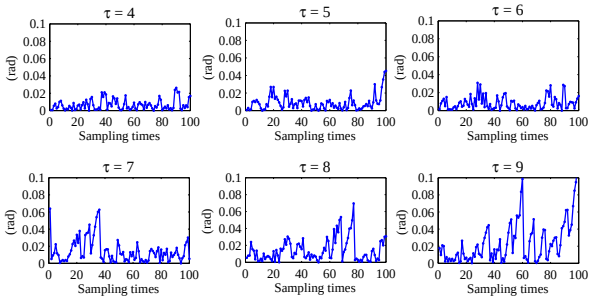


Fig. 16. The absolute value of angular error under different τ in $\rho = 15\%$.

control process to eliminate. Moreover, the increase in ρ makes multiple continuous loss periods too close to each other on the time scale, which will lead to the accumulation of the error not being eliminated and accumulated to the next loss period. In this way, the prediction error will keep increasing, and then the system gradually destabilizes over time.

V. Conclusion

In this paper, a novel co-design secure control method based on image attack detection and data compensation for NVCSs is proposed to address the incomplete and untrue data caused by cyber attacks. Firstly, a detector based on double-layer detection mechanism of timeout strategy and digital watermarking has been designed for image real-time and integrity discrimination to achieve image information protection and integrity detection simultaneously. Then, based on the detection results, an online compensation scheme based on cubic spline interpolation has been proposed to improve the control performance. Finally, the feasibility and effectiveness of the proposed method is confirmed on a practical platform. Limited by the high real-time requirement of system, the method proposed compensates the data in the perspective of non-visual information, so future research will be devoted to the efficiency improvement of the recovery algorithm on images.

References

[1] S. S. Beauchemin, M. A. Bauer, T. Kowsari, and J. Cho, "Portable and scalable vision-based vehicular instrumentation for the analysis of driver intentionality," *IEEE Transactions on Instrumentation and Measurement*, vol. 61, no. 2, pp. 391-401, 2012.

[2] S. Shirmohammadi and A. Ferrero, "Camera as the instrument: the rising trend of vision based measurement," *IEEE Instrumentation and Measurement Magazine*, vol. 17, no. 3, pp. 41-47, 2014.

[3] S. Yan, X. Tao and D. Xu, "Image-based visual servoing system for components alignment using point and line features," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1-11, 2022.

[4] R. C. Luo, C. W. Kuo, "Intelligent seven-DoF robot with dynamic obstacle avoidance and 3-D object recognition for industrial cyber-physical systems in manufacturing automation," *Proceedings of the IEEE*, vol. 104, no. 5, pp. 1102-1113, 2016.

[5] Y. Huang, M. Zhu, Z. Zheng and K. H. Low, "Linear velocity-free visual servoing control for unmanned helicopter landing on a ship with visibility constraint," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2021, DOI: 10.1109/TSMC.2021.3062712.

[6] H. P. Wang, Y. Tian and N. Christov, "Event-triggered observer based control of networked visual servoing control systems," *Control Engineering & Applied Informatics*, vol. 16, no. 1, pp. 22-30, 2014.

[7] D. Du, R. Chen, M. Fei and K. Li, "A novel networked online recursive identification method for multivariable systems with incomplete measurement information," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 3, no. 4, pp. 744-759, 2017.

[8] Y. Li, J. Zhou, "Fast and effective image copy-move forgery detection via hierarchical feature point matching," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 5, pp. 1307-1322, 2019.

[9] N. Sakr, N. D. Georganas and J. Zhao, " Human perception-based data reduction for haptic communication in six-DoF telepresence systems," *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 11, pp. 3534-3546, 2011.

[10] B. Lian, Q. Zhang and J. Li, " Integrated sliding mode control and neural networks based packet disordering prediction for nonlinear networked control systems," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 8, pp. 2324-2335, 2019.

[11] X. Jiang, G. Xia, Z. Feng and Z. Jiang, "Consensus tracking of data-sampled nonlinear multi-agent systems with packet loss and communication delay," *IEEE Transactions on Network Science and Engineering*, vol. 8, no. 1, pp. 126-137, 2021.

[12] X. Zhang, Q. Han and X. Ge, "A novel approach to H∞ performance analysis of discrete-time networked systems subject to network-induced delays and malicious packet dropouts," *Automatica*, vol. 136, 2022.

[13] M. Wolf, D. Serpanos, "Safety and security in cyber-physical systems and Internet-of-Things systems," *Proceedings of the IEEE*, vol. 106, no. 1, pp. 9-20, 2018.

[14] D. Pliatsios, P. Sarigiannidis, T. Lagkas and A. G. Sarigiannidis, "A survey on SCADA systems: Secure protocols, incidents, threats and tactics," *IEEE Communications Surveys Tutorials*, vol. 22, no. 3, pp. 1942-1976, 2020.

[15] E. Shereen, M. Delcourt, S. Barreto, G. Dan, J. Le Boudec and M. Paolone, "Feasibility of time-synchronization attacks against PMU-based state estimation," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 6, pp. 3412-3427, 2020.

[16] D. Du, C. Zhang, X. Li, M. Fei and H. Zhou, "Attack detection for networked control systems using event-triggered dynamic watermarking," *IEEE Transactions on Industrial Informatics*, 2022, DOI: 10.1109/TII.2022.3168868.

[17] S. Zhang, H. Gao and Q. Rao, "Defense against adversarial attacks by reconstructing images," *IEEE Transactions on Image Processing*, vol. 30, pp. 6117-6129, 2021.

[18] S. Hu, D. Yue, X. Xie, X. Chen and X. Yin, "Resilient event-triggered controller synthesis of networked control systems under periodic DoS jamming attacks," *IEEE Transactions on Cybernetics*, vol. 49, no. 12, pp. 4271-4281, 2019.

[19] T. M. Shashidhar, K. B. Ramesh, "Novel framework for optimized digital forensic for mitigating complex image attacks," *International Journal of Electrical and Computer Engineering*, vol. 10, no. 5, pp. 5198, 2020.

[20] A. Hambouz, Y. Shaheen, A. Manna, M. Al-Fayoumi and S. Tedmori, "Achieving data integrity and confidentiality using image steganography and hashing techniques," *2019 2nd International Conference on new Trends in Computing Sciences*, pp. 1-6, 2019.

[21] M. Barni, Q. T. Phan and B. Tondi, "Copy move source-target disambiguation through multi-branch CNNs," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 1825-1840, 2021.

[22] Y. Ma, W. Che, C. Deng and Z. Wu, "Observer-based event-triggered containment control for MASs under DoS attacks," *IEEE Transactions on Cybernetics*, 2021, DOI: 10.1109/TCYB.2021.3104178.

[23] Y. Ma, W. Che, C. Deng and Z. Wu, "Distributed model-free adaptive control for learning nonlinear MASs under DoS attacks," *IEEE Transactions on Neural Networks and Learning Systems*, 2021, DOI: 10.1109/TNNLS.2021.3104978.

[24] K. Pang, L. Ma, H. Bai, S. Xue, "Probability-guaranteed secure consensus control for time-varying stochastic multi-agent systems under mixed attacks," *Journal of the Franklin Institute*, vol. 359, no. 6, pp. 2541-2563, 2022.

[25] L. Su, D. Ye and X. Zhao, "Distributed secure state estimation for cyber-physical systems against replay attacks via multisensor method," *IEEE Systems Journal*, 2021, DOI: 10.1109/JSYST.2021.3123617.

[26] M. Li, Y. Chen, Y. Zhang and Y. Liu, "Adaptive sliding-mode tracking control of networked control systems with false data injection attacks," *Information Sciences*, vol. 585, pp. 194-208, 2022.

[27] A. Musleh, Chen. G and Z. Dong, "A survey on the detection algorithms for false data injection attacks in smart grids," *Journal of Hardware and Systems Security*, vol. 11, no. 3, pp. 2218-2234, 2020.

[28] D. Ding, Q. Han, X. Ge and J. Wang, "Secure state estimation and control of cyber-physical systems: A survey," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 1, pp. 176-190, 2021.

[29] X. Q. Fu, B. C. Liu, Y. Y. Xie, W. Li and Y. Liu, "Image encryption-then-transmission using DNA encryption algorithm and the double chaos," *IEEE Photonics Journal*, vol. 10, no. 3, pp. 1-15, 2018.

[30] M. Preishuber, T. Hutter, S. Katzenbeisser and A. Uhl, "Depreciating motivation and empirical security analysis of chaos-based image and video encryption," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 9, pp. 2137-2150, 2018.

[31] Y. Ding et al., "DeepEDN: A deep-learning-based image encryption and decryption network for internet of medical things," *IEEE Internet of Things Journal*, vol. 8, no. 3, pp. 1504-1518, 2021.

[32] A. T. S. Ho, X. Zhu, J. Shen and P. Marziliano, "Fragile watermarking based on encoding of the Zeroes of the Z-transform," *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 3, pp. 567-569, 2008.

[33] R. Sinhal, I. A. Ansari and C. W. Ahn, "Blind image watermarking for localization and restoration of color images," *IEEE Access*, vol. 8, pp. 200157-200169, 2020.

[34] S. Bhalerao, I. A. Ansari, A. Kumar, "A secure image watermarking for tamper detection and localization," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, pp. 1057-1068, 2021.

[35] M. Zhang and C. Desrosiers, "High-quality image restoration using low-rank patch regularization and global structure sparsity," *IEEE Transactions on Image Processing*, vol. 28, no. 2, pp. 868-879, 2019.

[36] T. Liu and X. Yuan, "Adaptive feature calculation and diagonal mapping for successive recovery of tampered regions," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 31, no. 7, pp. 2617-2630, 2021.

[37] Z. Jin, M. Z. Iqbal, D. Bobkov, W. Zou, X. Li and E. Steinbach, "A flexible deep CNN framework for image restoration," *IEEE Transactions on Multimedia*, vol. 22, no. 4, pp. 1055-1068, 2020.

[38] D. Du, C. Zhang, Y. Song, H. Zhou, X. Li, M. Fei, W. Li, "Real-time H∞ control of networked inverted pendulum visual servo systems," *IEEE Transactions on Cybernetics*, vol. 50, no. 12, pp. 5113-5126, 2020.

[39] D. Du, C. Zhang, X. Li, M. Fei, T. Yang and H. Zhou, "Secure control of networked nontrol systems using dynamic watermarking," *IEEE Transactions on Cybernetics*, 2021, DOI:10.1109/TCYB.2021.3110402.

[40] C. Qu, X. Liu , J. Wu , L. Yin, H. Li, C. Cheng and Q. Zhou , "Retransmission methods to improve voltage control of distributed generation system," *IEEE Communications Letters*, vol. 25, no. 6, pp. 1862-1866, 2021.

[41] T. Wang, H. Gao and J. Qiu, "A combined adaptive neural network and nonlinear model predictive control for multirate networked industrial process control," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 2, pp. 416-425, 2016.

[42] X. X. Meng, H. S. Yu, J. Zhang, K. J. Yan, "Optimized control strategy based on EPCH and DBMP algorithms for quadruple-tank liquid level system," *Journal of Process Control*, vol. 110, pp. 121-132, 2022.

[43] A. Ahmadi, F. R. Salmasi, M. Noori-Manzar, and T. A. Najafabadi, "Speed sensorless and sensor-fault tolerant optimal PI regulator for networked DC motor system with unknown time-delay and packet dropout," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 2, pp. 708-717, 2014.

[44] F. Mager, D. Baumann, R. Jacob, L. Thiele, and M. Zimmerling, "Feedback control goes wireless: Guaranteed stability over low-power multi-hop networks," *10th ACM/IEEE International Conference on Cyber-Physical Systems (ICCPS)*, ACM, 2019.

[45] M. Oikawa, T. Kusakabe, K. Kutsuzawa, S. Sakaino and T. Tsuji, "Reinforcement learning for robotic assembly using non-diagonal stiffness matrix," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 2737-2744, 2021.

[46] K. Li, T. Z. Huang, L. Li and S. Lanteri, "Non-intrusive reduced-order modeling of parameterized electromagnetic scattering problems using cubic spline interpolation," *Journal of Scientific Computing*, vol. 87, no. 52, 2021.

[47] Q. Y. Su, S. Q. Li, Y. C. Gao, X. Huang, and J. Li, "Observer-based detection and reconstruction of dynamic load altering attack in smart grid," *Journal of the Franklin Institute*, vol. 358, no. 7, pp. 4013-4027, 2021.

[48] A. Kusiak, Z. Zhang, A. Verma, "Prediction, operations, and condition monitoring in wind energy," *Energy*, vol. 60, pp. 1-12, 2013.

[49] M. Zhao, S. Zhong, X. Fu, B. Tang and M. Pecht, "Deep residual shrinkage networks for fault diagnosis," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 7, pp. 4681-4690, 2020.