Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

A robust tensor watermarking algorithm for diffusion-tensor images*#

Chengmeng LIU, Zhi LI‡, Guomei WANG, Long ZHENG

State Key Laboratory of Public Big Data, College of Computer Science and Technology, Guizhou University, Guiyang 550025, China

E-mail: 62377400@qq.com; zhili@gzu.edu.cn; 306252084@qq.com; zhenglong178@163.com

Received Dec. 8, 2022; Revision accepted Mar. 25, 2023; Crosschecked Mar. 4, 2024

Abstract: Watermarking algorithms that use convolution neural networks have exhibited good robustness in studies of deep learning networks. However, after embedding watermark signals by convolution, the feature fusion efficiency of convolution is relatively low; this can easily lead to distortion in the embedded image. When distortion occurs in medical images, especially in diffusion tensor images (DTIs), the clinical value of the DTI is lost. To address this issue, a robust watermarking algorithm for DTIs implemented by fusing convolution with a Transformer is proposed to ensure the robustness of the watermark and the consistency of sampling distance, which enhances the quality of the reconstructed image of the watermarked DTIs after embedding the watermark signals. In the watermark-embedding network, T1-weighted (T1w) images are used as prior knowledge. The correlation between T1w images and the original DTI is proposed to calculate the most significant features from the T1w images by using the Transformer mechanism. The maximum of the correlation is used as the most significant feature weight to improve the quality of the reconstructed DTI. In the watermark extraction network, the most significant watermark features from the watermarked DTI are adequately learned by the Transformer to robustly extract the watermark signals from the watermark features. Experimental results show that the average peak signal-to-noise ratio of the watermarked DTI reaches 50.47 dB, the diffusion characteristics such as mean diffusivity and fractional anisotropy remain unchanged, and the main axis deflection angle α_{AC} is close to 1. Our proposed algorithm can effectively protect the copyright of the DTI and barely affects the clinical diagnosis.

Key words: Robust watermarking algorithm; Transformer; Image reconstruction; Diffusion tensor images; Soft attention; Hard attention; T1-weighted images

https://doi.org/10.1631/FITEE.2200628 **CLC number:** TP391.4

1 Introduction

With the development of the information age, telemedicine technology has changed from simple video monitoring and remote telephone diagnosis to the integrated transmission of images, voice, and video over high-speed networks, enabling real-time voice and ultrahigh-resolution video communication for integrated diagnoses and providing a more convenient way for medical experts to deliver remote medicine. However, the transmission of unprotected medical images on the network is vulnerable to attacks, preventing doctors from making correct diagnoses. Furthermore, there is a risk that unprotected medical images will be accessed illegally by unauthorized persons, leading to a breach of patient privacy.

Diffusion tensor images (DTIs) are currently the only noninvasive method for visualizing the white matter fiber tracts of the living brain (Stejskal and Tanner, 1965; Le Bihan et al., 2001). DTIs are

‡ Corresponding author

* Project supported by the National Natural Science Foundation of China (No. 62062023), the Guizhou Science and Technology Plan Project of China (No. ZK[2021]-YB314), and the Stadholder Foundation of Guizhou Province, China (No. 2007(14))

Electronic supplementary materials: The online version of this article (https://doi.org/10.1631/FITEE.2200628) contains supplementary materials, which are available to authorized users

ORCID: Chengmeng LIU, https://orcid.org/0000-0003-0541-8329; Zhi LI, https://orcid.org/0000-0001-9813-4979

presented in five dimensions and can provide more detailed information about the brain than conventional imaging data. Medical staff can precisely locate the direction and distribution of brain nerve fiber conduction bundles, providing precise and powerful technical support for the target, enabling entry into the target and path of surgery, reducing damage to brain tissue, allowing medical personnel to observe neural development or damage to the brain, and promptly detecting subtle anomalous changes in the brains of patients with mental illness. Because DTIs have these important roles in medical care, it is necessary to propose a robust, blind watermarking algorithm to protect the reliability and integrity of DTIs. Currently, there is no robust digital watermarking algorithm for high-dimensional data to effectively protect DTIs.

Recent works on conventional watermarking algorithms provide a deeper study of the invisibility and robustness of watermarking algorithms (Amini et al., 2018; Liu et al., 2018; Huan et al., 2022; Su et al., 2022), but these methods are always for two-dimensional (2D) images and three-dimensional (3D) videos, and there are no watermarking algorithms for higher-dimensional data to protect DTIs directly. Amini et al. (2018) proposed a design for a multibit watermark blind decoder, combined with a wavelet-domain vector-based hidden Markov model, which embeds a random binary sequence in the wavelet domain of a 2D grayscale image and then extracts the watermark using a vector-based hidden Markov model method in the wavelet domain. Liu et al. (2018) proposed a blind double-watermarking mechanism for digital color images, which effectively protects digital color images. Su et al. (2022) proposed a new image-correction algorithm as a preprocessing method for watermark extraction, including fast detection of feature points, image correction, and jagged processing, which solves some problems in geometric feature based image-correction algorithms. Huan et al. (2022) performed chunked singular value decomposition (SVD) on the dual-tree cosine wavelet transformation (DTCWT) domain and extracted candidate coefficients, found two pairs of subbands with strong correlation among the candidate coefficients, and then embedded and extracted the watermark signals by embedding and extracting the watermark signals in the strongly correlated subbands. Deep learning algorithms for robust watermarking (Zhu et al., 2018; Wen BY and Aydore, 2019; Ahmadi et al., 2020; Luo et al., 2020) usually apply convolutions to embed watermark signals. Various attacks are simulated during training, so that the watermark extraction network can correctly extract watermark signals from the attacked watermarked images to ensure the robustness of the embedded watermarks. However, the current deep learning robust watermarking algorithms have two obvious problems: one is the limitation of the convolutional field in the process of embedding the watermark, which leads to a low feature fusion rate, and the other is that the visual quality of the reconstructed image containing the watermark signals is not elaborate, which affects the application of the watermarking algorithm for medical images, especially for DTIs. In this paper, we propose a robust watermarking algorithm for DTIs.

The specific contributions of this paper are as follows:

1. A tensor-based watermark-embedding network for DTIs is proposed. In this network, we embed many redundant watermark signals to ensure the robustness of the watermark.

2. T1-weighted (T1w) images are introduced to reduce the distortion level of tensor features because the T1w images have many structural features of the DTIs. The correlation of the T1w images with the tensor features is first calculated using the Transformer. Then, the most relevant features are fused with the tensor features, thus reducing the variation of tensor features.

3. A DTI watermark extraction network that can effectively extract the correct watermark signals from the watermarked DTIs is proposed. A decoding Transformer extracts the relevant watermarking features, and then the watermarking expression of the features is enhanced by deep residual convolution. The experimental results show that the proposed watermarking algorithm has excellent robustness and invisibility compared with state-of-the-art medical image watermarking algorithms.

2 Related works

2.1 Robust watermarking algorithms

There are a large number of watermarking algorithms (Anand and Niranjan, 1998; Wen Q et al.,

2003; Lee et al., 2005; Ni et al., 2006; Cintra and Cooklev, 2009; Singh et al., 2015; Wang N et al., 2018; Zhu et al., 2018; Wang CP et al., 2019; Wen BY and Aydore, 2019; Xia et al., 2019; Ahmadi et al., 2020; Luo et al., 2020), but watermarking algorithms for medical images need to ensure both invisibility and robustness of watermarks; however, these two traits are contradictory. Therefore, research on the methods for embedding robust watermarks in high-visual-quality medical images has become an important research topic. Wen Q et al. (2003) proposed the use of zero-watermarking to guarantee both the robustness of the watermarking and the visual quality of the image. In this method, the important robust features of the image were used to construct the watermark signals, but the watermark signals were not embedded into the image. The drawback of this method was that it required a large amount of storage space to protect the constructed watermark signals. Researchers found that embedding watermark signals in the frequency domain of medical images can effectively reduce the invisibility of the watermark while ensuring that the watermarked medical images are robust to specific attacks. For example, Singh et al. (2015) embedded watermarked images in a wavelet transform and discrete cosine transform iteratively. The patient-related information was embedded as a text watermarking using SVD. The watermarking algorithm was robust to some signal-processing attacks, but it was not able to resist any geometric attacks. The hiding data with deep networks (HiDDeN) agenda (Zhu et al., 2018) found that deep learning networks are susceptible to small perturbations. Therefore, HiDDeN researchers first proposed to use the susceptibility to embed the watermark signals in the image by using convolution. The watermarked image was attacked using various attacks. The watermark extraction network was programmed to extract the watermark signals from the attacked watermarked image. The watermark-embedding network and extraction network were trained together by using a multitask learning approach to make HiDDeN resistant to a variety of attacks. Subsequently, researchers conducted in-depth research on the HiDDeN framework, which further promotes the development of robust watermarking algorithms in deep learning. Residual diffusion watermarking (RedMark) (Ahmadi et al., 2020) introduced an attack layer consisting of a random combination of fixed distortions to improve the robustness of watermarking. The robust watermarking system ROMark (Wen BY and Aydore, 2019) implemented simple adversarial training with an adaptive selection of distortion type and distortion strength to minimize the accuracy of correct watermark extraction. Luo et al. (2020) used a channel coding strategy to encode the embedded watermarked string, and the resulting redundancy was used to correct the errors generated in the string during channel transmission, thus improving the robustness of the watermark.

Through analysis and research of the above-mentioned methods (Zhu et al., 2018; Wen BY and Aydore, 2019; Ahmadi et al., 2020; Luo et al., 2020), we find that although various methods have been implemented for improving the robustness of watermarking, they all use convolution to redundantly embed binary sequences as watermark signals in the image during image reconstruction; therefore, the image reconstruction method may affect the visual quality of the image after embedding. Since the embedded watermark signals constitute a binary sequence, the output of the watermark extraction network is consistent with the classification task; thus, the idea of an excellent classification task is helpful for designing the watermark extraction network. In the processing of research, we investigate many research works on image reconstruction and classification.

2.2 Image reconstruction algorithms

The image reconstruction algorithm in the watermark-embedding process affects the visual quality of the final image containing watermark signals (Zhu et al., 2018; Wen BY and Aydore, 2019; Ahmadi et al., 2020; Luo et al., 2020); hence, state-of-the-art image reconstruction algorithms are analyzed and studied, which motivates us to propose a watermark-embedding network for DTIs.

The reconstruction task (Ravishankar and Bresler, 2011; Shin et al., 2013; Lai et al., 2016; Zhan et al., 2016; Nakarmi et al., 2017; Zhou and Zhou, 2020) uses prior knowledge to overcome overlapping artifacts that violate the Shannon–Nyquist sampling theorem. Anctil-Robitaille et al. (2021) found that T1w images contain plenty of structural and diffusion information of the DTIs. The T1w images were added as prior knowledge to assist the generative adversarial network to reconstruct diffusion-weighted

images with high resolution from the T1w images. Texture Transformer network for image superresolution (TTSR) (Yang et al., 2020) used a Transformer as an attention mechanism, where the low-resolution image and the reference image were represented as query (q) and key (k) in the Transformer, respectively. When the extracted deep features focus on more accurate texture features, these features will help in better reconstruction of the image. Feng et al. (2021) proposed that the two tasks of reconstruction and superresolution reconstruction can complement and promote each other. A Transformer module was proposed to transfer the shared features, which enables the two tasks to share visual features during training by using the knowledge from one task to accelerate the learning process for the other task.

The proposed algorithm in this study adopts a Transformer to transfer significant features from the T1w images to assist watermarked DTI reconstruction. This proposed algorithm has two advantages over previous research works. One advantage is that the weights of the learned structural information are adaptively calculated by global information to achieve a larger range of fields of perception. The other is that the Transformer obtains more effective features compared to the convolution.

2.3 Classification tasks

In robust watermarking algorithms (Vaswani et al., 2017; Dosovitskiy et al., 2020; Wu et al., 2021) for deep learning, most algorithms embed binary sequences as watermark signals; therefore, the output of the watermark extraction network is consistent with that of the multi-classification task. The design of the watermark extraction network is inspired by some of the state-of-the-art classification tasks, thus improving the watermark extraction accuracy to some extent.

Recently, Transformers have achieved success in computer vision classification tasks by exploring the association between different regions of an image and thus learning to focus on the important image regions. The vision Transformer (ViT) (Dosovitskiy et al., 2020) process involves the following steps: slicing an image into a patch, encoding each image position, adding a learnable classification vector, tripling the parameters as the Transformer's query (q), key (k), and value (v) through a fully connected layer, performing global attention calculation and feature extraction, and finally extracting the classification vector for classification. Although ViT works well, it consumes more memory and time because it must triple the full parameters and calculate the association between each region. Convolutional vision Transformer (CvT) (Wu et al., 2021) was developed as an innovation based on ViT; instead of using fully connected layers to produce the query (q), key (k), and value (v) of the Transformer, the task was accomplished using convolution, which greatly reduces the number of the neural network parameters and effectively improves the performance. Inspired by previous works (Dosovitskiy et al., 2020; Wu et al., 2021), in the proposed algorithm the correlation of global information can be calculated by the Transformer to highlight the watermark signals, and the number of the neural network parameters can be effectively reduced by using convolution instead of a fully connected layer to ensure accuracy.

3 Algorithms

DTI and its clinical metrics are described in detail in Section 1 in the supplementary materials, and in the main text, we focus on our algorithm.

3.1 Algorithmic framework

The framework of the algorithm is shown in Fig. 1. It consists of three main parts:

1. Watermark-embedding network. The purpose of this network is to embed the watermark signals into the DTIs to obtain the watermarked DTIs $\boldsymbol{I}_{\mathrm{w\_DTI}}$ and guarantee that $\boldsymbol{I}_{\mathrm{w\_DTI}}$ still has medical value for clinical diagnosis.

2. Attack network. The main purposes of this network are as follows: (1) to perform various attacks on $\boldsymbol{I}_{\mathrm{w\_DTI}}$ to obtain $\boldsymbol{I}_{\mathrm{w\_DTI\_noise}}$ so that the watermark extraction network can still extract the watermark correctly from $\boldsymbol{I}_{\mathrm{w\_DTI\_noise}}$, and (2) to ensure that the watermark embedded in $\boldsymbol{I}_{\mathrm{w\_DTI}}$ has high robustness. Please refer to Section 3 in the supplementary materials for specific effects.

3. Watermark extraction network. This network performs mainly the extraction of watermark signals and correctly extracts the embedded watermark signals from $\boldsymbol{I}_{\mathrm{w\_DTI\_noise}}$.

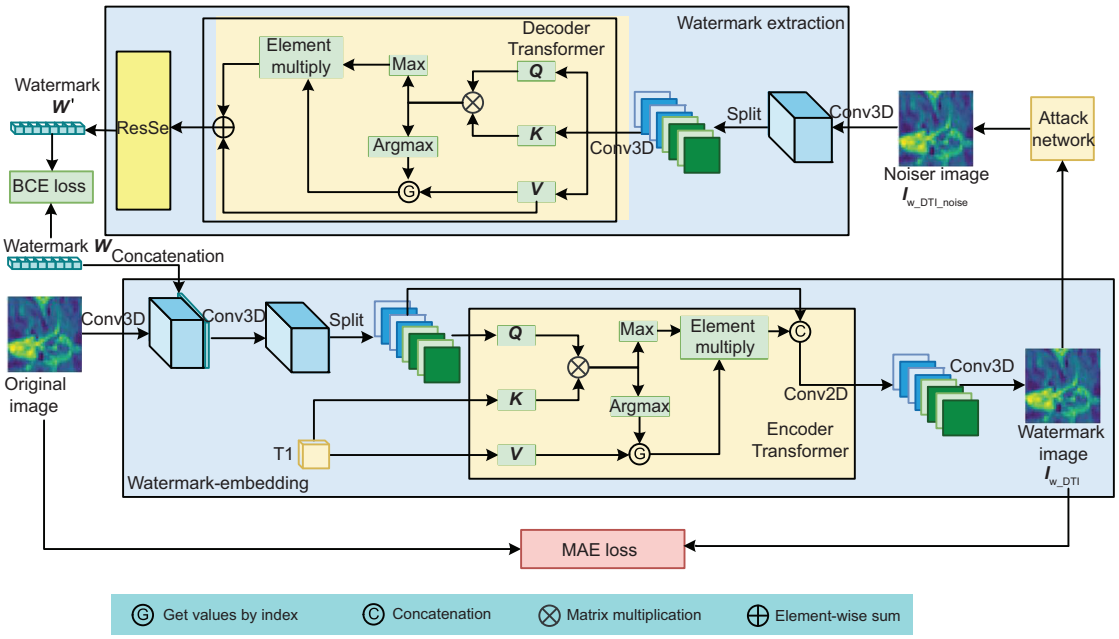

Fig. 1 Network diagram of the robust watermarking algorithm for diffusion tensor images

3.2 Watermark-embedding network

Robust watermarking algorithms based on neural networks (Zhu et al., 2018; Wen BY and Aydore, 2019; Ahmadi et al., 2020; Luo et al., 2020) usually extract image features from carrier images at first, and then reconstruct the watermarked image after fusing the image features with the watermark signals to accomplish the embedding of watermarks. However, these algorithms cannot directly embed watermark signals into high-dimensional tensor data of DTIs. To achieve the embedding of watermark signals in DTIs, we propose a watermark-embedding network for DTIs. The design of this network considers mainly two aspects: one is to ensure that the embedded watermarked DTIs still have clinical value, and the other is that the embedded watermark needs to have good robustness.

After analyzing magnetic resonance imaging (MRI) based reconstruction works, a few studies (Zhou and Zhou, 2020; Anctil-Robitaille et al., 2021) showed that the T1w images are fully sampled, which have more detailed structural information than the diffusion-weighted image (DWI). Furthermore, the dimensions of the T1w images are the same as the dimensions of the DTIs. In this proposed algorithm, the T1w images are used as prior knowledge to assist DTI reconstruction. The correlation between the DTIs and the T1w images can be calculated in the tensor channel dimension to effectively extract the structural information in the T1w images to assist in completing the reconstruction of the DTIs.

To ensure the image's visual quality after embedding the watermark signals, 3D convolution is used to obtain the features of the tensor of the DTI in the reconstruction process. Because the T1w images have plentiful structural information, the T1w images corresponding to the DTIs are introduced as the prior knowledge in the tensor channel reconstruction process. Due to the sensitivity of the diffusion tensor of DTIs, T1w images cannot be directly fused through convolution; this cannot effectively assist in DTI reconstruction and also has a serious negative impact on the robustness of watermark signals. To enhance the visual quality of the watermarked DTIs, the significant feature module is proposed to enable migration of the significant structural features from the T1w images to improve the confidence of the significant structural features. Furthermore, in the significant feature module, the significant structural features are combined with the tensor features by introducing a soft attention mechanism to effectively improve the visual quality of the watermarked DTIs.

The overall structure of the watermark-embedding network is shown in Fig. 2. First, the

global feature $\boldsymbol{F}_1$ is extracted from the DTIs by using 3D convolution. Then, the watermark signals are embedded into the feature $\boldsymbol{F}_1$ by channel concatenation. Then, the global feature is fused with the watermark signals by 3D convolution to obtain the global features of watermarked signals $\boldsymbol{F}_{\mathrm{e}}$. The specific calculation process is shown in Eq. (1), where "Cat" denotes the concatenation on the channel and "Conv3D" denotes 3D convolution.

$$\begin{cases} \boldsymbol{F}_1(\boldsymbol{x}) = \mathrm{Conv3D}(\boldsymbol{x}), \\ \boldsymbol{F}_2(\boldsymbol{x}) = \mathrm{Cat}\left(\boldsymbol{F}_1(\boldsymbol{x}), \boldsymbol{W}\right), \\ \boldsymbol{F}_{\mathrm{e}}(\boldsymbol{x}) = \mathrm{Conv3D}\left(\boldsymbol{F}_2(\boldsymbol{x})\right). \end{cases} \tag{1}$$

As shown in Fig. 2, the global features $\boldsymbol{F}_{\mathrm{e}}$ with watermark signals are divided into tensor features $\boldsymbol{F}_{\mathrm{e}\_i\_\mathrm{w}}$ ($i = 1, 2, \cdots, \mathrm{num}$, where num refers to the number of channels of 3D convolution). In the significant feature module, each tensor feature $\boldsymbol{F}_{\mathrm{e}\_i\_\mathrm{w}}$ combined with the T1w images is inputted into the encoder Transformer to calculate the most relevant significant feature of the T1w images and the tensor feature. The value of num can affect the quality of the reconstructed watermarked DTI to some extent.

The specific process is shown in Fig. 3. The encoder Transformer works by calculating the matrix multiplication of $\boldsymbol{Q}$ and $\boldsymbol{K}$ to obtain the correlation value $\boldsymbol{T}_{i,x,y}$, where x and y denote the position information after the image is converted into a vector, and i denotes the i^{th} tensor channel.

In each column, the maximum from the correlation value $\boldsymbol{T}_{i,x,y}$ is taken as the most relevant feature coefficient $\boldsymbol{S}_{i,x}$, and the index value of $\boldsymbol{S}_{i,x}$ is taken as the most relevant position information $\boldsymbol{P}_{i,x}$. Then, the significant feature $\boldsymbol{F}_{i,x}$ is obtained from $\boldsymbol{V}$ according to the most relevant position information $\boldsymbol{P}_{i,x}$. The calculation process is shown in Eqs. (2)–(4).

The significant feature $\boldsymbol{F}_{i,x}$ is selected from the tensor features $\boldsymbol{F}_{\mathrm{e}\_i\_\mathrm{w}}$ and the T1w images by using the idea of the hard attention mechanism. To

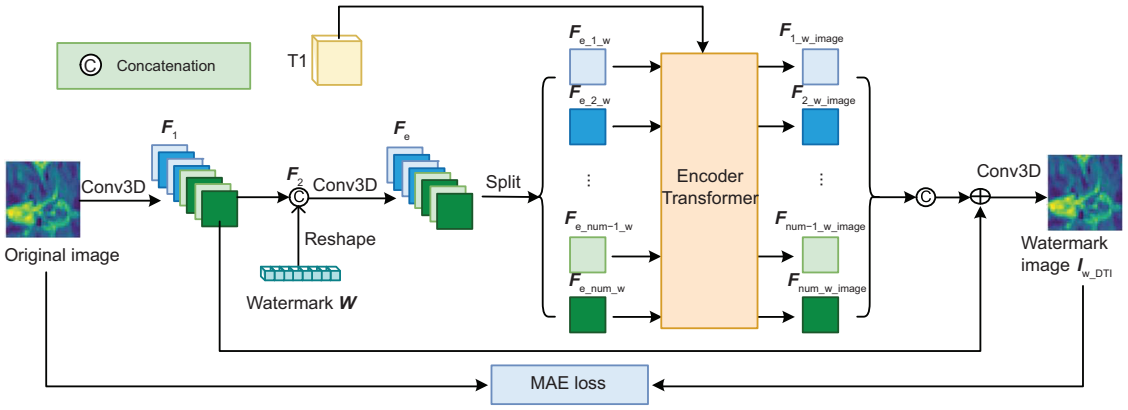

Fig. 2 Watermark-embedding network

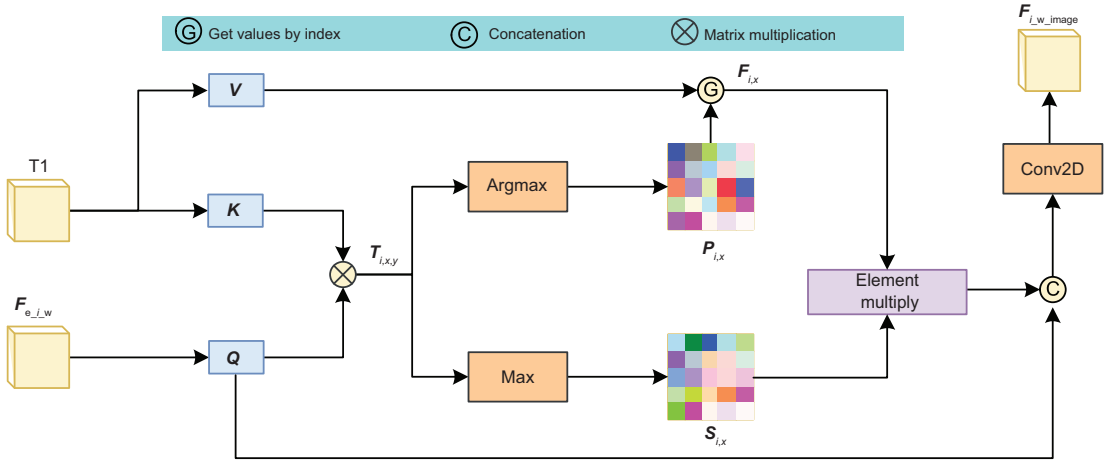

Fig. 3 Encoder Transformer

increase the confidence level of this feature, the idea of the soft attention mechanism is introduced, by multiplying the most relevant feature coefficient $\boldsymbol{S}_{i,x}$ as the weight of the significant feature $\boldsymbol{F}_{i,x}$ to obtain the most significant feature $\boldsymbol{F}\_\boldsymbol{S}_{i,x}$. Finally, the most significant feature $\boldsymbol{F}\_\boldsymbol{S}_{i,x}$ is fused with the tensor features $\boldsymbol{F}_{\mathrm{e}\_i\_\mathrm{w}}$ by convolution to obtain the reconstruction features $\boldsymbol{F}_{i\_\mathrm{w}\_\mathrm{image}}$. The specific calculation process is shown in Eqs. (4) and (5), where Conv2D denotes 2D convolution.

$$\boldsymbol{T}_{i,x,y}=\left\langle \frac{\boldsymbol{q}_x}{\|\boldsymbol{q}_x\|}, \frac{\boldsymbol{k}_y}{\|\boldsymbol{k}_y\|} \right\rangle, \tag{2}$$

$$\boldsymbol{P}_{i,x}=\operatorname{argmax}\left(\boldsymbol{T}_{i,x,y}\right),\ \boldsymbol{S}_{i,x}=\max\left(\boldsymbol{T}_{i,x,y}\right), \tag{3}$$

$$\boldsymbol{F}_{i,x}=\boldsymbol{V}[\boldsymbol{P}_{i,x}],\ \boldsymbol{F}\_\boldsymbol{S}_{i,x}=\boldsymbol{F}_{i,x}\boldsymbol{S}_{i,x}, \tag{4}$$

$$\boldsymbol{F}_{i\_\mathrm{w}\_\mathrm{image}}=\mathrm{Conv2D}\left(\mathrm{Cat}(\boldsymbol{F}\_\boldsymbol{S}_{i,x},\boldsymbol{F}_{\mathrm{e}\_i\_\mathrm{w}})\right), \tag{5}$$

where $\langle\cdot\rangle$ means calculating the cosine similarity, and $\boldsymbol{V}[\cdot]$ indicates the index of array $\boldsymbol{V}$.

Finally, the reconstruction features of each tensor channel are fused using 3D convolution to obtain the watermarked DTIs $\boldsymbol{I}_{\mathrm{w}\_\mathrm{DTI}}$, whose calculation process is shown in Eq. (6). The loss function is the mean absolute error (MAE) loss, shown in Eq. (7), which lets the network fully learn the image features of the original DTIs in the training process, where $\boldsymbol{I}_{\mathrm{o}}$ denotes the original DTIs.

$$\begin{aligned}\boldsymbol{I}_{\mathrm{w}\_\mathrm{DTI}}=&\mathrm{Conv3D}\big(\mathrm{Cat}(\boldsymbol{F}_{1\_\mathrm{w}\_\mathrm{image}},\boldsymbol{F}_{2\_\mathrm{w}\_\mathrm{image}},\\&\cdots,\boldsymbol{F}_{\mathrm{num}\_\mathrm{w}\_\mathrm{image}})\big).\end{aligned} \tag{6}$$

$$\begin{aligned}&L_{\mathrm{MAE}}\left(\boldsymbol{I}_{\mathrm{w}\_\mathrm{DTI}}\right)\\=&\frac{1}{CKB}\sum_{c=1}^{C}\sum_{k=1}^{K}\sum_{b=1}^{B}\left|\boldsymbol{I}_{\mathrm{w}\_\mathrm{DTI}(c,k,b)}-\boldsymbol{I}_{\mathrm{o}(c,k,b)}\right|.\end{aligned} \tag{7}$$

3.3 Watermark extraction network

Current robust watermarking algorithms based on deep learning usually target 2D or 3D images, and the watermark extraction networks of these algorithms cannot be directly used to extract watermark signals from the high-dimensional watermarked data. Through analysis of previous robust watermarking algorithms for deep learning, we find that most watermarking algorithms embed watermark signals as binary sequences. However, during training, those watermarking algorithms do not map the feature values in the interval [0, 1]. If the training uses the mean square error (MSE) loss with the output range of $(-\infty, +\infty)$, the extraction accuracy of the watermark is reduced. In our proposed algorithm, the output values are mapped to [0, 1] by the sigmoid function after extracting the watermark features. Furthermore, we propose to use cross-entropy loss to replace the MSE loss in the watermark extraction network. There are two main reasons. One is that MSE loss after the sigmoid function will generate a nonconvex optimization function, which easily leads to a locally optimal solution in the gradient descent process. The other is that cross-entropy loss decreases faster than the MSE loss gradient when the correct rate of the extracted watermark is low. Therefore, the cross-entropy loss makes the proposed network prefer to learn the high-dimensional features more adequately and robustly in DTI, which can improve the robustness of the proposed watermarking algorithm and the quality of the reconstructed DTI.

The multiheaded attention mechanism of Transformers can focus more effectively on the classified part of the image by calculating the correlation through global pixels. The idea of a multiheaded attention mechanism is introduced in the process of extracting image features. The decoder Transformer is proposed to extract the most significant features from the channels so that the watermark extraction network can pay more attention to the features embedded with watermark signals. Then, the semantic expression of the most significant features is enhanced to obtain watermark features through the convolution modules.

The specific extraction network framework is shown in Fig. 4, which is consistent with the idea of the watermark-embedding network. First, the global features $\boldsymbol{F}_{\mathrm{d}}$ are extracted from $\boldsymbol{I}_{\mathrm{w}\_\mathrm{DTI}}$ using Eq. (8). The features $\boldsymbol{F}_{\mathrm{d}}$ are divided into 3D channel features $\boldsymbol{F}_{\mathrm{d}\_i}$ from the tensor channels ($i=1,2,\cdots,\mathrm{num}$). Subsequently, the tensor features are inputted into the decoder Transformer, as shown in Fig. 5, to extract the most significant features.

$$\boldsymbol{F}_{\mathrm{d}}=\mathrm{Conv3D}(\boldsymbol{I}_{\mathrm{w}\_\mathrm{DTI}}). \tag{8}$$

First, the tensor channel features are projected by 3D convolution into three channels, $\boldsymbol{F}_{\mathrm{d}\_i\_1}$, $\boldsymbol{F}_{\mathrm{d}\_i\_2}$, and $\boldsymbol{F}_{\mathrm{d}\_i\_3}$. Then, the three channel features are linked using the decoder Transformer shown in Fig. 5; $\boldsymbol{F}_{\mathrm{d}\_i\_1}$, $\boldsymbol{F}_{\mathrm{d}\_i\_2}$, and $\boldsymbol{F}_{\mathrm{d}\_i\_3}$ are used for operations $\boldsymbol{Q}$, $\boldsymbol{K}$, and $\boldsymbol{V}$ in the Transformer

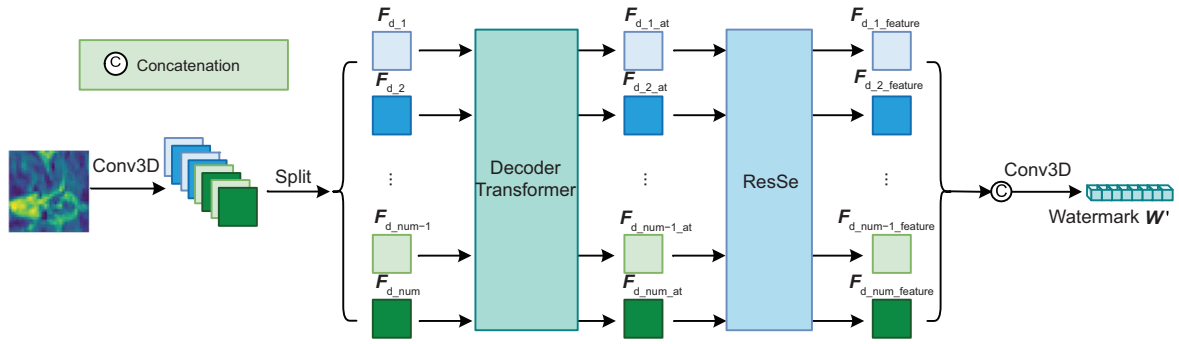

Fig. 4 Network extraction framework diagram

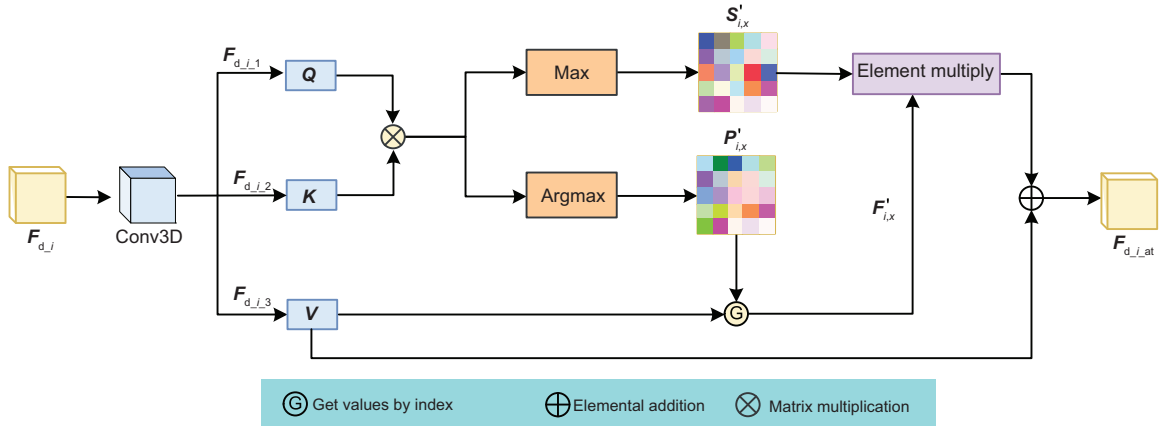

Fig. 5 Decoder Transformer

self-attention operation, respectively. Compared with ordinary Transformers, this method can make the sequence have spatial information in the convolution process.

Second, the Transformer block in Fig. 5 works as follows: the correlation $\boldsymbol{S}_{i,x}$ and the relevant position information $\boldsymbol{P}_{i,x}$ are first calculated using Eqs. (2)–(4). The significant feature is obtained from $\boldsymbol{F}_{\mathrm{d\_}i\mathrm{\_3}}$ using $\boldsymbol{P}_{i,x}$ as the index.

Third, to increase the confidence of the significant features, the most relevant scores $\boldsymbol{S}_{i,x}$ are calculated as the weights of $\boldsymbol{F}_{i,x}$ to obtain the most significant features.

Finally, to prevent feature loss, $\boldsymbol{F}_{\mathrm{d\_}i\mathrm{\_3}}$ is added to obtain the watermark feature map $\boldsymbol{F}_{\mathrm{d\_}i\mathrm{\_at}}$ that highlights the most significant watermark features, and its specific calculation process is shown in Eq. (9):

$$\boldsymbol{F}_{\mathrm{d\_}i\mathrm{\_at}} = \boldsymbol{F}'_{i,x}\boldsymbol{S}'_{i,x} + \boldsymbol{F}_{\mathrm{d\_}i\mathrm{\_3}}. \tag{9}$$

To obtain more watermark features, the semantic representation of image features needs to be improved by expanding the depth and width of the network. To effectively extract more semantic features that can express watermark signals from $\boldsymbol{F}_{\mathrm{d\_}i\mathrm{\_at}}$, we combine the residual mechanism of residual network (ResNet) (He et al., 2016), DenseNet (Huang et al., 2017), and SELayer (Hu et al., 2018) to form the ResSe module, which can extract the watermark features from $\boldsymbol{F}_{\mathrm{d\_}i\mathrm{\_at}}$ to obtain the watermark feature $\boldsymbol{F}_{\mathrm{d\_}i\mathrm{\_feature}}$ of the i^{th} tensor channel. The ResSe module consists of several ResSe blocks, and the ResSe block is shown in Fig. 6; here, $\boldsymbol{F}_{\mathrm{d\_}i\mathrm{\_at}}$ is the input, and $\boldsymbol{F}_{\mathrm{d\_}i\mathrm{\_feature}}$ is the output.

All watermark features $\boldsymbol{F}_{\mathrm{d\_}i\mathrm{\_feature}}$ are fused by 3D convolution, and the activating function is sigmoid. The calculation process is shown in Eq. (10). The proposed algorithm is inspired by the classification network, and it uses cross-entropy loss as the loss function of the watermark extraction network, which is calculated in Eq. (11). W_i represents the value of the i^{th} bit in the embedded watermark signal sequence, and W'_i represents the predicted value of the i^{th} bit in the extracted watermark signal sequence.

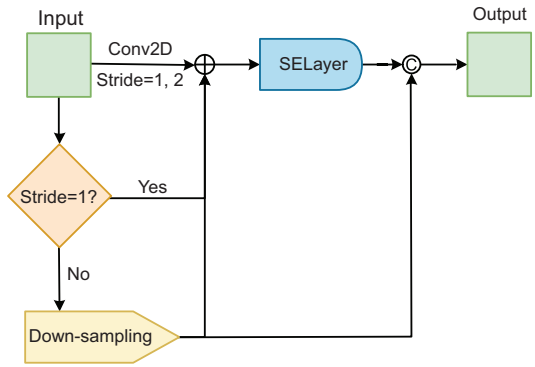

Fig. 6 ResSe block

$$\boldsymbol{W}' = \text{Sigmoid}\left(\text{Conv3D}(\text{Cat}(\boldsymbol{F}_{\text{d\_1\_feature}}, \boldsymbol{F}_{\text{d\_2\_feature}}, \cdots, \boldsymbol{F}_{\text{d\_num\_feature}}))\right), \tag{10}$$

$$L_{\text{BCE}}(\boldsymbol{W}') = \frac{1}{N}\sum_{i} -\left[W_i \log(W_i') + (1-W_i)\log(1-W_i')\right]. \tag{11}$$

3.4 Loss of network

This study focuses mainly on the end-to-end robustness of the watermarking algorithm by using neural networks. A large amount of watermark signals $\boldsymbol{W}$ is embedded through the watermark-embedding network to obtain the watermarked DTI $\boldsymbol{I}_{\text{w\_DTI}}$. Then, $\boldsymbol{I}_{\text{w\_DTI}}$ is attacked through the attack network to obtain $\boldsymbol{I}_{\text{w\_DTI\_noise}}$ to simulate the variety of attacks that may be encountered intentionally or unintentionally. The watermark signals $\boldsymbol{W}$ are extracted from $\boldsymbol{I}_{\text{w\_DTI\_noise}}$ through the watermark extraction network. The network is trained by using Eq. (12) as the total loss of the network, so that the network can guarantee both the visual reconstruction quality and the robustness of the watermarked DTI during the gradient descent process.

$$L_{\text{total}} = L_{\text{MAE}}(\boldsymbol{I}_{\text{w\_DTI}}) + L_{\text{BCE}}(\boldsymbol{W}'). \tag{12}$$

The purpose of $L_{\text{MAE}}(\boldsymbol{I}_{\text{w\_DTI}})$ is to make the watermarked DTI as similar as possible to the original image during the gradient descent, and the purpose of $L_{\text{BCE}}(\boldsymbol{W}')$ is to improve the accuracy of watermark extraction during the gradient descent. The two losses summed together for gradient descent can make the two networks compete with each other, thus effectively ensuring the robustness of the watermark.

4 Results of experiments

Data were provided by the WU-Minn Human Connectome Project (HCP) (van Essen et al., 2012) Consortium (principal investigators: David van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 National Institutes of Health (NIH) institutes and centers that support the NIH Blueprint for Neuroscience Research, and by the McDonnell Center for Systems Neuroscience at Washington University. From this HCP dataset, we used 185 patient DTIs as the training set and 16 patient DTIs as the test set. Because the complete DTI is a five-dimensional dataset of $145\times174\times145\times3\times3$ data points, it is difficult for the network to deal with such large images directly. Inspired by the cycle generative adversarial network (CycleGAN) (Anctil-Robitaille et al., 2021), each DTI is divided into 60 small-sized subimages of dimensions $32\times32\times32\times3\times3$ for training and testing. The watermark signals embedded in this study constitute a sequence of 8-bit binary values.

The experiments in this study are mainly to prove that the algorithm in this study can protect the copyright of DTIs without inflecting the clinical value of DTIs. Therefore, this study demonstrates that the algorithm can effectively protect the copyright of DTIs through experiments on watermarking robustness and then proves that watermarked DTIs have clinical value by comparing them with state-of-the-art DTI reconstruction works.

The experiments aim to demonstrate three main aspects: (1) verify the algorithm's effectiveness by ablation experiments to find the optimal network structure, (2) verify the clinical value of watermarked DTIs by our proposed algorithm, and (3) verify whether the embedded watermark is highly robust and whether the watermarked DTI can extract the embedded watermark signals through the extraction network after subjecting to various common image attacks.

4.1 Ablation experiments

4.1.1 Influence of the number of tensor feature channels on image quality

This subsection verifies whether the visual quality of the watermarked DTI can be improved by increasing the number of tensor feature channels. The experiments were performed in a setting without an

attack layer. The specific experimental results are shown in Table 1 (three sets of experiments whose corresponding numbers of channels num are 16, 32, and 64). It can be seen from Table 1 that as num increases, the network generates better-quality watermarked DTIs, but the time required for training is longer.

Table 1 Results of ablation experiments with network width

Num	Average PSNR (dB)	Training time (s)
16	47.93	43 500
32	49.75	81 542
64	50.47	144 836

Num: number of channels of 3D convolution; PSNR: peak signal-to-noise ratio

4.1.2 Fusion experiments with the most significant features

In this study, we use Eq. (5) to effectively fuse the significant features $\boldsymbol{F\_S}$ with the tensor features. Experiments were conducted using the network framework of Fig. 1, changing only the fusion of tensor features, and num is set to 16.

In this fusion process, we use two ways for experimental comparison; one is to use 3D convolution on the fusion features, as described in the previous-mentioned network, and the other is to directly use 2D convolution on the fusion features in the channel dimension. The specific experimental results are shown in Table 2. From the experimental results, it can be seen that using 2D convolution in the algorithm proposed in this study can yield better quality of reconstruction.

Table 2 Results of ablation experiments with two- and three-dimensional convolution

Convolution dimension	Average PSNR (dB)
3	39.58
2	41.72

PSNR: peak signal-to-noise ratio

4.1.3 Effectiveness of fusing T1w images by a Transformer

To verify the effectiveness of the most significant features and the soft attention mechanism in the proposed Transformer attention module on the reconstruction quality of watermarked DTIs, the ablation experiments shown in Table 3 are designed.

Through the comparative analysis of experiments 1 and 2, we find that directly fusing T1w images by convolution leads to a degradation in the visual quality of the watermarked DTIs. Through comparison of experiments 1, 3, 4, and 5, we find that using the Transformer of this study to fuse T1w images to generate the watermarked DTIs gives higher visual quality than the watermarked DTIs generated without introducing T1w images.

4.2 Medical value validation experiment of DTIs

Because there is no relevant robust blind DTI watermarking algorithm, we adopted only 3D convolution to substitute 2D convolution in the HiDDeN (Zhu et al., 2018). For comparison experiments, the DTIs and binary watermark signals were inputted into HiDDeN (Zhu et al., 2018), which is consistent with our proposed network. We compared HiDDeN (Zhu et al., 2018) with the proposed DTI watermarking algorithm. The visual quality of the DTIs is shown in Fig. 7. From the fourth and fifth columns of Fig. 7, it can be found that the watermarked DTI of HiDDeN (Zhu et al., 2018) changes more obviously in terms of the fractional anisotropy (FA) and tensor; therefore, it is necessary to pay attention to the visual quality of the tensor in the process of embedding the watermark signals. The experiments with complete DTIs are shown in Section 2 in the supplementary materials.

To better demonstrate whether watermarked DTIs have clinical value, we compared them with current state-of-the-art DTI reconstruction works on clinical metrics, as shown in Table 4. The SuperDTI (Li et al., 2021) reconstructs DTIs by using DWIs, which does not input binary watermark signals. We can see that the PSNRs of the FA image and the mean diffusivity (MD) image of HiDDeN (Zhu et al., 2018) are significantly smaller than the data from SuperDTI (Li et al., 2021). Therefore, the clinical metrics of HiDDeN (Zhu et al., 2018) are not up to the standard. Moreover, because our algorithm is more concerned about the transformation of tensor features, the PSNRs of our algorithm of both the FA image and the MD image are greater than the data of SuperDTI (Li et al., 2021). Based on the experimental results in this subsection, since HiDDeN

Table 3 Network width ablation experimental results

Experiment No.	Significant feature module	Soft attention	T1-weighted image	PSNR (dB)
1	×	×	×	39.38
2	×	×	✓	28.15
3	×	✓	✓	41.11
4	✓	×	✓	38.90
5	✓	✓	✓	41.72

PSNR: peak signal-to-noise ratio

Table 4 Comparison of the quality of watermarked DTIs

Serial number	Model	PSNR (dB)	nMSE		PSNR (dB)		α_{AC}
			FA	MD	FA	MD	
1	SuperDTI (Li et al., 2021)	–	0.021	0.001	38.9	38.39	–
2	HiDDeN (Zhu et al., 2018)	29.03	0.096	0.026	24.76	26.17	0.5605
3	Our-num(64)	42.57	0.012	0.00 015	43.44	55.27	0.9385

α_{AC}: main axis deflection angle; DTI: diffusion tensor image; FA: fractional anisotropy; MD: mean diffusivity; nMSE: normalized mean square error; PSNR: peak signal-to-noise ratio

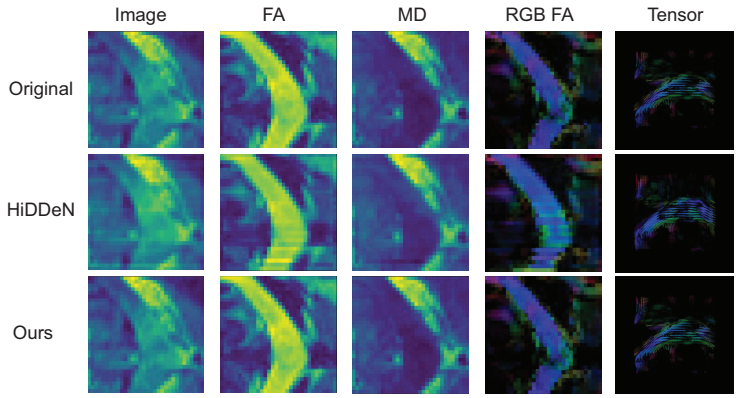

Fig. 7 Results of network width ablation experiments

(Zhu et al., 2018) is not a watermarking algorithm specifically for DTI, the watermarked DTIs generated by HiDDeN (Zhu et al., 2018) lose their clinical value. In our algorithm, we use the Transformer to fuse the DTI's structural information from the T1w image. Therefore, our algorithm can better protect the clinical value of the watermarked DTI.

4.3 Robustness of watermarking

In this study, the correctness of the watermark signals is calculated using a stricter criterion compared to the bit error rate (BER). In N test set images, for the i^{th} image-embedded watermark W_i, the extracted watermark signals are W_i'; when W_i and W_i' are exactly the same, it means that this image watermark extraction success count is +1, and finally the watermark extraction accuracy of this round of testing is the ratio of the total number of successful watermark signal extractions m to the number of test images N. The specific calculation formula is shown in Eq. (13):

$$\text{acc} = \frac{1}{N}\sum_{i}(W_i == W_i')\,. \tag{13}$$

In this study, the proposed algorithm was trained via two models to prove the robustness of watermarking against various attacks: one was trained by a single attack, and the other was trained by multiple attacks. The accuracy of watermark extraction is shown in Table 5. The proposed algorithm exhibited excellent robustness for various attacks. When the cropping attack left only 0.1 of the watermarked DTI, the accuracy of watermark extraction could reach 1.0; when the local cropout attack replaces the watermarked DTI with 0.8 pixels, i.e., 0.2 pixels are retained, the accuracy of watermark extraction can

Table 5 Accuracy of watermark extraction under signal attacks

Attack type	Accuracy (%)								
	p=0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Crop (single attack)	100	100	100	100	100	100	100	100	100
Cropout (single attack)	56.66	99.79	100	100	100	100	100	100	100
Dropout (single attack)	99.58	100	100	100	100	100	100	100	100
Crop (multiple attacks)	100	100	100	100	100	100	100	100	100
Cropout (multiple attacks)	40.20	98.95	100	100	100	100	100	100	100
Dropout (multiple attacks)	99.79	100	100	100	100	100	100	100	100

Attack type	Accuracy (%)			Attack type	Accuracy (%)		
	r=3	5	7		σ=0.01	0.02	0.03
Gauss blur (single attack)	100	99.37	99.37	Gauss noise (single attack)	100	100	99.79
Gauss blur (multiple attacks)	100	99.37	99.37	Gauss noise (multiple attacks)	100	100	100

be up to 0.9895. For Gaussian filtering, when the size of the Gaussian low-pass filter was 7, it could still accurately extract watermark signals. When the Gaussian noise was added into the watermarked DTI, the watermark signals could still be extracted accurately.

The training model of multiple attacks still exhibited excellent robustness for various attacks: for example, the watermark signals were still extracted accurately in the presence of Gaussian blur and Gaussian noise. Based on these two results in this subsection, we concluded that the watermark embedded by the proposed algorithm has extremely good robustness.

5 Conclusions

This paper solves the problem of robust watermarking algorithms not protecting DTIs. In the watermark-embedding network, there is a large amount of information in the T1w images which is not related to the tensor features. The tensor features in the T1w images need to be extracted by the coding Transformer proposed in this paper for fusion to bring a positive impact on the network. Cross-entropy loss is used in the watermark extraction network, instead of the MSE loss, to reduce the tolerance of the network to wrong watermark signals. Convolution is introduced in the Transformer to ensure the expressiveness of the feature space while extracting the most significant watermark features. The attention mechanism of the residual channel is used to improve the features' semantic expressiveness and the accuracy of watermark extraction.

The experimental results show that the watermarked DTI fully met the requirements of clinical medical diagnosis. The algorithm exhibits excellent robustness against several common attacks and can effectively protect the copyright information of DTIs. Since the DTI is a high-dimensional one, the algorithm in this paper takes up a lot of computational resources when the complete DTI is inputted during the training process. Therefore, we will do further research to solve this problem in future work.

Contributors

Chengmeng LIU designed the research. Long ZHENG processed the data. Chengmeng LIU drafted the paper. Zhi LI and Guomei WANG helped organize the paper. Zhi LI revised and finalized the paper.

Conflict of interest

All the authors declare that they have no conflict of interest.

Data availability

The medical image data used in this paper were obtained from the MGH-USC Human Connectome Project (HCP) database (https://ida.loni.usc.edu/login.jsp). The other data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Ahmadi M, Norouzi A, Karimi N, et al., 2020. ReDMark: framework for residual diffusion watermarking based on deep networks. *Exp Syst Appl*, 146:113157. https://doi.org/10.1016/j.eswa.2019.113157

Amini M, Ahmad MO, Swamy MNS, 2018. A robust multibit multiplicative watermark decoder using a vector-based hidden Markov model in wavelet domain. *IEEE Trans*

Circ Syst Video Technol, 28(2):402-413. https://doi.org/10.1109/tcsvt.2016.2607299

Anand D, Niranjan UC, 1998. Watermarking medical images with patient information. Proc 20th Annual Int Conf of the IEEE Engineering in Medicine and Biology Society, p.703-706. https://doi.org/10.1109/iembs.1998.745518

Anctil-Robitaille B, Desrosiers C, Lombaert H, 2021. Manifold-aware CycleGAN for high-resolution structural-to-DTI synthesis. Proc Int MICCAI Workshop on Computational Diffusion MRI, p.213-224. https://doi.org/10.1007/978-3-030-73018-5_17

Cintra RJ, Cooklev TV, 2009. Robust image watermarking using non-regular wavelets. *Signal Image Video Process*, 3(3):241-250. https://doi.org/10.1007/s11760-008-0070-7

Dosovitskiy A, Beyer L, Kolesnikov A, et al., 2020. An image is worth 16×16 words: Transformers for image recognition at scale. https://arxiv.org/abs/2010.11929

Feng CM, Yan Y, Fu H, et al., 2021. Task transformer network for joint MRI reconstruction and super-resolution. 24th Int Conf on Medical Image Computing and Computer-Assisted Intervention, p.307-317.

He KM, Zhang XY, Ren SQ, et al., 2016. Deep residual learning for image recognition. Proc IEEE Conf on Computer Vision and Pattern Recognition, p.770-778. https://doi.org/10.1109/CVPR.2016.90

Hu J, Shen L, Sun G, 2018. Squeeze-and-excitation networks. Proc IEEE/CVF Conf on Computer Vision and Pattern Recognition, p.7132-7141. https://doi.org/10.1109/cvpr.2018.00745

Huan WN, Li S, Qian ZX, et al., 2022. Exploring stable coefficients on joint sub-bands for robust video watermarking in DT CWT domain. *IEEE Trans Circ Syst Video Technol*, 32(4):1955-1965. https://doi.org/10.1109/tcsvt.2021.3092004

Huang G, Liu Z, van der Maaten L, et al., 2017. Densely connected convolutional networks. Proc IEEE Conf on Computer Vision and Pattern Recognition, p.2261-2269. https://doi.org/10.1109/cvpr.2017.243

Lai ZY, Qu XB, Liu YS, et al., 2016. Image reconstruction of compressed sensing MRI using graph-based redundant wavelet transform. *Med Image Anal*, 27:93-104. https://doi.org/10.1016/j.media.2015.05.012

Le Bihan D, Mangin JF, Poupon C, et al., 2001. Diffusion tensor imaging: concepts and applications. *J Magn Reson Imag*, 13(4):534-546. https://doi.org/10.1002/jmri.1076

Lee HK, Kim HJ, Kwon KR, et al., 2005. ROI medical image watermarking using DWT and bit-plane. Proc Asia-Pacific Conf on Communications, p.512-515. https://doi.org/10.1109/apcc.2005.1554112

Li HY, Liang ZF, Zhang CY, et al., 2021. SuperDTI: ultrafast DTI and fiber tractography with deep learning. *Magn Reson Med*, 86(6):3334-3347. https://doi.org/10.1002/mrm.28937

Liu XL, Lin CC, Yuan SM, 2018. Blind dual watermarking for color images' authentication and copyright protection. *IEEE Trans Circ Syst Video Technol*, 28(5):1047-1055. https://doi.org/10.1109/tcsvt.2016.2633878

Luo XY, Zhan RH, Chang HW, et al., 2020. Distortion agnostic deep watermarking. Proc IEEE/CVF Conf on Computer Vision and Pattern Recognition, p.13545-13554. https://doi.org/10.1109/cvpr42600.2020.01356

Nakarmi U, Wang Y, Lyu J, et al., 2017. A kernel-based low-rank (KLR) model for low-dimensional manifold recovery in highly accelerated dynamic MRI. *IEEE Trans Med Imag*, 36(11):2297-2307. https://doi.org/10.1109/tmi.2017.2723871

Ni ZC, Shi YQ, Ansari N, et al., 2006. Reversible data hiding. *IEEE Trans Circ Syst Video Technol*, 16(3):354-362. https://doi.org/10.1109/tcsvt.2006.869964

Ravishankar S, Bresler Y, 2011. MR image reconstruction from highly undersampled k-space data by dictionary learning. *IEEE Trans Med Imag*, 30(5):1028-1041. https://doi.org/10.1109/tmi.2010.2090538

Shin PJ, Larson PEZ, Ohliger MA, et al., 2013. Calibrationless parallel imaging reconstruction based on structured low-rank matrix completion. *Magn Reson Med*, 72(4):959-970. https://doi.org/10.1002/mrm.24997

Singh AK, Dave M, Mohan A, 2015. Hybrid technique for robust and imperceptible multiple watermarking using medical images. *Multim Tools Appl*, 75(14):8381-8401. https://doi.org/10.1007/s11042-015-2754-7

Stejskal EO, Tanner JE, 1965. Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *J Chem Phys*, 42(1):288-292. https://doi.org/10.1063/1.1695690

Su QT, Sun YH, Zhang XT, et al., 2022. A watermarking scheme for dual-color images based on URV decomposition and image correction. *Int J Intell Syst*, 37(10):7548-7570. https://doi.org/10.1002/int.22893

van Essen D, Ugurbil K, Auerbach E, et al., 2012. The Human Connectome Project: a data acquisition perspective. *NeuroImage*, 62(4):2222-2231. https://doi.org/10.1016/j.neuroimage.2012.02.018

Vaswani A, Shazeer N, Parmar N, et al., 2017. Attention is all you need. Proc 31st Int Conf on Neural Information Processing Systems, p.6000-6010. https://doi.org/10.5555/3295222.3295349

Wang CP, Wang XY, Xia ZQ, et al., 2019. Ternary radial harmonic Fourier moments based robust stereo image zero-watermarking algorithm. *Inform Sci*, 470:109-120. https://doi.org/10.1016/j.ins.2018.08.028

Wang N, Li Z, Cheng XY, et al., 2018. Diffusion weighted image reversible visible watermarking algorithm based on support vector regression. Proc 14th IEEE Int Conf on Signal Processing, p.1144-1148. https://doi.org/10.1109/icsp.2018.8652283

Wen BY, Aydore S, 2019. ROMark: a robust watermarking system using adversarial training. https://arxiv.org/abs/1910.01221

Wen Q, Sun TF, Wang SX, 2003. Concept and application of zero-watermark. *Acta Electon Sin*, 31(2):214-216 (in Chinese). https://doi.org/10.3321/j.issn:0372-2112.2003.02.015

Wu HP, Xiao B, Codella N, et al., 2021. CvT: introducing convolutions to vision Transformers. Proc IEEE/CVF Int Conf on Computer Vision, p.22-31. https://doi.org/10.1109/iccv48922.2021.00009

Xia ZQ, Wang XY, Li XX, et al., 2019. Efficient copyright protection for three CT images based on quaternion polar harmonic Fourier moments. *Signal Process*, 164:368-379. https://doi.org/10.1016/j.sigpro.2019.06.025

Yang FZ, Yang H, Fu JL, et al., 2020. Learning texture transformer network for image super-resolution. Proc IEEE/CVF Conf on Computer Vision and Pattern Recognition, p.5790-5799. https://doi.org/10.1109/cvpr42600.2020.00583

Zhan ZF, Cai JF, Guo D, et al., 2016. Fast multiclass dictionaries learning with geometrical directions in MRI reconstruction. *IEEE Trans Biomed Eng*, 63(9):1850-1861. https://doi.org/10.1109/tbme.2015.2503756

Zhou B, Zhou SK, 2020. DuDoRNet: learning a dual-domain recurrent network for fast MRI reconstruction with deep T1 prior. Proc IEEE/CVF Conf on Computer Vision and Pattern Recognition, p.4272-4281. https://doi.org/10.1109/cvpr42600.2020.00433

Zhu JR, Kaplan R, Johnson J, et al., 2018. HiDDeN: hiding data with deep networks. Proc 15th European Conf on Computer Vision, p.682-697.

List of supplementary materials