IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received August 18, 2020, accepted September 1, 2020, date of publication September 3, 2020, date of current version September 21, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3021523

Usage of Deep Learning and Blockchain in Compilation and Copyright Protection of Digital Music

ZHINI CAI
School of Arts, Hunan City University, Yiyang 413000, China
e-mail: czn830829@sina.com

ABSTRACT In order to explore the application of deep learning algorithms in arrangement and composition, and the role of blockchain in the protection of digital music copyright, a monophonic melody composition model based on the deep generative adversarial networks (DCGANs) is constructed firstly, and the composition performance of the model is analyzed using hymn as input sample in this study. Later, the multi-instrument co-arrangement (MICA) model based on the multi-task learning is proposed, and the composition performance is analyzed by taking the actual music as an input sample. Finally, the improved practical byzantine fault tolerance (IPBFT) algorithm is proposed, and a digital music copyright protection system is designed based on the blockchain in this study. The results indicate that the accuracies constructed DCGANs model in predicting the Soprano and Alto voice melody are higher than those of the DeepBatch model by 2.29% and 3.32%, respectively. The performance on the harmony score, note accuracy, levenshtein similarity (LS), notes distribution mean square error, and empty as well as the convergence speed of the constructed MICA model are better than those of other models. The average transaction per second (TPS) value of the proposed IPBFT algorithm in the real digital music copyright protection system is 3469, which is superior to other blockchain technologies. Finally, the digital music copyright protection system is achieved, the error rate of completing the request is 0% in the state of many users operating concurrently, and a high TPS value can be guaranteed. In short, the DCGANs and MICA models pointed out in this study can be used in the composition of monophonic melodies and complex melodies, and the digital music copyright protection system based on the blockchain has excellent performance in practical applications.

INDEX TERMS Deep generative adversarial networks, multi-instrument co-arrangement, practical byzantine fault tolerance, blockchain, digital music copyright protection system.

I. INTRODUCTION

Music has been deeply integrated into our daily life. The creation of music requires a lot of professional knowledge, so that it is difficult for ordinary people to complete the music creation, and professional musicians have to consume a lot of time and energy to create music [1]. Therefore, the use of computer technology to complete the music creation has become the focus of research. Occurrence of deep learning realizes the computer intelligent learning to a great content [2]. In addition, there have been many studies combining the art creation and artificial intelligence technology, and the application of deep learning algorithms in the field of music has gradually become a major trend of research [3], [4]. If the algorithms are used to compose music, it requires that the generated notes have to be correlated with each other to a certain extent [5]. The multi-track music requires cooperation of different instruments, so how to ensure the harmony of the multi-track melody is an important problem to be solved [6]. Existing music data can be divided into waveform music data (such as MP3, etc.) and musical instrument digital interface. In the generation and processing of the waveform music data, the signals obtained by simulation have lower accuracy, so the accuracy after composition can't be guaranteed [7], [8]. Therefore, the digital music data is selected in this study to solve the problem of sequence modeling.

The associate editor coordinating the review of this manuscript and approving it for publication was Mohammad Zia Ur Rahman.

With the continuous development of digital technology, the number of music distributions and transactions on the

Internet is higher and higher, but the current digital copyright protection mechanism is not perfect [9]. The key issue in the digital copyright protection technology is to solve the digital content dissemination and transaction process, and many companies have also proposed the system directing to the digital audio copyright protection [10]. Most solutions are based on the authorization and verification of digital content in third-party devices, or to prevent illegal operation of the digital content through some special techniques [11], [12]. In the research of copyright protection, blockchain has received extensive attention from experts and scholars, and there are also some studies that prove that the application of blockchain in the construction of digital copyright protection system can achieve excellent results [13], [14]. However, there is currently no research on applying blockchain to the copyright protection of digital music, so this article aims to improve the efficiency of deep learning algorithms used in music arrangement, enhance the music quality of algorithmic arrangement, and strengthen the security performance of digital music copyright protection.

The rest of this article is organized as follows: section 2 discusses the application of deep learning in music arrangement and the research progress of digital copyright protection; section 3 proposes the deep learning models for single-track and multi-track arrangement and a system for digital music copyright protection based on the blockchain; section 4 verifies the proposed arrangement model and digital music copyright protection system, which are compared with other advanced algorithms; and section 5 concludes the article with summary and future research directions.

II. LITERATURE REVIEW

A. APPLICATION PROGRESS OF DEEP LEARNING IN ARRANGEMENT

Music plays an important role in daily lives of human beings. With the rapid development of deep learning and modern generation technology, experts and scholars have done a lot of work in automatic music generation. Kereliuk *et al.* (2015) applied deep learning in the classification of music content, and found that the convolutional network used for individual classification of audio frames has stronger performance [15]. Ohtani *et al.* (2016) firstly used artificial intelligence technology and deep learning algorithms to re-synthesize the trajectories of guitar and drums in audio, and randomly optimized the synthesized parameters according to the personal preferences of the listeners; and the algorithm finally obtained an excellent supervisor evaluation score [16]. Briot *et al.* (2017) built a value music style automatic learning and generation model based on the deep learning, and found that deep learning algorithms can automatically generate music without human interaction, but there were still some problems of control, creativity, and interaction [17]. Dean *et al.* (2018) applied deep learning to generate music, and found that the synthesized music was roughly the same as the example [18]. Zhu *et al.* (2020) constructed a multi-instrument collaborative arrangement model by using the deep learning and multi-task learning methods, and verified the superiority and effectiveness of the model on real data sets [19]. However, most of the existing researches on music generation methods applied to the generation of multi-track music have great limitations. In addition, no better solutions have been found for factors related to the quality of the generated music, such as chords, rhythm, and music style.

B. RESEARCH STATUS OF DIGITAL COPYRIGHT PROTECTION

The overall scale of China's digital publishing industry has exceeded 750 billioYuan. The digitization of copyright can promote the industry reform, but it also increases the risk of infringement. The traditional data rights management (DRM) initially only focused on encryption and authorization, while the modern DRM technology has become a system engineering throughout the entire life cycle of digital works. Guo *et al.* (2016) indicated that strict copyright protection policies may lead to higher price settings, thereby reducing the company's profitability, but it can help improve the product quality [20]. Murali *et al.* (2018) proposed a digital copyright protection scheme based on Region of Interest and Singular Value Decomposition and Orthogonal Polynomials Transformation; it was found that the algorithm can replace the main sharing mode with secret images with the ownership sharing mode, and can improve the robustness of image manipulation and attack [11]. Devi *et al.* (2016) proposed a watermarking algorithm based on shuffled singular value decomposition and the visual cryptography, and applied it to the copyright protection of digital images; as a result, it was proved that the algorithm can clearly verify the copyright of digital images and can resist many kind of image processing attacks, with a better effect than the visual cryptography algorithm [21]. Liang *et al.* (2019) built a blockchain security system by using the distributed communication network technology and multimedia data security technology; after verification, it was found that this method had excellent data security protection capability [22]. Existing research mainly focuses on the copyright protection of digital images or data, and there are relatively few researches on the copyright protection of digital music. The research on blockchain technology is mostly focused on privacy protection, and the advantages of applying blockchain to digital music copyright protection still need to be explored.

Therefore, in order to make up for the above-mentioned gaps, a generation model of single-track melody based on the DCGAN and a multi-track melody generation model based on MICA are proposed to improve the generation quality of different track melody. Subsequently, the BPFT algorithm in the blockchain is improved and applied to the construction of the digital music copyright registration system, so as to improve the security performance of digital music copyright protection.

III. METHODOLOGY

Firstly, the DCGAN is constructed for the preparation of monophonic melody, and then the MICA model based on multi-task learning is constructed for the preparation of multi-tone melody. Subsequently, the PBFT consensus algorithm in the blockchain is improved, and a registration system for digital music copyright protection is proposed based on the improved blockchain. The overall structure of this research is shown in Figure 1.

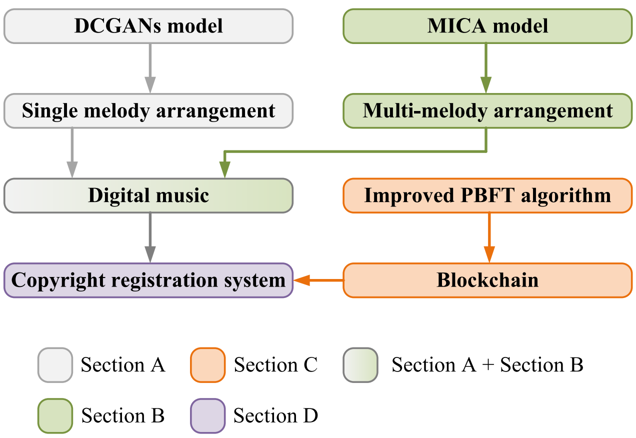

FIGURE 1. Overall structure of this research.

A. CONSTRUCTION OF COMPOSITION MODEL BASED ON MONOPHONIC MELODY UNDER THE DEEP CONVOLUTION GENERATIVE ADVERSARIAL NETWORKS

GANs include the generative model and discriminant model. DCGANs is a model generated by combing the structures of generative adversarial network (GAN) and convolutional neural network (CNN), which realizes the qualitative gap for the effect of GANs [23]. The generative model G can learn the true distribution probability of the data in the training set, mainly to convert the random noise input into the model into "music", so that the generated music is more similar to the music in the training set. The discriminant model D is used to judge whether the generated music is the real music, mainly to distinguish the generated music from the real music in the training set.

The relative pitch is used to encode the notes and each part is modeled separately in this study. However, when using DCGANs for algorithm composition, only single-track problems are considered, and a data set containing 6 lists $(v_1, v_2, v_3, v_4, s, f)$ is used to express the hymn. Where, v is the voice part (1 is soprano, 2 is alto, 3 is tenor, and 4 is bass), s is the beat, and f is the stress. These 6 lists have the same size when they are indexed based on the time t.

In order to simply the model generation process, it is assumed that the data set contains only 1 hymn, and the duration is T, and the voice part is v. Then a mathematical expression of the conditional probability distribution that can be parameterized by θ can be defined as:

$$\left\{p_{n.t}\left(v_n^t \left| v'_{n.t}, s, f, \theta_{n,t}\right.\right)\right\}_{n\in[4],t\in[T]} \quad (1)$$

In the above equation, v_n^t is the note of the n part at the time t; $v'_{n.t}$ is the other variable except $v_{n.t}$ in v.

In order to ensure the invariance principle of the time in the model and guarantee the model adapts to sequences of different lengths, the conditional probability models of all the same voices have to share the same parameters, that is, the parameters of the same voices at different times are the same [24]. Then, the conditions for the parameter θ and conditional probability distribution p are as follows:

$$\theta_n = \theta_{n,t} \quad p_n = p_{n,t}\ \forall t \in [T], \quad i \in [4] \quad (2)$$

In order to obtain the optimal parameters, the log likelihood of the voice can be maximized, but the weight is shared, so the mathematical expression of the voice can be written as:

$$\max_{\theta_n} \sum_t \log p_n\left(v_t^n \left| v'_{n,t}, s, f, \theta_n\right.\right) \quad (3)$$

Through the above equation, the melody information around the current note can be used as the notes to predict the current state and time, and this range is given according to the style of the music.

With the purpose of obtaining more accurate prediction results, the time sequence of the music data has to be considered. In this study, four neural networks are the generation models of the four voices. For the Bach-style automatic composition, the note at time t is not only related to the music information of the four voices around the time, and the music information of the other three parts at this time also affect the notes at other times. Thus, the basic structure of the generative model G formed in this study is shown in Fig. 2. The neural network in this model is composed of 4 Dense layers, which can map to the space with the same dimension based on the features of the current time, the previous time, and the latter time. The features of the current time had to go through 2 Dense layers, and the output results of the subsequent 3 parts are connected into a sequence. After obtaining the sequence, the model is required to predict the notes at the current time, and the two-way mode has to be considered. Thus, the long-short term memory (LSTM) [25] is selected as the basic structure of the second part of the network in this study, then the output at a single time during the pre-training is $p_i\left(v_n^t \left| v'_{t,n}, s, f, \theta\right.\right)$, and the output of all the information of a voice with a length of T after training GANs is $p_i\left(v_n \left| v'_n, s, f, \theta\right.\right)$. The length is set to 32 (2 music fragments) in this study, the LSTM recurrent network is used to analyze the music at current time, previous time, and later time, and then the current time information after prediction is obtained finally. The Dense of part 3 and 4 is used to predict the notes obtained during the pre-training.

The process of generating the melody requires to use the Gibbs sampling method to sample from the generated model G, and the process of generating the model G using the Gibbs algorithm is shown in Fig. 3. The implementation process is

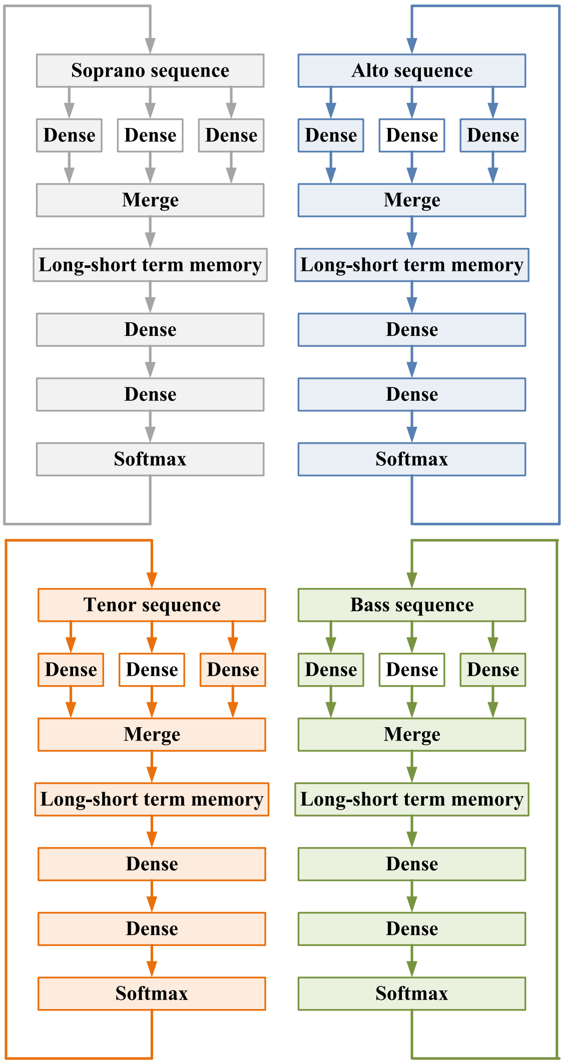

FIGURE 2. Basic structure of melody generative model.

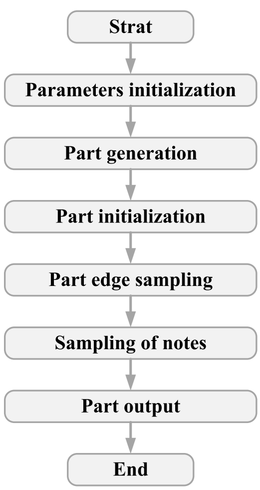

FIGURE 3. Implementation flow of the algorithm for the generative model based on the Gibbs sampling.

as follows: initialization of each parameter; generation of the hymn part V of the length D; initialization of V and sampling from the edge of each note; random selection of a voice and a time, and sampling the notes in the conditional probability distribution data set; outputting the V after completion.

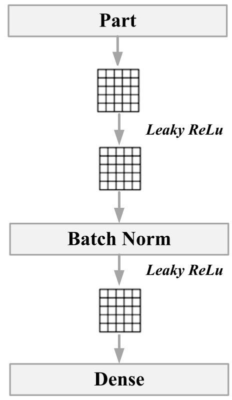

FIGURE 4. Basic structure of melody discriminant model.

The discriminant model D constructed in this study is composed of 3 layers of DNN and 1 layer of Dense (fully connected layer), and the input dimensions of different discriminant models D are different. The basic structure of the specific discriminant model D is shown in Fig. 4. It can be seen that the basic structure of the model contains three layers of CNN, and the connection activation functions between the CNN layers are all Leaky ReLus. The second layer of CNN has to use the batch normalization (BN) for processing [26], [27]. The final layer of CNN inputs the output structure to the Dense layer, and the activation function in the Dense is the ReLu function.

The input to the discriminant model includes the real melodies and the melodies generated by the generative model G. If the input sample is judged as a true melody by the discriminant model D, the output label is defined as 1; if the input sample is judged as a generated melody by the discriminant model D, the output label is defined as 0.

The goal of the generative model G is to convert the random noise sampled from the prior distribution into the distribution state of the real data, while the goal of the discriminant model D is to distribute the real data and the generated data. During the training, the model G and D have to be updated by the minimum and maximization equation, and the mathematical expression of which is given as below:

$$\min_G \max_D E_{v_n \sim p_{data}} \left[\log \left(D_n \left(v_n\right)\right)\right] + E_{\overline{v_n} \sim p_{G_n}} \left[1 - \log \left(D_n \left(\overline{v}_n\right)\right)\right] \quad (4)$$

In the above equation, v_n is the distribution of real data in a melody of a voice obtained by sampling in p_{data}; and $\overline{v}_n$ is the distribution obtained by sampling in p_{Gn}.

After treatment with above equation (4), the distribution of p_{data} and p_{Gn} can can become more similar, so that the Nash balance between the two can be achieved. In addition, the implementation process of the discriminant model algorithm is shown in Fig. 5, and the specific process is: initialization is achieved by using the real melody, predicted melody, and

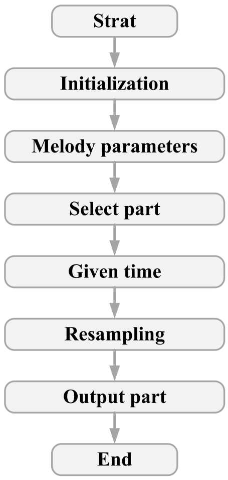

FIGURE 5. Implementation flows of discriminant model algorithm.

the distribution of note data of each part; the time, beat, and stress of the hymns are input; randomly selecting a part, and giving the sequence V with timing length as L; *vn* is sampled from *pn* again; the operations are finished and the result V is output.

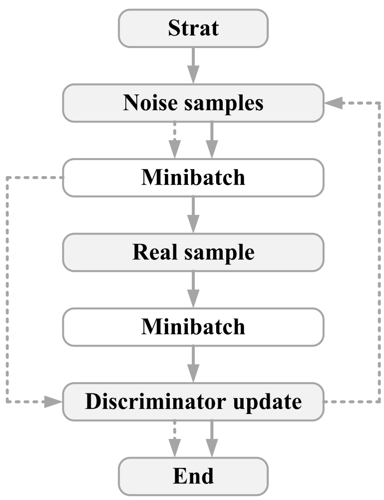

FIGURE 6. Training flows of DCGANs model (the solid and dashed line represents a calculating flow respectively).

The determined model D is trained together with the model G. In this study, k is defined as 1 (the two models have to be updated after each training until they are converged). The training process during the DCGANs training is shown in Fig. 6. The specific process of DCGANs training is as follows: m noise samples of the $p_{Gn}(z_n)$ are distributed and the minibatch is selected; m real samples of the p_{data} are distributed based on the data, and the minibatch is selected; the discriminator is updated with the grid search algorithm, and the update equation is given as $\nabla_{\theta_{D_n}} \frac{1}{m} \sum_{i=1}^{m} \left[\log D_n\left(V_n^{(i)}\right) + \log\left(1 + D_n\left(G_n\left(z_n^{(i)}\right)\right)\right)\right]$; m noise samples of $p_{Gn}(z_n)$ are distributed based on the noise, and the minibatch is selected; the generator is updated with the grid search algorithm, and the update equation is $\nabla_{\theta_{G_n}} \frac{1}{m} \sum_{i=1}^{m} \log\left(1 - D_n\left(G_n\left(z_n^{(i)}\right)\right)\right)$.

B. CONSTRUCTION OF MUTI-INSTRUMENT CO-ARRGANGEMENT MODEL BASED ON MULTI-TASK LEARNING

Popular music often contains multiple tracks and instruments, and each track or instrument can be regarded as a sequence [6]. Thus, a MICA model based on multi-task learning is proposed in this study based on the characteristic relationship among the learning tracks in the process of music generation. The model contains two types of cells, namely attention cell and multlayer perception (MLP) cell. The basic structure of the model is shown in Fig. 7 below.

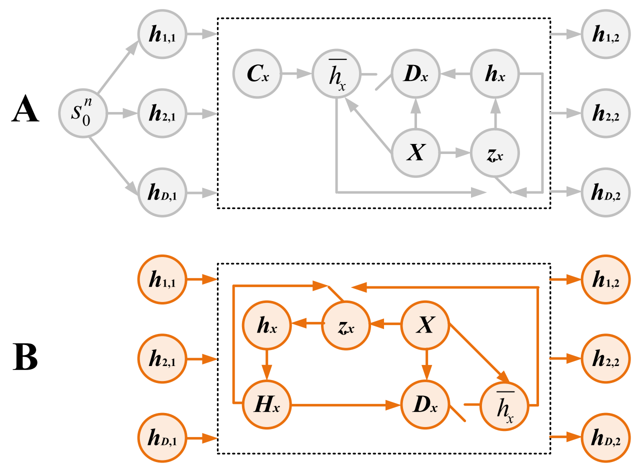

FIGURE 7. Basic structure of MICA model based on multi-task learning. A is attention cell, and B is MLP cell.

Attention cell is used to capture the most closely related relationship between status of the current task and the status of other tasks. The mathematical representation of the attention mechanism can be expressed as equation (5) below:

$$\begin{aligned} \alpha_{d,n}^{x} &= \sum_{y=1}^{D} \alpha_{d,x,y} h_{d,n-1}^{y} \\ e_{n,x,y} &= v^{D} \tanh\left(W h_{d,n-1}^{y} + V h_{d,n-1}^{y}\right) \\ \alpha_{n,x,y} &= \frac{\exp\left(e_{n,x,y}\right)}{\sum_{s=1}^{D} \exp\left(e_{n,x,s}\right)} \end{aligned} \tag{5}$$

In the above equation, $\alpha_{d,n}^{x}$ is the cooperation vector of the n^{th} step in the x tasks in the section d; and $h_{d,n-1}^{y}$ is the hidden state of the n-1^{th} step in the task y in the d section d.

The gated recurrent unit (GRU) is revised based on the obtained cooperation vector, so that the impacts of other track information can be considered when the current track is generated, and then the revised GRU unit is:

$$\begin{aligned} r_{d,n}^{x} &= \sigma\left(W_r^x u_{d,n}^x + V_r^x h_{d,n-1}^y + A_r^x a_{d,n}^x + b_r^x\right) \\ z_{d,n}^{x} &= \sigma\left(W_z^x u_{d,n}^x + V_z^x h_{d,n-1}^y + A_z^x a_{d,n}^x + b_z^x\right) \\ \overline{h_{d,n}^{x}} &= \sigma\left(W^x u_{d,n}^x + V^x \left[r_{d,n}^x h_{d,n-1}^y\right] + A^x a_{d,n}^x + b^x\right) \\ h_{d,n}^{x} &= \left(1 - z_{d,n}^x\right) h_{d,n-1}^y + z_{d,n}^x \overline{h_{d,n}^x} \end{aligned} \tag{6}$$

In the above equations, σ is activation function; W_r^x, V_r^x, A_r^x, b_r^x, W_z^x, V_z^x, A_z^x, b_z^x, W^x, V^x, A^x, and b^x are corresponding parameters of the x^{th} task.

The MLP cell can control information interaction of multiple sequences through the gate unit to obtain the most closely related information in each track, thereby improving the overall effect. The mathematical expression of the gate unit can be written as:

$$\begin{aligned} r_{d,n}^{x} &= \sigma\left(W_r^x u_{d,n}^x + V_r^x H_{d,n-1}^y + b_r^x\right) \\ z_{d,n}^{x} &= \sigma\left(W_z^x u_{d,n}^x + V_z^x H_{d,n-1}^y + b_z^x\right) \\ \overline{h_{d,n}^{x}} &= \sigma\left(W^x u_{d,n}^x + V^x \left[r_{d,n}^x H_{d,n-1}^y\right]\right) \\ h_{d,n}^{x} &= \left(1 - z_{d,n}^x\right) H_{d,n-1}^y + z_{d,n}^x \overline{h_{d,n}^x} \\ H_{d,n-1}^{y} &= \sigma\left(W^x \left[h_{d,n-1}^1, \cdots, h_{d,n-1}^N\right] + b^x\right) \end{aligned} \tag{7}$$

In the above equations, $H_{d,n-1}^y$ is the hidden state of the *n-1*$^{\text{th}}$ steps of the x task in the section d; $h_{d,n-1}^1, \cdots, h_{d,n-1}^N$ is the key information of other tasks obtained by the gate unit.

The constructed MICA model based on the multi-task learning is taken with the joint training of the loss function directly in this study to realize the multi-task learning. The sum of conditional probability items to optimize the multi-tasks through an encoder can be calculated with below equation:

$$P(\theta) = \arg\max_{\theta} \left(\sum_{D_n}\left(\frac{1}{N_p}\sum_{x}^{N_p}\log p\left(R_x^{D_n} \middle| U_x^{D_n}; \theta\right)\right)\right) \tag{8}$$

In the above equation, $\theta = \{\theta_{src}, \theta_{trgD_d} | D_n = 1, 2, \cdots, D_m\}$, m is the total number of tasks, θ_{src} is the parameter set of the original encoder, θ_{trgD_d} is the parameter set of the n^{th} track; and N_p is the size of the parallel training corpus on the p^{th} sequence.

The chord based the rhythm and melody cross-generation model is used to generate the melody $M = \{m_1, m_2, \cdots, m_{l_n}\}$, and the GRU is used to process the melody sequence to obtain the initial state of the multi-sequence decoder:

$$\overline{m_i} = E_n m_i \quad E_n \in R^{V_n * d} \tag{9}$$

$$s_{l_n}^n = GRU_n\left(\overline{m_{l_n}}, s_{l_n-1}^n\right) \tag{10}$$

In the above equation, $s_{l_n}^n$ is the initial state of the multi-sequence decoder.

The final output of the multi-sequence decoder is the track of the instrument, so it can be expressed as:

$$\begin{cases} s_y^x = CellOperation\left(u_{y-1}^x, s_{y-1}^x\right) \\ CellOperation \in \{AttentionCell, MLPCell\}, t > 0 \\ u_y^x = Softmax\left(s_y^x\right) \end{cases} \tag{11}$$

In the above equations, s_y^x is the hidden state of the x^{th} task in step y; the Softmax layer can obtain the track sequence of the x^{th} instrument; attention cell and MLP cell are used to maintain the harmonious state of all tracks.

In the final evaluation, harmony score (HS) is used to analyze the harmony between multiple tracks:

$$HS = \sum_{p=1}^{P} \delta\left(\bigcap_{k=1}^{K} C_p^k\right) \tag{12}$$

In the above equation, P is the number of sections of the generated music; K is the number of tracks of the generated music; C_p^k is the chord corresponding to the p^{th} section of the k^{th} track; and δ is defined as 0 or 1.

The note accuracy (NA) is used to analyze the probability that the generated note is the same as the real note:

$$NA = \sum_{n=1}^{N} \frac{e\left(v_n, \overline{v_n}\right)}{N} \tag{13}$$

In the above equation, v_n is the n^{th} source note; and $\overline{v_n}$ is the generated note.

The LS is used to analyze and measure the similarity between the length of generated note and the actual note:

$$LS = 1 - \frac{LevenshteinDistance}{N + \overline{N}} \tag{14}$$

In the above equation, $\overline{N}$ is the length of generated note, and N is the length of the source note.

The mean square error of notes distribution (NDMSE) is used to analyze the distribution difference between the generated note and the source note:

$$NDMSE = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}\left(\frac{y_i}{N} - \frac{\overline{y_i}}{\overline{N}}\right)}{MN} \tag{15}$$

In the above equation, M is the number of music categories, and N is the number of note categories.

Empty is used to evaluate the stability of the generated results. The lower the Empty score, the better the performance of the model generation.

C. IMPROVEMENT OF CONSENSUS ALGORITHM OF DIGITAL MUSIC COPYRIGHT PROTECTION SYSTEM BASED ON THE BLOCKCHAIN

Based on the structure of data, the blockchain is a chain-like data structure that connects a series of block information in the order of timestamps, where block is the most basic structure in blockchain [28]. In the blockchain layer, the commonly used technologies include Peer-to-Peer network, Merkle Tree, and consensus algorithm, etc. [29]. At present, the equation algorithm can no longer be well applied to the new blockchain environment, so the existing equation algorithm has to be improved. The practical byzantine fault tolerance (PBFT) algorithm can reduce the complexity of the original byzantine fault tolerance algorithm while improving the calculation efficiency without influence on the security of the algorithm [30], [31]. However, the PBFT algorithm needs to give a limited condition to each node of the copy. For example, the execution results of the nodes under the same state and parameter settings are the same, and all nodes have to start executing tasks from the same state.

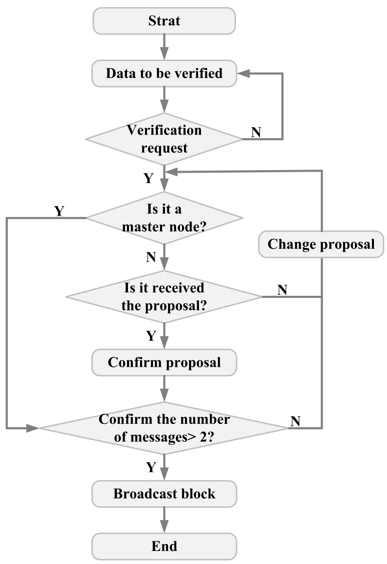

FIGURE 8. Calculating flowchart of IPBFT.

The IPBFT algorithm is proposed to enable the PBFT algorithm to solve the large throughput and achieve the fast confirmation speed. The calculation flow of the IPBFT algorithm is shown in Fig. 8, and it can be seen that the IPBFT algorithm includes a node change module, a configuration change module, and a new block generation module.

Then, the TPS is used to test the throughput of the IPBFT algorithm. The calculation equation of TPS is given as below:

$$TPS = \frac{Transactions}{\Delta} \tag{16}$$

In the above equation, *Transactions* is the total amount of data in the block; and Δ is the consensus interval.

D. DESIGN OF DIGITAL MUSIC COPYRIGHT PROTECTION SYSTEM BASED ON THE BLOCKCHAIN

The most important feature of applying the blockchain to the design of digital music copyright protection system is that the traditional storage method is replaced by distributed storage structure of the blockchain. Therefore, the basic structure of the digital music copyright protection system based on blockchain constructed in this study is shown in Fig. 9.

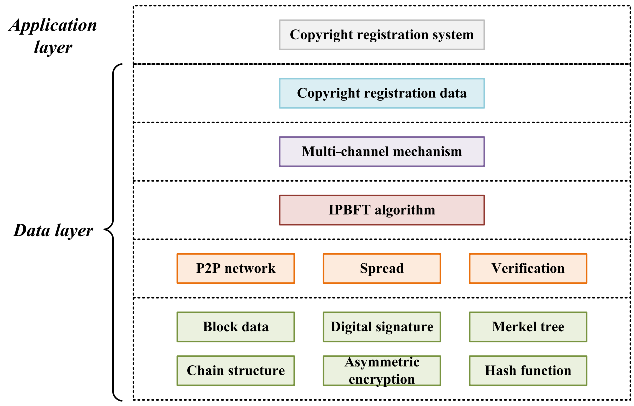

FIGURE 9. Basic structure of the digital music copyright protection system.

The application layer is the digital copyright registration business system, and the data layer is the data interaction and storage model of IPBFT algorithm and digital copyright registration.

In order to meet the demand, the digital music copyright protection system based on blockchain constructed in this study contains a total of four modules: user management, registration management, authority management, and log management, and specific sub-function modules are included under different modules. The function modules are shown in Fig. 10 as below.

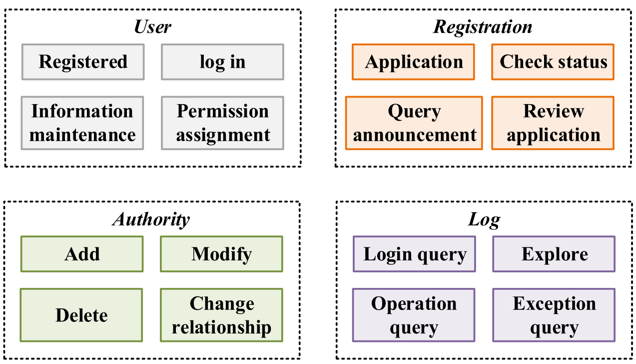

FIGURE 10. Basic functions of the digital music copyright protection system.

The IntelliJ IDEA 2017 development tools and Java are used to complete the development on Window7 64bit system, the framework used for the development is Spring Boot 1.5.9, and the container is Tomcat 8.

IV. RESULTS AND DISCUSSION

A. IMPLEMENTATION OF SINGLE MELODY COMPOSITION MODEL BASED ON DCGANS

The model is implemented on MXNet, Bach's public praise comes from music21 software, and the DCGANs model constructed in this article is mainly to solve the monophonic problem, so the selected 4 voices are all monophonic, with a total of 350 songs. All the music is transposed into the relative pitch range, and finally 2,500 melodies are obtained, which are then divided into a testing set and a training set at the ratio of 1:4.

All acyclic networks in Figure 2 are fully connected layers and the number of neurons in the hidden layer is 200. The activation function is ReLu, the recurrent neural network (RNN) is set to a 2-layer LSTM, and the number of neurons is 200.

In the discriminant model D, the number of CNN convolution kernels in the first layer is 64 and the step size is 2; the number of CNN convolution kernels in the second layer is 128 and the step size is 2; the activation functions of both these two CNN layers are Leaky ReLu, and the second layer of the CNN has to be processed by BN. The number of CNN convolution kernels in the third layer is 2, and the step size is 1. The last layer is Dense, and the number of neurons is 1.

Before the experiment, it is necessary to perform rearrangement and harmony processing of different parts, and then calculate the accuracy of the predicted notes. In the formal experiment, 150 hymns are selected, and the generative model G is used to sample a total of 2000 times for each part, and the sampling is performed every 100 times. The generated melody is compared with the real melody, and the accuracy of each note is calculated based on the statistic of note one by one. The average accuracy of these 150 hymns is used as the final result.

1) IMPACTS OF OPTIMIZING ALGORITHM WITH DIFFERENT PARAMETERS ON THE MODEL's PERFORMANCE

Firstly, impacts of optimizing algorithms with different parameters are compared in CNN parameter optimizing of the discriminant model and the prediction accuracy of melody (taking Soprano voice as an ample), and these algorithms include Gradient Descent (GD), Batch Gradient Descent (BGD), Stochastic Gradient Descent (SGD), Newton's Method (NM), Quasi-Newton Methods (QNM), Grid Search (GS), and Non-dominated Sorting Genetic Algorithm (NSGA).

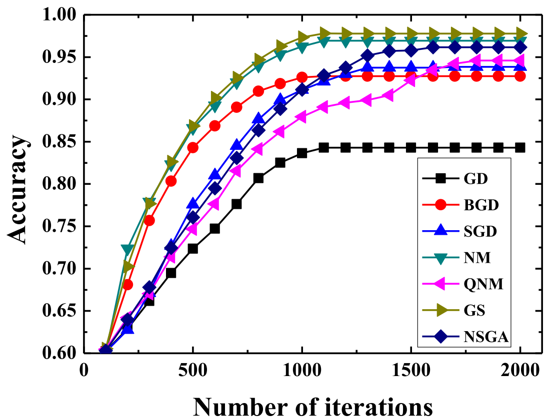

FIGURE 11. Impacts of optimizing algorithms with different parameters in prediction accuracy of DCGANs model.

Figure 11 indicates that in the parameter optimization of the constructed model by using the GS algorithm, the prediction accuracy of the constructed model is the highest (97.78%), so the GS algorithm is selected to optimize the model parameters.

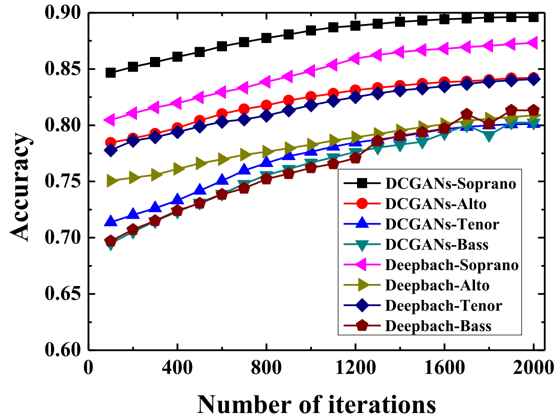

FIGURE 12. Predication accuracy of DCGANs model in different voice melodies.

2) THE PREDICTION ACCURACY OF DIFFERENT VOICE MELODIES BASED ON DCGANS

The prediction results of the DeepBatch model [32] are used as references to compare the accuracy of the DCGANs model constructed in this study in the Soprano, Alto, Tenor, and Bass melody. The results are shown in Fig. 12. As the number of iterations increases gradually, the prediction accuracy of different voice training also increases. The accuracy of the DCGANs model constructed in this study in Soprano and Alto voice melody prediction is significantly higher than that of the DeepBatch model by 2.29% and 3.32%, respectively. The accuracy of the DeepBatch model in Tenor and Bass melody prediction is slightly higher than that of the DCGANs model by 3.99% and 1.09%, respectively. In the melody prediction of Tenor and Bass parts, the accuracy of the DeepBatch model and the DCGANs model is not much different, and the prediction accuracy of both is higher than 80%. It shows that the DCGANs model constructed in this study has a good performance in the prediction of melody in different parts of monophonic music.

B. COMPOSITION GENERATION OF MICA MODEL BASED ON MULTI-TASK LEARNING

9000 digital music files are included in the training set and 2000 digital music files are included in the testing set. The latest multi-track generation model hierarchical recurrent neural network (HRNN) model [33] is taken as a control to compare the HS, NA, LS, NDMSE, and Empty. It can be seen from Table 1 that the MICA model has the highest HS, and is 25.1% higher than that of the HRNN model. This indicates that the MICA model can make full use of the useful information of other tasks, thereby improving the harmony of multi-track music composition. In addition, the MICA model has achieved the best results on NA, LS, and NDMSE, and the MICA model results are better than the results of

TABLE 1. Composition performance of MICA model based on the multi-task learning.

Melody		Drum	Bass	String	Guitar
HRNN	HC	0.211	0.326	0.355	0.354
	NA	0.263	0.062	0.018	0.046
	LS	0.275	0.281	0.182	0.178
	NDMSE	0.676	0.307	0.177	0.232
	Empty	0.375	0.365	0.579	0.275
MICA	HC	0.305	0.423	0.426	0.403
	NA	0.351	0.081	0.028	0.082
	LS	0.318	0.293	0.198	0.211
	NDMSE	0.112	0.098	0.032	0.182
	Empty	0.095	0.312	0.537	0.180

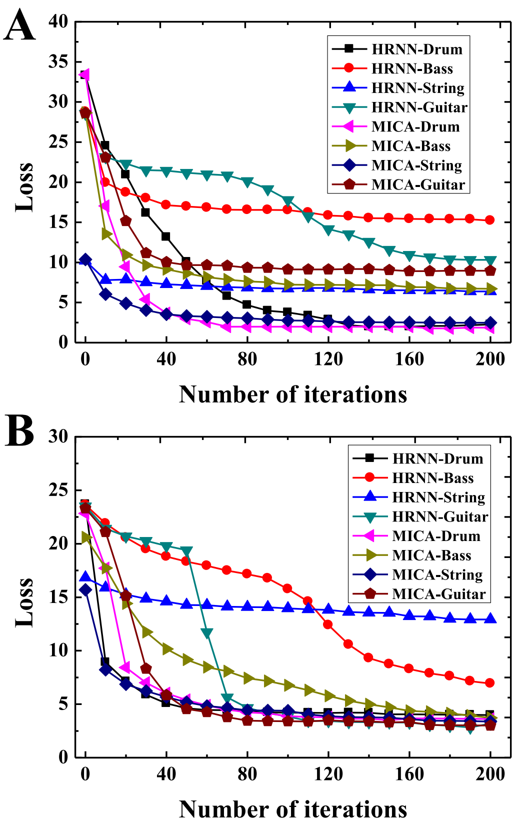

FIGURE 13. Analysis on multi-task learning effects of MICA model (A indicated the learning effect of the music sequence with common length; B gave the learning effect of the music sequence with long length).

HRNN model, which suggests that the MICA model has good composition performance and good stability.

The multi-task learning effects of different models are compared, and the results are shown in Fig. 13. As can be seen from Fig. 13A, the MICA model has a faster convergence speed on drum, guitar, bass, and string tasks, and the learning effect of the MICA model is significantly better than that of the HRNN model. The performance of different models in the learning of long-sequence tasks of drums, guitars, basses, and strings is compared, and the results can be seen from Fig. 13B. The MICA model constructed in this study is still better than the HRNN model. In addition, facing with complex learning tasks, the MICA model still has good performance.

C. SIMULATION TEST OF IMPROVED CONSENSUS ALGORITHM FOR DIGITAL MUSIC COPYRIGHT PROTECTION SYSTEM BASED ON THE BLOCKCHAIN

The IPBFT algorithm is implemented based on the Java language. A blockchain network consisting of 4 consensus nodes (1 master node and 3 backup nodes) and 2 ordinary nodes is built in a single machine environment, and data is formatted using the JSON data format transmission. 50 threads are used to send requests simultaneously to simulate the actual application environment, and the real registration data are selected for testing.

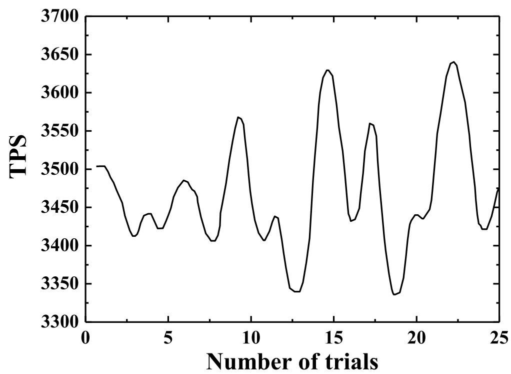

FIGURE 14. Test results of throughput of IPBFT algorithm.

It is assumed that $\Delta = 10s$, it is tested for 25 consecutive times, and each test contains a total of 20 consensus nodes. The performance of the IPBFT algorithm is evaluated by calculating the average throughput. The results are shown in Fig. 14. It can be seen that the IPBFT algorithm proposed in this study has an average TPS value of 3469 during the experimental test of registration data in a real music digital copyright protection system, and the throughput of the IPBFT algorithm is relatively stable during the test, so it can meet the actual application requirements [9].

The IPBFT algorithm proposed in this study is compared with other blockchain platforms, and the results are shown in Table 2. It can be seen that the TPS value of the proposed IPBFT algorithm is significantly higher than that of other blockchain platforms, so the IPBFT algorithm has higher performance and can avoid the state of low throughput to a certain extent, so it is more suitable for application in the digital copyrights protection system.

TABLE 2. Comparison on TPSs of different consensus algorithm.

Blockchain	Factom	Bitshare	Ripple	NXT	IPBFT
TPS	53	510	1108	1009	3469

D. TEST OF DIGITAL MUSIC COPYRIGHT PROTECTION SYSTEM BASED ON THE BLOCKCHAIN

Taking the digital copyright registration interface in the system as an example, it can be seen from Fig. 15 that the

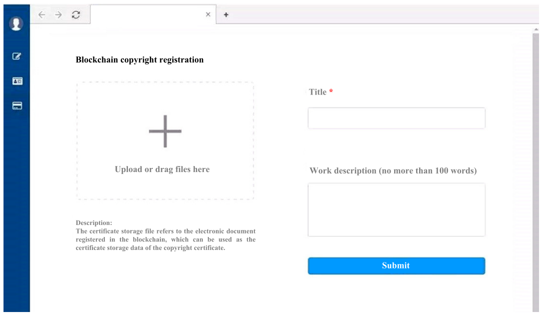

FIGURE 15. Digital copyright registration interface.

logged-in user can fill in and query the digital music copyright registration information through the system constructed in this study. At the same time, it is also possible to inquire about the published registration information.

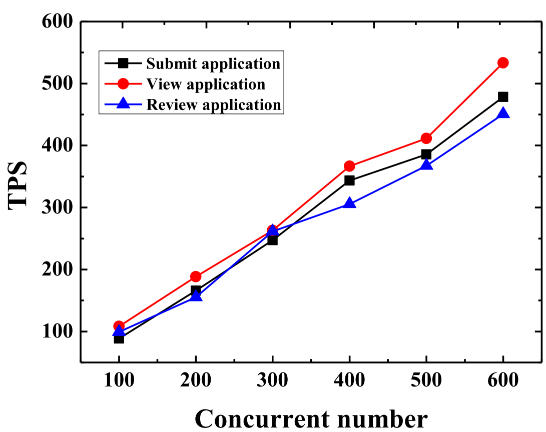

FIGURE 16. Digital copyright registration interface.

The Apache JMeter is used for test of the digital music copyright protection system based on the blockchain designed in this study under the Centos 6.9 operating system. TPS is still used for the performance test of the system, and the virtual users using 100, 200, 300, 400, 500, and 600 are compared for the system TPS test. The results are shown in Fig. 16. It can be seen that the error rate of virtual users under different numbers of submission, review, and check is 0%. As the number of virtual users continues to increase, the TPS value under submission, review, and check of applications also gradually increases. It shows that the digital music copyright protection system based on the blockchain proposed in this study can accurately complete the request under the concurrent operation of multiple users, and at the same time, it can also ensure a high throughput, which is consistent with the research results [34].

V. CONCLUSION

The DCGANs model constructed in this study can accurately complete the arrangement of Soprano and Alto in monophonic melody, with an accuracy rate of higher than 80%. The MICA model based on multi-tasking can efficiently complete the arrangement task of complex melodies. The melodies compiled by this model have high harmony, and the generated melodies have high similarity with the real notes. Finally, a digital music copyright protection system based on the blockchain is proposed. The system can still accurately complete the requests under the condition of multiple users operating concurrently, and ensure high throughput. However, this article only explores the performances of constructed system through simulation test, and the constructed system has to be applied in practice in future for long-term monitor of the digital music copyright registration system to protection effect and security. In short, the results of this study can provide a theoretical basis for improving the efficiency of algorithmic composition and the digital music copyright protection.

REFERENCES

[1] S. H. Didwania, "Copyright protection and cumulative creation: Evidence from early twentieth-century music," *J. Legal Stud.*, vol. 47, no. 2, pp. 235–268, Jun. 2018, doi: 10.1086/698923.

[2] C. Guo, J. Lu, Z. Tian W. Guo, and A. Darvishan, "Optimization of critical parameters of pem fuel cell using tlbo-de based on elman neural network," *Energy Convers. Manage.*, vol. 183, pp. 149–158, Mar. 2019, doi: 10.1016/j.enconman.2018.12.088.

[3] D. Herremans and C.-H. Chuan, "The emergence of deep learning: New opportunities for music and audio technologies," *Neural Comput. Appl.*, vol. 32, no. 4, pp. 913–914, Apr. 2019, doi: 10.1007/s00521-019-04166-0.

[4] E. Coutinho and B. Schuller, "Shared acoustic codes underlie emotional communication in music and speech—Evidence from deep transfer learning," *PLoS ONE*, vol. 12, no. 6, Jun. 2017, Art. no. e0179289, doi: 10.1371/journal.pone.0179289.

[5] R. P. Oliveira and T. F. Tavares, "Impact of algorithmic composition on player immersion in computer games: A case study using Markov chains," *Revista Música Hodie*, vol. 18, no. 1, pp. 61–73, Jun. 2018, doi: 10.5216/mh.v18i1.53571.

[6] B. Li, X. Liu, K. Dinesh, Z. Duan, and G. Sharma, "Creating a multitrack classical music performance dataset for multimodal music analysis: Challenges, insights, and applications," *IEEE Trans. Multimedia*, vol. 21, no. 2, pp. 522–535, Jul. 2018, doi: 10.1109/TMM.2018.2856090.

[7] Y. Wu and W. Li, "Automatic audio chord recognition with MIDI-trained deep feature and BLSTM-CRF sequence decoding model," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 27, no. 2, pp. 355–366, Feb. 2019, doi: 10.1109/TASLP.2018.2879399.

[8] B. G. Schultz, "The schultz MIDI benchmarking toolbox for MIDI interfaces, percussion pads, and sound cards," *Behav. Res. Methods*, vol. 51, no. 1, pp. 204–234, Apr. 2018, doi: 10.3758/s13428-018-1042-7.

[9] U. F. Ugwu, "Reconciling the right to learn with copyright protection in the digital age: Limitations of contemporary copyright treaties," *Law Develop. Rev.*, vol. 12, no. 1, pp. 41–77, May 2018, doi: 0.1515/ldr-2018-0007.

[10] V. Sunita, R. S. Dhavale, Deodhar, and L. M. Patnaik, "Walsh Hadamard transform based robust blind watermarking for digital audio copyright protection," *Commun. Comput. Inf. Sci.*, vol. 250, pp. 469–475, Jan. 2011, doi: 10.1007/978-3-642-25734-6_77.

[11] P. Murali and V. Sankaradass, "An efficient roi based copyright protection scheme for digital images with svd and orthogonal polynomials transformation," *Optik*, vol. 170, pp. 242–264, Oct. 2018, doi: 10.1016/j.ijleo.2018.04.050.

[12] S. Sinha Roy, A. Basu, and A. Chattopadhyay, "On the implementation of a copyright protection scheme using digital image watermarking," *Multimedia Tools Appl.*, vol. 79, nos. 19–20, pp. 13125–13138, Jan. 2020, doi: 10.1007/s11042-020-08652-9.

[13] Y. J. Wu, S.-C. Chen, and C.-I. Pan, "Entrepreneurship in the Internet age: Internet, entrepreneurs, and capital resources," *Int. J. Semantic Web Inf. Syst.*, vol. 15, no. 4, pp. 21–30, Oct. 2019, doi: 10.4018/IJSWIS.2019100102.

[14] W. Liang, D. Zhang, X. Lei, M. Tang, K.-C. Li, and A. Zomaya, "Circuit copyright blockchain: Blockchain-based homomorphic encryption for IP circuit protection," *IEEE Trans. Emerg. Topics Comput.*, early access, May 11, 2020, doi: 10.1109/TETC.2020.2993032.
[15] C. Kereliuk, B. L. Sturm, and J. Larsen, "Deep learning and music adversaries," *IEEE Trans. Multimedia*, vol. 17, no. 11, pp. 2059–2071, May 2015, doi: 10.1109/TMM.2015.2478068.
[16] K. Ohtani, K. Niwa, and K. Takeda, "Ai framework to arrange audio objects according to listener preferences," *Acoust. Soc. Amer. J.*, vol. 140, no. 4, p. 3060, Oct. 2016, doi: 10.1121/1.4969523.
[17] J. P. Briot and P. François, "Deep learning for music generation: Challenges and directions," *Neural Comput. Appl.*, vol. 32, no. 4, pp. 981–993, 2020, doi: 10.1007/s00521-018-3813-6.
[18] R. T. Dean and J. Forth. "Towards a deep improviser: A prototype deep learning post-tonal free music generator," *Neural Comput. Appl.*, vol. 32, pp. 969–979, Oct. 2018, doi: 10.1007/s00521-018-3765-x.
[19] H. Zhu, Q. Liu, N. J. Yuan, K. Zhang, G. Zhou, E. Chen, "Pop music generation," *ACM Trans. Knowl. Discovery Data*, vol. 14, no. 5, pp. 1–31, Jul. 2020, doi: 10.1145/3374915.
[20] L. Guo, X. Meng, "Digital content provision and optimal copyright protection," *Oper. Res.*, vol. 56, nos. 5–6, pp. 545–547, Apr. 2015, doi: 10.2139/ssrn.2263258.
[21] B. P. Devi, K. M. Singh, S. Roy, "A copyright protection scheme for digital images based on shuffled singular value decomposition and visual cryptography," *SpringerPlus*, vol. 5, no. 1, p. 1091, Jul. 2016, doi: 10.1186/s40064-016-2733-0.
[22] G. Liang, S. R. Weller, F. Luo, J. Zhao, and Z. Y. Dong, "Distributed blockchain-based data protection framework for modern power systems against cyber attacks," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 3162–3173, May 2019, doi: 10.1109/TSG.2018.2819663.
[23] K. T. Chui, R. W. Liu, M. Zhao, Patricia Ordóez De Pablos, "Predicting students' performance with school and family tutoring using generative adversarial network-based deep support vector machine," *IEEE Access*, vol. 8, pp. 86745–86752, 2020, doi: 10.1109/ACCESS.2020.2992869.
[24] A.-C. Guei and M. Akhloufi, "Deep learning enhancement of infrared face images using generative adversarial networks," *Appl. Opt.*, vol. 57, no. 18, p. D98, Jun. 2018, doi: 10.1364/AO.57.000D98.
[25] D. Lee, P. H. Seong, and J. Kim, "Autonomous operation algorithm for safety systems of nuclear power plants by using long-short term memory and function-based hierarchical framework," *Ann. Nucl. Energy*, vol. 119, pp. 287–299, Sep. 2018, doi: 10.1016/j.anucene.2018.05.020.
[26] H. Luo, W. Jiang, Y. Gu, F. Liu, X. Liao, S. Lai, and J. Gu, "A strong baseline and batch normalization neck for deep person re-identification," *IEEE Trans. Multimedia*, early access, Dec. 9,2019, doi: 10.1109/TMM.2019.2958756.
[27] R. Rajeev, J. A. Samath, and N. K. Karthikeyan, "An intelligent recurrent neural network with long short-term memory (LSTM) BASED batch normalization for medical image denoising," *J. Med. Syst.*, vol. 43, no. 8, p. 324, Aug. 2019, doi: 10.1007/s10916-019-1371-9.
[28] Z. Xiong, Y. Zhang, D. Niyato, P. Wang, and Z. Han, "When mobile blockchain meets edge computing," *IEEE Commun. Mag.*, vol. 56, no. 8, pp. 33–39, Aug. 2018, doi: 10.1109/MCOM.2018.1701095.
[29] H. Subramanian, "Decentralized blockchain-based electronic marketplaces," *Commun. ACM*, vol. 61, no. 1, pp. 78–84, Dec. 2017, doi: 10.1145/3158333.
[30] B. Yuan, H. Jin, D. Zou, L. T. Yang, and S. Yu, "A practical byzantine-based approach for faulty switch tolerance in software-defined networks," *IEEE Trans. Netw. Service Manage.*, vol. 15, no. 2, pp. 825–839, Jun. 2018, doi: 10.1109/TNSM.2018.2822668.
[31] Y. Wang, S. Cai, C. Lin, Z. Chen, T. Wang, Z. Gao, and C. Zhou, "Study of blockchains's consensus mechanism based on credit," *IEEE Access*, vol. 7, pp. 10224–10231, 2019, doi: 10.1109/ACCESS.2019.2891065.
[32] F. Colombo, A. Seeholzer, and W. Gerstner, "Deep artificial composer: A creative neural network model for automated melody generation," *Int. Evol. Biologically Inspired Music Art*, vol. 10198, pp. 81–96, Mar. 2017, doi: 10.1007/978-3-319-55750-2_6.
[33] J. Wu, C. Hu, Y. Wang, X. Hu, and J. Zhu, "A hierarchical recurrent neural network for symbolic melody generation," *IEEE Trans. Cybern.*, vol. 50, no. 6, pp. 2749–2757, Jun. 2020, doi: 10.1109/TCYB.2019.2953194.
[34] A. Savelyev, "Copyright in the blockchain era: Promises and challenges," *Comput. Law Secur. Rev.*, vol. 34, no. 3, pp. 550–561, Jun. 2018, doi: 10.1016/j.clsr.2017.11.008.